

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 006 444**

51 Int. Cl.:

C07D 405/04 (2006.01)
C07D 413/04 (2006.01)
C07D 417/04 (2006.01)
C07F 9/44 (2006.01)
C12Q 1/68 (2008.01)
C07F 9/6558 (2006.01)
C07F 9/6561 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **16.10.2018** **PCT/GB2018/052971**
87 Fecha y número de publicación internacional: **25.04.2019** **WO19077331**
96 Fecha de presentación y número de la solicitud europea: **16.10.2018** **E 18788878 (9)**
97 Fecha y número de publicación de la concesión europea: **04.12.2024** **EP 3697778**

54 Título: **Compuestos de cumarina sustituidos por aminas secundarias y sus usos como etiquetas fluorescentes**

30 Prioridad:

16.10.2017 GB 201716931

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
18.03.2025

73 Titular/es:

ILLUMINA CAMBRIDGE LIMITED (100.00%)
19 Granta Park Great Abington
Cambridge CB21 6DF, GB

72 Inventor/es:

ROMANOV, NIKOLAI NIKOLAEVICH;
MCCAULEY, PATRICK;
ANASTASI, CAROLE;
WU, XIAOLIN y
HYNES, NIALL

74 Agente/Representante:

DEL VALLE VALIENTE, Sonia

ES 3 006 444 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Compuestos de cumarina sustituidos por aminas secundarias y sus usos como etiquetas fluorescentes

- 5 La presente descripción se refiere a compuestos de cumarina sustituidos con amina secundaria y a su uso como marcadores fluorescentes. En particular, los compuestos pueden utilizarse como etiquetas fluorescentes para nucleótidos en aplicaciones de secuenciación de ácidos nucleicos.

Antecedentes

- 10 En esta solicitud se hace referencia a varias publicaciones y memorias de patente para describir con más detalle el estado de la técnica al que pertenece esta descripción.

- 15 La detección no radiactiva de ácidos nucleicos que llevan etiquetas fluorescentes es una tecnología importante en la biología molecular. Numerosos procedimientos empleados en la tecnología de ADN recombinante se basaron previamente en el uso de nucleótidos o polinucleótidos etiquetados radioactivamente con, por ejemplo, ^{32}P . Los compuestos radiactivos permiten la detección sensible de ácidos nucleicos y otras moléculas de interés. Sin embargo, existen graves limitaciones en el uso de isótopos radiactivos, tal como su gasto, su vida útil limitada, su sensibilidad insuficiente y, más importante, las razones de seguridad. La eliminación de la necesidad de etiquetas radiactivas reduce tanto los riesgos para la seguridad como el impacto medioambiental y los costes asociados, por ejemplo, con la eliminación de los reactivos. Los métodos susceptibles de detección fluorescente no radiactiva incluyen, a modo de ejemplos no limitantes, la secuenciación automatizada de ADN, los métodos de hibridación, la detección en tiempo real de productos de reacción de la cadena de la polimerasa, e inmunoensayos.

- 25 Para muchas aplicaciones es deseable emplear varias etiquetas fluorescentes espectralmente distinguibles para lograr la detección independiente de una pluralidad de analitos superpuestos espacialmente. En tales métodos multiplexados, la cantidad de recipientes puede reducirse para simplificar los protocolos experimentales y facilitar la producción de kits de reactivos específicos de la aplicación. En los sistemas de secuenciación de ADN automatizada multicolor, por ejemplo, la detección fluorescente multiplexada permite el análisis de múltiples bases de nucleótidos en una única línea de electroforesis, aumentando de este modo el rendimiento en comparación con los métodos de un solo color, y reduciendo las incertidumbres asociadas con las variaciones de movilidad electroforética entre líneas.

- 35 Sin embargo, la detección fluorescente multiplexada puede ser problemática y hay una serie de factores importantes que restringen la selección de las etiquetas fluorescentes apropiadas. En primer lugar, puede ser difícil encontrar compuestos de colorante con espectros de absorción y emisión resueltos adecuadamente en una aplicación determinada. Además, cuando se utilizan varios colorantes fluorescentes juntos, la generación de señales de fluorescencia en regiones espectrales distinguibles mediante excitación simultánea puede resultar complicada porque las bandas de absorción de los colorantes suelen estar ampliamente separadas, por lo que es difícil lograr eficiencias de excitación de fluorescencia comparables incluso para dos colorantes. Muchos métodos de excitación utilizan fuentes de luz de alta potencia, como el láser, y, por lo tanto, el colorante debe tener una fotostabilidad suficiente para resistir tal excitación. Una consideración final de particular importancia para los métodos de biología molecular es el grado de compatibilidad de los colorantes fluorescentes con los químicos de los reactivos, tales como por ejemplo, los disolventes y reactivos de síntesis de ADN, reguladores, enzimas polimerasas, y enzimas ligasas.

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

- 50 Las moléculas de colorante fluorescente con propiedades de fluorescencia mejoradas, tales como intensidad de fluorescencia adecuada, forma, y longitud de onda máxima de fluorescencia, pueden mejorar la velocidad y la precisión de la secuenciación de ácidos nucleicos. Las señales de fluorescencia intensas son especialmente importantes cuando las mediciones se realizan en reguladores biológicos a base de agua y a temperaturas más elevadas ya que las intensidades de fluorescencia de la mayoría de los colorantes son significativamente menores en tales condiciones. Además, la naturaleza de la base a la que se une un colorante también afecta el máximo de la fluorescencia, la intensidad de la fluorescencia, y a otras propiedades espectrales del colorante. Las interacciones específicas de la secuencia entre las nucleobases y los colorantes fluorescentes pueden adaptarse mediante el diseño específico de los colorantes fluorescentes. La optimización de la estructura de los colorantes fluorescentes puede mejorar la eficacia de la incorporación de nucleótidos, reducir el nivel de errores de secuenciación, y disminuir el uso de reactivos en la secuenciación de ácidos nucleicos y, por lo tanto, sus costes.

- 65 Algunos avances ópticos y técnicos ya han permitido mejorar considerablemente la calidad de la imagen, pero en última instancia se vieron limitados por la mala resolución óptica. En general, la resolución óptica de la microscopía óptica se limita a objetos espaciados aproximadamente a la mitad de la longitud de onda de la luz utilizada. En términos prácticos, entonces, solo los objetos que se encuentran bastante separados (al menos 200 a 350 nm) podrían resolverse mediante microscopía óptica. Una forma de mejorar la resolución de la imagen y aumentar la cantidad de

objetos que se pueden resolver por unidad de área de superficie es utilizar luz de excitación de longitud de onda más corta. Por ejemplo, si la longitud de onda de la luz se acorta en $\Delta\lambda \sim 100$ nm con la misma óptica, la resolución será mejor (alrededor de $\Delta 50$ nm/(alrededor del 15 %)), se grabarán imágenes menos distorsionadas, y la densidad de los objetos en el área reconocible aumentará aproximadamente un 35 %.

Ciertos métodos de secuenciación de ácidos nucleicos emplean luz láser para excitar y detectar nucleótidos etiquetados con colorante. Estos instrumentos utilizan luz de longitud de onda más larga, tales como láseres rojos, junto con colorantes apropiados que son excitables a 660 nm. Para detectar grupos de secuenciación de ácidos nucleicos más densamente empaquetados, mientras se mantiene una resolución útil, se puede utilizar una fuente de luz azul de longitud de onda más corta (450-460 nm). En este caso, la resolución óptica estará limitada no por la longitud de onda de emisión de los colorantes fluorescentes rojos de longitud de onda más larga, sino más bien por la emisión de colorantes excitables por la siguiente fuente de luz de longitud de onda más larga, por ejemplo, mediante un "láser verde" a 532 nm. Por lo tanto, existe la necesidad de etiquetas de colorante azul para utilizar en la detección de fluorescencia en aplicaciones de secuenciación.

Aunque la química de los colorantes azules y las tecnologías láser asociadas han mejorado, por ejemplo, para producir colorantes para discos DVD y Blu-ray, estos compuestos no son apropiados para el etiquetado biológico y no pueden utilizarse como biomarcadores.

Los colorantes azules conocidos se describen a continuación: Kamaldeep P y col. (2013) Synthesis of new conjugated coumarin-benzimidazole hybrids and their anticancer activity. Bioorganic & Medicinal Chemistry Letters, Pergamon, Amsterdam, NL 23(12), 3667-3672; US 2013/059321 A1; Srinivasan R y col. (1974) Anomalous fluorescence and laser emission from 7-alkylamino coumarins in acid solutions. Chemical Physical Letters 25(4), 537-540; CN 103 740 136 B; JP 2005 007851 A; JP 2002 278066 A; EP 0 730 007 A1; JP H07 43896 A; JP H05 263072 A, y; EP 0 608 737 A1.

Desafortunadamente, los colorantes azules disponibles comercialmente con una fuerte fluorescencia adecuados para el etiquetado de nucleótidos son todavía bastante raros. En la presente memoria se describen nuevos compuestos fluorescentes adecuados para el etiquetado de nucleótidos con fluorescencia fuerte bajo excitación con luz azul.

Resumen

La presente invención se refiere a derivados de cumarina sustituidos con amina secundaria. Los compuestos pueden utilizarse como etiquetas fluorescentes, particularmente para el etiquetado de nucleótidos en aplicaciones de secuenciación de los ácidos nucleicos. En algunos aspectos, los colorantes absorben la luz óptimamente a una longitud de onda de 450-460 nm y son particularmente ventajosos en situaciones donde se utilizan fuentes de excitación de longitud de onda azul que tienen una longitud de onda de 450-460 nm. La excitación de longitud de onda azul permite la detección y resolución de una mayor densidad de características por unidad de área debido a la longitud de onda más corta de la emisión de fluorescencia. Cuando tales colorantes se utilizan en conjugados con nucleótidos, se pueden observar mejoras en la longitud, intensidad, precisión, y calidad de las lecturas de secuenciación obtenidas durante los métodos de secuenciación de ácidos nucleicos.

La invención se describe en las reivindicaciones adjuntas.

Breve descripción de los dibujos

Las **figuras 1A, 1B y 1C** muestran gráficos de dispersión que demuestran la usabilidad de los nuevos colorantes descritos en la presente memoria en la secuenciación utilizando la química de secuenciación por síntesis (SBS) de Illumina. Los colorantes se detectan como cuatro áreas distintas resueltas espectralmente, como se describe en el ejemplo 5.

Descripción detallada

Esta descripción proporciona compuestos de cumarina sustituidos con aminas secundarias particularmente adecuados para los métodos de detección de fluorescencia y secuenciación por síntesis. Las realizaciones descritas en la presente memoria se refieren a colorantes y sus derivados de la estructura de fórmula (I).

Según un primer aspecto, la descripción proporciona compuestos de fórmula (I) o sales de los mismos.

Las limitaciones particulares para los diversos sustituyentes se muestran a continuación. Cada grupo individual se puede combinar con cualquier otra limitación individual a menos que se especifique lo contrario.

Para mejorar las propiedades fluorescentes de los biomarcadores y especialmente de sus bioconjugados en soluciones a base de agua, el compuesto de fórmula (I) es un compuesto en el que:

i) R^2 es $-\text{SO}_3\text{H}$; y/o

ii) R^4 es $-\text{SO}_3\text{H}$; y/o

iii) R^5 es $-\text{SO}_3\text{H}$ o $-\text{SO}_2\text{NH}_2$.

5 En algunos aspectos, X es O o S. En algunos aspectos, X es O. En algunos aspectos, X es S. En algunos aspectos, X es NR^n , donde R^n es H o alquilo C_{1-6} , y en algunos aspectos, R^n es H.

En algunos aspectos, R^3 es alquilo C_{2-6} sustituido con $-\text{CO}_2\text{H}$.

10 Cuando el acoplamiento a un enlazador o nucleótido es a través de R^3 , R^3 debe tener una longitud suficiente para permitir el acoplamiento a un grupo funcional unido al mismo. En algunos aspectos, R^3 no es $-\text{CH}_2\text{COOH}$ ni $-\text{CH}_2\text{COO}$.

Opcionalmente, R^3 es $-(\text{CH}_2)_n\text{COOH}$, donde n es 2-6. En algunos aspectos, n es 2, 3, 4, 5 o 6. En otros aspectos, n es 2 o 5. En algunos aspectos, n es 2. En algunos aspectos, n es 5.

15 Opcionalmente, R^3 es $-(\text{CH}_2)_n\text{SO}_3\text{H}$, donde n es 2-6. En algunos aspectos, n es 2, 3, 4, 5 o 6. En otros aspectos, n es 2 o 5. En algunos aspectos, n es 2. En algunos aspectos, n es 5.

20 El anillo de benceno del resto de indol está opcionalmente sustituido en cualquiera de las posiciones una, dos, tres, o cuatro por un sustituyente mostrado como R^5 . Cuando m es cero, el anillo de benceno no se sustituye. Cuando m es mayor que 1, cada R^5 puede ser igual o diferente. En algunos aspectos, m es 0. En otros aspectos, m es 1. En otros aspectos, m es 2. En algunos aspectos, m es 1, 2, o 3, y cada R^5 es independientemente halo, $-\text{CN}$, $-\text{CO}_2\text{H}$, amino, $-\text{OH}$, $-\text{SO}_3\text{H}$, o $-\text{SO}_2\text{NH}_2$. En algunos aspectos, R^5 es $-(\text{CH}_2)_x\text{COOH}$, donde x es 2-6. En algunos aspectos, x es 2, 3, 4, 5 o 6. En otros aspectos, x es 2 o 5. En algunos aspectos, x es 2. En algunos aspectos, x es 5.

25 En algunos aspectos, R^5 es halo, $-\text{CN}$, $-\text{CO}_2\text{H}$, $-\text{SO}_3\text{H}$, $-\text{SO}_2\text{NH}_2$, o alquilo C_{1-6} opcionalmente sustituido. En algunos aspectos, R^5 es halo, $-\text{CO}_2\text{H}$, $-\text{SO}_3\text{H}$, o $-\text{SO}_2\text{NH}_2$. En algunos aspectos, R^5 es alquilo C_{2-6} sustituido con $-\text{CO}_2\text{H}$, $-\text{SO}_3\text{H}$, o $-\text{SO}_2\text{NH}_2$. En algunos aspectos, cada R^5 es independientemente alquilo C_{1-6} opcionalmente sustituido, halo, $-\text{CN}$, $-\text{CO}_2\text{H}$, amino, $-\text{OH}$, $-\text{SO}_3\text{H}$, o $-\text{SO}_2\text{NH}_2$.

30 En algunos aspectos, R^1 es H. En algunos aspectos, R^1 es halo.

35 En algunos aspectos, R es H. En algunos aspectos, R es halo. En algunos aspectos, R es alquilo C_{1-6} . En algunos aspectos, R no es $-\text{CN}$. En algunos aspectos, R es H, halo, amino, $-\text{OH}$, C-amido, N-amido, $-\text{NO}_2$, $-\text{SO}_3\text{H}$, $-\text{SO}_2\text{NH}_2$, alquilo opcionalmente sustituido, alqueno opcionalmente sustituido, alquino opcionalmente sustituido, alcoxi opcionalmente sustituido, aminoalquilo opcionalmente sustituido, carbociclilo opcionalmente sustituido, heterociclilo opcionalmente sustituido, arilo opcionalmente sustituido, o heteroarilo opcionalmente sustituido.

40 En algunos aspectos, R^2 es H. En algunos aspectos, R^2 es opcionalmente alquilo sustituido. En algunos aspectos, R^2 es alquilo C_{1-4} opcionalmente sustituido con $-\text{CO}_2\text{H}$ o $-\text{SO}_3\text{H}$. En algunos aspectos, R^2 es $-\text{SO}_3\text{H}$. En algunos aspectos, R^2 es H, alquilo opcionalmente sustituido, alquilo C_{1-4} opcionalmente sustituido con $-\text{CO}_2\text{H}$ o $-\text{SO}_3\text{H}$, o $-\text{SO}_3\text{H}$. En algunos aspectos, R^2 es H o $-\text{SO}_3\text{H}$.

45 En algunos aspectos, R^4 es H. En algunos aspectos, R^4 es alquilo opcionalmente sustituido. En algunos aspectos, R^4 es alquilo C_{1-4} opcionalmente sustituido con $-\text{CO}_2\text{H}$ o $-\text{SO}_3\text{H}$. En algunos aspectos, R^4 es $-\text{SO}_3\text{H}$.

Los ejemplos particulares de un compuesto de fórmula (I) incluyen cuando X es O o S; R es H; R^1 es H; R^3 es $-(\text{CH}_2)_n\text{COOH}$, donde n es 2-6; R^5 es H, $-\text{SO}_3\text{H}$, o $-\text{SO}_2\text{NH}_2$; R^2 es H o $-\text{SO}_3\text{H}$; y R^4 es H o $-\text{SO}_3\text{H}$.

50 Los ejemplos particulares de un compuesto de fórmula (I) incluyen cuando X es O o S; R es H; R^1 es H; R^3 es $-(\text{CH}_2)_2\text{COOH}$; R^5 es H, $-\text{SO}_3\text{H}$, o $-\text{SO}_2\text{NH}_2$; R^2 es H o $-\text{SO}_3\text{H}$; y R^4 es H o $-\text{SO}_3\text{H}$.

Los ejemplos particulares de un compuesto de fórmula (I) incluyen cuando X es O o S; R es H; R^1 es H; R^3 es $-(\text{CH}_2)_5\text{COOH}$; R^5 es H, $-\text{SO}_3\text{H}$, o $-\text{SO}_2\text{NH}_2$; R^2 es H o $-\text{SO}_3\text{H}$; y R^4 es H o $-\text{SO}_3\text{H}$.

55 En algunos aspectos de fórmula (II), X' es O. En algunos aspectos, X' es S. En algunos aspectos, X' es NR^p , donde R^p es H o alquilo C_{1-6} . En algunos aspectos, X' es NR^p , donde R^p es H.

En algunos aspectos, R^6 es H. En algunos aspectos, R^6 es alquilo C_{1-4} .

60 En algunos aspectos, R^7 es H. En algunos aspectos, R^7 es alquilo C_{1-4} , $-\text{CO}_2\text{H}$, $-\text{SO}_3\text{H}$, $-\text{SO}_2\text{NH}_2$, $-\text{SO}_2\text{NH}(\text{alquilo } \text{C}_{1-4})$, o $-\text{SO}_2\text{N}(\text{alquilo } \text{C}_{1-4})_2$. En algunos aspectos, R^7 es alquilo C_{1-4} opcionalmente sustituido con $-\text{CO}_2\text{H}$.

En algunos aspectos, R^8 es H. En algunos aspectos, R^8 es $-\text{CO}_2\text{H}$, $-\text{SO}_3\text{H}$, o $-\text{SO}_2\text{NH}_2$. En algunos aspectos, R^8 es $-\text{SO}_3\text{H}$.

65 En algunos aspectos, R^{10} es H. En algunos aspectos, R^{10} es $-\text{CO}_2\text{H}$, $-\text{SO}_3\text{H}$, o $-\text{SO}_2\text{NH}_2$. En algunos aspectos, R^{10} es $-\text{SO}_3\text{H}$. En algunos aspectos, R^8 es H y R^{10} es $-\text{SO}_3\text{H}$. En algunos aspectos, R^8 es $-\text{SO}_3\text{H}$ y R^{10} es H.

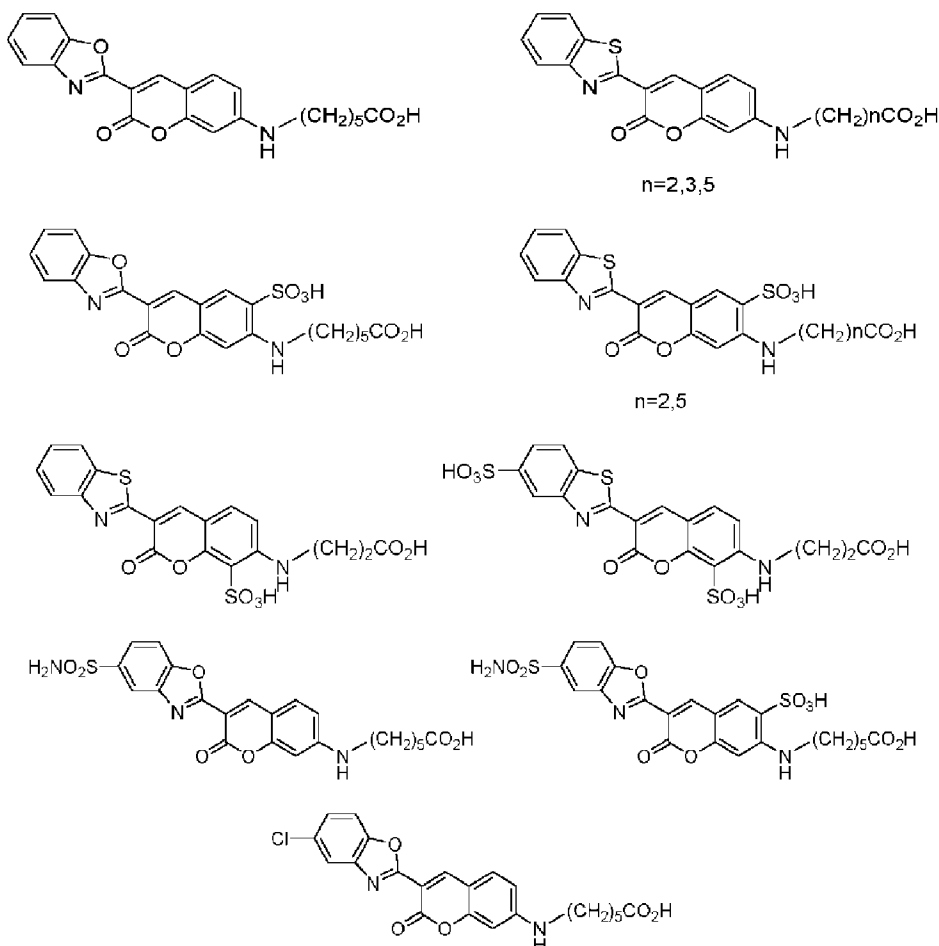
En algunos aspectos, uno de R^8 y R^{10} es H, halo, -CN, -CO₂H, amino, -OH, -SO₃H, -SO₂NH₂, -SO₂NH(alquilo C₁₋₄), -SO₂N(alquilo C₁₋₄)₂, alquilo C₁₋₆ opcionalmente sustituido, alqueno C₁₋₆ opcionalmente sustituido, alquino C₂₋₆ opcionalmente sustituido, o alcoxi C₁₋₆ opcionalmente sustituido, y el otro de R^8 y R^{10} se toma con R^9 para formar un anillo heterocíclico de 4 a 7 miembros opcionalmente sustituido.

En algunos aspectos, R^9 es alquilo C₂₋₆ sustituido con -CO₂H, -CO₂alquilo C₁₋₄, -CONH₂, -CONH(alquilo C₁₋₄), -CON(alquilo C₁₋₄)₂, -CN, -SO₃H, -SO₂NH₂, -SO₂NH(alquilo C₁₋₄), o -SO₂N(alquilo C₁₋₄)₂. En algunos aspectos, R^9 es -(CH₂)_y-CO₂H, donde y es 2, 3, 4, o 5.

En algunos aspectos, cada R^{11} es independientemente halo -CO₂H, -SO₃H, -SO₂NH₂, -SO₂NH(alquilo C₁₋₄), -SO₂N(alquilo C₁₋₄)₂, o alquilo opcionalmente sustituido. En otros aspectos, cada R^{11} es independientemente halo, -CO₂H, -SO₃H o -SO₂NH₂.

En algunos aspectos, q es 0. En otros aspectos, q es 1. En otros aspectos más, q es 2.

Los ejemplos específicos de colorantes de cumarina sustituidos con amina secundaria incluyen:



y sales del mismo.

Un compuesto particularmente útil es un nucleótido u oligonucleótido etiquetado con un colorante tal como se describe en la presente memoria. El nucleótido u oligonucleótido etiquetado puede tener la etiqueta unida al átomo de nitrógeno de la molécula de cumarina mediante un grupo alquilcarboxi para formar una alquilamida. El nucleótido u oligonucleótido etiquetado puede tener la etiqueta unida a la posición C5 de una base de pirimidina o la posición C7 de una base de 7-deaza purina a través de un resto enlazador.

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

En la presente memoria se proporcionan kits que incluyen dos o más nucleótidos en donde al menos un nucleótido es un nucleótido etiquetado con un compuesto de la presente descripción. El kit puede incluir dos o más nucleótidos etiquetados. Los nucleótidos pueden etiquetarse con dos o más etiquetas fluorescentes. Dos o más de las etiquetas pueden excitarse utilizando una única fuente de excitación, que puede ser un láser. Por ejemplo, las bandas de excitación para las dos o más etiquetas pueden superponerse al menos parcialmente, de modo que la excitación en la región de superposición del espectro haga que ambas etiquetas emitan fluorescencia. En realizaciones particulares, la emisión a partir de las dos o más etiquetas se producirá en diferentes regiones del espectro, de modo que la presencia de al menos una de las etiquetas se puede determinar distinguiendo ópticamente la emisión.

El kit puede contener cuatro nucleótidos etiquetados, donde el primero de los cuatro nucleótidos está etiquetado con un compuesto tal como se describe en la presente memoria. En tal kit, cada uno de los cuatro nucleótidos se puede etiquetar con un compuesto que es igual o diferente de la etiqueta de los otros tres nucleótidos. Por lo tanto, uno o más de los compuestos pueden tener un máximo de absorción y/o un máximo de emisión distintos, de modo que el (los) compuesto(s) sea(n) distinguible(s) de otros compuestos. Por ejemplo, cada compuesto puede tener un máximo de absorción y/o un máximo de emisión distintos, de modo que cada uno de los compuestos sea distinguible de los otros tres compuestos. Se entenderá que partes del espectro de absorción y/o espectro de emisión distintas de los máximos pueden diferir y estas diferencias pueden aprovecharse para distinguir los compuestos. El kit puede ser tal que dos o más de los compuestos tengan un máximo de absorción distinto. Los compuestos de la invención absorben típicamente la luz en la región por debajo de 500 nm.

Los compuestos, nucleótidos, o kits que se exponen en la presente memoria pueden usarse para detectar, medir o identificar un sistema biológico (incluidos, por ejemplo, los procesos o componentes del mismo). Las técnicas ilustrativas que pueden emplear los compuestos, nucleótidos o kits incluyen la secuenciación, el análisis de expresión, el análisis de hibridación, el análisis genético, el análisis de ARN, el ensayo celular (p. ej., el análisis de unión celular o función celular) o el ensayo de proteínas (p. ej., el ensayo de unión a proteínas o el ensayo de actividad de las proteínas). El uso puede ser en un instrumento automatizado para llevar a cabo una técnica particular, tal como un instrumento de secuenciación automática. El instrumento de secuenciación puede contener dos láseres que funcionan a diferentes longitudes de onda.

En la presente memoria se describen métodos para sintetizar los compuestos de la descripción. Los colorantes según la presente descripción pueden sintetizarse a partir de una variedad de diferentes materiales de partida adecuados. Los métodos para preparar colorantes de cumarina son bien conocidos en la técnica.

Definiciones

Los encabezados de secciones que se utilizan en la presente memoria tienen únicamente fines organizativos y no deben interpretarse como una limitación de la materia descrita.

Cabe señalar que, como se utilizan en esta especificación y las reivindicaciones adjuntas, las formas singulares “un”, “uno”, “una”, “el” y “la” incluyen referentes plurales a menos que se limiten expresa e inequívocamente a un referente.

A menos que se defina de cualquier otra manera, todos los términos técnicos y científicos utilizados en la presente memoria tienen el significado que entiende comúnmente un experto en la técnica. Además, el uso del término “incluyendo”, así como otras formas, como “incluyen”, “incluye” e “incluido”, no es limitativo. Además, el uso del término “que tiene”, así como otras formas, como “tienen”, “tiene” y “tenido”, no es limitativo. Como se utiliza en la presente memoria descriptiva, ya sea en una expresión de transición o en el cuerpo de la reivindicación, se ha de interpretar que las expresiones “que consta de” y “que constan de” tienen un significado abierto. Es decir, las expresiones se han de interpretar como sinónimos de las expresiones “que tiene(n) al menos” o “que incluye(n) al menos”. Por ejemplo, cuando se utiliza en el contexto de un proceso, la expresión “que comprende” significa que el proceso incluye al menos las etapas citadas, pero puede incluir etapas adicionales. Cuando se utiliza en el contexto de un compuesto, composición o dispositivo, la expresión “que consta de” significa que el compuesto, la composición o el dispositivo incluye al menos las características o los componentes citados, pero que también puede incluir características o componentes adicionales.

Como se utiliza en la presente memoria, la expresión “unido covalentemente” o “enlazado covalentemente” se refiere a la conformación de un enlace químico que se caracteriza por compartir pares de electrones entre átomos. Por ejemplo, una “lámina de polímero unida covalentemente”, cuando se utiliza en referencia a la superficie de un sustrato, se refiere a una lámina de polímero que forma enlaces químicos con la superficie funcionalizada de un sustrato, en comparación con la unión a la superficie a través de otros medios, por ejemplo, la adhesión o la interacción electrostática. Se apreciará que los polímeros que se unen covalentemente a una superficie también se pueden unir a través de distintos medios, además de la unión covalente.

El término “halógeno” o “halo”, como se utiliza en la presente memoria, significa uno cualquiera de los átomos radioestables de la columna 7 de la Tabla periódica de los elementos, p. ej., flúor, cloro, bromo o yodo, prefiriéndose el flúor y el cloro.

Como se utiliza en la presente memoria, “alquilo” se refiere a una cadena de hidrocarburos recta o ramificada que está completamente saturada (es decir, que no contiene enlaces dobles ni triples). El grupo alquilo puede tener de 1 a 20 átomos de carbono (siempre que aparezca en la presente memoria, un intervalo numérico tal como “de 1 a 20” se refiere a cada número entero en el intervalo dado; *p. ej.*, “de 1 a 20 átomos de carbono” significa que el grupo alquilo puede consistir en 1 átomo de carbono, 2 átomos de carbono, 3 átomos de carbono, *etc.*, hasta e inclusive 20 átomos de carbono, aunque la presente definición también cubre la aparición del término “alquilo” donde no se designa un intervalo numérico). El grupo alquilo también puede ser un alquilo de tamaño medio que tenga de 1 a 9 átomos de carbono. El grupo alquilo podría ser un alquilo inferior que tenga de 1 a 6 átomos de carbono. El grupo alquilo puede designarse como “alquilo C₁₋₄” o designaciones similares. Únicamente a modo de ejemplo, “alquilo C₁₋₆” indica que hay de uno a cuatro átomos de carbono en la cadena alquilo, es decir, la cadena alquilo se selecciona en el grupo que consiste en metilo, etilo, propilo, iso-propilo, n-butilo, iso-butilo, sec-butilo y t-butilo. Los grupos alquilo típicos incluyen, aunque no de forma limitativa de ninguna manera, metilo, etilo, propilo, isopropilo, butilo, isobutilo, butilo terciario, pentilo, hexilo y similares.

Como se emplea en la presente memoria, “alcoxi” se refiere a la fórmula -OR en donde R es un alquilo como se definió anteriormente, tal como “alcoxi C₁₋₉”, que incluye, entre otros, metoxi, etoxi, n-propoxi, 1-metiletoxí (isopropoxi), n-butoxi, iso-butoxi, sec-butoxi y terc-butoxi, y similares.

Como se utiliza en la presente memoria, “alquenilo” se refiere a una cadena de hidrocarburos recta o ramificada que contiene uno o más enlaces dobles. El grupo alquenilo puede tener de 2 a 20 átomos de carbono, aunque la presente definición también abarca la aparición del término “alquenilo” cuando no se designa ningún intervalo numérico. El grupo alquenilo también puede ser un alquenilo de tamaño medio que tenga de 2 a 9 átomos de carbono. El grupo alquenilo podría ser un alquenilo inferior que tenga de 2 a 6 átomos de carbono. El grupo alquenilo puede designarse como “alquenilo C₂₋₆” o designaciones similares. Únicamente a modo de ejemplo, “alquenilo C₂₋₆” indica que hay de dos a seis átomos de carbono en la cadena de alquenilo, es decir, la cadena de alquenilo se selecciona en el grupo que consta de etenilo, propen-1-ilo, propen-2-ilo, propen-3-ilo, buten-1-ilo, buten-2-ilo, buten-3-ilo, buten-4-ilo, 1-metil-propen-1-ilo, 2-metil-propen-1-ilo, 1-etil-eten-1-ilo, 2-metil-propen-3-ilo, buta-1,3-dienilo, buta-1,2-dienilo, y buta-1,2-dien-4-ilo. Los grupos alquenilo típicos incluyen, aunque no de forma limitativa de ninguna manera, etenilo, propenilo, butenilo, pentenilo y hexenilo, y similares.

Como se utiliza en la presente memoria, “alquinilo” se refiere a una cadena de hidrocarburos recta o ramificada que contiene uno o más enlaces triples. El grupo alquinilo puede tener de 2 a 20 átomos de carbono, aunque la presente definición también abarca la aparición del término “alquinilo” cuando no se designa ningún intervalo numérico. El grupo alquinilo también puede ser un alquinilo de tamaño medio que tenga de 2 a 9 átomos de carbono. El grupo alquinilo podría ser un alquinilo inferior que tenga de 2 a 6 átomos de carbono. El grupo alquinilo puede designarse como “alquinilo C₂₋₆” o designaciones similares. Únicamente a modo de ejemplo, “alquinilo C₂₋₆” indica que hay de dos a seis átomos de carbono en la cadena de alquinilo, es decir, la cadena de alquinilo se selecciona en el grupo que consiste en etinilo, propin-1-ilo, propin-2-ilo, butin-1-ilo, butin-3-ilo, butin-4-ilo, y 2-butinilo. Los grupos alquinilo típicos incluyen, aunque no de forma limitativa de ninguna manera, etinilo, propinilo, butinilo, pentinilo y hexinilo, y similares.

Como se emplea en la presente memoria, “heteroalquilo” se refiere a una cadena de hidrocarburo recta o ramificada que contiene uno o más heteroátomos, es decir, un elemento distinto del carbono, que incluye, entre otros, nitrógeno, oxígeno y azufre, en la estructura molecular básica de la cadena. El grupo heteroalquilo puede tener de 1 a 20 átomos de carbono, aunque la presente definición también abarca la aparición del término “heteroalquilo” cuando no se designa ningún intervalo numérico. El grupo heteroalquilo también puede ser un heteroalquilo de tamaño medio que tenga de 1 a 9 átomos de carbono. El grupo heteroalquilo podría ser un heteroalquilo inferior que tenga de 1 a 6 átomos de carbono. El grupo heteroalquilo puede designarse como “heteroalquilo C₁₋₆” o designaciones similares. El grupo heteroalquilo puede contener uno o más heteroátomos. Únicamente a modo de ejemplo, “heteroalquilo C₄₋₆” indica que hay de cuatro a seis átomos de carbono en la cadena de heteroalquilo y adicionalmente uno o más heteroátomos en la estructura molecular básica de la cadena.

El término “aromático” se refiere a un anillo o sistema de anillos que tiene un sistema de electrones pi conjugado e incluye grupos aromáticos carbocíclicos (por ejemplo, fenilo) y heterocíclicos (por ejemplo, piridina). El término incluye grupos monocíclicos o policíclicos de anillos condensados (es decir, anillos que comparten pares de átomos adyacentes), siempre que todo el sistema de anillos sea aromático.

Como se utiliza en la presente memoria, “arilo” se refiere a un anillo o sistema de anillos aromáticos (es decir, dos o más anillos fusionados que comparten dos átomos de carbono adyacentes) que solo contienen carbono en la estructura molecular básica del anillo. Cuando el arilo es un sistema de anillos, cada anillo del sistema es aromático. El grupo arilo puede tener de 6 a 18 átomos de carbono, aunque la presente definición también abarca la aparición del término “arilo” cuando no se designa ningún intervalo numérico. En algunas realizaciones, el grupo arilo tiene de 6 a 10 átomos de carbono. El grupo arilo puede designarse como “arilo C₆₋₁₀”, “arilo C₆” o “arilo C₁₀” o designaciones similares. Los ejemplos de grupos arilo incluyen, entre otros, fenilo, naftilo, fluorenilo, azulenilo, antrilo, fenantrilo, pirenilo, bifenilo y terfenilo.

Un “aralquilo” o “arilalquilo” es un grupo arilo conectado, como sustituyente, a través de un grupo alquileo, tal como “aralquilo C₇₋₁₄” y similares, que incluyen, entre otros, bencilo, 2-feniletilo, 3-fenilpropilo y naftilalquilo. En algunos casos, el grupo alquileo es un grupo alquileo inferior (es decir, un grupo alquileo C₁₋₆).

Como se utiliza en la presente memoria, “heteroarilo” se refiere a un anillo o sistema de anillos aromáticos (es decir, dos o más anillos fusionados que comparten dos átomos adyacentes) que contiene/n uno o más heteroátomos, es decir, un elemento que no es carbono, incluidos, entre otros, nitrógeno, oxígeno y azufre, en la estructura molecular básica del anillo. Cuando el heteroarilo es un sistema de anillos, cada anillo en el sistema es aromático. El grupo heteroarilo puede tener 5-18 miembros de anillo (es decir, el número de átomos que forman la estructura molecular básica del anillo, que incluye átomos de carbono y heteroátomos), aunque la presente definición también abarca la aparición del término “heteroarilo” cuando no se designa ningún rango numérico. En algunas realizaciones, el grupo heteroarilo tiene 5 a 10 miembros del anillo o 5 a 7 miembros del anillo. El grupo heteroarilo puede designarse como “heteroarilo de 5-7 miembros”, “heteroarilo de 5-10 miembros” o designaciones similares. Los ejemplos de anillos de heteroarilo incluyen, entre otros, furilo, tienilo, ftalazinilo, pirrolilo, oxazolilo, tiazolilo, imidazolilo, pirazolilo, isoxazolilo, isotiazolilo, triazolilo, tiadiazolilo, piridinilo, piridazinilo, pirimidinilo, pirazinilo, quinolinilo, isoquinilo, bencimidazolilo, benzoxazolilo, benzotiazolilo, indolilo, isoindolilo y benzotienilo.

Un “heteroaralquilo” o “heteroarilalquilo” es un grupo heteroarilo conectado, como un sustituyente, mediante un grupo alquileo. Los ejemplos incluyen, entre otros, 2-tienilmetilo, 3-tienilmetilo, furilmetilo, tieniletilo, pirrolilalquilo, piridilalquilo, isoxazolilalquilo e imidazolilalquilo. En algunos casos, el grupo alquileo es un grupo alquileo inferior (es decir, un grupo alquileo C₁₋₆).

Como se utiliza en la presente memoria, “carbociclilo” hace referencia un anillo o sistema de anillos cíclicos no aromático/s que solo contiene átomos de carbono en la estructura molecular básica del sistema de anillo. Cuando el carbociclilo es un sistema de anillo, se pueden unir dos o más anillos de forma fusionada, por puentes o conectada en espiral. Los carbociclicos pueden tener cualquier grado de saturación, siempre que al menos un anillo de un sistema de anillos no sea aromático. Por lo tanto, los carbociclicos incluyen cicloalquilos, cicloalquenos y cicloalquinos. El grupo carbociclilo puede tener de 3 a 20 átomos de carbono, aunque la presente definición también abarca la aparición del término “carbociclilo” cuando no se designa ningún intervalo numérico. El grupo carbociclilo también puede ser un carbociclilo de tamaño medio que tenga de 3 a 10 átomos de carbono. El grupo carbociclilo podría ser un carbociclilo que tenga de 3 a 6 átomos de carbono. El grupo carbociclilo puede designarse como “carbociclilo C₃₋₆” o designaciones similares. Los ejemplos de anillos de carbociclilo incluyen, entre otros, ciclopropilo, ciclobutilo, ciclopentilo, ciclohexilo, ciclohexenilo, 2,3-dihidro-indeno, biciclo[2.2.2]octanilo, adamantilo y espiro[4.4]nonanilo.

Como se utiliza en la presente memoria, “cicloalquilo” significa un anillo o sistema de anillos carbociclilo completamente saturado. Los ejemplos incluyen ciclopropilo, ciclobutilo, ciclopentilo y ciclohexilo.

Como se utiliza en la presente memoria, “heterociclilo” hace referencia a un anillo o sistema de anillos cíclico no aromático que contiene al menos un heteroátomo en la estructura molecular básica del anillo. Los heterociclicos pueden unirse entre sí de forma fusionada, por puentes o conectada a espiro. Los heterociclicos pueden tener cualquier grado de saturación siempre que al menos un anillo en el sistema de anillos no sea aromático. El/los heteroátomo/s puede/n estar presente/s en un anillo aromático o no aromático dentro del sistema de anillos. El grupo heterociclilo puede tener 3 a 20 miembros del anillo (es decir, el número de átomos que forman la estructura molecular básica del anillo, que incluye átomos de carbono y heteroátomos), aunque la presente definición también abarca la aparición del término “heterociclilo” cuando no se designa ningún rango numérico. El grupo heterociclilo también puede ser un heterociclilo de tamaño medio que tenga de 3 a 10 miembros del anillo. El grupo heterociclilo también podría ser un heterociclilo que tenga de 3 a 6 miembros del anillo. El grupo heterociclilo puede designarse como “heterociclilo de 3-6 miembros” o designaciones similares. En los heterociclicos monocíclicos preferidos de seis miembros, el heteroátomo o los heteroátomos se seleccionan en uno hasta tres de O, N o S, y entre los heterociclicos monocíclicos preferidos de cinco miembros, el heteroátomo o los heteroátomos se seleccionan entre uno o dos heteroátomos seleccionados de O, N o S. Los ejemplos de anillos heterociclicos incluyen, entre otros, azepinilo, acridinilo, carbazolilo, cinnolinilo, dioxolanilo, imidazolinilo, imidazolidinilo, morfolinilo, oxiranilo, oxeanilo, tiepanilo, piperidinilo, piperazinilo, dioxoperazinilo, pirrolidinilo, pirrolidonilo, pirrolidionilo, 4-piperidonilo, pirazolinilo, pirazolidinilo, 1,3-dioxinilo, 1,3-dioxanilo, 1,4-dioxinilo, 1,4-dioxanilo, 1,3-oxatiano, 1,4-oxatiinilo, 1,4-oxatiano, 2H-1,2-oxazinilo, trioxanilo, hexahidro-1,3,5-triazinilo, 1,3-dioxolilo, 1,3-dioxolanilo, 1,3-ditiolilo, 1,3-ditioanilo, isoxazolinilo, isoxazolidinilo, oxazolinilo, oxazolidinilo, oxazolidinonilo, tiazolinilo, tiazolidinilo, 1,3-oxatiolanilo, indolinilo, isoindolinilo, tetrahidrofurano, tetrahidropirano, tetrahidrotiofeno, tetrahidrotiopirano, tetrahidro-1,4-tiazinilo, tiamorfolinilo, dihidrobenzofuranilo, bencimidazolidinilo y tetrahidroquinolina.

Un grupo “O-carboxi” se refiere a un grupo “-OC(=O)R” en el que R se selecciona entre hidrógeno, alquilo C₁₋₆, alqueno C₂₋₆, alquino C₂₋₆, carbociclilo C₃₋₇, arilo C₆₋₁₀, heteroarilo de 5-10 miembros y heterociclilo de 3-10 miembros, como se define en la presente memoria.

Un grupo “C-carboxi” se refiere a un grupo “-C(=O)OR” en el que R se selecciona entre hidrógeno, alquilo C₁₋₆, alqueno C₂₋₆, alquino C₂₋₆, carbociclilo C₃₋₇, arilo C₆₋₁₀, heteroarilo de 5-10 miembros y heterociclilo de 3-10 miembros, como se define en la presente memoria. Un ejemplo no limitante incluye el carboxilo (es decir, -C(=O)OH).

Un grupo "sulfonilo" se refiere a un grupo " $-\text{SO}_2\text{R}$ " en el que R se selecciona entre hidrógeno, alquilo C₁₋₆, alquenilo C₂₋₆, alquinilo C₂₋₆, carbociclilo C₃₋₇, arilo C₆₋₁₀, heteroarilo de 5-10 miembros y heterociclilo de 3-10 miembros, como se define en la presente memoria.

Un grupo "sulfino" se refiere a un grupo " $-\text{S}(=\text{O})\text{OH}$ ".

Un grupo "S-sulfonamido" se refiere a un grupo " $-\text{SO}_2\text{NR}_\text{A}\text{R}_\text{B}$ " en el que R_A y R_B se seleccionan cada uno independientemente entre hidrógeno, alquilo C₁₋₆, alquenilo C₂₋₆, alquinilo C₂₋₆, carbociclilo C₃₋₇, arilo C₆₋₁₀, heteroarilo de 5-10 miembros y heterociclilo de 3-10 miembros, como se define en la presente memoria.

Un grupo "N-sulfonamido" se refiere a un grupo " $-\text{N}(\text{R}_\text{A})\text{SO}_2\text{R}_\text{B}$ " en el que R_A y R_B se seleccionan cada uno independientemente entre hidrógeno, alquilo C₁₋₆, alquenilo C₂₋₆, alquinilo C₂₋₆, carbociclilo C₃₋₇, arilo C₆₋₁₀, heteroarilo de 5-10 miembros y heterociclilo de 3-10 miembros, como se define en la presente memoria.

Un grupo "C-amido" se refiere a un grupo " $-\text{C}(=\text{O})\text{NR}_\text{A}\text{R}_\text{B}$ " en el que R_A y R_B se seleccionan cada uno independientemente entre hidrógeno, alquilo C₁₋₆, alquenilo C₂₋₆, alquinilo C₂₋₆, carbociclilo C₃₋₇, arilo C₆₋₁₀, heteroarilo de 5-10 miembros y heterociclilo de 3-10 miembros, como se define en la presente memoria.

Un grupo "N-amido" se refiere a un grupo " $-\text{N}(\text{R}_\text{A})\text{C}(=\text{O})\text{R}_\text{B}$ " en el que R_A y R_B se seleccionan cada uno independientemente entre hidrógeno, alquilo C₁₋₆, alquenilo C₂₋₆, alquinilo C₂₋₆, carbociclilo C₃₋₇, arilo C₆₋₁₀, heteroarilo de 5-10 miembros, y heterociclilo de 3-10 miembros, como se define en la presente memoria.

Un grupo "amino" se refiere a un grupo " $-\text{NR}_\text{A}\text{R}_\text{B}$ " en el que R_A y R_B se seleccionan cada uno independientemente entre hidrógeno, alquilo C₁₋₆, alquenilo C₂₋₆, alquinilo C₂₋₆, carbociclilo C₃₋₇, arilo C₆₋₁₀, heteroarilo de 5-10 miembros y heterociclilo de 3-10 miembros, como se define en la presente memoria. Un ejemplo no limitativo incluye amino libre (es decir, $-\text{NH}_2$).

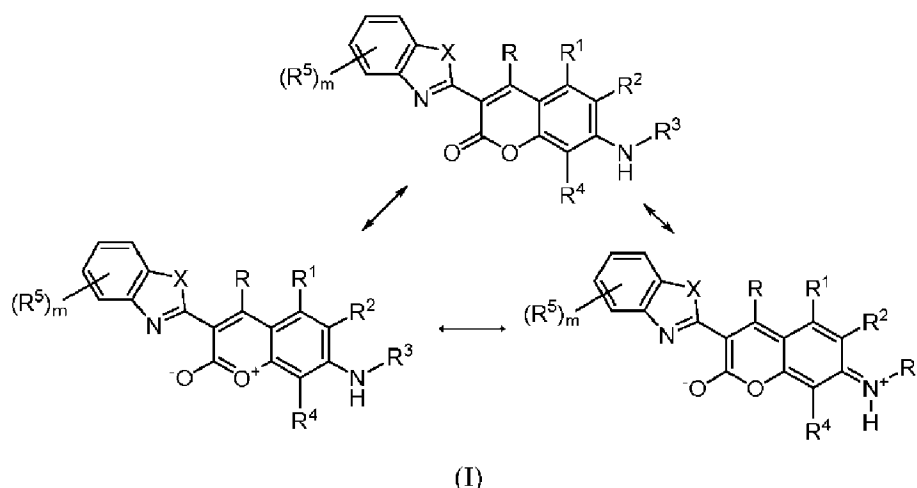
Un grupo "aminoalquilo" se refiere a un grupo amino conectado mediante un grupo alquileo.

Un grupo "alcoxialquilo" se refiere a un grupo alcoxi conectado mediante un grupo alquileo, como un "alcoxialquilo C₂₋₈" y similares.

Como se utiliza en la presente memoria, un grupo sustituido deriva del grupo original no sustituido en el que ha habido un intercambio de uno o más átomos de hidrógeno por otro átomo o grupo. A menos que se indique lo contrario, cuando se considera que un grupo es "sustituido", se entiende que el grupo se sustituye con uno o más sustituyentes seleccionados independientemente de alquilo C₁₋₆, alquenilo C₁₋₆, alquinilo C₁₋₆, heteroalquilo C₁₋₆, carbociclilo C₃₋₇ (opcionalmente sustituido con halo, alquilo C₁₋₆, alcoxi C₁₋₆, haloalquilo C₁₋₆ y haloalcoxi C₁₋₆), carbociclil-C₃₋₇-alquilo-C₁₋₆ (opcionalmente sustituido con halo, alquilo C₁₋₆, alcoxi C₁₋₆, haloalquilo C₁₋₆ y haloalcoxi C₁₋₆), heterociclilo de 3-10 miembros (opcionalmente sustituido con halo, alquilo C₁₋₆, alcoxi C₁₋₆, haloalquilo C₁₋₆ y haloalcoxi C₁₋₆), heterociclil-alquilo-C₁₋₆ de 3-10 miembros (opcionalmente sustituido con halo, alquilo C₁₋₆, alcoxi C₁₋₆, haloalquilo C₁₋₆ y haloalcoxi C₁₋₆), arilo (opcionalmente sustituido con halo, alquilo C₁₋₆, alcoxi C₁₋₆, haloalquilo C₁₋₆ y haloalcoxi C₁₋₆), aril-alquilo (C₁₋₆) (opcionalmente sustituido con halo, alquilo C₁₋₆, alcoxi C₁₋₆, haloalquilo C₁₋₆ y haloalcoxi C₁₋₆), heteroarilo de 5-10 miembros (opcionalmente sustituido con halo, alquilo C₁₋₆, alcoxi C₁₋₆, haloalquilo C₁₋₆ y haloalcoxi C₁₋₆), heteroarilo de 5-10 miembros alquilo (C₁₋₆) (opcionalmente sustituido con halo, alquilo C₁₋₆, alcoxi C₁₋₆, haloalquilo C₁₋₆ y haloalcoxi C₁₋₆), halo, ciano, hidroxilo, alcoxi C₁₋₆, alcoxi C₁₋₆-alquilo (C₁₋₆) (es decir, éter), ariloxi, sulfhidrilo (mercapto), halo-alquilo (C₁₋₆) (p. ej., $-\text{CF}_3$), halo-alcoxi (C₁₋₆) (p. ej., $-\text{OCF}_3$), alquiltio C₁₋₆, ariltio, amino, amino-alquilo (C₁₋₆), nitro, O-carbamilo, N-carbamilo, O-tiocarbamilo, N-tiocarbamilo, C-amido, N-amido, S-sulfonamido, N-sulfonamido, C-carboxi, O-carboxi, acilo, cianato, isocianato, tiocianato, isotiocianato, sulfinilo, sulfonilo, SO_3H , sulfino, $-\text{OSO}_2$ alquilo C₁₋₄ y oxo (=O). Siempre que un grupo se describa como "opcionalmente sustituido", ese grupo puede estar sustituido con los sustituyentes anteriores.

En algunas realizaciones, los grupos alquilo, alquenilo o alquinilo sustituidos se sustituyen con uno o más sustituyentes seleccionados del grupo que consiste en halo, $-\text{CN}$, SO_3^- , SR^a , OR^a , $\text{NR}^\text{b}\text{R}^\text{c}$, oxo, $\text{CONR}^\text{b}\text{R}^\text{c}$, COOH , y COOR^b , donde R^a, R^b y R^c se seleccionan cada uno independientemente de H, alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alquinilo, alquinilo sustituido, arilo y arilo sustituido.

Los compuestos descritos en la presente memoria pueden representarse como varias formas mesoméricas. Cuando se dibuja una única estructura, se entiende cualquiera de las formas mesoméricas relevantes. Los compuestos de cumarina descritos en la presente memoria se representan mediante una única estructura, pero pueden mostrarse igualmente como cualquiera de las formas mesoméricas relacionadas. Las estructuras mesoméricas ilustrativas se muestran a continuación para la fórmula (I):

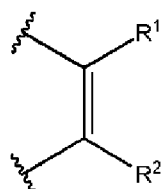


En cada caso en donde se muestra una única forma mesomérica de un compuesto descrito en la presente memoria, se contemplan igualmente las formas mesoméricas alternativas.

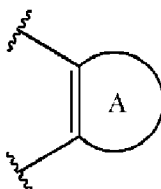
Como entenderá un experto en la técnica, un compuesto descrito en la presente memoria puede existir en forma ionizada, p. ej., $-\text{CO}_2^-$ o $-\text{SO}_3^-$. Si un compuesto contiene un grupo sustituyente cargado positiva o negativamente, por ejemplo, SO_3^- , también puede contener un contraión cargado negativa o positivamente, de modo que el compuesto en su conjunto sea neutro. En otros aspectos, el compuesto puede existir en forma de sal, donde el contraión lo proporciona un ácido o base conjugado.

Debe entenderse que ciertas convenciones de denominación de los radicales pueden incluir un monorradicale o un dirradical, dependiendo del contexto. Por ejemplo, cuando un sustituyente requiere dos puntos de unión al resto de la molécula, se entiende que el sustituyente es un dirradical. Por ejemplo, un sustituyente identificado como el alquilo que requiere dos puntos de unión incluye dirradicales como $-\text{CH}_2-$, $-\text{CH}_2\text{CH}_2-$, $-\text{CH}_2\text{CH}(\text{CH}_3)\text{CH}_2-$ y similares. Otras convenciones de denominación de los radicales indican claramente que el radical es un dirradical, como “alquilenos” o “alquenileno”.

Cuando se dice que dos grupos R “adyacentes” forman un anillo “junto con el átomo al que están unidos”, significa que la unidad colectiva de los átomos, enlaces intermedios y los dos grupos R son el anillo mencionado. Por ejemplo, cuando está presente la siguiente subestructura:



y R^1 y R^2 se definen como seleccionados en el grupo que consta de hidrógeno y alquilo, o R^1 y R^2 , junto con los átomos a los que están unidos, forman un anillo o carbociclilo, se entiende que R^1 y R^2 pueden seleccionarse entre hidrógeno o alquilo o, alternativamente, la subestructura tiene estructura:



Donde A es un anillo arilo o un carbociclilo que contiene el enlace doble representado.

Nucleótidos etiquetados

Según un aspecto de la descripción, se proporcionan compuestos de colorante adecuados para la unión a los restos del sustrato, que comprenden particularmente grupos enlazadores para permitir la unión a los restos del sustrato. Los restos de sustratos pueden ser prácticamente cualquier molécula o sustancia a la que puedan conjugarse los colorantes de la descripción y, a modo de ejemplo no limitante, pueden incluir nucleósidos, nucleótidos, polinucleótidos, carbohidratos, ligandos, partículas, superficies sólidas, polímeros orgánicos e inorgánicos,

cromosomas, núcleos, células vivas y combinaciones o ensamblajes de los mismos. Los colorantes pueden conjugarse con un enlazador opcional mediante una serie de medios que incluyen la atracción hidrófoba, la atracción iónica, y la unión covalente. En algunos aspectos, los colorantes se conjugan con el sustrato mediante la unión covalente. Más particularmente, la unión covalente se realiza por medio de un grupo enlazador. En algunos casos, dichos nucleótidos etiquetados también se denominan “nucleótidos modificados”.

La presente descripción proporciona además conjugados de nucleósidos y nucleótidos etiquetados con uno o más de los colorantes expuestos en la presente memoria (nucleótidos modificados). Los nucleósidos y nucleótidos etiquetados son útiles para etiquetar polinucleótidos formados por síntesis enzimática, como, a modo de ejemplo no limitante, en la amplificación por PCR, la amplificación isotérmica, la amplificación en fase sólida, la secuenciación de polinucleótidos (p. ej., la secuenciación en fase sólida), las reacciones de traslación de cortes y similares.

La unión a las biomoléculas puede ser mediante el R, R¹, R², R³, R⁴, R⁵ o la posición X del compuesto de fórmula (I). En algunos aspectos, la conexión se realiza a través del grupo R³ o R⁵ de fórmula (I). Para la fórmula (II), la unión puede estar en cualquier posición R⁶⁻¹¹ o X'. En algunos casos, el grupo sustituyente es un alquilo sustituido, por ejemplo, un alquilo sustituido con -CO₂H o una forma activada del grupo carboxilo, por ejemplo, una amida o un éster, que puede usarse para unirse al grupo amino o hidroxilo de las biomoléculas. En una realización, el grupo R, R¹, R², R³, R⁴, R⁵ o X de fórmula (I) o los grupos R⁶⁻¹¹ o X' de fórmula (II) pueden contener un residuo de éster o amida activado más adecuado para la formación adicional de enlaces amida/péptido. El término “éster activado”, como se emplea en la presente memoria, se refiere a un derivado del grupo carboxilo que es capaz de reaccionar en condiciones suaves, por ejemplo, con un compuesto que contiene un grupo amino. Los ejemplos no limitantes de ésteres activados incluyen, entre otros, ésteres de p-nitrofenilo, pentafluorofenilo y succinimido.

En algunas realizaciones, los compuestos de colorante pueden unirse covalentemente a oligonucleótidos o nucleótidos mediante la base de nucleótidos. Por ejemplo, el nucleótido u oligonucleótido etiquetado puede tener la etiqueta unida a la posición C5 de una base de pirimidina o la posición C7 de una base de 7-deaza purina a través de un resto enlazador. El nucleótido u oligonucleótido etiquetado también puede tener un grupo bloqueante 3'-OH unido covalentemente al azúcar ribosa o desoxirribosa del nucleótido.

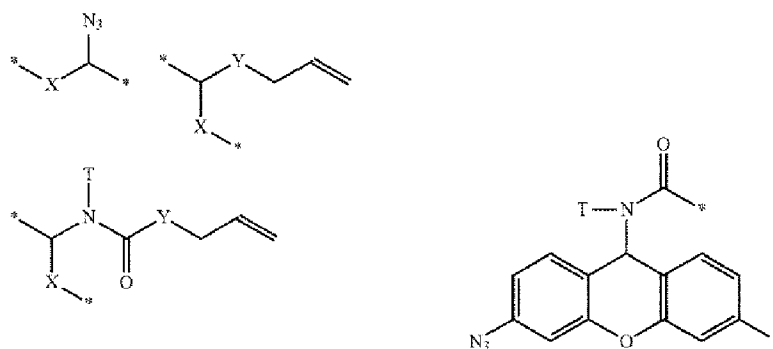
Una aplicación particularmente útil de los nuevos colorantes fluorescentes, como se describe en la presente memoria, es para etiquetar las biomoléculas, por ejemplo, los nucleótidos o los oligonucleótidos. Algunas realizaciones de la presente solicitud están dirigidas a un nucleótido u oligonucleótido etiquetado con los nuevos compuestos fluorescentes, como se describe en la presente memoria.

Enlazadores

Los compuestos de colorante, como se describen en la presente memoria, pueden incluir un grupo enlazador reactivo en una de las posiciones de los sustituyentes de la unión covalente del compuesto a un sustrato o a otra molécula. Los grupos de enlace reactivos son restos capaces de formar un enlace (p. ej., un enlace covalente o no covalente), en particular un enlace covalente. En una realización particular, el enlazador puede ser un enlazador escindible. El uso del término “enlazador escindible” no pretende implicar que se requiera eliminar el enlazador completo. La zona de escisión puede ubicarse en una posición del conector que asegure que parte del enlazador permanezca unida al colorante y/o al resto del sustrato después de la escisión. Los enlazadores escindibles pueden ser, a modo de ejemplo no limitante, enlazadores escindibles electrófilos, conectores escindibles nucleófilos, enlazadores fotoescindibles, escindibles en condiciones reductoras (por ejemplo, enlazadores que contienen disulfuro o azida), condiciones oxidativas, escindibles mediante el uso de enlazadores de captura de seguridad y pueden escindirse mediante mecanismos de eliminación. El uso de un enlazador escindible para unir el compuesto de colorante a un resto de sustrato asegura que la etiqueta pueda, si es necesario, eliminarse después de la detección, por lo que se evitará cualquier señal interferente en etapas posteriores.

Los grupos enlazadores útiles pueden incluir enlazadores que pueden escindirse utilizando fosfinas solubles en agua o catalizadores de metales de transición solubles en agua formados a partir de un metal de transición y ligandos solubles en agua, al menos parcialmente. En solución acuosa, este último forma, al menos parcialmente, complejos de metales de transición solubles en agua. Tales enlazadores escindibles se pueden utilizar para conectar bases de nucleótidos a etiquetas tales como los colorantes expuestos en la presente memoria.

Los enlazadores particulares incluyen restos de las fórmulas:



(en donde X se selecciona del grupo que comprende O, S, NH y NQ, en donde 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-C₁₀ sustituido o no sustituido y * indica dónde el resto se conecta al resto del nucleótido o nucleósido). En algunos aspectos, los enlazadores conectan las bases de los nucleótidos a etiquetas tales como, por ejemplo, los compuestos de colorante descritos en la presente memoria.

En realizaciones particulares, la longitud del enlazador entre un colorante fluorescente (fluoróforo) y una base de guanina puede alterarse, por ejemplo, introduciendo un grupo espaciador de polietilenglicol, aumentando de este modo la intensidad de fluorescencia en comparación con el mismo fluoróforo unido a la base de guanina mediante otros enlaces conocidos en la técnica. El diseño de enlazadores, y especialmente su longitud aumentada, pueden permitir mejorar el brillo de los fluoróforos unidos a las bases de guanina de los nucleótidos guanosina cuando se incorporan a polinucleótidos tales como ADN. Por lo tanto, cuando el colorante se emplea en cualquier método de análisis que requiere la detección de una etiqueta de colorante fluorescente unida a un nucleótido que contiene guanina, es ventajoso si el enlazador comprende un grupo espaciador de fórmula $-(CH_2)_nO-$, en donde n es un número entero entre 2 y 50.

Los nucleósidos y nucleótidos pueden etiquetarse en zonas del azúcar o la nucleobase. Como se conoce en la técnica, un “nucleótido” consiste en una base nitrogenada, un azúcar y uno o más grupos fosfato. En el ARN, el azúcar es una ribosa, y en el ADN es una desoxirribosa, es decir, un azúcar que carece de un grupo hidroxilo que está presente en la ribosa. La base nitrogenada es un derivado de la purina o la pirimidina. Las purinas son adenina (A) y guanina (G), y las pirimidinas son citosina (C) y timina (T) o en el contexto de ARN, uracilo (U). El átomo C-1 de la desoxirribosa se une al N-1 de una pirimidina o al N-9 de una purina. Un nucleótido también es un éster de fosfato de un nucleósido, teniendo la esterificación en el grupo hidroxilo unido a la C-3 o C-5 del azúcar. Los nucleótidos generalmente son mono-, di- o trifosfatos.

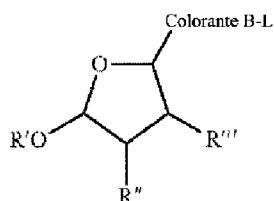
Un “nucleósido” es estructuralmente similar a un nucleótido, pero le faltan los restos de fosfato. Un ejemplo de un análogo de nucleósido sería uno en el que la etiqueta está unida a la base y no hay un grupo fosfato unido a la molécula de azúcar.

Aunque la base generalmente se denomina “purina” o “pirimidina”, el experto apreciará que los derivados y análogos disponibles no alteran la capacidad del nucleótido o nucleósido de someterse al emparejamiento de las bases de Watson-Crick. “Derivado” o “análogo” significa que hace referencia a un compuesto o una molécula cuya estructura central es la misma que la de un compuesto original, o se asemeja mucho a ella, pero que tiene una modificación química o física, como, por ejemplo, un grupo lateral diferente o adicional, que permite que el nucleótido derivado o el nucleósido se conecte a otra molécula. Por ejemplo, la base puede ser una desazopurina. En realizaciones particulares, los derivados deben ser capaces de experimentar el emparejamiento de Watson-Crick. “Derivado” y “análogo” también incluyen, por ejemplo, un nucleótido sintético o el derivado de un nucleósido que tiene restos de base modificados y/o restos de azúcar modificados. De dichos derivados y análogos se habla, por ejemplo, en Scheit, Nucleotide analogs (John Wiley & Son, 1980) y Uhlman et al., Chemical Reviews 90:54 3-584, 1990. Los análogos de nucleótidos también pueden constar de enlaces fosfodiéster modificados que incluyen enlaces de fosforotioato, fosforoditioato, alquilo-fosfonato, fosforanilida, fosforamidato y similares.

Un colorante se puede unir a cualquier posición en la base nucleotídica, por ejemplo, a través de un enlazador. En realizaciones particulares, el apareamiento de bases de Watson-Crick todavía se puede llevar a cabo para el análogo resultante. Determinadas zonas de etiquetado de nucleobase incluyen la posición C5 de una base de pirimidina o la posición C7 de una base de 7-deaza purina. Como se describió anteriormente, un grupo enlazador puede utilizarse para unir covalentemente un colorante al nucleósido o nucleótido.

En realizaciones particulares, el nucleósido o nucleótido etiquetado puede incorporarse enzimáticamente y ser enzimáticamente extensible. Por consiguiente, un resto enlazador puede ser de longitud suficiente para conectar el nucleótido al compuesto, de modo que el compuesto no interfiera significativamente en la unión general y el reconocimiento del nucleótido por una enzima de replicación de ácido nucleico. Por lo tanto, el conector también puede constar de una unidad separadora. Las distancias espaciadoras, por ejemplo, la base de los nucleótidos de una zona o una etiqueta de escisión.

Los nucleósidos o nucleótidos etiquetados con los colorantes descritos en la presente memoria pueden tener la fórmula:



donde

El colorante es un compuesto de colorante,

B es una nucleobase, tal como, por ejemplo, uracilo, timina, citosina, adenina, guanina y similares,

L es un grupo enlazador opcional que puede estar presente o no,

R' puede ser H, monofosfato, difosfato, trifosfato, tiofosfato, un análogo de éster de fosfato, -O- unido a un grupo que contiene fósforo reactivo u -O- protegido por un grupo bloqueante,

R'' puede ser H, OH, una fosforamidita o un grupo bloqueante 3'-OH y

R''' es H o OH.

Donde R'' es fosforamidita, R' es un grupo protector hidroxilo escindible con ácido que permite el acoplamiento de monómero posterior en condiciones de síntesis automatizadas.

En una realización particular, el grupo bloqueante está separado e independiente del compuesto de colorante, es decir, no está unido a él. Por otra parte, el colorante puede constar de todo o parte del grupo bloqueante 3'-OH. Por lo tanto, R'' puede ser un grupo bloqueante 3'-OH que puede constar o no del compuesto colorante.

En otra realización alternativa, no hay ningún grupo bloqueante en el carbono 3' del azúcar pentosa y el colorante (o la construcción de colorante y enlazador) unido a la base, por ejemplo, puede tener un tamaño o estructura suficiente para actuar como bloqueante de la incorporación de otro nucleótido. Por lo tanto, el bloqueo puede deberse a un impedimento estérico o puede deberse a una combinación de tamaño, carga y estructura, independientemente de que el colorante esté unido o no a la posición 3' del azúcar.

En otra realización alternativa, el grupo bloqueante está presente en el carbono 2' o 4' del azúcar pentosa y puede tener un tamaño o estructura suficientes para actuar como un bloque para la incorporación de un nucleótido adicional.

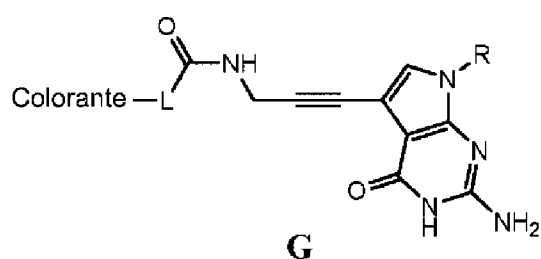
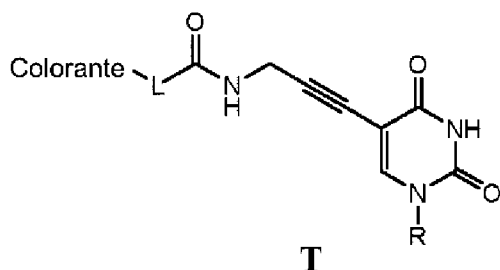
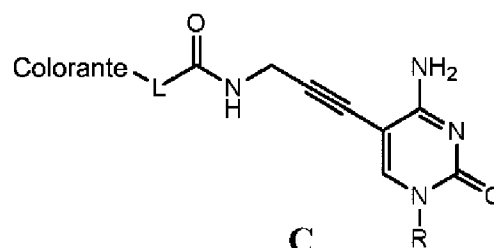
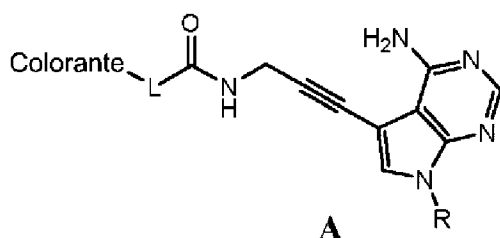
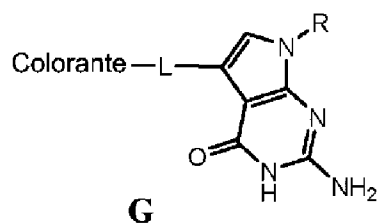
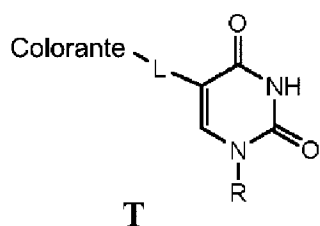
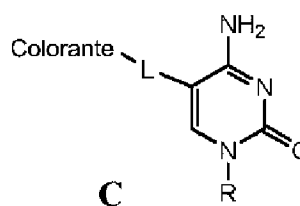
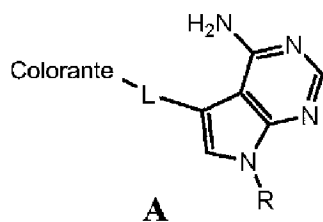
El uso de un grupo bloqueante permite controlar la polimerización, como detener la extensión cuando se incorpora un nucleótido modificado. Si el efecto bloqueante es reversible, por ejemplo, a modo de ejemplo no limitante, cambiando las condiciones químicas o mediante la retirada de un bloqueo químico, la extensión puede detenerse en ciertos puntos y luego dejar que continúe.

En otra realización particular, el grupo bloqueante puede ser azidometilo (-CH₂N₃) o azidometilo sustituido (p. ej., -CH(CHF₂)N₃ o CH(CH₂F)N₃) o alilo.

En una realización particular, el enlazador (entre el colorante y el nucleótido) y el grupo bloqueante están presentes y son restos separados. En realizaciones particulares, el enlazador y el grupo bloqueante son ambos escindibles en condiciones sustancialmente similares. Por lo tanto, los procesos de desprotección y desbloqueo pueden ser más eficientes, porque solo se requerirá un único tratamiento para eliminar tanto el compuesto de colorante como el grupo bloqueante. Sin embargo, en algunas realizaciones, un enlazador y un grupo bloqueante no necesitan ser escindibles en condiciones similares, sino que pueden escindirse individualmente en condiciones distintas.

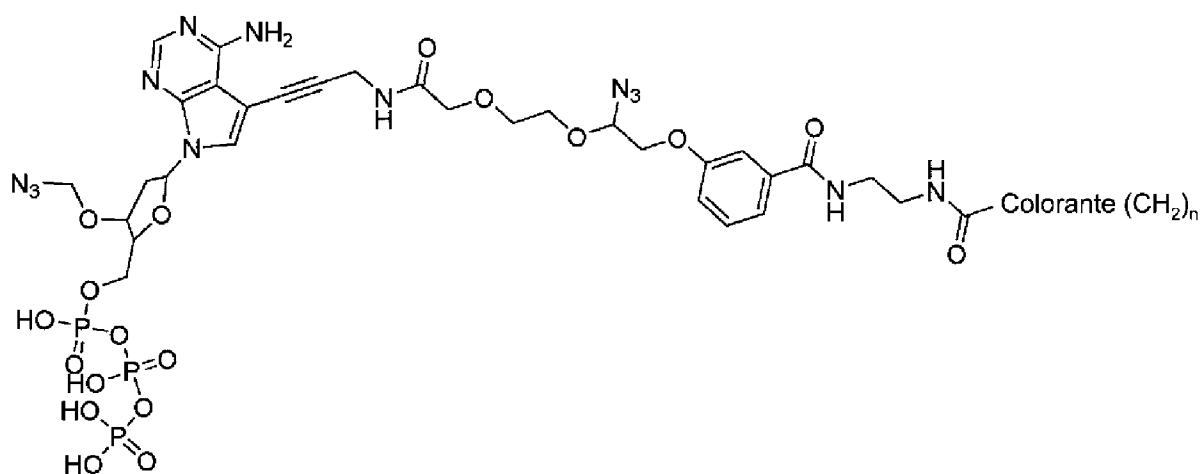
La descripción también abarca polinucleótidos que incorporan compuestos de colorante. Dichos polinucleótidos pueden ser ADN o ARN que consten de, respectivamente de, desoxirribonucleótidos o ribonucleótidos unidos en un enlace fosfodiéster. Los polinucleótidos pueden comprender nucleótidos naturales, nucleótidos no naturales (o modificados) distintos de los nucleótidos etiquetados descritos en la presente memoria o cualquier combinación de los mismos, en combinación con al menos un nucleótido modificado (p. ej., etiquetado con un compuesto de colorante) tal como se expone en la presente memoria. Los polinucleótidos según la descripción también pueden incluir enlaces principales no naturales y/o modificaciones químicas no nucleotídicas. También se contemplan estructuras químicas compuestas por mezclas de ribonucleótidos y desoxirribonucleótidos que constan de al menos un nucleótido etiquetado.

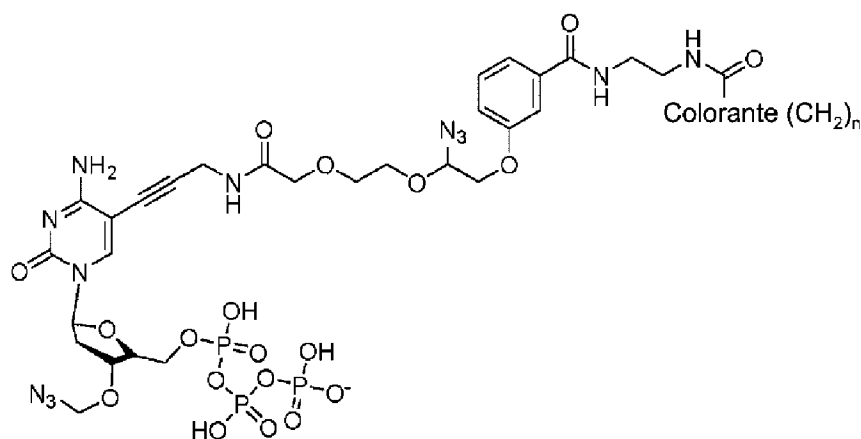
Los nucleótidos etiquetados de ejemplo no limitantes, como se describen en la presente memoria incluyen:



en donde L representa un enlazador y R representa un residuo de azúcar, como se describió anteriormente.

En algunas realizaciones, se muestran a continuación conjugados de colorante fluorescente ilustrativos no limitantes:



**Colorante ffC-LN3****Kits**

La presente descripción también proporciona kits que incluyen nucleósidos modificados y/o nucleótidos etiquetados con colorantes. Tales kits generalmente incluirán al menos un nucleótido modificado o nucleósido etiquetado con un colorante tal como se expone en la presente memoria junto con al menos un componente adicional. El (Los) componente(s) pueden(n) ser uno o más de los componentes identificados en un método expuesto en la presente memoria o en la sección de Ejemplos a continuación. A continuación, se exponen algunos ejemplos no limitante de componentes que se pueden combinar en un kit de la presente descripción.

En una realización particular, un kit puede incluir al menos un nucleótido o nucleósido modificado etiquetado con un colorante expuesto en la presente memoria junto con nucleótidos o nucleósidos modificados o no modificados. Por ejemplo, los nucleótidos modificados etiquetados con colorantes según la descripción pueden suministrarse en combinación con nucleótidos no etiquetados o nativos, y/o con nucleótidos etiquetados con fluorescencia o cualquier combinación de los mismos. Por consiguiente, los kits pueden comprender nucleótidos modificados etiquetados con colorantes según la descripción y nucleótidos modificados etiquetados con otros, por ejemplo, compuestos de colorante de la técnica anterior. Las combinaciones de nucleótidos se pueden proporcionar como componentes individuales separados (p. ej., un tipo de nucleótido por recipiente o tubo) o como mezclas de nucleótidos (p. ej., dos o más nucleótidos mezclados en el mismo recipiente o tubo).

Cuando los kits comprenden una pluralidad, particularmente dos, o tres, o más particularmente cuatro, nucleótidos modificados etiquetados con un compuesto de colorante, los diferentes nucleótidos pueden etiquetarse con diferentes compuestos de colorante, o uno puede ser oscuro, sin compuestos de colorante. Cuando los diferentes nucleótidos están etiquetados con diferentes compuestos de colorante, es una característica de los kits que los compuestos de colorante sean colorantes fluorescentes espectralmente distinguibles. Como se utiliza en la presente memoria, el término “colorantes fluorescentes espectralmente distinguibles” se refiere a los colorantes fluorescentes que emiten energía fluorescente a longitudes de onda que pueden distinguirse por el equipo de detección fluorescente (por ejemplo, una plataforma comercial de secuenciación de ADN basada en capilares) cuando dos o más de dichos colorantes están presentes en una muestra. Cuando dos nucleótidos modificados etiquetados con compuestos de colorante fluorescente se suministran en forma de kit, es una característica de algunas realizaciones que los colorantes fluorescentes espectralmente distinguibles pueden ser excitados a la misma longitud de onda, tal como, por ejemplo, por el mismo láser. Cuando cuatro nucleótidos modificados etiquetados con compuestos de colorante fluorescente se suministran en forma de kit, es una característica de algunas realizaciones que dos de los colorantes fluorescentes espectralmente distinguibles pueden ser excitados ambos a una longitud de onda y los otros dos colorantes espectralmente distinguibles pueden ser excitados ambos a otra longitud de onda. Las longitudes de onda de excitación específicas son 488 nm y 532 nm.

En una realización, un kit incluye un nucleótido modificado etiquetado con un compuesto de la presente descripción y un segundo nucleótido modificado etiquetado con un segundo colorante en donde los colorantes tienen una diferencia en el máximo de absorbancia de al menos 10 nm, particularmente 20 nm a 50 nm. Más particularmente, los dos compuestos de colorante tienen desplazamientos de Stokes de entre 15 y 40 nm, donde el “desplazamiento de Stokes” es la distancia entre las longitudes de onda de absorción máxima y de emisión máxima.

En una realización adicional, un kit puede incluir además otros dos nucleótidos modificados etiquetados con colorantes fluorescentes en donde los colorantes se excitan por el mismo láser a 532 nm. Los colorantes pueden tener una diferencia en la absorbancia máxima de al menos 10 nm, particularmente 20 nm a 50 nm. Más particularmente, los dos compuestos de colorante tienen desplazamientos de Stokes de entre 20-40 nm. Los colorantes particulares que se distinguen espectralmente de los colorantes de la presente descripción y que cumplen los criterios anteriores son

los análogos de polimetina, por ejemplo, Cy3, Alexa 532, Molecular Probes A20106 o polimetinas asimétricas. Los colorantes alternativos incluyen análogos de rodamina, por ejemplo, tetrametilrodamina y análogos de la misma.

En una realización alternativa, los kits de la descripción pueden contener nucleótidos donde la misma base está etiquetada con dos compuestos diferentes. Un primer nucleótido puede etiquetarse con un compuesto de la descripción. Un segundo nucleótido puede etiquetarse con un compuesto espectralmente distinto, por ejemplo, un colorante “verde” que absorbe menos de 600 nm. Un tercer nucleótido puede etiquetarse como una mezcla del compuesto de la descripción y el compuesto espectralmente distinto, y el cuarto nucleótido puede ser “oscuro” y no contiene ninguna etiqueta. En términos simples, por lo tanto, los nucleótidos 1-4 pueden etiquetarse como “azul”, “verde”, “azul/verde” y oscuro. Para simplificar el instrumental, se pueden etiquetar cuatro nucleótidos con dos colorantes excitados con un único láser y, por lo tanto, el etiquetado de los nucleótidos 1-4 puede ser “azul 1”, “azul 2”, “azul 1/azul 2” y oscuro.

Los nucleótidos pueden contener dos colorantes de la presente descripción. Un kit puede contener dos o más nucleótidos etiquetados con los colorantes de la descripción. Los kits pueden contener un nucleótido adicional en donde el nucleótido está etiquetado con un colorante que se absorbe en la región de 520 nm a 560 nm. Los kits pueden contener además un nucleótido no etiquetado.

Aunque los kits se ejemplifican en la presente memoria con respecto a configuraciones que tienen diferentes nucleótidos que están etiquetados con diferentes compuestos de colorante, se entenderá que los kits pueden incluir 2, 3, 4 o más nucleótidos diferentes que tienen el mismo compuesto de colorante.

En realizaciones particulares, un kit puede incluir una enzima polimerasa capaz de catalizar la incorporación de los nucleótidos modificados a un polinucleótido. Otros componentes que se incluirán en dichos kits pueden incluir búferes y similares. Los nucleótidos modificados etiquetados con colorantes según la descripción, y cualquier otro componente nucleotídico, que incluyen mezclas de nucleótidos diferentes, pueden suministrarse en el kit en forma concentrada para diluirlos antes de usarse. En tales realizaciones, también puede incluirse un regulador de dilución adecuado. De nuevo, uno o más de los componentes identificados en un método expuesto en la presente memoria pueden incluirse en un kit de la presente descripción.

Métodos de secuenciación

Los nucleótidos (o nucleósidos) modificados que comprenden un compuesto de colorante según la presente descripción pueden utilizarse en cualquier método de análisis como los que incluyen la detección de una etiqueta fluorescente unida a un nucleótido o nucleósido, ya sea por sí sola o incorporada o asociada a una estructura molecular o conjugado mayor. En este contexto, el término “incorporado en un polinucleótido” puede significar que el fosfato 5' está unido en un enlace fosfodiéster al grupo hidroxilo 3' de un segundo nucleótido (modificado o no modificado), que a su vez puede formar parte de una cadena polinucleotídica más larga. El extremo 3' de un nucleótido modificado expuesto en la presente memoria puede estar unido o no a un enlace fosfodiéster al fosfato 5' de otro nucleótido (modificado o no modificado). Por lo tanto, en una realización no limitante, la descripción proporciona un método para detectar un nucleótido modificado incorporado en un polinucleótido que comprende: (a) incorporar al menos un nucleótido modificado de la descripción en un polinucleótido y (b) detectar el (los) nucleótido(s) modificado(s) incorporado(s) al polinucleótido mediante la detección de la señal fluorescente del compuesto de colorante unido a dicho(s) nucleótido(s) modificado(s).

Este método puede incluir: una etapa de síntesis (a) en la que uno o más nucleótidos modificados según la descripción se incorporan a un polinucleótido y una etapa de detección (b) en la que uno o más nucleótidos modificados incorporados al polinucleótido se detectan detectando o midiendo cuantitativamente su fluorescencia.

Algunas realizaciones de la presente solicitud se dirigen a métodos de secuenciación que incluyen: (a) incorporar al menos un nucleótido etiquetado, como se describe en la presente memoria, a un polinucleótido; y (b) detectar el o los nucleótidos etiquetados incorporados al polinucleótido detectando la señal fluorescente del nuevo colorante fluorescente unido a dicho nucleótido o nucleótidos modificados.

En una realización, al menos un nucleótido modificado se incorpora a un polinucleótido en la etapa de síntesis por la acción de una enzima polimerasa. Sin embargo, pueden utilizarse otros métodos de unión de nucleótidos modificados a polinucleótidos, tales como, por ejemplo, la síntesis química de oligonucleótidos o la ligadura de oligonucleótidos etiquetados a oligonucleótidos no etiquetados. Por lo tanto, el término “incorporar”, cuando se utiliza en referencia a un nucleótido y polinucleótido, puede abarcar la síntesis de polinucleótidos por métodos químicos, así como por métodos enzimáticos.

En una realización específica, se lleva a cabo una etapa de síntesis y, opcionalmente, puede comprender incubar una cadena de polinucleótidos molde con una mezcla de reacción que comprende los nucleótidos modificados etiquetados con fluorescencia de la descripción. También se puede proporcionar una polimerasa en condiciones que permitan la formación de un enlace fosfodiéster entre un grupo hidroxilo 3' libre en una cadena de polinucleótido hibridada con la cadena de polinucleótido molde y un grupo fosfato 5' en el nucleótido modificado. Por lo tanto, una etapa de síntesis

puede incluir la formación de una cadena de polinucleótidos dirigida por el apareamiento de bases complementarias de nucleótidos con una cadena molde.

En todas las realizaciones de los métodos, la etapa de detección puede llevarse a cabo mientras la cadena de polinucleótidos a la que se incorporan los nucleótidos etiquetados se hibrida con una cadena molde, o después de una etapa de desnaturalización en la que las dos cadenas se separan. Pueden incluirse etapas adicionales, por ejemplo, etapas de reacción química o enzimática o etapas de purificación entre la etapa de síntesis y la etapa de detección. En particular, la cadena diana que incorpora el o los nucleótidos etiquetados puede aislarse o purificarse y luego procesarse adicionalmente o utilizarse en un análisis posterior. A modo de ejemplo, los polinucleótidos diana etiquetados con nucleótidos modificados como se describe en la presente memoria en una etapa sintética pueden utilizarse posteriormente como sondas o cebadores etiquetados. En otras realizaciones, el producto de la etapa de síntesis expuesto en la presente memoria puede someterse a etapas de reacción adicionales y, si se desea, el producto de estas etapas posteriores se purifica o aísla.

Las condiciones adecuadas para la etapa de síntesis serán bien conocidas por quienes estén familiarizados con las técnicas estándar de la biología molecular. En una realización, una etapa de síntesis puede ser análoga a una reacción de extensión de cebador estándar utilizando precursores de nucleótidos, incluidos los nucleótidos modificados tal como se expone en la presente memoria, para formar una cadena diana extendida complementaria a la cadena molde en presencia de una enzima polimerasa adecuada. En otras realizaciones, la etapa de síntesis puede formar parte de una reacción de amplificación que genera un producto de amplificación bicatenario etiquetado compuesto por cadenas complementarias hibridadas derivadas de la copia de las cadenas de polinucleótidos diana y molde. Otras etapas sintéticas ilustrativas incluyen la traslación de cortes, la polimerización por desplazamiento de la cadena, el etiquetado de ADN con cebado aleatorio, etc. Una enzima polimerasa particularmente útil para una etapa de síntesis es aquella que es capaz de catalizar la incorporación de nucleótidos modificados tal como se expone en la presente memoria. Se puede utilizar una variedad de polimerasas de origen natural o modificadas. A modo de ejemplo, puede utilizarse una polimerasa termoestable para una reacción sintética que se lleva a cabo utilizando condiciones de termociclado, mientras que puede no desearse una polimerasa termoestable para las reacciones de extensión de cebadores isotérmicas. En las reacciones sintéticas que se llevan a cabo a temperaturas más bajas, tal como 37 °C, las enzimas polimerasa no necesitan necesariamente ser las polimerasas termoestables; por lo tanto, la elección de la polimerasa dependerá de una serie de factores, como la temperatura de reacción, el pH, la actividad de desplazamiento de la cadena y similares.

En realizaciones específicas no limitantes, la descripción abarca métodos de secuenciación de ácidos nucleicos, resecuenciación, secuenciación del genoma completo, puntuación de polimorfismos de un único nucleótido, y cualquier otra aplicación que implique la detección del nucleótido o nucleósido modificado etiquetado con los colorantes expuestos en la presente memoria cuando se incorpora a un polinucleótido. Cualquiera de una variedad de otras aplicaciones que benefician el uso de polinucleótidos etiquetados con los nucleótidos modificados que comprenden colorantes fluorescentes puede usar nucleótidos modificados o nucleósidos con los colorantes expuestos en la presente memoria.

En una realización particular, la descripción proporciona el uso de nucleótidos modificados que comprenden compuestos de colorante según la descripción en una reacción de secuenciación por síntesis de polinucleótidos. La secuenciación por síntesis generalmente implica la adición secuencial de uno o más nucleótidos u oligonucleótidos a una cadena polinucleotídica en crecimiento en la dirección 5' a 3' utilizando una polimerasa o ligasa para formar una cadena polinucleotídica extendida complementaria al ácido nucleico molde que se va a secuenciar. La identidad de la base presente en uno o más de los nucleótidos añadidos se puede determinar en una etapa de detección o "toma de imágenes". La identidad de la base añadida puede determinarse después de cada etapa de incorporación de los nucleótidos. La secuencia del molde se puede inferir usando reglas convencionales de emparejamiento de bases Watson-Crick. El uso de los nucleótidos modificados etiquetados con colorantes expuestos en la presente memoria para la determinación de la identidad de una sola base puede ser útil, por ejemplo, en la puntuación de los polimorfismos de un solo nucleótido, y tales reacciones de extensión de una sola base están dentro del ámbito de esta descripción.

En una realización de la presente descripción, la secuencia de un polinucleótido molde se determina detectando la incorporación de uno o más nucleótidos a una cadena naciente complementaria al polinucleótido molde que se va a secuenciar a través de la detección de etiquetas fluorescentes unidas al nucleótido o nucleótidos incorporados. La secuenciación del polinucleótido molde se puede cebar con un cebador adecuado (o preparado como un constructo de horquilla que contendrá el cebador como parte de la horquilla), y la cadena naciente se extiende de modo escalonado mediante la adición de nucleótidos al extremo 3' del cebador en una reacción catalizada por polimerasa.

En realizaciones particulares, cada uno de los diferentes trifosfatos de los nucleótidos (A, T, G y C) puede etiquetarse con un fluoróforo único y también consta de un grupo bloqueante en la posición 3' para prevenir la polimerización descontrolada. Por otra parte, uno de los cuatro nucleótidos puede no estar etiquetado (oscuro). La enzima polimerasa incorpora un nucleótido en la cadena naciente complementaria al polinucleótido molde, y el grupo bloqueante impide la incorporación adicional de nucleótidos. Los nucleótidos no incorporados pueden eliminarse por lavado y la señal fluorescente de cada nucleótido incorporado puede "leerse" ópticamente por medios adecuados, tal como un dispositivo de carga acoplada que utilice excitación láser y filtros de emisión adecuados. A continuación, el grupo de

bloqueo 3' y los compuestos de colorante fluorescente pueden eliminarse (desprotegerse) (simultánea o secuencialmente) para exponer la cadena naciente de una incorporación posterior de nucleótidos. Típicamente, la identidad del nucleótido incorporado se determinará después de cada etapa de incorporación, pero esto no es estrictamente esencial. De forma similar, la patente estadounidense U.S. Pat. No. 5,302,509 describe un método para secuenciar polinucleótidos inmovilizados sobre un soporte sólido.

El método, como se ha ilustrado anteriormente, utiliza la incorporación de nucleótidos A, G, C y T etiquetados con fluorescencia y bloqueados en 3' en una cadena en crecimiento complementaria al polinucleótido inmovilizado en presencia de ADN polimerasa. La polimerasa incorpora una base complementaria al polinucleótido diana, pero se evitan más adiciones por el grupo bloqueante 3'. La etiqueta del nucleótido incorporado puede, luego, determinarse y el grupo bloqueante eliminarse por escisión química para permitir que se produzca la polimerización adicional. El molde de ácido nucleico que se va a secuenciar en una reacción de secuenciación por síntesis puede ser cualquier polinucleótido que se desee secuenciar. El molde de ácido nucleico para una reacción de secuenciación constará típicamente de una región bicatenaria que tiene un grupo hidroxilo 3' libre que sirve como un cebador o punto de iniciación para la adición de más nucleótidos en la reacción de secuenciación. La región del molde que se va a secuenciar colgará este grupo hidroxilo 3' libre en la cadena complementaria. La región sobresaliente del molde que secuenciar puede ser monocatenaria, pero puede ser bicatenaria, siempre que un "corte esté presente" en la cadena complementaria a la cadena molde que se vaya a secuenciar para proporcionar un grupo 3' OH libre al inicio de la reacción de secuenciación. En tales realizaciones, la secuenciación puede continuar mediante el desplazamiento de cadenas. En ciertas realizaciones, un cebador que porta el grupo hidroxilo 3' libre puede añadirse como un componente aparte (p. ej., un oligonucleótido corto) que se hibrida con una región monocatenaria del molde que se va a secuenciar. Por otra parte, el cebador y la cadena molde que se va a secuenciar pueden formar parte de una cadena de ácido nucleico parcialmente autocomplementaria capaz de formar un dúplex intramolecular, como por ejemplo, una estructura en bucle de horquilla. Los polinucleótidos de horquilla y los métodos por los que pueden unirse a soportes sólidos se describen en las publicaciones PCT de las patentes WO0157248 y WO2005/047301. Los nucleótidos pueden añadirse sucesivamente a un cebador en crecimiento, dando lugar a la síntesis de una cadena polinucleotídica en la dirección 5' a 3'. La naturaleza de la base que se ha añadido puede determinarse, particularmente pero no necesariamente, después de cada adición de nucleótidos, lo que así proporciona información de la secuencia del molde de ácido nucleico. Por lo tanto, un nucleótido se incorpora a una cadena de ácido nucleico (o polinucleótido) mediante la unión del nucleótido al grupo hidroxilo 3' libre de la cadena de ácido nucleico mediante la formación de un enlace fosfodiéster con el grupo fosfato 5' del nucleótido.

El molde de ácido nucleico que se va a secuenciar puede ser ADN o ARN, o incluso una molécula híbrida compuesta de desoxinucleótidos y ribonucleótidos. El molde de ácido nucleico puede constar de nucleótidos de origen natural y/o no natural y enlaces del segmento principal naturales o no naturales, siempre que estos no impidan copiar el molde en la reacción de secuenciación.

En ciertas realizaciones, el molde de ácido nucleico que se va a secuenciar puede unirse a un soporte sólido a través de cualquier método de enlace adecuado conocido en la técnica, por ejemplo, mediante unión covalente. En ciertas realizaciones, los polinucleótidos molde pueden unirse directamente a un soporte sólido (p. ej., un soporte a base de sílice). Sin embargo, en otras realizaciones de la descripción, la superficie del soporte sólido puede modificarse de algún modo para permitir una unión covalente directa de los polinucleótidos del molde, o inmovilizar los polinucleótidos molde a través de un hidrogel o una multicapa polielectrolítica, que puede estar unido de forma no covalente al soporte sólido.

Las matrices en las que los polinucleótidos se han unido directamente a los soportes a base de sílice son las que se describen, por ejemplo, en la patente WO00006770, en donde los polinucleótidos se inmovilizan sobre un soporte de vidrio mediante reacción entre un grupo epóxido colgante en el vidrio con un grupo amino interno en el polinucleótido. Además, los polinucleótidos se pueden unir a un soporte sólido mediante la reacción de un nucleófilo a base de azufre con el soporte sólido, por ejemplo, como se describe en la patente WO2005/047301.

Otro ejemplo adicional de polinucleótidos molde con soporte sólido es cuando los polinucleótidos molde se unen a un hidrogel soportado sobre soportes a base de sílice u otros soportes sólidos, por ejemplo, como se describe en las patentes WO00/31148, WO01/01143, WO02/12566, WO03/014392, la patente US-6.465.178 y la patente WO00/53812.

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

Los hidrogeles específicos que pueden utilizarse en la presente solicitud incluyen los descritos en la patente WO 2005/065814 y la publicación de la patente US-2014/0079923. En una realización, el hidrogel es PAZAM (poli (N-(5-azidoacetamidil) acrilamida-co-acrilamida)).

Las moléculas molde de ADN pueden unirse a microesferas o micropartículas, por ejemplo, como se describe en la patente US-6.172.218. La unión a microesferas o micropartículas puede ser útil para aplicaciones de secuenciación. Se pueden preparar bibliotecas de microesferas en donde cada perla contiene diferentes secuencias de ADN. Las bibliotecas ilustrativas y los métodos para su creación se describen en Nature, 437, 376-380 (2005); Science, 309,

5741, 1728-1732 (2005). La secuenciación de las matrices de tales microesferas utilizando nucleótidos expuestos en la presente memoria, está dentro del ámbito de la presente descripción.

El molde o los moldes que se van a secuenciar pueden formar parte de una “matriz” sobre un soporte sólido, en cuyo caso la matriz puede tomar cualquier forma conveniente. Por lo tanto, el método de la descripción se aplica a todos los tipos de matrices de alta densidad, que incluyen matrices de moléculas individuales, matrices agrupadas y matrices de microesferas. Los nucleótidos modificados etiquetados con los compuestos de colorante de la presente descripción pueden utilizarse para secuenciar moldes en prácticamente cualquier tipo de matriz, que incluye, aunque no de forma limitante, al as formadas por la inmovilización de moléculas de ácido nucleico sobre un soporte sólido.

Sin embargo, los nucleótidos modificados etiquetados con los compuestos de colorante de la descripción son particularmente ventajosos en el contexto de la secuenciación de las matrices agrupadas. En las matrices agrupadas, las distintas regiones de la matriz (frecuentemente denominadas sitios o características) comprenden múltiples moléculas molde de polinucleótidos. En general, las múltiples moléculas de polinucleótidos no se pueden resolver individualmente por medios ópticos y, en cambio, se detectan como un ensamblaje. Dependiendo de cómo se forme la matriz, cada zona en la matriz puede comprender varias copias de una molécula de polinucleótido individual (p. ej., la zona es homogénea para una especie ácida nucleica monocateriana o bicateriana) o incluso de varias copias de un pequeño número de moléculas de polinucleótidos diferentes (p. ej., varias copias de dos cadenas de ácido nucleico complementarias). Las matrices agrupadas de moléculas de ácido nucleico pueden producirse mediante el uso de técnicas generalmente conocidas en la técnica. A modo de ejemplo, las patentes WO 98/44151 y WO00/18957 describen métodos de amplificación de ácidos nucleicos en donde tanto los productos molde como los de amplificación permanecen inmovilizados sobre un soporte sólido para formar matrices compuestas por agrupaciones o “colonias” de moléculas de ácido nucleico inmovilizadas. Las moléculas de ácido nucleico presentes en las matrices agrupadas preparadas según estos métodos son moldes adecuados para la secuenciación utilizando los nucleótidos modificados etiquetados con compuestos de colorante de la descripción.

Los nucleótidos modificados etiquetados con compuestos de colorante de la presente descripción también son útiles en la secuenciación de moldes en las matrices de moléculas individuales. El término “matriz de moléculas individuales” o “SMA”, en inglés, como se utiliza en la presente memoria, se refiere a una población de moléculas de polinucleótidos, distribuidas (o dispuestas) sobre un soporte sólido, en donde la separación de cualquier polinucleótido individual de todos los demás de la población es tal que es posible resolver individualmente las moléculas polinucleotídicas individuales. Por lo tanto, las moléculas de ácido nucleico diana inmovilizadas sobre la superficie del soporte sólido pueden ser capaces de resolverse por medios ópticos en algunas realizaciones. Esto significa que se producirán una o más señales distintas, cada una de las cuales representa un polinucleótido, dentro del área resoluble del dispositivo de obtención de imágenes particular utilizado.

La detección de moléculas individuales puede lograrse en donde la separación entre moléculas de polinucleótidos adyacentes en una matriz es al menos 100 nm, más particularmente al menos 250 nm, aún más particularmente al menos 300 nm, incluso más particularmente de al menos 350 nm. Por lo tanto, cada molécula puede resolverse individualmente y detectarse como un punto fluorescente de una sola molécula, y la fluorescencia de dicho punto fluorescente de molécula única también exhibe un fotoblanqueamiento de una única etapa.

Los términos “resueltas individualmente” y “resolución individual” se utilizan en la presente memoria para especificar que, cuando se visualiza, es posible distinguir una molécula en la matriz de sus moléculas vecinas. La separación entre moléculas individuales en la matriz se determinará, en parte, por la técnica particular utilizada para resolver las moléculas individuales. Las características generales de las matrices de moléculas individuales se entenderán con referencia a las solicitudes de patentes publicadas WO00/06770 y WO01/57248. Aunque un uso de los nucleótidos modificados de la descripción está en las reacciones de secuenciación por síntesis, la utilidad de los nucleótidos modificados no se limita a tales métodos. De hecho, los nucleótidos pueden utilizarse ventajosamente en cualquier metodología de secuenciación que requiera la detección de etiquetas fluorescentes unidas a nucleótidos incorporados a un polinucleótido.

En particular, los nucleótidos modificados etiquetados con los compuestos de colorante de la descripción pueden utilizarse en protocolos automatizados de secuenciación fluorescente, particularmente, en la secuenciación del ciclo terminador del colorante fluorescente en función del método de secuenciación de la terminación de la cadena de Sanger y colaboradores. Dichos métodos emplean generalmente enzimas y la secuenciación del ciclo para incorporar didesoxinucleótidos marcados con fluorescencia en una reacción de secuenciación de extensión del cebador. Los llamados métodos de secuenciación de Sanger, y los protocolos relacionados (tipo Sanger), utilizan la terminación aleatoria de la cadena con dideoxinucleótidos etiquetados.

Por lo tanto, la presente descripción también abarca nucleótidos modificados etiquetados con compuestos de colorante que son dideoxinucleótidos que carecen de grupos hidroxilo en ambas posiciones 3' y 2', tales dideoxinucleótidos modificados son adecuados para usarse en métodos de secuenciación de tipo Sanger y similares.

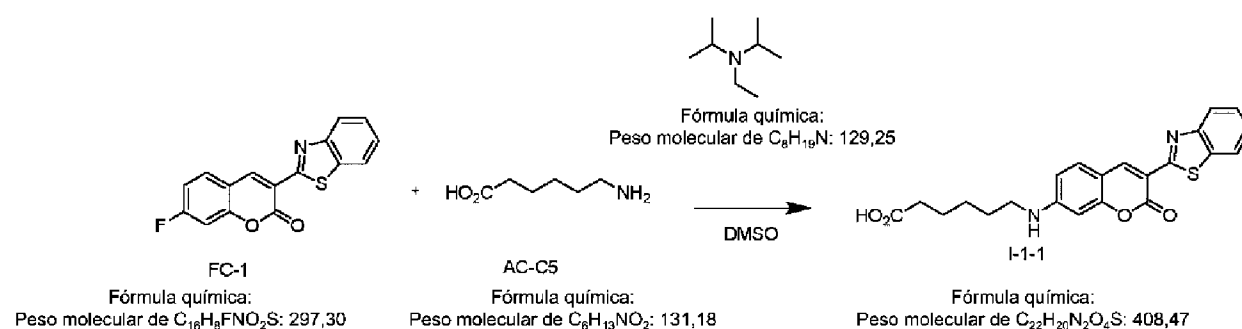
Se reconocerá que los nucleótidos modificados etiquetados con los compuestos de colorante de la presente descripción que incorporan grupos de bloqueo 3' también pueden ser útiles en los métodos de Sanger y protocolos relacionados ya que el mismo efecto conseguido utilizando dideoxinucleótidos modificados puede lograrse utilizando

nucleótidos modificados que tienen grupos de bloqueo de 3'-OH: ambos impiden la incorporación de nucleótidos posteriores. Cuando se vayan a utilizar nucleótidos según la presente descripción que tengan un grupo bloqueante 3' en métodos de secuenciación de tipo Sanger, se apreciará que los compuestos de colorante o las etiquetas detectables unidas a los nucleótidos no necesitan conectarse mediante enlazadores escindibles, ya que en cada caso en donde se incorpora un nucleótido etiquetado de la descripción; no es necesario incorporar ningún nucleótido posteriormente y, por lo tanto, no es necesario eliminar la etiqueta del nucleótido.

Ejemplos

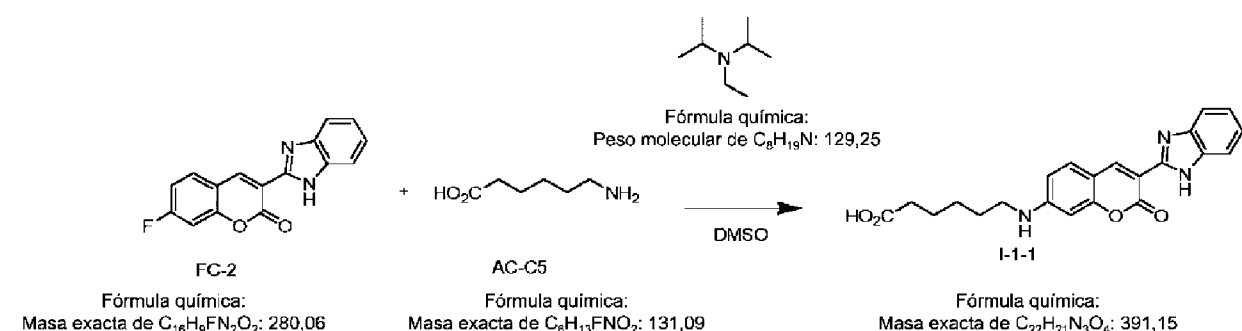
Se describen otras realizaciones con más detalle en los siguientes ejemplos, que no pretenden limitar el alcance de las reivindicaciones. En la Tabla 2 se resumen las propiedades espectrales de los nuevos colorantes fluorescentes de cumarina descritos en los ejemplos. La Tabla 3 resume la estructura y las propiedades espectrales de diversos nucleótidos etiquetados con los nuevos colorantes descritos en la presente memoria.

Ejemplo 1-1-1: 7-(5-carboxipentil)amino-3-(benzotiazol-2-il)cumarina



Se añadió el derivado 3-(benzotiazol-2-il)-7-fluoro-cumarina (FC-1, 0,4 g, 1,345 mmol, 1 eq.) y ácido 6-aminohexanoico (AC-C5, 0,25 g, 1,906 mmol, 1,417 eq.) a dimetilsulfóxido anhidro (DMSO, 3 ml). Una vez completada la adición, la mezcla se agitó durante unos minutos a temperatura ambiente y después se añadió N,N-diisopropil-N-etilamina (DIPEA, 0,25 g, 2 mmol, 2 eq.) a esta mezcla. La mezcla de reacción se agitó durante 3 horas a 120 °C. Tras reposar a temperatura ambiente durante 1 hora, la mezcla de reacción semisólida amarilla se diluyó con agua (5 ml) y se agitó durante la noche. El precipitado resultante se recogió mediante filtración por succión. Rendimiento 0,36 g (65,5 %). MS (DUIS): PM calculado 408,47. Encontrado m/z: (+) 409 (M+1)⁺; (-), 407 (M-1)⁻. ¹H NMR (400 MHz, DMSO-*d*₆) δ: 12,03 (m, 2H), 9,00 (s, 1H), 8,12 (d, *J* = 7,9 Hz, 1H), 7,99 (d, *J* = 8,1 Hz, 1H), 6,73 (dd, *J* = 8,8, 2,1 Hz, 1H), 6,54 (d, *J* = 2,0 Hz, 1H), 3,18 (q, *J* = 6,5 Hz, 2H), 2,23 (t, *J* = 7,3 Hz, 2H), 1,57 (dp, *J* = 14,7, 7,2 Hz, 4H), 1,39 (dq, *J* = 9,2, 4,5, 3,5 Hz, 2H).

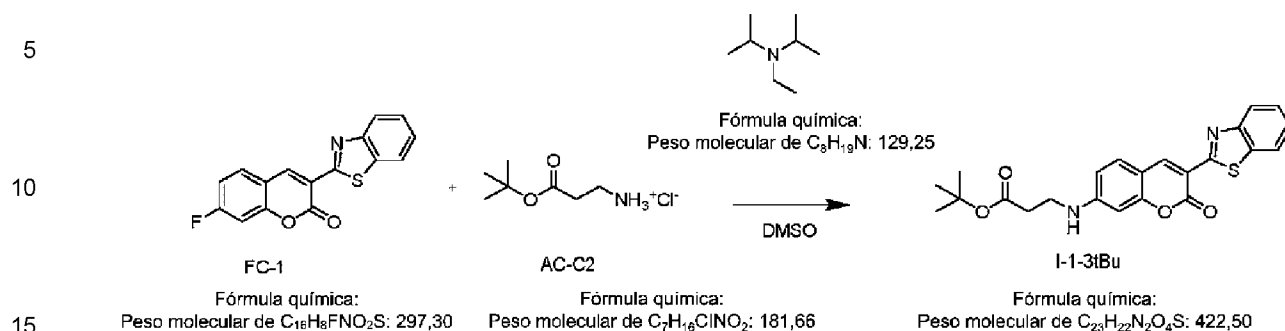
Ejemplo 1-1-2: 7-(5-carboxipentil)amino-3-(bencimidazol-2-il)cumarina



Se añadió 3-(Benzimidazol-2-il)-7-fluoro-cumarina (FC-2, 0,28 g, 1 mmol, 1 eq.) y ácido 6-aminohexanoico (AC-C5, 0,13 g, 1 mmol, 1 eq.) a dimetilsulfóxido anhidro (DMSO, 2 ml). La mezcla resultante se agitó durante unos minutos a temperatura ambiente y después se añadió DIPEA (0,25 g, 2 mmol, 2 eq.). La mezcla de reacción se agitó durante 4 horas a una temperatura de 130 °C. Se añadieron porciones adicionales de ácido 6-aminohexanoico (AC-1, 0,13 g, 1 mmol, 1 eq.) y DIPEA (0,26 g, 2 mmol, 2 eq.) a la mezcla de reacción y el calentamiento continuó a 130 °C durante 5 horas. Tras reposar a temperatura ambiente durante 1 hora, la mezcla de reacción de color amarillo pálido se diluyó con agua (5 ml) y se agitó durante la noche. El precipitado resultante se recogió mediante filtración por succión. Rendimiento 0,26 g (68,5 %). La pureza, la estructura y la composición del producto se confirmaron mediante HPLC, RMN y LCMS. MS (DUIS): PM calculado 391,15. Encontrado m/z: (+) 392 (M+1)⁺; (-) 390 (M-1)⁻, 781 (2M-1)⁻.

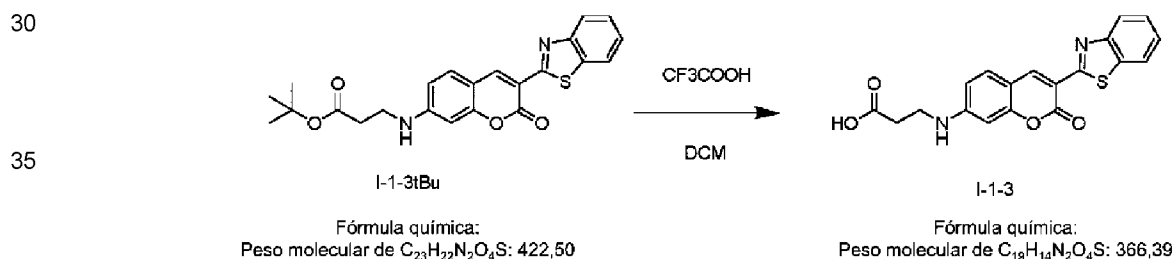
Ejemplo 1-1-3: 7-(2-carboxietil)amino-3-(benzotiazol-2-il)cumarina

EtapA A: 7-[2-(t-butiloxicarbonil)etil]amino-3-(benzotiazol-2-il)cumarina.



Se añadieron 3-(benzotiazol-2-il)-7-fluorocumarina (FC-1, 0,3 g, 1,01 mmol, 1 eq.) y clorhidrato de 3-aminopropionato de t-butilo (AC-C2, 0,2 g, 1,1 mmol, 1,09 eq.) a sulfóxido de dimetilo anhidro (DMSO, 2 ml) y la mezcla resultante se agitó durante unos minutos a temperatura ambiente y luego se añadió DIPEA (0,26 g, 2 mmol, 2 eq.). La mezcla resultante se agitó durante 2 horas a 100 °C. Tras reposar a temperatura ambiente durante 1 hora, la mezcla de reacción amarilla se diluyó con agua (7 ml) y se agitó durante la noche. El precipitado resultante se recogió mediante filtración por succión. Rendimiento 0,38 g (69 %). MS (DUIS): PM calculado 422,13. Encontrado m/z: (+) 423 (M+1)⁺; (-), 421 (M-1)⁻. ¹H NMR (400 MHz, DMSO-*d*₆) δ: 9,28 (s, 1H), 9,01 (s, 1H), 8,27 - 8,16 (m, 1H), 8,10 (tt, *J* = 8,3, 0,9 Hz, 2H), 8,05 - 7,92 (m, 1H), 7,72 (d, *J* = 8,8 Hz, 1H), 7,66-7,55 (m, 1H), 7,51 (dddd, *J* = 11,4, 8,2, 7,1, 1,3 Hz, 2H), 7,46 - 7,32 (m, 2H), 6,74 (dd, *J* = 8,7, 2,1 Hz, 1H), 6,58 (d, *J* = 2,1 Hz, 1H), 3,41 (q, *J* = 6,3 Hz, 2H), 2,55 (t, *J* = 6,4 Hz, 2H), 1,41 (s, 9H).

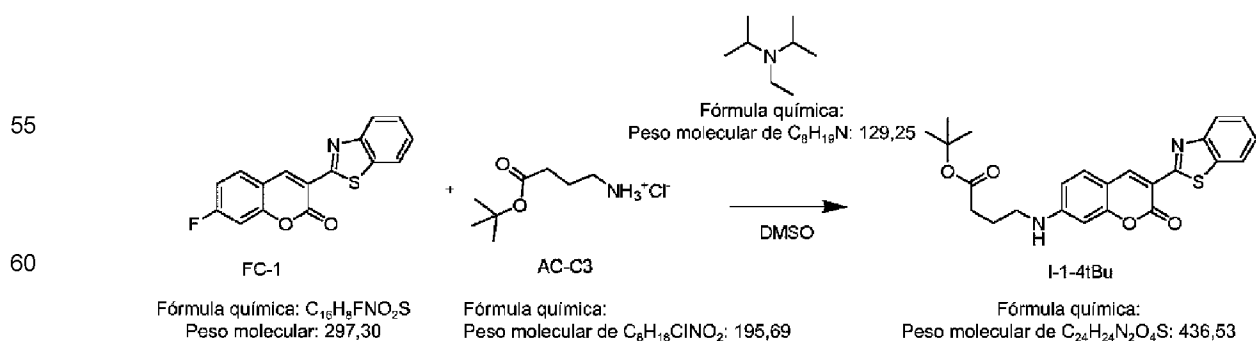
EtapA B.



Una solución de 7-[2-(t-butiloxicarbonil)etil]amino-3-(benzotiazol-2-il)cumarina (I-1-3tBu, 0,2 g, 0,473 mmol) en diclorometano anhidro (20 ml) se trató con ácido trifluoroacético (0,5 ml) y la mezcla resultante se agitó durante 24 horas a temperatura ambiente. Los disolventes se separaron por destilación y el residuo se trituró con agua (10 ml). El precipitado resultante se recogió mediante filtración por succión. Rendimiento 0,15 g (86 %). Se confirmaron la pureza, la estructura y la composición mediante HPLC, RMN y LCMS. MS (DUIS): PM calculado 366,39. Encontrado m/z: (+) 367 (M+1)⁺; (-), 365 (M-1)⁻.

Ejemplo 1-1-4: 7-(3-carboxipropil) amino-3-(benzotiazol-2-il)cumarina

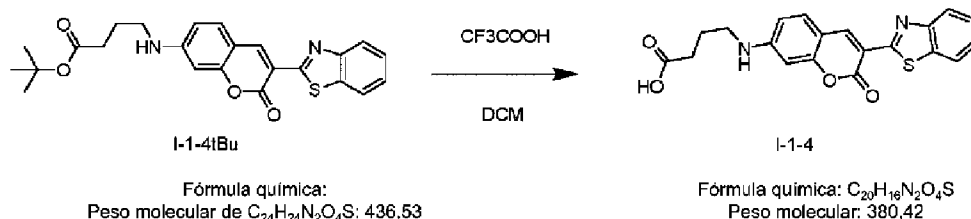
EtapA A: 7-[3-(t-butiloxicarbonil)propil]amino-3-(benzotiazol-2-il)cumarina.



Se añadieron 3-(benzotiazol-2-il)-7-fluoro-cumarina (FC-1, 0,6 g, 2,02 mmol, 1 eq.) y clorhidrato de 4-aminobutanoato de t-butilo (AC-C3, 0,5 g, 2,56 mmol, 1,27 eq.) a sulfóxido de dimetilo anhidro (DMSO, 5 ml). Una vez completada la

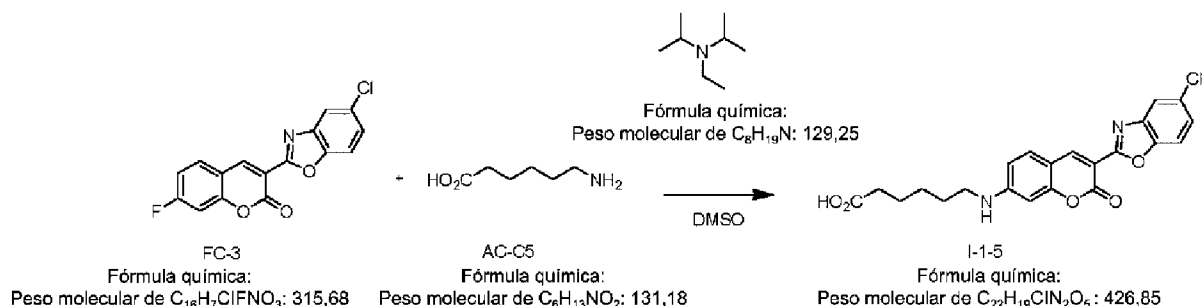
adición, la mezcla se agitó durante unos minutos a temperatura ambiente y después se añadió DIPEA (0,65 g, 5 mmol, 4 eq.). La mezcla de reacción se agitó durante 3 horas a una temperatura de 100 °C. Tras reposar a temperatura ambiente durante 1 hora, la mezcla de reacción semisólida amarilla se diluyó con agua (10 ml) y se dejó agitar durante la noche. El precipitado resultante se recogió mediante filtración por succión. Rendimiento 0,7 g (79 %). Se confirmaron la pureza, la estructura y la composición mediante HPLC, RMN y LCMS. MS (DUIS): PM calculado 436,53. Encontrado m/z: (+) 437 (M+1)⁺; (-), 435 (M-1)⁻.

Etapa B.



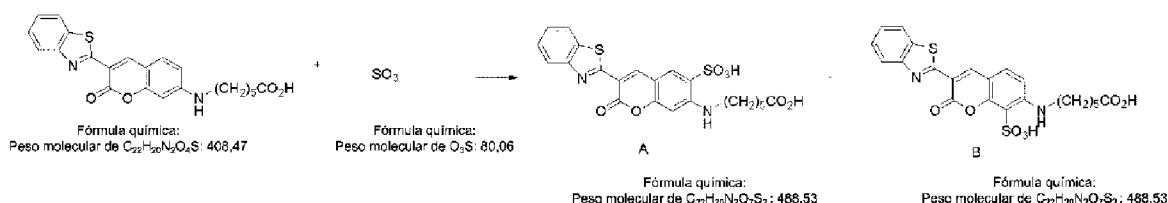
Una solución de 7-[3-(t-butiloxycarbonil)propil]amino-3-(benzotiazol-2-il)cumarina (I-1-4tBu, 0,7 g, 1,604 mmol) en diclorometano anhidro (25 ml) se trató con ácido trifluoroacético (1 ml) y la mezcla de reacción se agitó durante 24 horas a temperatura ambiente. Los disolventes se separaron por destilación y el residuo se trituró con agua (10 ml). El precipitado resultante se recogió mediante filtración por succión. Rendimiento 0,59 g (97 %). Se confirmaron la pureza, la estructura y la composición mediante HPLC, RMN y LCMS. MS (DUIS): PM calculado 366,39. Encontrado m/z: (+) 381 (M+1)⁺; (-), 379 (M-1)⁻. ¹H NMR (400 MHz, DMSO-*d*₆) δ: 12,17 (s, 1H), 9,01 (s, 1H), 8,12 (d, *J* = 8,0 Hz, 1H), 7,99 (d, *J* = 8,1 Hz, 1H), 7,71 (d, *J* = 8,8 Hz, 1H), 7,48 - 7,30 (m, 2H), 6,73 (dd, *J* = 8,8, 2,1 Hz, 1H), 6,57 (d, *J* = 2,1 Hz, 1H), 3,21 (q, *J* = 6,6 Hz, 2H), 2,36 (d, *J* = 7,3 Hz, 2H), 1,80 (p, *J* = 7,3 Hz, 2H).

Ejemplo 1-1-5: 7-(5-carboxipentil)amino-3-(5-cloro-benzoxazol-2-il)cumarina



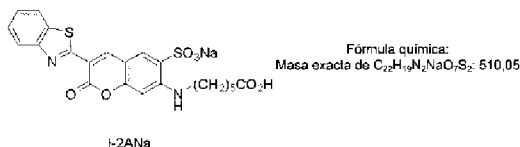
Se añadieron 3-(5-cloro-benzoxazol-2-il)-7-fluoro-cumarina (FC-3, 0,32 g, 1 mmol, 1 eq.) y ácido 6-aminohexanoico (AC-C5, 0,26 g, 2 mmol, 2 eq.) a sulfóxido de dimetilo anhidro (DMSO, 5 ml) en un matraz de fondo redondo. Una vez completada la adición, la mezcla se agitó durante unos minutos a temperatura ambiente y después se añadió DIPEA (0,52 g, 4 mmol, 2 eq.). La mezcla de reacción se agitó durante 7 horas a una temperatura de 135 °C. Se añadieron porciones adicionales de ácido 6-aminohexanoico (AC-1, 0,13 g, 1 mmol, 1 eq.) y DIPEA (0,26 g, 2 mmol, 2 eq.) y se continuó calentando a 135 °C durante 5 horas. Tras reposar a temperatura ambiente durante 1 hora, la mezcla de reacción de color amarillo pálido se diluyó con agua (15 ml) y se agitó durante la noche. El precipitado resultante se recogió mediante filtración por succión. Rendimiento 0,09 g (21 %). La pureza, la estructura y la composición del producto se confirmaron mediante HPLC, RMN y LCMS. MS (DUIS): PM calculado 426,10. Encontrado m/z: (+) 427 (M+1)⁺; (-) 425 (M-1)⁻, 851 (2M-1)⁻.

Ejemplo 2: Ácido 7-(5-carboxipentil)amino-3-(benzotiazol-2-il)cumarina-6-sulfónico (2A) y ácido 7-(5-carboxipentil)amino-3-(benzotiazol-2-il)cumarina-8-sulfónico (2B)



El compuesto 1-1-1 (0,1 g, 0,245 mmol) se añadió en pequeñas porciones con agitación a ácido sulfúrico fumante al 20 % (1 ml) que se enfrió en un baño de hielo seco/acetona. Una vez completada la adición, la mezcla se agitó durante

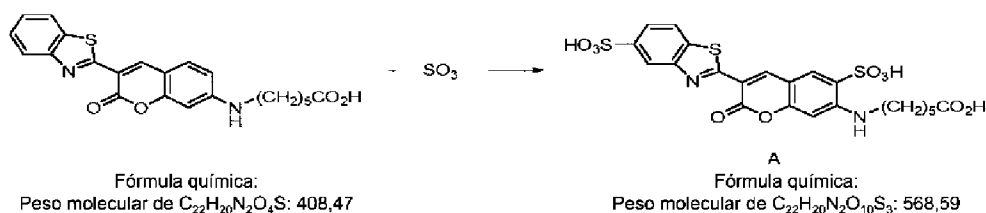
1 hora a 0 °C, se calentó hasta temperatura ambiente, y después se agitó durante 2 horas a temperatura ambiente. La solución se vertió en éter anhidro (25 ml). Tras dejar reposar a temperatura ambiente durante 1 hora, el precipitado resultante se recogió mediante filtración por succión. Rendimiento 78 mg (65 %). ¹H NMR (d₆-DMSO) mostró el compuesto 2A más una pequeña cantidad (~4 %) del compuesto 2B.



Ejemplo 2A, sal de sodio: El precipitado anterior se volvió a suspender en agua (2 ml) y el pH de la suspensión se ajustó a ~5 mediante la adición de una solución de 5 M NaOH. La mezcla resultante se vertió en 10 ml de metanol y la suspensión se filtró. El filtrado se evaporó hasta sequedad para dar el colorante en forma de sal de sodio (I-2A-Na). Se confirmaron la pureza, la estructura y la composición mediante HPLC, RMN y LCMS. MS (DUIS): PM calculado 488,07. Encontrado m/z: (+) 489 (M+1)⁺; (-) 243 (M-1)²⁻, 487 (M-1)⁻.

Preparación de sales de trietilamonio de 2A y 2B: El compuesto I-1-1 (0,41 g, 1 mmol) se añadió en pequeñas porciones con agitación a ácido sulfúrico fumante al 20 % (5 ml) que se enfrió en un baño de hielo seco/acetona. Una vez completada la adición, la mezcla se agitó durante 1 hora a 0 °C, se calentó hasta temperatura ambiente, y después se agitó durante 2 horas a temperatura ambiente. La solución se vertió en éter anhidro (50 ml). Tras reposar a temperatura ambiente durante 1 hora, la capa de disolvente orgánico se decanta y la capa inferior semisólida se disuelve en acetonitrilo-agua (1:1, 10 ml). El pH de la solución se ajustó a ~7,0 mediante la adición de una solución de 2M TEAB en agua. La solución resultante se filtró a través de un filtro de nailon de 20 µm y los isómeros se separaron mediante HPLC preparativa. La solución de los isómeros se concentró *al vacío*, después se volvió a disolver en agua (20 ml) y el disolvente se retiró *al vacío* hasta sequedad para dar los colorantes en forma de sales de trietilamonio. Se confirmaron la pureza y la composición mediante HPLC y LCMS.

Ejemplo 3: Sal de trietilamonio de 7-(5-carboxipentil)amino-3-[5-sulfonato(benzotiazol-2-il)-cumarina-6-sulfonato



El compuesto I-1-1 (0,08 g, 0,2 mmol) se añadió en pequeñas porciones con agitación a ácido sulfúrico fumante al 20 % (2 ml) que se enfrió en un baño de hielo seco/acetona. Una vez completada la adición, la mezcla se agitó durante 1 hora a 0 °C, se calentó hasta temperatura ambiente, y después se agitó durante 2 horas a 70 °C. La mezcla se agitó después durante la noche a temperatura ambiente. La solución se vertió en éter anhidro (30 ml). Tras agitar a temperatura ambiente durante 1 hora, el precipitado resultante se recogió mediante filtración por succión. Rendimiento 43 mg (38 %).

El precipitado se volvió a suspender en agua (2 ml) y el pH de la suspensión se ajustó a ~7,5 mediante la adición de una solución de 2M TEAB en agua. La mezcla resultante se filtró a través de un filtro de nailon de 20 µm y se purificó mediante HPLC preparativa. La fracción de colorante se concentró *al vacío*, después se volvió a disolver en agua (20 µl) y el disolvente se retiró *al vacío* hasta sequedad para dar el colorante en forma de la sal de bis-trietilamonio. Se confirmaron la pureza y la composición mediante HPLC y LCMS. MS (DUIS): PM calculado 568,03. Encontrado m/z: (+) 569 (M+ 1)⁺.

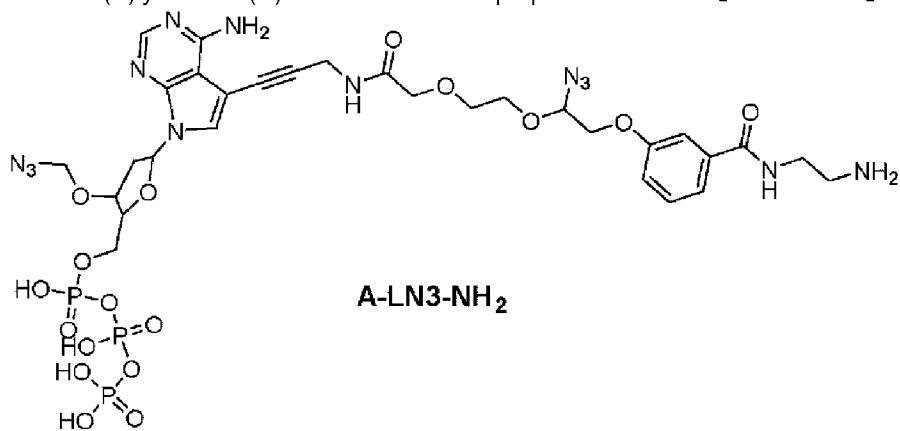
Las intensidades de fluorescencia de las soluciones de colorante ilustrativas se compararon con las de un colorante comercial para la misma región espectral. Los resultados se muestran en la Tabla 2 y demuestran las ventajas significativas de los colorantes ilustrativas para aplicaciones analíticas basadas en fluorescencia.

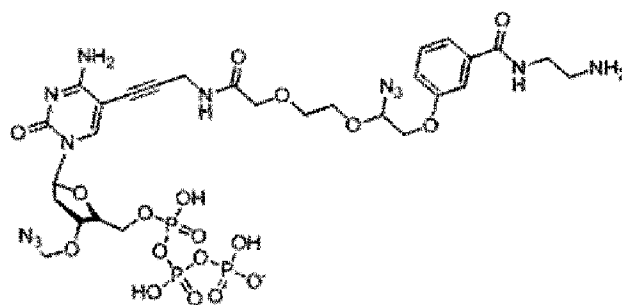
Tabla 2. Propiedades espectrales de los nuevos colorantes fluorescentes descritos en los ejemplos.

Número	Estructura	Propiedades espectrales en EtOH-agua 1:1		
		Absorción máxima nm	Fluorescencia máxima nm	Flúor relativo.* Intensidad, %
I-1-1		460	499	275
I-1-2		437	488	175
I-1-3		453	499	230
I-1-4		455	500	220
I-1-5		430	490	200
I-2A		465	503	395
Estándar	Atto465 de AttoTec	455	508	100
*Excitación de fluorescencia a 460 nm				

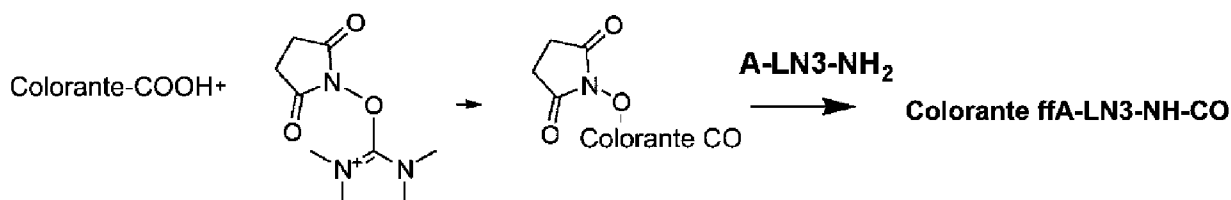
Ejemplo 4: Procedimiento general para la síntesis de conjugados de nucleótidos completamente funcionales con nuevos colorantes fluorescentes

Los colorantes fluorescentes de cumarina descritos en la presente memoria se acoplaron con los derivados nucleotídicos de adenina (A) y citosina (C) amino-sustituídos apropiados A-LN3-NH₂ o C-LN3-NH₂:

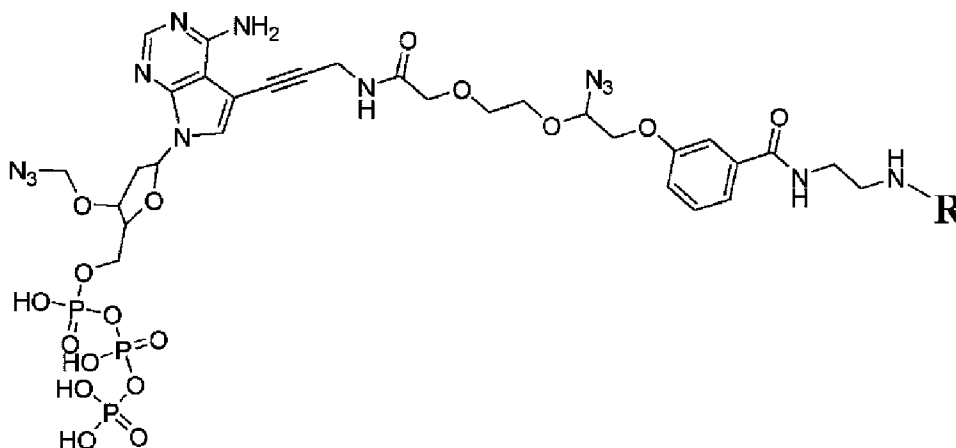


**C-LN3-NH₂**

después de la activación del grupo carboxílico de un colorante con los reactivos apropiados según el siguiente esquema ilustrativo de adenina:



El producto general para el acoplamiento de adenina es el que se muestra a continuación:

**Colorante ffA-LN3**

El colorante ffA-LN3 se refiere a un nucleótido A completamente funcionalizado con un enlazador LN3 y etiquetado con un colorante de cumarina descrito en la presente memoria. El grupo R de cada una de las estructuras se refiere al resto del colorante a base de cumarina después de la conjugación.

El colorante (10 μ mol) se seca colocándolo en un matraz de fondo redondo de 5 ml y se disuelve en dimetilformamida anhidra (DMF, 1 ml) y después el disolvente se separa por destilación *al vacío*. Este procedimiento se repite dos veces. El colorante seco se disuelve en N,N-dimetilacetamida (DMA, 0,2 ml) anhidra a temperatura ambiente. Se añade tetrafluoroborato de N,N,N',N'-Tetrametil-O-(N-succinimidil)uronio (TSTU, 1,5 eq., 15 μ mol, 4,5 mg) a la solución de colorante y, a continuación, se añade DIPEA (3 eq., 30 μ mol, 3,8 mg, 5,2 μ l) mediante micropipeta a esta solución. El matraz de la reacción se cierra herméticamente con gas nitrógeno. El progreso de la reacción se controla mediante TLC (eluyente Acetonitrilo-Agua 1:9) y HPLC. Mientras tanto, se concentra *al vacío* una solución del derivado nucleotídico amino-sustituido apropiado (A-LN3-NH₂, 20 mM, 1,5 eq., 15 μ mol, 0,75 ml) y se vuelve a disolver en agua (20 μ l). Una solución del colorante activado en DMA se transfirió al matraz que contenía la solución de N-LN3-NH₂. Más DIPEA (3 eq., 30 μ mol, 3,8 mg, 5,2 μ l) se añade junto con trietilamina (1 μ l). El progreso del acoplamiento se monitoriza cada hora mediante TLC, HPLC y LCMS. Cuando se completa la reacción, se añade un regulador de bicarbonato de trietilamina (TEAB, 0,05 M ~ 3 ml) a la mezcla de reacción mediante una pipeta. La purificación inicial del nucleótido completamente funcionalizado se lleva a cabo haciendo pasar la mezcla de reacción inactivada a través de una columna de DEAE-Sephadex® para eliminar la mayor parte del colorante restante sin reaccionar. Por ejemplo, el Sephadex se vierte en un cartucho Biotage vacío de 25 g, sistema disolvente TEAB/MeCN. Se concentra la solución

de la columna Sephadex *al vacío*. El material restante se volvió a disolver en el volumen mínimo de agua y acetonitrilo, antes de filtrarlo a través de un filtro de nailon de 20 µm. La solución filtrada se purifica mediante HPLC preparativa. La composición de los compuestos preparados se confirmó mediante LCMS.

- 5 Tabla 3. Estructura y propiedades espectrales de diversos nucleótidos etiquetados con nuevos colorantes a base de cumarina descritos en la presente memoria.

Compuesto	Propiedades espectrales en SRE		
	Absorción, nm	Fluorescencia, nm	Flúor relativo. Intensidad, %
ffA-I-1-1	448	505	480
ffA-I-1-3	454	499	500
ffA-I-2A	475	510	575
ffA-Estándar	465	504	100

Una comparación de las intensidades de fluorescencia en solución de nucleótidos etiquetados con colorantes descritos en la presente memoria con los datos apropiados para los nucleótidos etiquetados con un colorante comercial para la misma región espectral (Atto465 de AttoTec GmbH) demuestra la ventaja de los colorantes descritos en la presente memoria para el etiquetado de biomoléculas para usarse en aplicaciones analíticas basadas en la fluorescencia.

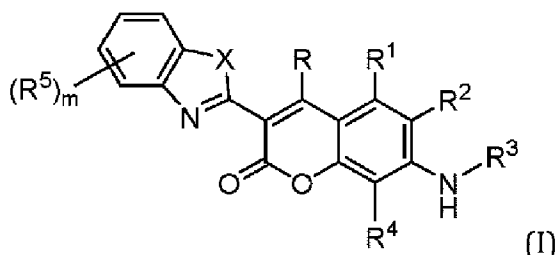
Ejemplo 5: Análisis de secuenciación

Algunos ffA basados en colorantes descritos en la presente memoria se han probado para la secuenciación (usando química de secuenciación por síntesis) y las imágenes de las Figuras 1A, 1B y 1C demuestran su utilidad. Los nucleótidos A, C, G, y T se etiquetaron del modo siguiente:

A: BL-NR550S0; C: NR440; G: Oscuro; y T: AF550POPOS0. Instrumento: M615, temperatura de escaneo 22 °C, tiempos de exposición: azul 500 ms, verde 500 ms; Celda de flujo SFA. Los números entre paréntesis en los gráficos siguientes indican la relación entre el nuevo ffA azul y el verde A (A-BL-NR550S0).

REIVINDICACIONES

1. Un compuesto de fórmula (I), o una sal del mismo:



en donde:

X es O , S , Se o NR^n , donde R^n es H o alquilo C_{1-6} ;

R es H , halo, $-CN$, $-NR_AR_B$, $-C(=O)NR_AR_B$ o $-N(R_A)C(=O)R_B$, en los que R_A y R_B se seleccionan cada uno independientemente de hidrógeno, alquilo C_{1-6} , alquenilo C_{2-6} , alquinilo C_{2-6} , carbociclilo C_{3-7} , arilo C_{6-10} , heteroarilo de 5-10 miembros y heterociclilo de 3-10 miembros, $-OH$, $-NO_2$, $-SO_3H$, $-SO_2NH_2$, alquilo opcionalmente sustituido, alquenilo opcionalmente sustituido, alquinilo opcionalmente sustituido, alcoxi opcionalmente sustituido, aminoalquilo opcionalmente sustituido, carbociclilo opcionalmente sustituido, heterociclilo opcionalmente sustituido, arilo opcionalmente sustituido o heteroarilo opcionalmente sustituido;

R^1 es H , halo, $-CN$, $-CO_2H$, NR_AR_B , $-C(=O)NR_AR_B$, o $-N(R_A)C(=O)R_B$, en los que R_A y R_B se seleccionan cada uno independientemente de hidrógeno, alquilo C_{1-6} , alquenilo C_{2-6} , alquinilo C_{2-6} , carbociclilo C_{3-7} , arilo C_{6-10} , heteroarilo de 5-10 miembros y heterociclilo de 3-10 miembros, $-OH$, $-NO_2$, $-SO_3H$, $-SO_2NH_2$, alquilo opcionalmente sustituido, alquenilo opcionalmente sustituido, alquinilo opcionalmente sustituido, alcoxi opcionalmente sustituido, aminoalquilo opcionalmente sustituido, carbociclilo opcionalmente sustituido, heterociclilo opcionalmente sustituido, arilo opcionalmente sustituido o heteroarilo opcionalmente sustituido;

R^2 y R^4 son cada uno independientemente H , halo, $-CN$, $-CO_2H$, NR_AR_B , $-C(=O)NR_AR_B$, o $-N(R_A)C(=O)R_B$, en los que R_A y R_B se seleccionan cada uno independientemente de hidrógeno, alquilo C_{1-6} , alquenilo C_{2-6} , alquinilo C_{2-6} , carbociclilo C_{3-7} , arilo C_{6-10} , heteroarilo de 5-10 miembros y heterociclilo de 3-10 miembros, $-OH$, $-NO_2$, $-SO_3H$, $-SO_2NH_2$, alquilo opcionalmente sustituido, alquenilo opcionalmente sustituido, alquinilo opcionalmente sustituido, alcoxi opcionalmente sustituido, aminoalquilo opcionalmente sustituido, carbociclilo opcionalmente sustituido, heterociclilo opcionalmente sustituido, arilo opcionalmente sustituido o heteroarilo opcionalmente sustituido;

R^3 es alquilo C_{2-6} sustituido con $-CO_2H$ o SO_3H , o heterociclilo opcionalmente sustituido;

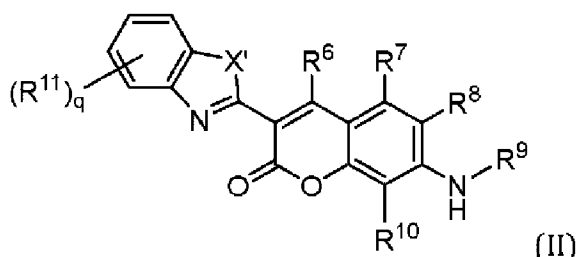
cada R^5 es independientemente halo, $-CN$, $-CO_2H$, NR_AR_B , $-C(=O)NR_AR_B$, o $-N(R_A)C(=O)R_B$, en los que R_A y R_B se seleccionan cada uno independientemente de hidrógeno, alquilo C_{1-6} , alquenilo C_{2-6} , alquinilo C_{2-6} , carbociclilo C_{3-7} , arilo C_{6-10} , heteroarilo de 5-10 miembros y heterociclilo de 3-10 miembros, $-OH$, $-NO_2$, $-SO_3H$, $-SO_2NH_2$, alquilo opcionalmente sustituido, alquenilo opcionalmente sustituido, alquinilo opcionalmente sustituido, alcoxi opcionalmente sustituido, aminoalquilo opcionalmente sustituido, carbociclilo opcionalmente sustituido, heterociclilo opcionalmente sustituido, arilo opcionalmente sustituido o heteroarilo opcionalmente sustituido; y m es 0, 1, 2, 3 o 4.

2. El compuesto de la reivindicación 1, en donde R^2 y R^4 son cada uno independientemente H , halo, $-CN$, $-CO_2H$, NR_AR_B , $-C(=O)NR_AR_B$, o $-N(R_A)C(=O)R_B$, en los que R_A y R_B se seleccionan cada uno independientemente de hidrógeno, alquilo C_{1-6} , alquenilo C_{2-6} , alquinilo C_{2-6} , carbociclilo C_{3-7} , arilo C_{6-10} , heteroarilo de 5-10 miembros y heterociclilo de 3-10 miembros, $-OH$, $-NO_2$, $-SO_3H$, $-SO_2NH_2$, alquilo opcionalmente sustituido, alquenilo opcionalmente sustituido, alquinilo opcionalmente sustituido, alcoxi opcionalmente sustituido, aminoalquilo opcionalmente sustituido, carbociclilo opcionalmente sustituido, heterociclilo opcionalmente sustituido, arilo opcionalmente sustituido o heteroarilo opcionalmente sustituido.
3. El compuesto de una cualquiera de la reivindicación 1 o la reivindicación 2, en donde X es O o S , o en donde X es NR^n , donde R^n es H o alquilo C_{1-6} .
4. El compuesto de una cualquiera de las reivindicaciones 1 a 3, en donde R^3 es $-(CH_2)_nCOOH$, donde n es 2, 3, 4, 5 o 6.
5. El compuesto de una cualquiera de las reivindicaciones 1 a 4, en donde R^3 es $-(CH_2)_nSO_3H$ en donde n es 2, 3, 4, 5 o 6.

6. El compuesto de una cualquiera de las reivindicaciones 1-5, en donde m es 0, 1 o 2.
7. El compuesto de una cualquiera de las reivindicaciones 1-6, en donde cada R⁵ es independientemente alquilo C₁₋₆ opcionalmente sustituido, halo, -CN, -CO₂H, NR_AR_B, en los que R_A y R_B se seleccionan cada uno independientemente de hidrógeno, alquilo C₁₋₆, alqueniilo C₂₋₆, alquinilo C₂₋₆, carbociclilo C₃₋₇, arilo C₆₋₁₀, heteroarilo de 5-10 miembros y heterociclilo de 3-10 miembros, -OH, -SO₃H, o -SO₂NH₂, opcionalmente en donde cada R⁵ es halo, -CN, -CO₂H, -SO₃H, -SO₂NH₂, o alquilo C₁₋₆ opcionalmente sustituido, opcionalmente en donde R⁵ es alquilo C₂₋₆ sustituido con -CO₂H, -SO₃H, o -SO₂NH₂, opcionalmente en donde R⁵ es -(CH₂)_xCOOH, donde x es 2, 3, 4, 5 o 6, u opcionalmente, en donde R⁵ es -SO₃H o -SO₂NH₂.
8. El compuesto de una cualquiera de las reivindicaciones 1-7, en donde R¹ es H, halo o alquilo C₁₋₆, opcionalmente en donde R¹ es H o Cl.
9. El compuesto de una cualquiera de las reivindicaciones 1-8, en donde R es H, halo o alquilo C₁₋₆, opcionalmente en donde R es H o Cl, u opcionalmente en donde R es H.
10. El compuesto de una cualquiera de las reivindicaciones 1-9, en donde a) R² es H, alquilo opcionalmente sustituido, alquilo C₁₋₄ opcionalmente sustituido con -CO₂H o -SO₃H, o -SO₃H, u opcionalmente en donde R² es H o -SO₃H.
11. El compuesto de una cualquiera de las reivindicaciones 1-10, en donde R⁴ es H, alquilo opcionalmente sustituido, alquilo C₁₋₄ opcionalmente sustituido con -CO₂H o -SO₃H o -SO₃H.
12. El compuesto de la reivindicación 1, en donde X es 0 o S; R es H; R¹ es H; R³ es -(CH₂)_nCOOH donde n es 2, 3, 4, 5, o 6; R⁵ es H, -SO₃H, o -SO₂NH₂; R² es H o -SO₃H; y R⁴ es H o -SO₃H, opcionalmente

a) en donde R³ es -(CH₂)₂COOH, o
b) en donde R³ es -(CH₂)₅COOH.

13. Un compuesto de fórmula (II), o una sal del mismo:



en donde:

- X' se selecciona de 0, S y NRP, donde RP es H o alquilo C₁₋₆;
R⁶ es H o alquilo C₁₋₄;
R⁷ es H, halo, -CN, -OH, alquilo C₁₋₄ opcionalmente sustituido, alqueniilo C₁₋₄ opcionalmente sustituido, alquinilo C₂₋₄ opcionalmente sustituido, -CO₂H, -SO₃H, -SO₂NH₂, -SO₂NH(alquilo C₁₋₄), -SO₂N(alquilo C₁₋₄)₂ y alcoxi C₁₋₄ opcionalmente sustituido;
R⁸ y R¹⁰ son cada uno independientemente H, halo, -CN, -CO₂H, NR_AR_B en los que R_A y R_B se seleccionan cada uno independientemente de hidrógeno, alquilo C₁₋₆, alqueniilo C₂₋₆, alquinilo C₂₋₆, carbociclilo C₃₋₇, arilo C₆₋₁₀, heteroarilo de 5-10 miembros y heterociclilo de 3-10 miembros, -OH, -SO₃H, -SO₂NH₂, -SO₂NH(alquilo C₁₋₄), -SO₂N(alquilo C₁₋₄)₂, alquilo C₁₋₆ opcionalmente sustituido, alqueniilo C₁₋₆ opcionalmente sustituido, alquinilo C₂₋₆ opcionalmente sustituido o alcoxi C₁₋₆ opcionalmente sustituido;
R⁹ es alquilo C₂₋₆ sustituido con -CO₂H, -CO₂alquilo C₁₋₄, -CONH₂, -CONH(C₁₋₄), -CON(alquilo C₁₋₄)₂, -CN, -SO₃H, -SO₂NH₂, -SO₂NH(alquilo C₁₋₄) o -SO₂N(alquilo C₁₋₄)₂;
cada R¹¹ es independientemente halo, -CN, carboxi, NR_AR_B, -C(=O)NR_AR_B o -N(R_A)C(=O)R_B, en los que R_A y R_B se seleccionan cada uno independientemente de hidrógeno, alquilo C₁₋₆, alqueniilo C₂₋₆, alquinilo C₂₋₆, carbociclilo C₃₋₇, arilo C₆₋₁₀, heteroarilo de 5-10 miembros y heterociclilo de 3-10 miembros, -OH, nitro, -SO₃H, -SO₂NH₂, -SO₂NH(alquilo C₁₋₄), -SO₂N(alquilo C₁₋₄)₂, alquilo opcionalmente sustituido, alqueniilo opcionalmente sustituido, alquinilo opcionalmente sustituido y alcoxi C₁₋₆ opcionalmente sustituido; y q es 0, 1 o 2.
14. El compuesto de la reivindicación 13, en donde X' es 0 o X' es S, o en donde X' es NRP, en donde RP es H o alquilo C₁₋₆.

15. El compuesto de la reivindicación 13, en donde al menos uno de:

i) R^6 es H o R^6 es alquilo C_{1-4} ; y

ii) R^7 es H o R^7 es alquilo C_{1-4} opcionalmente sustituido, $-CO_2H$, $-SO_3H$, $-SO_2NH_2$, $-SO_2NH$ (alquilo C_{1-4}), o $-SO_2N$ (alquilo C_{1-4})₂, opcionalmente en donde R^7 es alquilo C_{1-4} opcionalmente sustituido con $-CO_2H$.

16. El compuesto de la reivindicación 13, en donde al menos uno de:

i) R^8 es H o R^8 es $-CO_2H$, $-SO_3H$, o $-SO_2NH_2$;

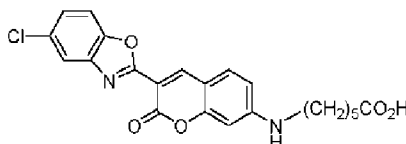
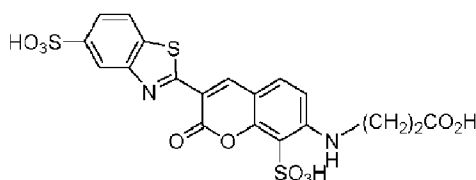
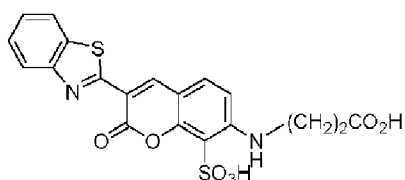
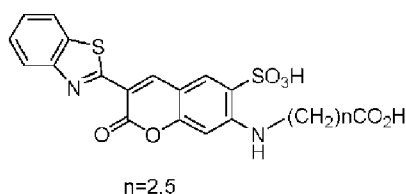
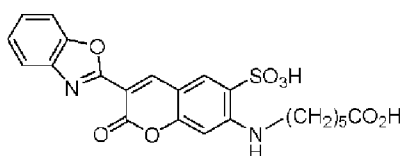
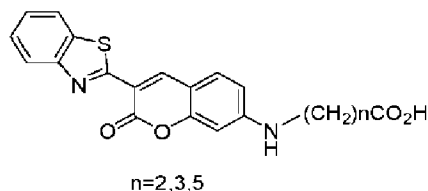
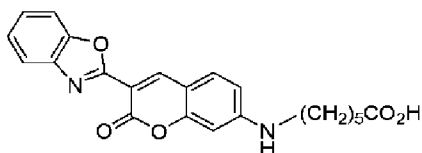
ii) R^{10} es H o es $-CO_2H$, $-SO_3H$, o $-SO_2NH_2$; y

iii) (a) R^8 es H y R^{10} es $-SO_3H$ o (b) R^8 es $-SO_3H$ y R^{10} es H.

17. El compuesto de la reivindicación 13, en donde R^9 es alquilo C_{2-6} sustituido con $-CO_2H$, opcionalmente donde R^9 es $-(CH_2)_y-CO_2H$, donde y es 2, 3, 4 o 5.

18. El compuesto de una cualquiera de las reivindicaciones 13-17, en donde cada R^{11} es independientemente halo, $-CO_2H$, $-SO_3H$, $-SO_2NH_2$, $-SO_2NH$ (alquilo C_{1-4}), $-SO_2N$ (alquilo C_{1-4})₂ o alquilo opcionalmente sustituido.

19. Un compuesto de la reivindicación 1, seleccionado del grupo que consiste en:



y sales del mismo.

20. Un nucleótido u oligonucleótido etiquetado con un compuesto según una cualquiera de las reivindicaciones 1-19.

21. El nucleótido u oligonucleótido etiquetado según la reivindicación 20, que comprende opcionalmente al menos uno de:

(A) en donde el compuesto está unido mediante R^3 de fórmula (I) o R^9 de fórmula (II), comprendiendo además opcionalmente en donde R^3 de fórmula (I) es $-(CH_2)_nCOOH$ donde n es 2, 3, 4, 5 o 6, o R^9 de fórmula (II) es $-(CH_2)_yCOOH$ donde y es 2, 3, 4, 5 o 6, y la unión forma un amida que utiliza el grupo $-COOH$,

(B) en donde la etiqueta está unida a la posición C5 de una base pirimidínica o a la posición C7 de una base de 7-deaza purina a través de un resto enlazador y

(C) un grupo bloqueante 3' OH unido covalentemente al azúcar ribosa o desoxirribosa del nucleótido.

22. Un kit que comprende un primer nucleótido etiquetado según la reivindicación 20 o la reivindicación 21 y un segundo nucleótido etiquetado, opcionalmente

i) en donde el segundo nucleótido etiquetado está etiquetado con un compuesto diferente al del primer nucleótido etiquetado o no está etiquetado, opcionalmente en donde el primer y el segundo nucleótidos etiquetados son excitables utilizando una única longitud de onda láser y/o,

ii) que comprende además un tercer nucleótido y un cuarto nucleótido, en donde cada uno del segundo, tercer y cuarto nucleótidos está etiquetado con un compuesto diferente o no está etiquetado, en donde cada etiqueta tiene un máximo de absorbancia distinto que es distinguible de las otras etiquetas, o

iii) en donde el kit comprende cuatro nucleótidos, en donde un primero de los cuatro nucleótidos es un nucleótido etiquetado según una cualquiera de las reivindicaciones 18 o 19, un segundo de los cuatro nucleótidos lleva una segunda etiqueta, un tercer nucleótido lleva una tercera etiqueta, y un cuarto nucleótido no está etiquetado (oscuro), o

iv) en donde el kit comprende cuatro nucleótidos, en donde un primero de los cuatro nucleótidos es un nucleótido etiquetado según una cualquiera de las reivindicaciones 18 o 19, un segundo de los cuatro nucleótidos lleva una segunda etiqueta, un tercer nucleótido lleva una mezcla de dos etiquetas y un cuarto nucleótido no está etiquetado (oscuro).

23. Uso de un nucleótido según una cualquiera de las reivindicaciones 1 a 19, un nucleótido o un oligonucleótido según una cualquiera de las reivindicaciones 20-21, o un kit según la reivindicación 22 en la secuenciación, el análisis de expresión, el análisis de hibridación, el análisis genético, el análisis de ARN o un ensayo de unión a proteínas,

24. Uso según la reivindicación 23 en un instrumento de secuenciación automática, en donde dicho instrumento de secuenciación automática comprende dos láseres que funcionan a diferentes longitudes de onda.

25. Un método de secuenciación que comprende incorporar un nucleótido según una cualquiera de las reivindicaciones 20-21 en un ensayo de secuenciación, que comprende opcionalmente además detectar el nucleótido, opcionalmente en donde el ensayo de secuenciación se realiza en un instrumento de secuenciación automática y en donde el instrumento de secuenciación automática comprende dos fuentes de luz que funcionan a diferentes longitudes de onda.

FIGURA 1A

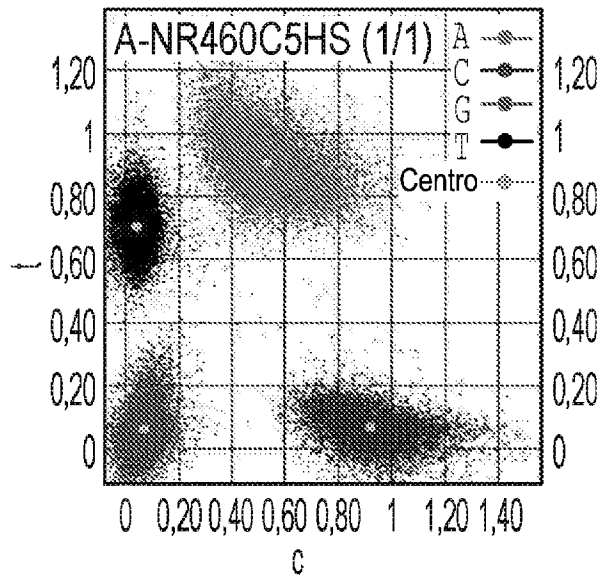


FIGURA 1B

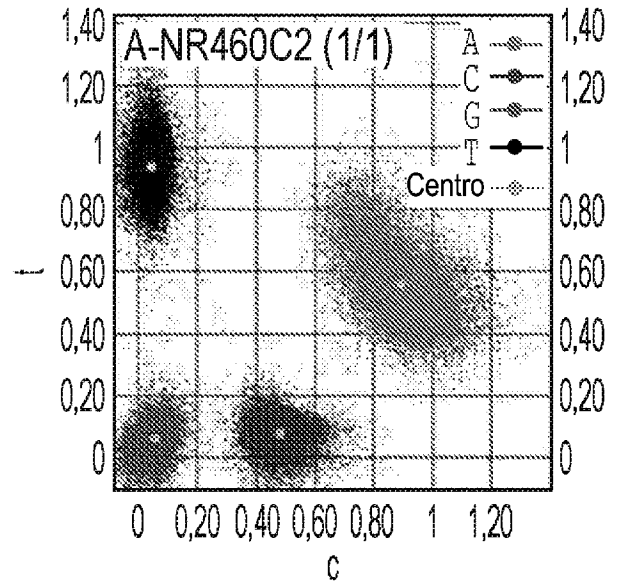


FIGURA 1C

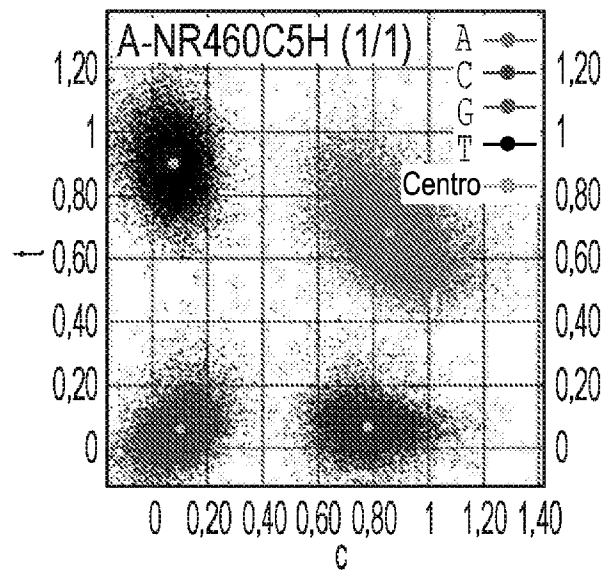


FIGURA 1