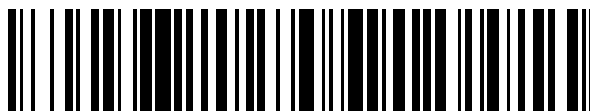


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 641 871**

51 Int. Cl.:

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **16.12.2011 PCT/US2011/065640**

87 Fecha y número de publicación internacional: **21.06.2012 WO12083249**

96 Fecha de presentación y número de la solicitud europea: **16.12.2011 E 11848220 (7)**

97 Fecha y número de publicación de la concesión europea: **05.07.2017 EP 2652153**

54 Título: **Secuenciación de ADN mediante síntesis usando nucleótidos modificados y detección con nanoporos**

30 Prioridad:

09.11.2011 US 201161557558 P
17.12.2010 US 201061424480 P

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

14.11.2017

73 Titular/es:

**THE TRUSTEES OF COLUMBIA UNIVERSITY IN
THE CITY OF NEW YORK (100.0%)
West 116th Street and Broadway
New York, NY 10027, US**

72 Inventor/es:

**JU, JINGYUE;
KUMAR, SHIV;
LI, ZENGMIN;
TAO, CHUANJUAN;
CHIEN, MINCHEN;
RUSSO, JAMES, J.;
KALACHIKOV, SERGEY;
SHEPARD, KEN y
ROSENSTEIN, JACOB, KARL**

74 Agente/Representante:

ARIAS SANZ, Juan

ES 2 641 871 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

Secuenciación de ADN mediante síntesis usando nucleótidos modificados y detección con nanoporos

Antecedentes de la invención

- 5 La secuenciación de ADN es una tecnología fundamental para biología. Se han detectado varios métodos analíticos para detectar ADN o ARN a nivel de moléculas individuales usando tecnologías microscópicas químicas o físicas [Perkins *et al.* 1994, Rief *et al.* 1999, Smith *et al.* 1996 y Vercoutere *et al.* 2001].
- 10 En los últimos años, se han explorado tecnologías de detección de iones tales como canal de iones, que se basa en la detección de ion hidrógeno (H^+) liberado cuando se incorpora un nucleótido en una cadena de ADN mediante una polimerasa [Rothberg *et al.* 2011], para detectar cadenas de ADN o ARN individuales [Kasianowicz 2003 y 2004, Chandler *et al.* 2004, Deamer *et al.* 2002, Berzukov *et al.* 2001, y Henrickson *et al.* 2000].
- 15 Se ha demostrado que puede usarse un canal de α -hemolisina, una exotoxina secretada por una bacteria, para detectar ácidos nucleicos a nivel de moléculas individuales [Kasianowicz *et al.* 1996]. Una proteína α -hemolisina es un polipéptido monomérico que se autoensambla en una membrana de bicapa lipídica para formar un poro heptamérico, con un vestíbulo de 2,6 nm de diámetro y una abertura limitante de 1,5 nm de diámetro (el punto más estrecho del poro) [Meller *et al.* 2000, Akesson *et al.* 1999, y Deamer *et al.* 2002]. La abertura limitante del nanoporo permite que moléculas de ácido nucleico (diámetro \sim 2,0 nm) monocatenarias lineales, pero no bicatenarias, pasen a través de la misma. En una disolución acuosa de sal iónica tal como KCl, cuando se aplica una tensión apropiada a través de la membrana, el poro formado por un canal de α -hemolisina transporta una corriente iónica suficientemente intensa y estable. Los ácidos nucleicos polianiónicos se impulsan a través del poro mediante el campo eléctrico aplicado, bloqueando o reduciendo por tanto la corriente iónica que de lo contrario o se vería impedida. Este proceso de paso genera una firma electrónica (figura 1) [Vercoutere *et al.* 2001 y Deamer *et al.* 2002]. Una molécula de ácido nucleico particular, cuando entra y pasa a través del nanoporo, genera una firma característica que la distingue de otras moléculas de ácido nucleico. La duración del bloqueo es proporcional a la longitud de ácido nucleico, y la intensidad de señal está relacionada con las propiedades estéricas y electrónicas de los nucleótidos, concretamente la identidad de las cuatro bases (A, C, G y T). Por tanto, se obtiene un diagrama de acontecimientos específicos, que es una representación gráfica de tiempo de translocación frente a corriente de bloqueo, y se usa para distinguir la longitud y la composición de polinucleótidos mediante técnicas de grabación de un único canal basadas en parámetros característicos tales como corriente de translocación, duración de translocación y sus correspondientes dispersiones en el diagrama [Meller *et al.* 2000].
- 20 También se ha mostrado que un nanoporo proteico con un adaptador unido de manera covalente puede identificar con precisión 5'-monofosfatos de nucleósidos sin marcar (dAMP, dGMP, dCMP y dTMP) con alta precisión [Clarke *et al.* 2009]. Por ejemplo, se ha unido de manera covalente adaptador de aminociclodextrina dentro del poro de α -hemolisina de manera satisfactoria. Cuando se capta un dNMP y se impulsa a través del poro en una membrana de bicapa lipídica, la corriente iónica a través del poro se reduce hasta uno de cuatro niveles, representando, cada uno, uno de los cuatro dNMP (A, G, C o T). Además, Robertson *et al.* [2007] han demostrado recientemente que cuando una molécula de polietilenglicol (PEG) entra en un único poro de α -hemolisina, provoca estados de conductancia dependientes de la masa definidos con tiempos de residencia medios característicos. El espectro de masas basado en conductancia resuelve de manera clara las unidades de repetición de etilenglicol, y el tiempo de residencia aumenta con la masa del PEG.
- 30 John Eid *et al.* dan a conocer datos de secuenciación de moléculas individuales, en tiempo real, obtenidos a partir de una ADN polimerasa que realiza una síntesis dirigida por molde, sin interrupciones, usando cuatro trifosfatos de desoxirribonucleósido marcados de manera fluorescente, distinguibles (dNTP). ("Real-Time DNA Sequencing from Single Polymerase Molecules", Science, vol. 23, n.º 5910, 2008, páginas 133-138).
- 35 Anup Sood *et al.* dan a conocer tetrafosfatos pentafosfatos de nucleótidos marcados con colorante de γ -fosfato. ("Terminal phosphate-labeled nucleotides with improved substrate properties for homogeneous nucleic acid assays" JACS, vol. 127, n.º 8, 2005, páginas 2394-2395).
- 40 B. A. Mulder *et al.* dan a conocer que la fidelidad de la transcriptasa inversa se mejora usando nucleótidos modificados ("Nucleotide modification at the γ -phosphate leads to the improved fidelity of HIV-1 reverse transcriptase", Nucleic Acids Research, vol. 33, n.º 15, 2005, páginas 4865-4873).
- 45 Shiv Kumar *et al.* dan a conocer polifosfatos de nucleósidos que contienen 3 o más fosfatos con el colorante fluorescente unido al fosfato terminal. ("Terminal Phosphate Labeled Nucleotides: Synthesis, Applications, and Linker Effect on Incorporation by DNA Polymerases" Nucleosides, Nucleotides and Nucleic Acids, vol. 24, n.º 5-7, 2005, páginas 401-408).
- 50 Advanced Res. & Tech. Institute da a conocer un trifosfato de nucleósido (NTP) que comprende α , β y γ -fosfatos y un grupo funcional unido a fosfoéster de γ -fosfato; (documento WO 01/48235 A2, 5 de julio de 2001).
- 55

Andrzej Guranowski *et al.* dan a conocer la hidrólisis de diversos polifosfatos de diadenosina. ("Selective Degradation of 2'-Adenylated Diadenosine Tri- and Tetraphosphates, Ap3A and Ap4A, by Two Specific Human Dinucleoside Poliphosphate Hydrolases", Archives of Biochemistry and Biophysics, vol. 373, n.º 1, 2000, páginas 218-224).

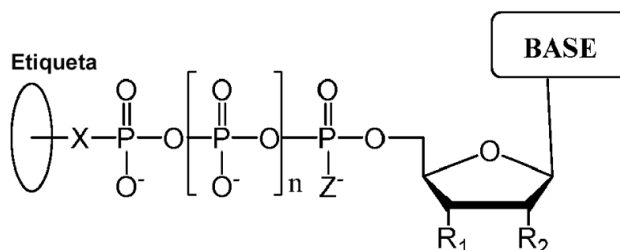
- 5 Bambi Reynolds *et al.* dan a conocer el uso de determinados nucleótidos marcados con fosfato terminal en experimentos de secuenciación de moléculas individuales. ("Synthesis and Stability of Novel Terminal Phosphate-Labeled Nucleotides", Nucleosides, Nucleotides and Nucleic Acids vol. 27, n.º 1, 2008, páginas 18-30).

Aunque el enfoque de nanoporos actual muestra ser prometedor como método de detección de ADN, aún no se ha cumplido el objetivo más desafiante de una secuenciación precisa de base a base.

10 Sumario de la invención

Método para determinar la secuencia de nucleótidos de un ADN monocatenario que comprende:

- (a) poner en contacto el ADN monocatenario, en el que el ADN monocatenario está en una disolución de electrolito en contacto con un nanoporo en una membrana y en el que el ADN monocatenario tiene un cebador hibridado con una parte del mismo, con una ADN polimerasa y cuatro análogos de polifosfato de desoxirribonucleótido (dNPP) al menos uno de los cuales puede hibridarse con cada uno de un nucleótido A, T, G o C en el ADN que está secuenciándose en condiciones que permiten que la ADN polimerasa catalice la incorporación de uno de los análogos de dNPP en el cebador si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-terminal del cebador, para formar un producto de extensión de ADN, en el que cada uno de los cuatro análogos de dNPP tiene la estructura:



- en la que la base es adenina, guanina, citosina, timina o uracilo, o un derivado de una o más de estas bases, en la que R_1 es OH, en la que R_2 es H, en la que X es O, NH, S o CH_2 , en la que n es 1, 2, 3 ó 4, en la que Z es O, S o BH_3 , y con la condición de que (i) el tipo de base en cada análogo de dNPP es diferente del tipo de base en cada uno de los otros tres análogos de dNPP, y (ii) o bien el valor de n de cada análogo de dNPP es diferente del valor de n de cada uno de los otros tres análogos de dNPP, o bien el valor de n de cada uno de los cuatro análogos de dNPP es el mismo y el tipo de etiqueta en cada análogo de dNPP es diferente del tipo de etiqueta en cada uno de los otros tres análogos de dNPP, en el que la incorporación del análogo de dNPP da como resultado la liberación de un polifosfato que tiene la etiqueta unida al mismo; y

- (b) identificar qué análogo de dNPP se ha incorporado en el cebador para formar un producto de extensión de ADN en la etapa (a) aplicando una tensión a través de la membrana y midiendo un cambio electrónico a través del nanoporo que resulta de que el polifosfato que tiene la etiqueta unida al mismo generado en la etapa (a) se transloca a través del nanoporo, en el que el cambio electrónico es diferente para cada valor de n, o para cada tipo de etiqueta diferente, lo que sea aplicable, permitiendo así identificar el residuo de nucleótido en el ADN monocatenario complementario al análogo de dNPP incorporado; y

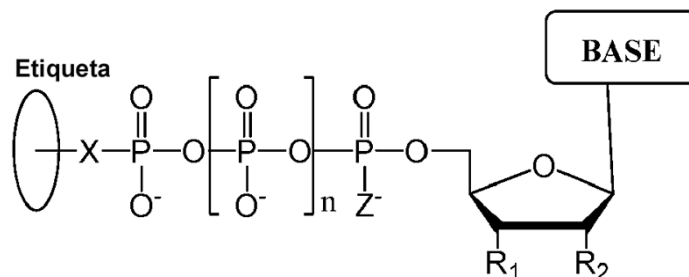
- (c) realizar de manera repetida la etapa (b) para cada residuo de nucleótido del ADN monocatenario que está secuenciándose, en el que en cada repetición de la etapa (b) se incorpora el análogo de dNPP en el producto de extensión de ADN si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-terminal del producto de extensión de ADN,

determinando así la secuencia de nucleótidos del ADN monocatenario.

Método para determinar la secuencia de nucleótidos de un ADN monocatenario que comprende:

- (a) poner en contacto el ADN monocatenario, en el que el ADN monocatenario está en una disolución de electrolito en contacto con un nanoporo en una membrana y en el que el ADN monocatenario tiene un cebador hibridado con una parte del mismo, una ADN polimerasa y un análogo de polifosfato de desoxirribonucleótido (dNPP) en condiciones que permiten que la ADN polimerasa catalice la incorporación del análogo de dNPP en el cebador si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-terminal del cebador, para

formar un producto de extensión de ADN, en el que el análogo de dNPP tiene la estructura:



en la que la base es adenina, guanina, citosina, uracilo o timina, o un derivado de cada una de las mismas, en la que R_1 es $-OH$, $-O-CH_2N_3$ o $-O-2$ -nitrobenilo, en la que R_2 es H , en la que X es O , NH , S o CH_2 , en la que n es 1, 2, 3 ó 4, en la que Z es O , S o BH_3 , y

en el que si no se incorpora el análogo de dNPP, repetir de manera iterativa la puesta en contacto con un análogo de dNPP diferente hasta que se incorpora un análogo de dNPP, con la condición de que (1) el tipo de base en cada análogo de dNPP es diferente del tipo de base en cada uno de los otros análogos de dNPP, y (2) o bien el valor de n de cada análogo de dNPP es diferente del valor de n de cada uno de los otros tres análogos de dNPP, o bien el valor de n de cada uno de los cuatro análogos de dNPP es el mismo y el tipo de etiqueta en cada análogo de dNPP es diferente del tipo de etiqueta en cada uno de los otros tres análogos de dNPP,

en el que la incorporación de un análogo de dNPP da como resultado la liberación de un polifosfato que tiene la etiqueta unida al mismo;

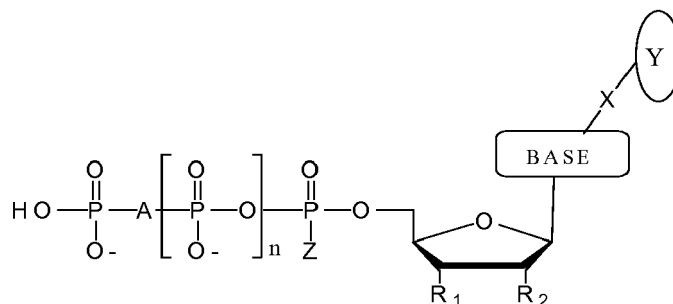
(b) determinar qué análogo de dNPP se ha incorporado en el cebador para formar un producto de extensión de ADN en la etapa (a) aplicando una tensión a través de la membrana y midiendo un cambio electrónico a través del nanoporo que resulta de que el polifosfato que tiene la etiqueta unida al mismo generado en la etapa (a) se transloca a través del nanoporo, en el que el cambio electrónico es diferente para cada valor de n , o para cada tipo de etiqueta diferente, según sea apropiado, identificando de ese modo el residuo de nucleótido en el ADN monocatenario complementario al análogo de dNPP incorporado;

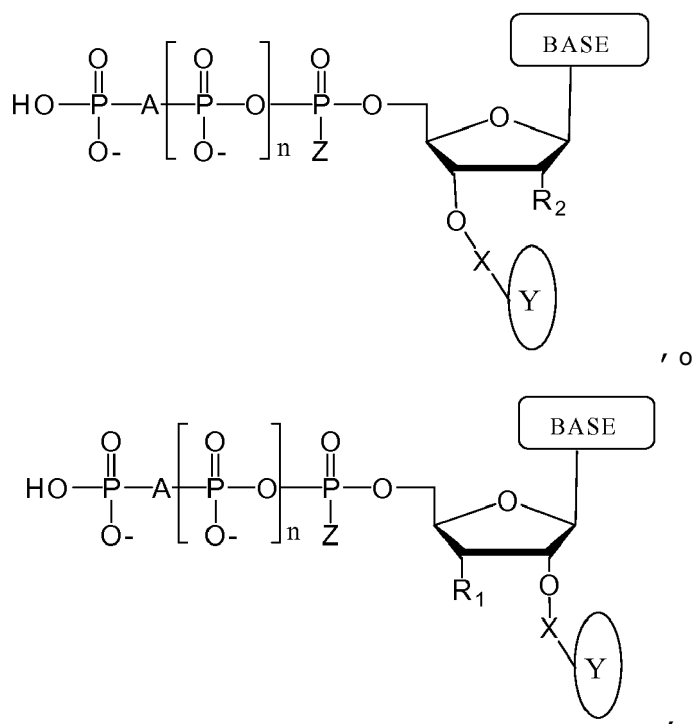
(c) realizar de manera repetida las etapas (a) y (b) para cada residuo de nucleótido del ADN monocatenario que está secuenciándose, en el que en cada repetición de la etapa (a) se incorpora el análogo de dNPP en el producto de extensión de ADN si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-terminal del producto de extensión de ADN,

determinando así la secuencia de nucleótidos del ADN monocatenario.

Método para determinar la secuencia de nucleótidos de un ADN monocatenario que comprende:

(a) poner en contacto el ADN monocatenario, en el que el ADN monocatenario está en una disolución de electrolito en contacto con un nanoporo en una membrana y en el que el ADN monocatenario tiene un cebador hibridado con una parte del mismo, con una ADN polimerasa y al menos cuatro análogos de polifosfato de desoxirribonucleótido (dNPP) en condiciones que permiten que la ADN polimerasa catalice la incorporación de uno de los análogos de dNPP en el cebador si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-terminal del cebador, para formar un producto de extensión de ADN, en el que cada uno de los cuatro análogos de dNPP tiene una estructura elegida de las siguientes:



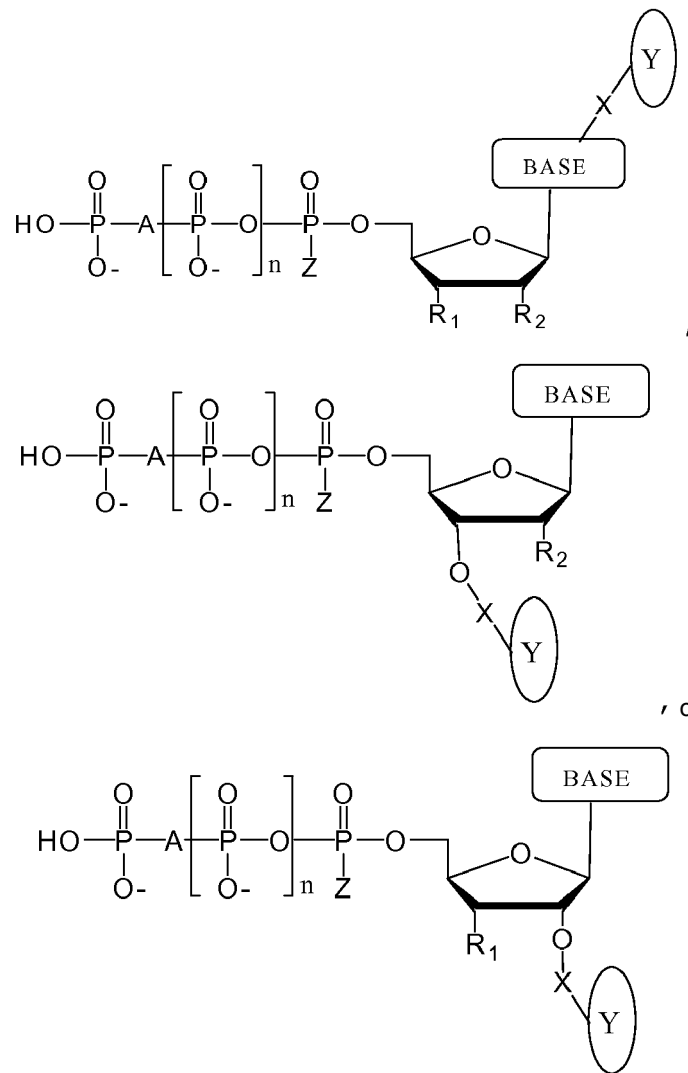


5 en las que la base es adenina, guanina, citosina, timina o uracilo, o un derivado de cada una de las mismas, en las que Y es una etiqueta, en las que R₁, si está presente, es OH, en las que R₂, si está presente, es H, en las que X es un grupo de unión escindible, en las que Z es O, S o BH₃, en las que n es 1, 2, 3 ó 4, en las que A es O, S, CH₂, CHF, CFF o NH, y con la condición de que (i) el tipo de base en cada análogo de dNPP es diferente del tipo de base en cada uno de los otros tres análogos de dNPP, y (ii) el tipo de etiqueta en cada análogo de dNPP es diferente del tipo de etiqueta en cada uno de los otros tres análogos de dNPP;

- (b) escindir la etiqueta del análogo de dNPP incorporado en la etapa (a); y
- 10 (c) determinar qué análogo de dNPP se incorporó en la etapa (a) aplicando una tensión a través de la membrana y midiendo un cambio electrónico a través del nanoporo que resulta de que la etiqueta escindida en la etapa (b) se transloca a través del nanoporo, en el que el cambio electrónico es diferente para cada tipo diferente de etiqueta, identificando de ese modo el residuo de nucleótido en el ADN monocatenario complementario al análogo de dNPP incorporado; y
- 15 (d) realizar de manera repetida las etapas (a), (b) y (c) para cada residuo de nucleótido del ADN monocatenario que está secuenciándose, en el que en cada repetición de la etapa (a) se incorpora el análogo de dNPP en el producto de extensión de ADN resultante de la iteración anterior de la etapa (a) si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-terminal del producto de extensión de ADN,
- 20 determinando así la secuencia de nucleótidos del ADN monocatenario.

Método para determinar la secuencia de nucleótidos de un ADN monocatenario que comprende:

- (a) poner en contacto el ADN monocatenario, en el que el ADN monocatenario está en una disolución de electrolito en contacto con un nanoporo en una membrana, en el que el ADN monocatenario tiene un cebador hibridado con una parte del mismo, una ADN polimerasa y un análogo de polifosfato de desoxirribonucleótido (dNPP) en
- 25 condiciones que permiten que la ADN polimerasa catalice la incorporación del análogo de dNPP en el cebador si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-terminal del cebador, para formar un producto de extensión de ADN, en el que el análogo de dNPP tiene la estructura:



5 en las que la base es adenina, guanina, citosina, uracilo o timina, o un derivado de cada una de las mismas, en las que Y es una etiqueta, y en las que R₁ si está presente es OH, -OCH₂N₃ o -O-2-nitrobencilo, R₂ si está presente es H, en las que X es un grupo de unión escindible, en las que Z es O, S o BH₃, en las que n es 1, 2, 3 ó 4, en las que A es O, S, CH₂, CHF, CFF o NH,

10 y si no se incorpora el análogo de dNPP, repetir de manera iterativa la puesta en contacto con un análogo de dNPP diferente hasta que se incorpora un análogo de dNPP, con la condición de que (1) el tipo de base en cada análogo de dNPP es diferente del tipo de base en cada uno de los otros análogos de dNPP, y (2) el tipo de etiqueta en cada análogo de dNPP es diferente del tipo de etiqueta en cada uno de los otros análogos de dNPP, en el que la incorporación de un análogo de dNPP da como resultado la liberación de un polifosfato que tiene la etiqueta unida al mismo;

(b) escindir la etiqueta del análogo de dNPP incorporado en la etapa (a); y

15 (c) determinar qué análogo de dNPP se incorporó en la etapa (a) para formar un producto de extensión de ADN aplicando una tensión a través de la membrana y midiendo un cambio electrónico a través del nanoporo que resulta de que la etiqueta escindida en la etapa (b) se transloca a través del nanoporo, en el que el cambio electrónico es diferente para cada tipo de etiqueta, identificando de ese modo el residuo de nucleótido en el ADN monocatenario complementario al análogo de dNPP incorporado;

20 (d) realizar de manera iterativa las etapas (a) a (c) para cada residuo de nucleótido del ADN monocatenario que está secuenciándose, en el que en cada iteración de la etapa (a) se incorpora el análogo de dNPP en el producto de extensión de ADN resultante de la iteración anterior de la etapa (a) si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-terminal del producto de extensión de ADN,

25 determinando así la secuencia de nucleótidos del ADN monocatenario.

Procedimiento para producir un análogo de trifosfato de nucleótido, en el que el análogo de trifosfato de nucleótido difiere de un trifosfato de nucleótido por tener una etiqueta unida al fosfato terminal del mismo, que comprende:

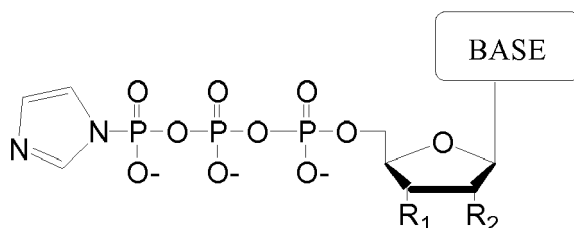
- a) poner en contacto un trifosfato de nucleótido con diciclohexilcarbodiimida/dimetilformamida en condiciones que permiten la producción de un trimetafosfato cíclico;
- 5 b) poner en contacto el producto resultante de la etapa a) con una etiqueta que tiene un grupo hidroxilo o amino unido a la misma en condiciones que permiten la apertura nucleófila del trimetafosfato cíclico para unir la etiqueta a un fosfato terminal formando así el análogo de trifosfato de nucleótido.

Procedimiento para producir un análogo de trifosfato de nucleótido, en el que el análogo de trifosfato de nucleótido difiere de un trifosfato de nucleótido por tener una etiqueta unida al fosfato terminal del mismo, que comprende:

- 10 a) poner en contacto un trifosfato de nucleótido con diciclohexilcarbodiimida/dimetilformamida en condiciones que permiten la producción de un trimetafosfato cíclico;
- b) poner en contacto el producto resultante de la etapa a) con un nucleófilo para formar un compuesto funcionalizado con -OH o -NH₂;
- 15 c) hacer reaccionar el producto de la etapa b) con una etiqueta que tiene un grupo -COR unido a la misma en condiciones que permiten que la etiqueta se una indirectamente a un fosfato terminal formando así el análogo de trifosfato de nucleótido.

Procedimiento para producir un análogo de tetrafosfato de nucleótido, en el que el análogo de tetrafosfato de nucleótido difiere de un tetrafosfato de nucleótido por tener una etiqueta unida al fosfato terminal del mismo, que comprende:

- 20 a) poner en contacto un trifosfato de nucleótido con 1,1'-carbonildiimidazol/dimetilformamida en condiciones que permiten la formación de la siguiente estructura:

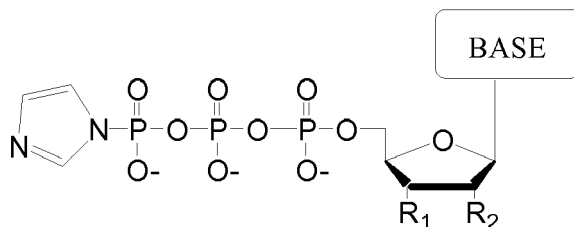


en la que R₁ es OH, en la que R₂ es H u OH, en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina;

- 25 b) poner en contacto el producto resultante de la etapa a) con una etiqueta que tiene un grupo monosfosfato unido a la misma en condiciones que permiten la formación del análogo de tetrafosfato de nucleótido.

Procedimiento para producir un análogo de tetrafosfato de nucleótido, en el que el análogo de tetrafosfato de nucleótido difiere de un tetrafosfato de nucleótido por tener una etiqueta unida al fosfato terminal del mismo, que comprende:

- 30 a) poner en contacto un trifosfato de nucleótido con 1,1'-carbonildiimidazol/dimetilformamida en condiciones que permiten la formación de la siguiente estructura:



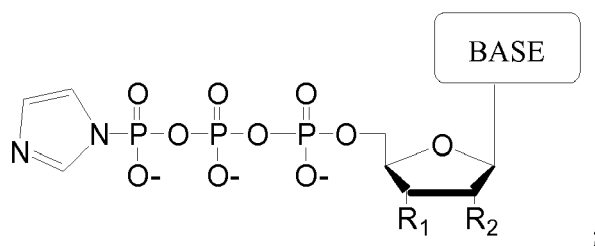
en la que R₁ es OH, en la que R₂ es H u OH, en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina;

- 35 b) poner en contacto el producto resultante de la etapa a) con ácido fosfórico en condiciones que permiten la formación de un tetrafosfato de nucleótido;
- c) poner en contacto el tetrafosfato de nucleótido con 1) carbonildiimidazol/dimetilformamida; 2) un nucleófilo y después 3) NH₄OH para formar un compuesto funcionalizado con -OH o -NH₂;

d) poner en contacto el producto de la etapa c) con una etiqueta que tiene un grupo -COR unido a la misma en condiciones que permiten que la etiqueta se una indirectamente a un fosfato terminal formando así el análogo de tetrafosfato de nucleótido.

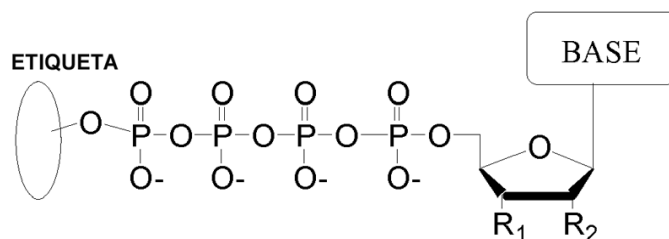
5 Procedimiento para producir un análogo de tetrafosfato de nucleótido, en el que el análogo de tetrafosfato de nucleótido difiere de un tetrafosfato de nucleótido por tener una etiqueta unida al fosfato terminal del mismo, que comprende:

a) poner en contacto un trifosfato de nucleótido con 1,1'-carbonildiimidazol/dimetilformamida en condiciones que permiten la formación de la siguiente estructura:



10 b) poner en contacto el producto resultante de la etapa a) con ácido fosfórico en condiciones que permiten la formación de un tetrafosfato de nucleótido;

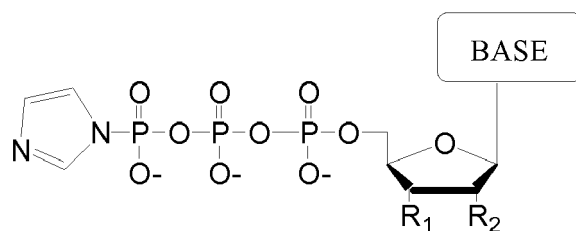
c) poner en contacto el tetrafosfato de nucleótido con carbonildiimidazol/dimetilformamida y una etiqueta que tiene un grupo hidroxilo o amino unido a la misma para formar un compuesto que tiene la estructura:



15 en la que R₁ es OH, en la que R₂ es H u OH, en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina.

Procedimiento para producir un análogo de pentafofosfato de nucleótido, en el que el análogo de pentafofosfato de nucleótido difiere de un pentafofosfato de nucleótido por tener una etiqueta unida al fosfato terminal del mismo, que comprende:

20 a) poner en contacto un trifosfato de nucleótido con 1,1'-carbonildiimidazol/dimetilformamida en condiciones que permiten la formación de la siguiente estructura:

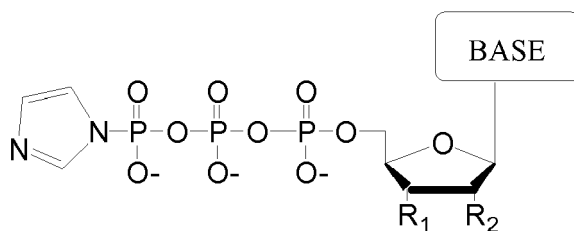


en la que R₁ es OH, en la que R₂ es H u OH, en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina;

25 b) poner en contacto el producto resultante de la etapa a) con una etiqueta que tiene un grupo pirofosfato unido a la misma en condiciones que permiten la formación del análogo de pentafofosfato de nucleótido.

Procedimiento para producir un análogo de pentafofosfato de nucleótido, en el que el análogo de pentafofosfato de nucleótido difiere de un pentafofosfato de nucleótido por tener una etiqueta unida al fosfato terminal del mismo, que comprende:

30 a) poner en contacto un trifosfato de nucleótido con 1,1'-carbonildiimidazol/dimetilformamida en condiciones que permiten la formación de la siguiente estructura:



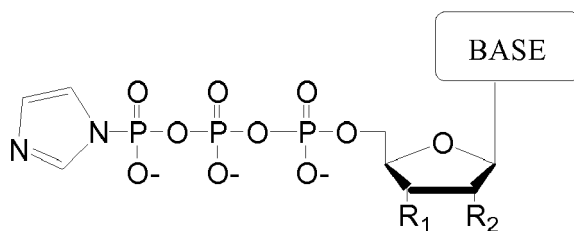
en la que R_1 es OH, en la que R_2 es H u OH, en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina;

5 b) poner en contacto el producto resultante de la etapa a) con un grupo pirofosfato en condiciones que permiten la formación de un pentafofosfato de nucleótido;

c) poner en contacto el pentafofosfato de nucleótido con carbonildiimidazol/dimetilformamida y una etiqueta que tiene un grupo hidroxilo o amino unido a la misma para formar el análogo de pentafofosfato de nucleótido.

10 Procedimiento para producir un análogo de hexafofosfato de nucleótido, en el que el análogo de hexafofosfato de nucleótido difiere de un hexafofosfato de nucleótido por tener una etiqueta unida al fosfato terminal del mismo, que comprende:

a) poner en contacto un trifosfato de nucleótido con 1,1'-carbonildiimidazol/dimetilformamida en condiciones que permiten la formación de la siguiente estructura:

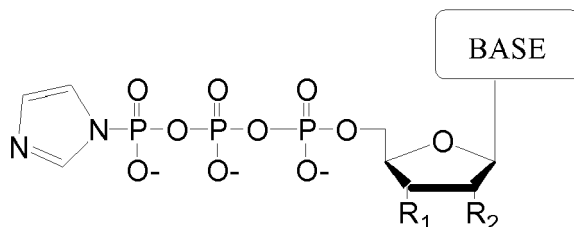


15 en la que R_1 es OH, en la que R_2 es H u OH, en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina;

b) poner en contacto el producto resultante de la etapa a) con una etiqueta que tiene un grupo trifosfato unido a la misma en condiciones que permiten la formación del análogo de hexafofosfato de nucleótido.

20 Procedimiento para producir un análogo de hexafofosfato de nucleótido, en el que el análogo de hexafofosfato de nucleótido difiere de un hexafofosfato de nucleótido por tener una etiqueta unida al fosfato terminal del mismo, que comprende:

a) poner en contacto un trifosfato de nucleótido con 1,1'-carbonildiimidazol/dimetilformamida en condiciones que permiten la formación de la siguiente estructura:

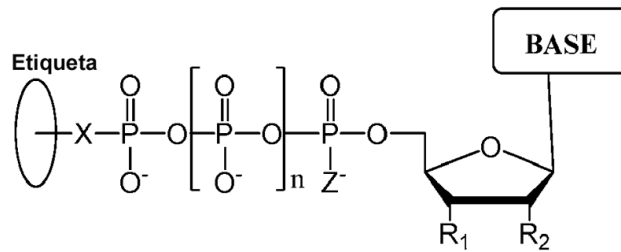


25 en la que R_1 es OH, en la que R_2 es H u OH, en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina;

b) poner en contacto el producto resultante de la etapa a) con un grupo trifosfato en condiciones que permiten la formación de un hexafofosfato de nucleótido;

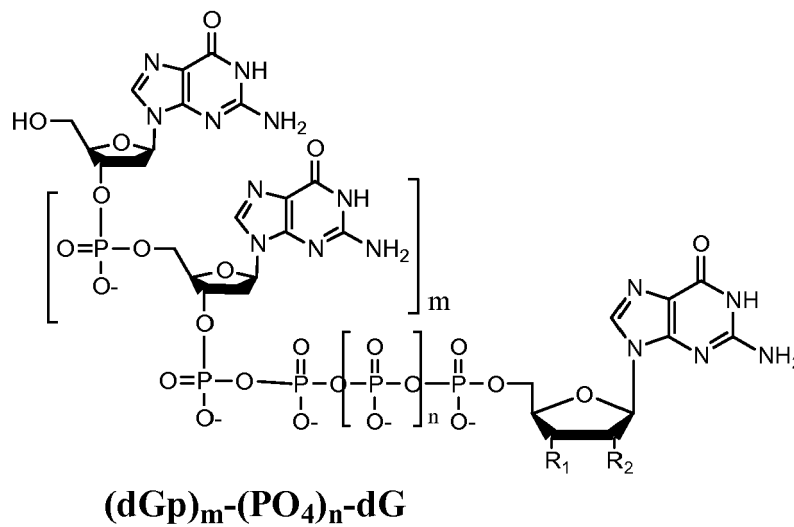
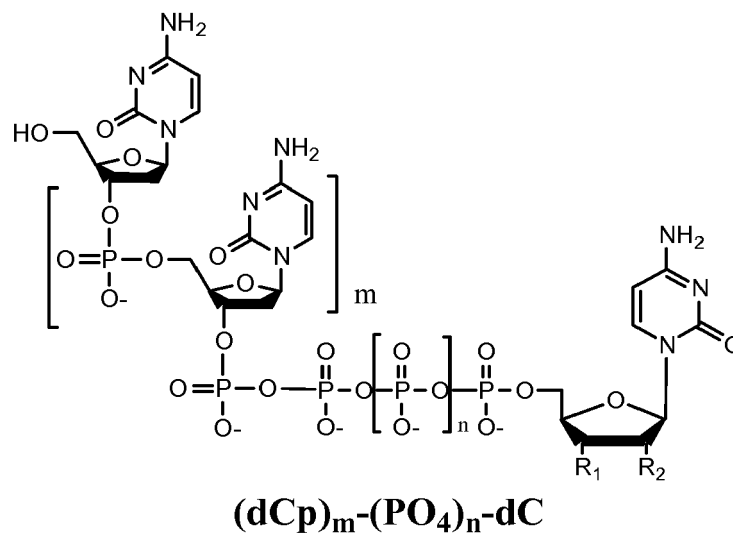
c) poner en contacto el hexafofosfato de nucleótido con carbonildiimidazol/dimetilformamida y una etiqueta que tiene un grupo hidroxilo o amino unido a la misma para formar el análogo de hexafofosfato de nucleótido.

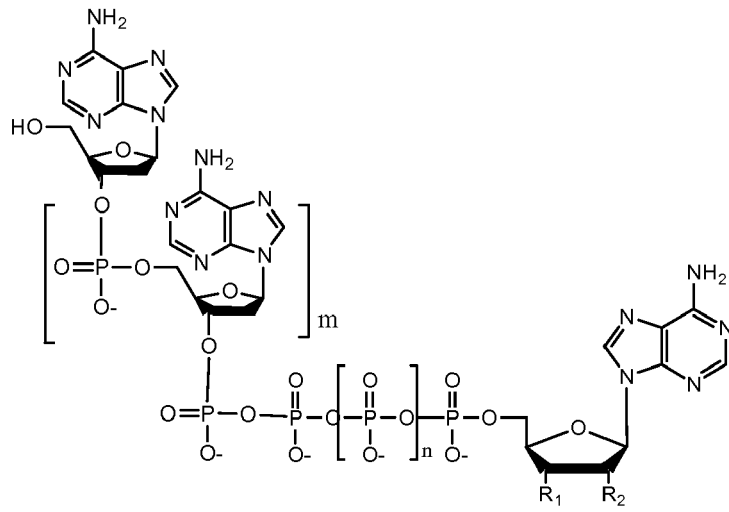
30 Compuesto que tiene la estructura:



5 en la que la etiqueta es etilenglicol, un aminoácido, un hidrato de carbono, un colorante, mononucleótido, dinucleótido, trinucleótido, tetranucleótido o hexanucleótido, en la que R_1 es OH, en la que R_2 es H u OH, en la que X es O, NH, S o CH_2 , en la que Z es O, S o BH_3 , en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina, y en la que n es 1, 2, 3 ó 4.

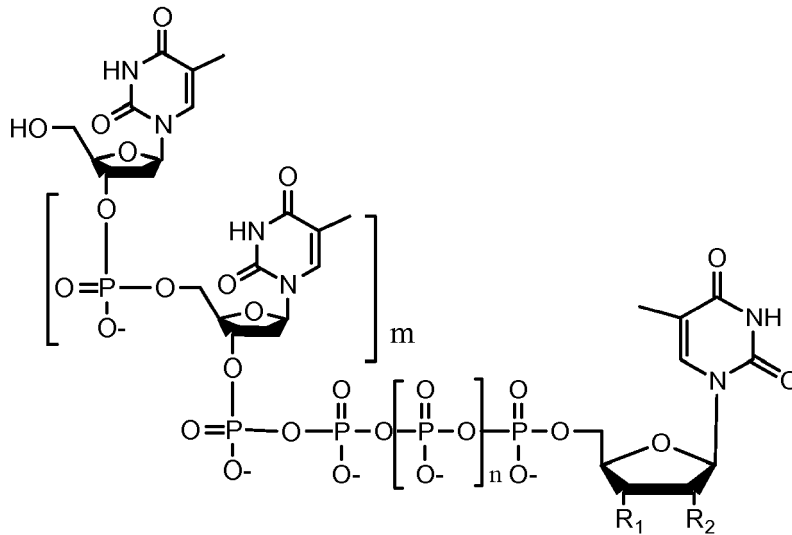
Compuesto que tiene la estructura:





(dAp)_m-(PO₄)_n-dA

10

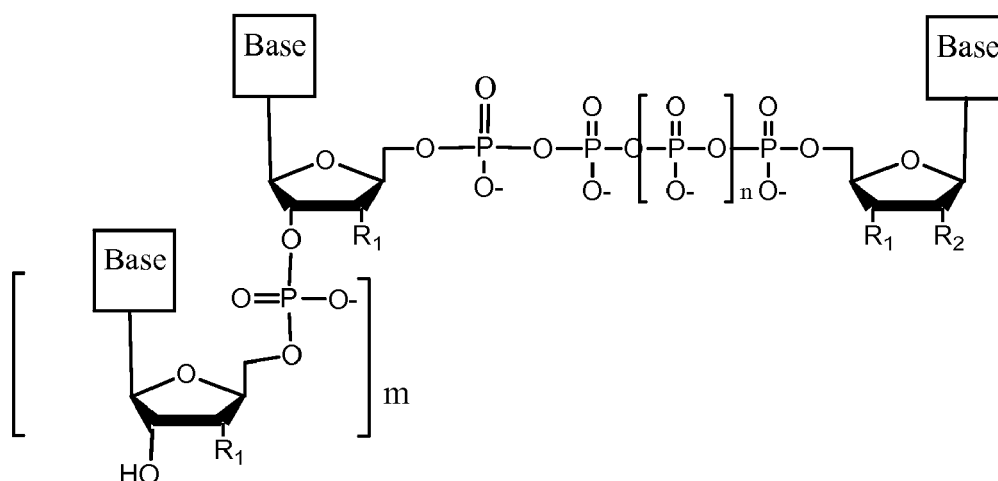


(dTp)_m-(PO₄)_n-dT

5 o en el que en cada estructura n es, independientemente, 1, 2, 3 ó 4, y m es, independientemente, un número entero de desde 0 hasta 100, y en el que cuando m es 0 el fosfato terminal del dNPP está unido directamente al átomo de O en 3' del nucleósido mostrado en el lado izquierdo de la estructura, en las que R₁ es -OH, o -O-CH₂N₃, y R₂ es H u OH.

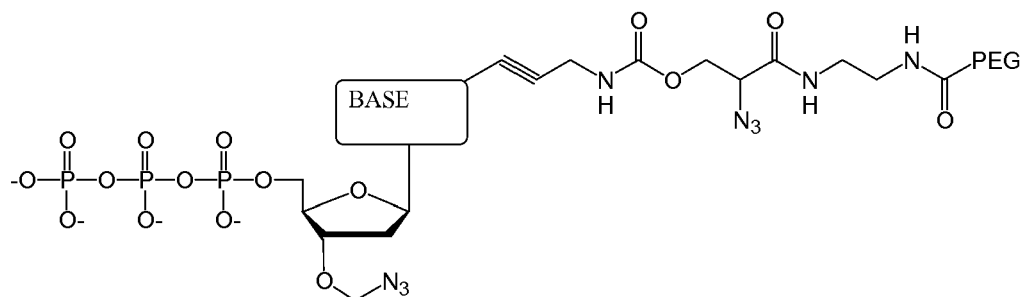
10 Composición que comprende al menos cuatro análogos de polifosfato de desoxinucleótido (dNPP), que tienen cada uno una estructura seleccionada de las estructuras expuestas en las reivindicaciones 74 y 75, en la que cada uno de los cuatro análogos de dNPP comprende un tipo de base diferente del tipo de base de los otros tres análogos de dNPP.

Compuesto que tiene la estructura:



en la que m es un número entero de desde 0 hasta 100, y en el que el compuesto comprende un único tipo de base, y en el que la base es adenina, guanina, citosina, uracilo o timina o un derivado de cada una de las mismas.

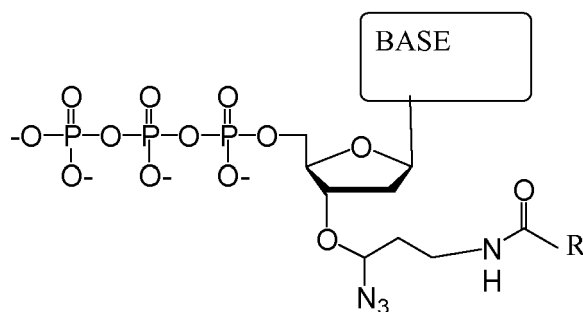
Compuesto que tiene la estructura:



5

en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina.

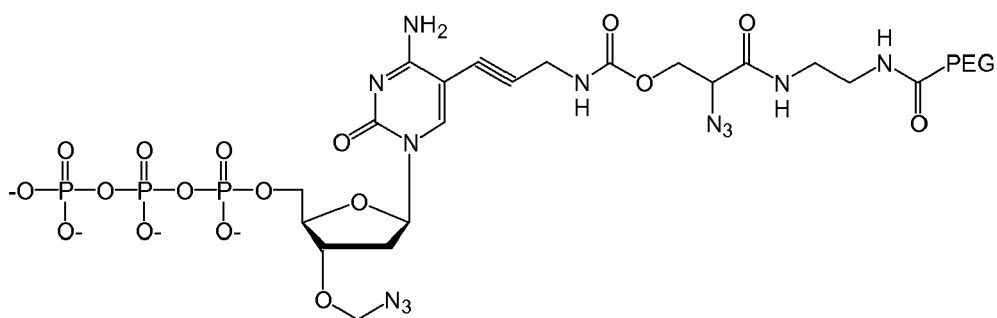
Compuesto que tiene la estructura:



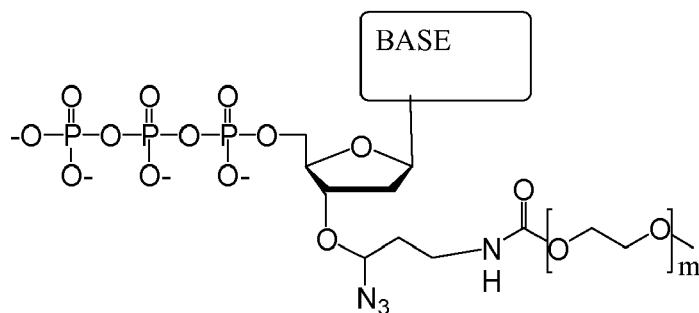
10

en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina, y R es un hidrocarbilo sustituido o no sustituido, de hasta 3000 Dalton.

Compuesto que tiene la estructura:

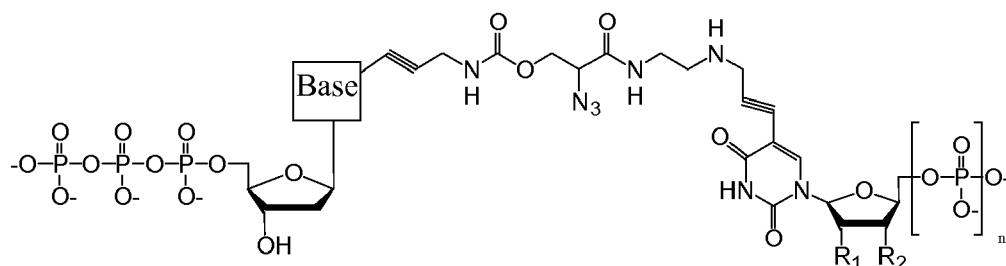


Compuesto que tiene la estructura:



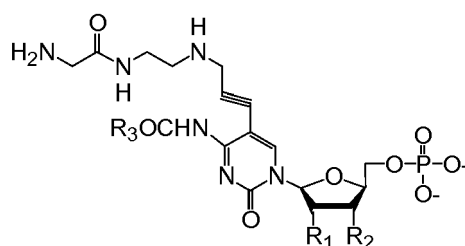
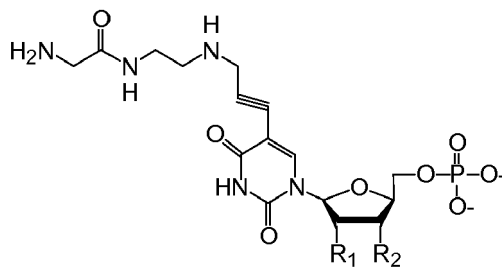
5 en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina, y m es un número entero de desde 1-50.

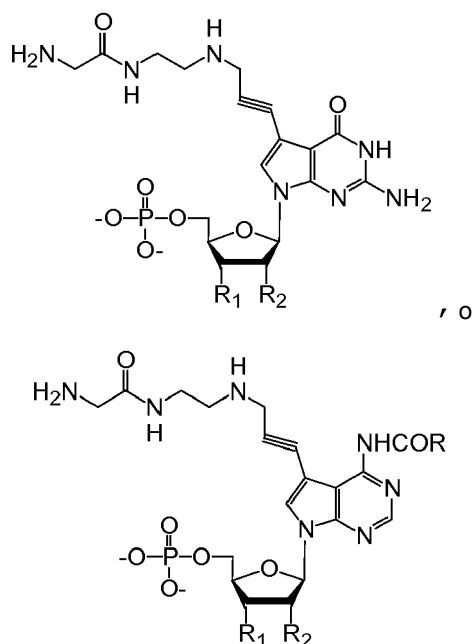
Compuesto que tiene la estructura:



en la que n es 1 ó 2 y la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina.

10 Compuesto que tiene la estructura:

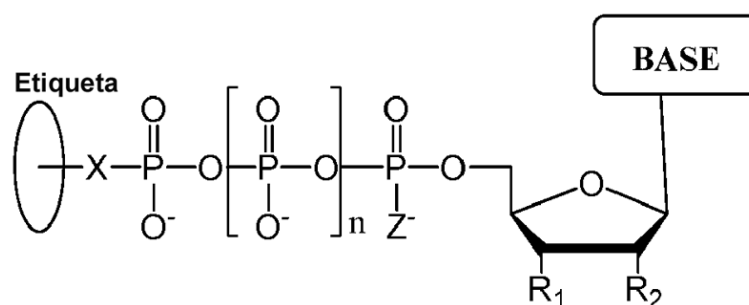




en las que R_1 es $-OH$, o $-O-CH_2N_3$, y R_2 es H u OH .

Método para determinar la secuencia de nucleótidos de un ARN monocatenario que comprende:

- 5 (a) poner en contacto el ARN monocatenario, en el que el ARN monocatenario está en una disolución de electrolito en contacto con un nanoporo en una membrana, en el que el ARN monocatenario tiene un cebador hibridado con una parte del mismo, con una ARN polimerasa y al menos cuatro análogos de polifosfato de ribonucleótido (rNPP) en condiciones que permiten que la ARN polimerasa catalice la incorporación de uno de los análogos de rNPP en el cebador si es complementario al residuo de nucleótido del ARN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ARN monocatenario hibridado con el residuo de nucleótido 3'-terminal del cebador, para formar un producto de extensión de ARN, en el que cada uno de los cuatro análogos de rNPP tiene la estructura:



- 15 en la que la base es adenina, guanina, citosina, timina o uracilo, o un derivado de cada una de las mismas, en la que R_1 es OH , en la que R_2 es OH , en la que X es O , NH , S o CH_2 , en la que n es 1, 2, 3 ó 4, en la que Z es O , S o BH_3 , y con la condición de que (i) el tipo de base en cada análogo de rNPP es diferente del tipo de base en cada uno de los otros tres análogos de rNPP, y (ii) o bien el valor de n de cada análogo de rNPP es diferente del valor de n de cada uno de los otros tres análogos de rNPP, o bien el valor de n de cada uno de los cuatro análogos de rNPP es el mismo y el tipo de etiqueta en cada análogo de rNPP es diferente del tipo de etiqueta en cada uno de los otros tres análogos de rNPP,

en el que la incorporación del análogo de rNPP da como resultado la liberación de un polifosfato que tiene la etiqueta unida al mismo; y

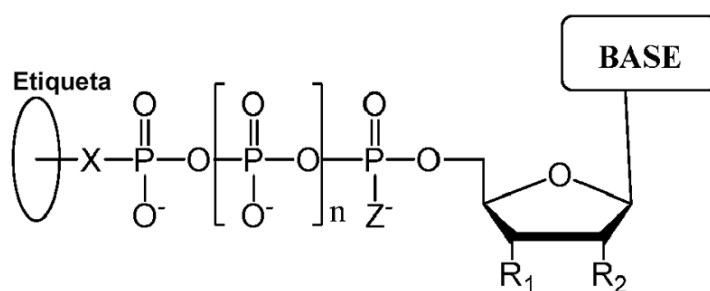
- 25 (b) determinar qué análogo de rNPP se ha incorporado en el cebador para formar un producto de extensión de ARN en la etapa (a) aplicando una tensión a través de la membrana y midiendo un cambio electrónico a través del nanoporo que resulta de que el polifosfato que tiene la etiqueta unida al mismo generado en la etapa (a) se transloca a través del nanoporo, en el que el cambio electrónico es diferente para cada valor de n , o para cada tipo de etiqueta diferente, según sea apropiado, identificando de ese modo el residuo de nucleótido en el ARN monocatenario complementario al análogo de rNPP incorporado; y

(c) realizar de manera iterativa las etapas (a) y (b) para cada residuo de nucleótido del ARN monocatenario que está secuenciándose, en el que en cada iteración de la etapa (a) el análogo de rNPP se incorpora en el producto de extensión de ARN resultante de la iteración anterior de la etapa (a) si es complementario al residuo de nucleótido del ARN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ARN monocatenario hibridado con el residuo de nucleótido 3'-terminal del producto de extensión de ARN,

determinando así la secuencia de nucleótidos del ARN monocatenario.

Método para determinar la secuencia de nucleótidos de un ARN monocatenario que comprende:

(a) poner en contacto el ARN monocatenario, en el que el ARN monocatenario está en una disolución de electrolito en contacto con un nanoporo en una membrana y en el que el ARN monocatenario tiene un cebador hibridado con una parte del mismo, una ARN polimerasa y un análogo de polifosfato de ribonucleótido (rNPP) en condiciones que permiten que la ARN polimerasa catalice la incorporación del análogo de rNPP en el cebador si es complementario al residuo de nucleótido del ARN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ARN monocatenario hibridado con el residuo de nucleótido 3'-terminal del cebador, para formar un producto de extensión de ARN, en el que el análogo de rNPP tiene la estructura:



en la que la base es adenina, guanina, citosina, uracilo o timina, en la que R_1 es $-OH$, $-O-CH_2N_3$ o $-O-2$ -nitrobencilo, en la que R_2 es $-OH$, en la que X es O, NH, S o CH_2 , en la que n es 1, 2, 3 ó 4, en la que Z es O, S o BH_3 ,

y en el que si no se incorpora el análogo de rNPP, repetir de manera iterativa la puesta en contacto con un análogo de rNPP diferente hasta que se incorpora un análogo de rNPP, con la condición de que (1) el tipo de base en cada análogo de rNPP es diferente del tipo de base en cada uno de los otros análogos de rNPP, y (2) o bien el valor de n de cada análogo de rNPP es diferente del valor de n de cada uno de los otros tres análogos de rNPP, o bien el valor de n de cada uno de los cuatro análogos de rNPP es el mismo y el tipo de etiqueta en cada análogo de rNPP es diferente del tipo de etiqueta en cada uno de los otros tres análogos de rNPP,

en el que la incorporación de un análogo de rNPP da como resultado la liberación de un polifosfato que tiene la etiqueta unida al mismo;

(b) determinar qué análogo de rNPP se ha incorporado en el cebador para formar un producto de extensión de ARN en la etapa (a) aplicando una tensión a través de la membrana y midiendo un cambio electrónico a través del nanoporo que resulta de que el polifosfato que tiene la etiqueta unida al mismo generado en la etapa (a) se transloca a través del nanoporo, en el que el cambio electrónico es diferente para cada valor de n, o diferente para cada tipo de etiqueta, según sea apropiado, identificando de ese modo el residuo de nucleótido en el ARN monocatenario complementario al análogo de dNPP incorporado;

(c) realizar de manera iterativa las etapas (a) y (b) para cada residuo de nucleótido del ARN monocatenario que está secuenciándose, en el que en cada iteración de la etapa (a) el análogo de rNPP se incorpora en el producto de extensión de ARN resultante de la iteración anterior de la etapa (a) si es complementario al residuo de nucleótido del ARN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ARN monocatenario hibridado con el residuo de nucleótido 3'-terminal del producto de extensión de ARN,

determinando así la secuencia de nucleótidos del ARN monocatenario.

Breve descripción de las figuras

Figura 1. La proteína α -hemolisina se autoensambla para dar una bicapa lipídica para formar un canal de iones y un tramo de ácido nucleico pasa a través del mismo (parte superior), con las firmas electrónicas correspondientes generadas (parte inferior) [Vercutere *et al.* 2001 y Deamer *et al.* 2002].

Figura 2. Estructuras de nucleótidos: desoxirribonucleótido adenosina trifosfato, desoxirribonucleótido guanosina trifosfato, desoxirribonucleótido citosina trifosfato y desoxirribonucleótido timidina trifosfato.

Figura 3. Mecanismo de extensión de cebador y liberación de polifosfato marcado con etiqueta para su detección.

- Figura 4. Estructura de cuatro 5'-polifosfatos de nucleósidos marcados con etiqueta de fosfato.
- Figura 5. Síntesis de 5'-trifosfatos de nucleósidos marcados con etiqueta de fosfato.
- Figura 6. Síntesis de 5'-tetrafosfatos de nucleósidos marcados con etiqueta de fosfato.
- Figura 7. Síntesis de 5'-pentafosfatos de nucleósidos marcados con etiqueta de fosfato terminal.
- 5 Figura 8. a) unión de oligo-3' a 5'-fosfato, b) unión de oligo-5' a 5'-fosfato, c) resto detectable tras reacción con polimerasa.
- Figura 9(A). Síntesis de 5'-trifosfatos de nucleósidos con base modificada. Figura 9(B). Escisión de 5'-trifosfato de nucleósido con base modificada y escisión con TCEP.
- 10 Figura 10. Síntesis de 5'-trifosfatos de nucleósidos modificados en 3'-O. A. dNTP unidos a 3'-O-2-nitrobencilo; B. dNTP unidos a 3'-O-azidometilo; C. Resto detectable tras la extensión con polimerasa y escisión con TCEP; y D. Resto detectable tras la extensión con polimerasa y escisión por UV.
- Figura 11. Reacción de extensión de ADN usando análogos de nucleótidos modificados con fosfato.
- Figura 12. Reacción de extensión de ADN usando análogos de nucleótidos marcados con etiqueta de base.
- Figura 13. Reacción de extensión de ADN usando análogos de nucleótidos marcados con 2' o 3'-OH.
- 15 Figura 14. Esquema de secuenciación de ADN mediante nanoporo con nucleótidos modificados, particularmente aplicable a la secuenciación de moléculas individuales en tiempo real que implica la adición de los 4 nucleótidos y polimerasa al mismo tiempo para entrar en contacto con una única molécula de molde.
- Figura 15. Fosfatos de nucleósidos modificados con fosfato, base, en 2' y en 3' con posibles grupos de unión y etiquetas
- 20 BASE = adenina, guanina, timina, citosina, uracilo, 5-metil-C, 7-desaza-A, 7-desaza-G o derivados de las mismas;
 R_1 y R_2 = H, OH, F, NH_2 , N_3 u OR' ;
 n = 1-5;
 A = O, S, CH_2 , CHF, CFF, NH;
 Z = O, S, BH_3 ;
- 25 X = grupo de unión que une fosfato a 2'-O o 3'-O o la base al resto detectable y puede contener átomos de O, N o S, P. (El grupo de unión también puede ser un resto detectable, directa o indirectamente, tal como aminoácidos, péptidos, proteínas, hidratos de carbono, PEG de diferentes longitudes y pesos moleculares, colorantes orgánicos o inorgánicos, colorantes fluorescentes y fluorogénicos, fármacos, oligonucleótidos, etiquetas de masa, etiquetas quimioluminiscentes y puede contener cargas positivas o negativas);
- 30 Y = etiquetas o resto detectable, tales como compuestos aromáticos orgánicos o alifáticos con uno o más anillos, colorantes, proteínas, hidratos de carbono, PEG de diferentes longitudes y pesos moleculares, fármacos, oligonucleótidos, etiquetas de masa, etiquetas fluorescentes, etiquetas quimioluminiscentes y pueden contener carga positiva o negativa.
- 35 Figura 16. Estructuras de nucleótidos marcados con PEG-fosfato y ejemplos de posibles PEG con diferentes grupos reactivos para reaccionar con grupos funcionales.
- Figura 17. Ejemplos específicos, no limitativos, de grupos reactivos en los fosfatos terminales, que también pueden unirse con cambios apropiados a un resto de base de nucleósido, y grupos con los que pueden reaccionar los grupos para formar etiquetas.
- Figura 18. Un esquema de matriz de nanoporos para secuenciación de ADN en paralelo, masiva, mediante síntesis.
- 40 Figura 19. Síntesis de nucleótidos marcados con PEG-fosfato.
- Figura 20. Espectros de masas de MALDI-TOF de los productos de extensión de ADN generados mediante incorporación de análogos de nucleótidos marcados con PEG-fosfato (dG4P-PEG). Los productos individuales mostrados en los espectros indican que dG4P-PEG24 y dG4P-PEG37 se incorporan con una eficacia de casi el 100%.
- 45 Figura 21. Las distribuciones de profundidad de bloqueo relativo para nanoporo de α -hemolisina en presencia de PEG que contienen 49, 37, 24 ó 16 monómeros de óxido de etileno a un potencial aplicado de +40 mV. Las cuatro

especies se identifican fácilmente.

Figura 22. (A) Separación y distribución de masas de unidades de polietilenglicol (PEG) mixtas a través de un único nanoporo; y (B) selección de 4 unidades de PEG diferenciadas con separación de línea base como etiquetas para las 4 bases, A, C, G, y T. También se muestran las estructuras de PEG lineales y ramificados.

5 Figura 23. Síntesis de PEG-trifosfatos cargados (la carga puede ajustarse basándose en los requisitos).

Figura 24. Síntesis de 5'-trifosfatos de nucleósidos marcados con etiqueta de fosfato.

Figura 25. Síntesis de 5'-tetrafosfatos de nucleósidos marcados con etiqueta de fosfato.

Figura 26. Síntesis de 5'-pentafosfatos de nucleósidos marcados con etiqueta de fosfato terminal.

10 Figura 27. Plataforma de medición de nanoporos con CMOS integrado: (A) una microfotografía del chip preamplificador de CMOS de ocho canales con una imagen de un canal amplificador con el electrodo en el lado de cis integrado; (B) diagrama que muestra la integración de dos chips con un nanoporo en estado sólido; (C) diagrama que muestra la sección transversal del chip y cómo se graba el nanoporo directamente en el chip en la implementación de un chip; el empaquetamiento se produce con un pocillo independiente en el lado de cis; y una imagen de TEM de un nanoporo con un diámetro de 3,5 nm.

15 Figura 28. Rendimiento eléctrico de la electrónica de nanoporos con CMOS integrado (A) Espectro de ruido de corriente de línea base relacionado con la entrada para $C_F=0,15$ pF, filtro de Bessel de 4 polos de 1 MHz, $f_s=4$ MS/s. También se muestra el cabezal abierto medido de un instrumento Axopatch 200B en modo de célula completa con $\beta=1$, filtro de Bessel de 4 polos de 100 kHz, $f_s=250$ kS/s. (B) Umbral mínimo de ruido del nuevo amplificador con un nanoporo unido en comparación con el mismo nanoporo medido mediante el instrumento Axopatch 200B.

20 Figura 29. Anclaje de la polimerasa en las proximidades del nanoporo. Un pocillo ayuda a limitar la difusión. L indica la distancia crítica desde la abertura de poro a la que los movimientos moleculares debidos a difusión y electroforesis son iguales.

Figura 30. Síntesis de 5'-polifosfatos de nucleósidos marcados con etiqueta.

Figura 31. Síntesis de PEG-nucleótidos bloqueados en 3'-O.

25 Figura 32. Secuenciación por síntesis con PEG-nucleótidos y detección con nanoporos (muchas copias de la misma molécula de ADN inmovilizadas sobre una perla y adición de un PEG-nucleótido cada vez). Se usa el mismo PEG unido a los cuatro nucleótidos. Se añade un PEG-nucleótido cada vez, se lee al menos una base por ciclo si se incorpora el nucleótido correcto.

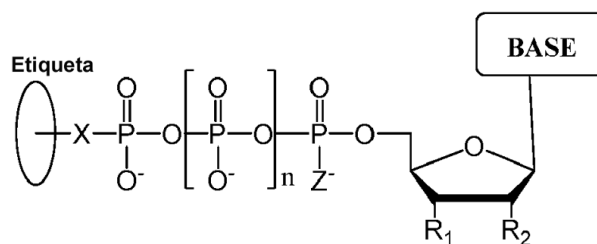
30 Figura 33. Secuenciación por síntesis con PEG-nucleótidos bloqueados en 3'-O y detección con nanoporos (muchas copias de la misma molécula de ADN inmovilizadas sobre una perla y adición de los cuatro PEG-nucleótidos bloqueados en 3'-O). Se añaden los cuatro nucleótidos unidos a PEG de diferente tamaño, bloqueados en 3' (dNTP-PEG bloqueados en 3') juntos. Detección del nucleótido incorporado basándose en la señal de bloqueo de los PEG liberados. El grupo de bloqueo en 3' se retira mediante tratamiento con TECP y se continúa el ciclo para secuenciar correctamente el molde incluyendo regiones homopoliméricas.

35 Figura 34. Un esquema para una matriz de micropocillos de alta densidad, en paralelo, masiva, para realizar el procedimiento bioquímico. Cada pocillo puede contener un molde de ADN diferente y dispositivo de nanoporos.

Descripción detallada de la invención

Método para determinar la secuencia de nucleótidos de un ADN monocatenario que comprende:

40 (a) poner en contacto el ADN monocatenario, en el que el ADN monocatenario está en una disolución de electrolito en contacto con un nanoporo en una membrana y en el que el ADN monocatenario tiene un cebador hibridado con una parte del mismo, con una ADN polimerasa y al menos cuatro análogos de polifosfato de desoxirribonucleótido (dNPP) en condiciones que permiten que la ADN polimerasa catalice la incorporación de uno de los análogos de dNPP en el cebador si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-
45 terminal del cebador, para formar un producto de extensión de ADN, en el que cada uno de los cuatro análogos de dNPP tiene la estructura:



en la que la base es adenina, guanina, citosina, timina o uracilo, o un derivado de cada una de las mismas, en la que R_1 es OH, en la que R_2 es H, en la que X es O, NH, S o CH_2 , en la que n es 1, 2, 3 ó 4, en la que Z es O, S o BH_3 , y con la condición de que (i) el tipo de base en cada análogo de dNPP es diferente del tipo de base en cada uno de los otros tres análogos de dNPP, y (ii) o bien el valor de n de cada análogo de dNPP es diferente del valor de n de cada uno de los otros tres análogos de dNPP, o bien el valor de n de cada uno de los cuatro análogos de dNPP es el mismo y el tipo de etiqueta en cada análogo de dNPP es diferente del tipo de etiqueta en cada uno de los otros tres análogos de dNPP,

en el que la incorporación del análogo de dNPP da como resultado la liberación de un polifosfato que tiene la etiqueta unida al mismo; y

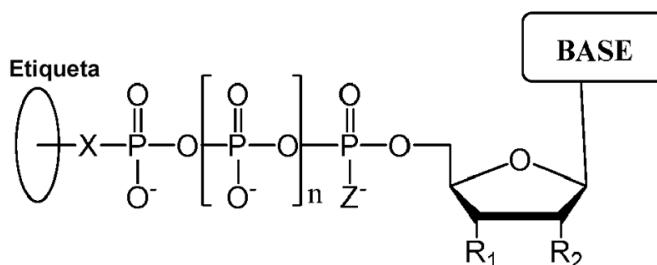
(b) determinar qué análogo de dNPP se ha incorporado en el cebador para formar un producto de extensión de ADN en la etapa (a) aplicando una tensión a través de la membrana y midiendo un cambio electrónico a través del nanoporo que resulta de que el polifosfato que tiene la etiqueta unida al mismo generado en la etapa (a) se transloca a través del nanoporo, en el que el cambio electrónico es diferente para cada valor de n, o para cada tipo de etiqueta diferente, según sea apropiado, identificando de ese modo el residuo de nucleótido en el ADN monocatenario complementario al análogo de dNPP incorporado; y

(c) realizar de manera iterativa las etapas (a) y (b) para cada residuo de nucleótido del ADN monocatenario que está secuenciándose, en el que en cada iteración de la etapa (a) se incorpora el análogo de dNPP en el producto de extensión de ADN resultante de la iteración anterior de la etapa (a) si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-terminal del producto de extensión de ADN,

determinando así la secuencia de nucleótidos del ADN monocatenario.

Método para determinar la secuencia de nucleótidos de un ADN monocatenario que comprende:

(a) poner en contacto el ADN monocatenario, en el que el ADN monocatenario está en una disolución de electrolito en contacto con un nanoporo en una membrana y en el que el ADN monocatenario tiene un cebador hibridado con una parte del mismo, una ADN polimerasa y un análogo de polifosfato de desoxirribonucleótido (dNPP) en condiciones que permiten que la ADN polimerasa catalice la incorporación del análogo de dNPP en el cebador si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-terminal del cebador, para formar un producto de extensión de ADN, en el que el análogo de dNPP tiene la estructura:



en la que la base es adenina, guanina, citosina, uracilo o timina, o un derivado de cada una de las mismas, en la que R_1 es -OH, -O- CH_2N_3 o -O-2-nitrobencilo, en la que R_2 es H, en la que X es O, NH, S o CH_2 , en la que n es 1, 2, 3 ó 4, en la que Z es O, S o BH_3 ,

y en el que si no se incorpora el análogo de dNPP, repetir de manera iterativa la puesta en contacto con un análogo de dNPP diferente hasta que se incorpora un análogo de dNPP, con la condición de que (1) el tipo de base en cada análogo de dNPP es diferente del tipo de base en cada uno de los otros análogos de dNPP, y (2) o bien el valor de n de cada análogo de dNPP es diferente del valor de n de cada uno de los otros tres análogos de dNPP, o bien el valor de n de cada uno de los cuatro análogos de dNPP es el mismo y el tipo de etiqueta en cada análogo de dNPP es diferente del tipo de etiqueta en cada uno de los otros tres análogos de dNPP,

en el que la incorporación de un análogo de dNPP da como resultado la liberación de un polifosfato que tiene la

etiqueta unida al mismo;

5 (b) determinar qué análogo de dNPP se ha incorporado en el cebador para formar un producto de extensión de ADN en la etapa (a) aplicando una tensión a través de la membrana y midiendo un cambio electrónico a través del nanoporo que resulta de que el polifosfato que tiene la etiqueta unida al mismo generado en la etapa (a) se transloca a través del nanoporo, en el que el cambio electrónico es diferente para cada valor de n, o para cada tipo de etiqueta diferente, según sea apropiado, identificando de ese modo el residuo de nucleótido en el ADN monocatenario complementario al análogo de dNPP incorporado;

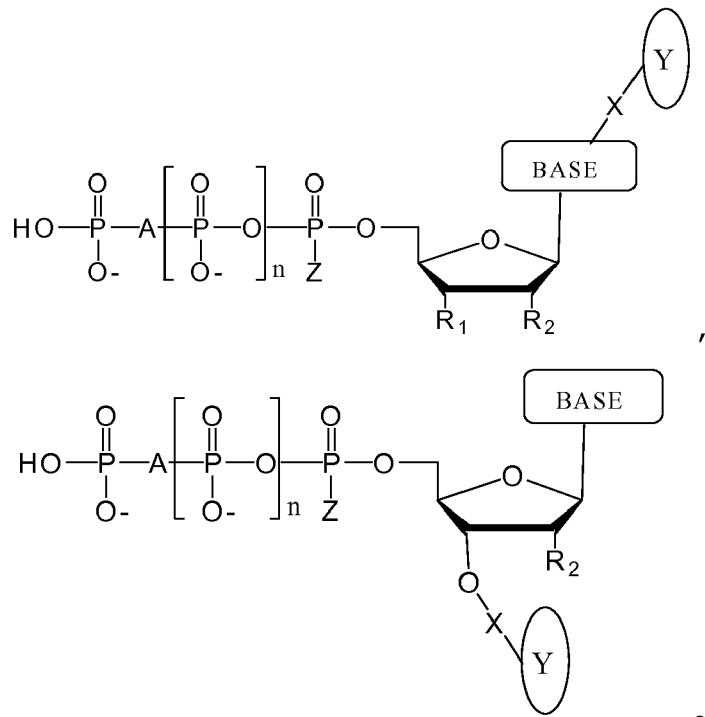
10 (c) realizar de manera iterativa las etapas (a) y (b) para cada residuo de nucleótido del ADN monocatenario que está secuenciándose, en el que en cada iteración de la etapa (a) se incorpora el análogo de dNPP en el producto de extensión de ADN resultante de la iteración anterior de la etapa (a) si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-terminal del producto de extensión de ADN,

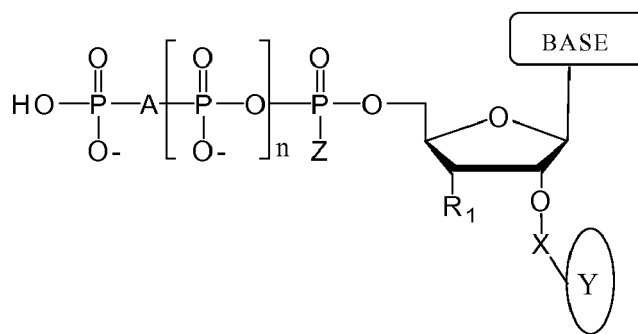
determinando así la secuencia de nucleótidos del ADN monocatenario.

Método para determinar la secuencia de nucleótidos de un ADN monocatenario que comprende:

15 (a) poner en contacto el ADN monocatenario, en el que el ADN monocatenario está en una disolución de electrolito en contacto con un nanoporo en una membrana y en el que el ADN monocatenario tiene un cebador hibridado con una parte del mismo, con una ADN polimerasa y al menos cuatro análogos de polifosfato de desoxirribonucleótido (dNPP) en condiciones que permiten que la ADN polimerasa catalice la incorporación de uno de los análogos de dNPP en el cebador si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-terminal del cebador, para formar un producto de extensión de ADN, en el que cada uno de los cuatro análogos de dNPP tiene una estructura elegida de las siguientes:

20





5 en las que la base es adenina, guanina, citosina, uracilo o timina, o un derivado de cada una de las mismas, en las que Y es una etiqueta, en las que R₁, si está presente, es OH, en las que R₂, si está presente, es H, en las que X es un grupo de unión escindible, en las que Z es O, S o BH₃, en las que n es 1, 2, 3 ó 4, en las que A es O, S, CH₂, CHF, CFF o NH, y con la condición de que (i) el tipo de base en cada análogo de dNPP es diferente del tipo de base en cada uno de los otros tres análogos de dNPP, y (ii) el tipo de etiqueta en cada análogo de dNPP es diferente del tipo de etiqueta en cada uno de los otros tres análogos de dNPP;

(b) escindir la etiqueta del análogo de dNPP incorporado en la etapa (a); y

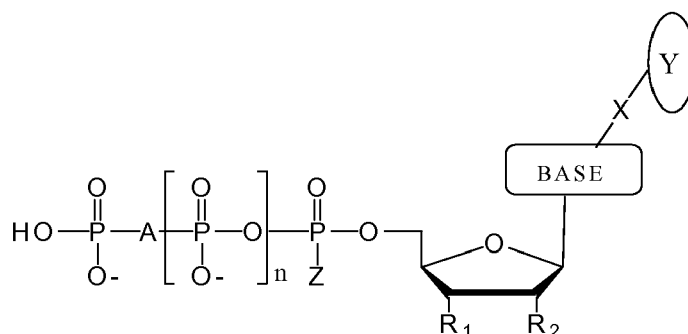
10 (c) determinar qué análogo de dNPP se incorporó en la etapa (a) aplicando una tensión a través de la membrana y midiendo un cambio electrónico a través del nanoporo que resulta de que la etiqueta escindida en la etapa (b) se transloca a través del nanoporo, en el que el cambio electrónico es diferente para cada tipo diferente de etiqueta, identificando de ese modo el residuo de nucleótido en el ADN monocatenario complementario al análogo de dNPP incorporado; y

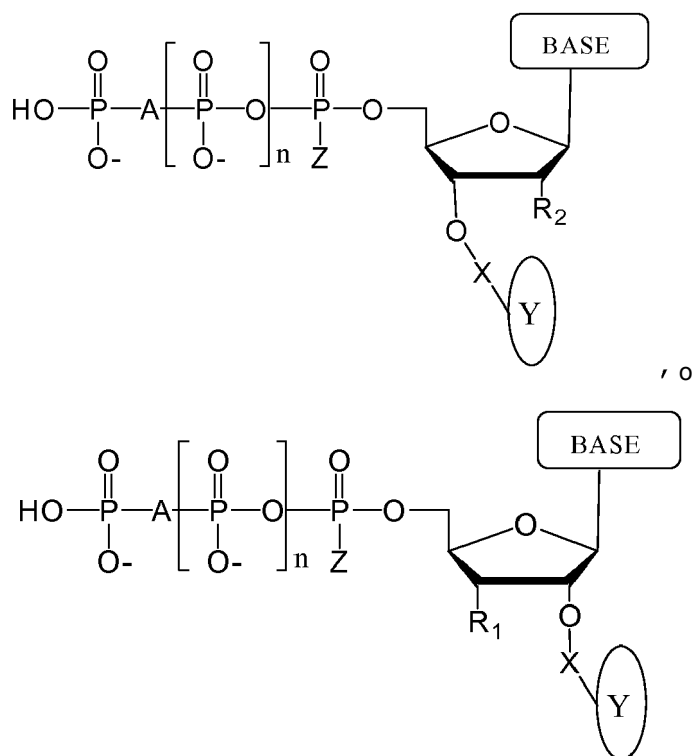
15 (d) realizar de manera iterativa las etapas (a), (b) y (c) para cada residuo de nucleótido del ADN monocatenario que está secuenciándose, en el que en cada iteración de la etapa (a) se incorpora el análogo de dNPP en el producto de extensión de ADN resultante de la iteración anterior de la etapa (a) si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-terminal del producto de extensión de ADN,

determinando así la secuencia de nucleótidos del ADN monocatenario.

20 Método para determinar la secuencia de nucleótidos de un ADN monocatenario que comprende:

25 (a) poner en contacto el ADN monocatenario, en el que el ADN monocatenario está en una disolución de electrolito en contacto con un nanoporo en una membrana, en el que el ADN monocatenario tiene un cebador hibridado con una parte del mismo, una ADN polimerasa y un análogo de polifosfato de desoxirribonucleótido (dNPP) en condiciones que permiten que la ADN polimerasa catalice la incorporación del análogo de dNPP en el cebador si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-terminal del cebador, para formar un producto de extensión de ADN, en el que el análogo de dNPP tiene la estructura:





o

5 en las que la base es adenina, guanina, citosina, uracilo o timina, o derivado de cada una de las mismas, en las que Y es una etiqueta, y en las que R₁ si está presente es OH, -OCH₂N₃ o -O-2-nitrobencilo, R₂ si está presente es H, en las que X es un grupo de unión escindible, en las que Z es O, S o BH₃, en las que n es 1, 2, 3 ó 4, en las que A es O, S, CH₂, CHF, CFF o NH,

10 y si no se incorpora el análogo de dNPP, repetir de manera iterativa la puesta en contacto con un análogo de dNPP diferente hasta que se incorpora un análogo de dNPP, con la condición de que (1) el tipo de base en cada análogo de dNPP es diferente del tipo de base en cada uno de los otros análogos de dNPP, y (2) el tipo de etiqueta en cada análogo de dNPP es diferente del tipo de etiqueta en cada uno de los otros análogos de dNPP,

en el que la incorporación de un análogo de dNPP da como resultado la liberación de un polifosfato que tiene la etiqueta unida al mismo;

(b) escindir la etiqueta del análogo de dNPP incorporado en la etapa (a); y

15 (c) determinar qué análogo de dNPP se incorporó en la etapa (a) para formar un producto de extensión de ADN aplicando una tensión a través de la membrana y midiendo un cambio electrónico a través del nanoporo que resulta de que la etiqueta escindida en la etapa (b) se transloca a través del nanoporo, en el que el cambio electrónico es diferente para cada tipo de etiqueta, identificando de ese modo el residuo de nucleótido en el ADN monocatenario complementario al análogo de dNPP incorporado;

20 (d) realizar de manera iterativa las etapas (a) a (c) para cada residuo de nucleótido del ADN monocatenario que está secuenciándose, en el que en cada iteración de la etapa (a) se incorpora el análogo de dNPP en el producto de extensión de ADN resultante de la iteración anterior de la etapa (a) si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-terminal del producto de extensión de ADN,

determinando así la secuencia de nucleótidos del ADN monocatenario.

25 En una realización de los métodos, la etiqueta es etilenglicol, un aminoácido, un hidrato de carbono, un colorante, un mononucleótido, un dinucleótido, un trinucleótido, un tetranucleótido, un pentanucleótido o un hexanucleótido, un colorante fluorescente, un compuesto quimioluminiscente, un aminoácido, un péptido, un hidrato de carbono, un monofosfato de nucleótido, un difosfato de nucleótido, un ácido alifático o un ácido aromático o un alcohol o un tiol no sustituido o sustituido con uno o más halógenos, un grupo ciano, un grupo nitro, un grupo alquilo, un grupo alquenilo, un grupo alquinilo, un grupo azido.

30 En una realización de los métodos la base se selecciona del grupo que consiste en adenina, guanina, citosina, timina, 7-desazaguanina, 7-desazaadenina o 5-metilcitosina.

En una realización los métodos comprenden además una etapa de lavado tras cada iteración de la etapa (b) para retirar análogos de dNPP no incorporados del contacto con el ADN monocatenario.

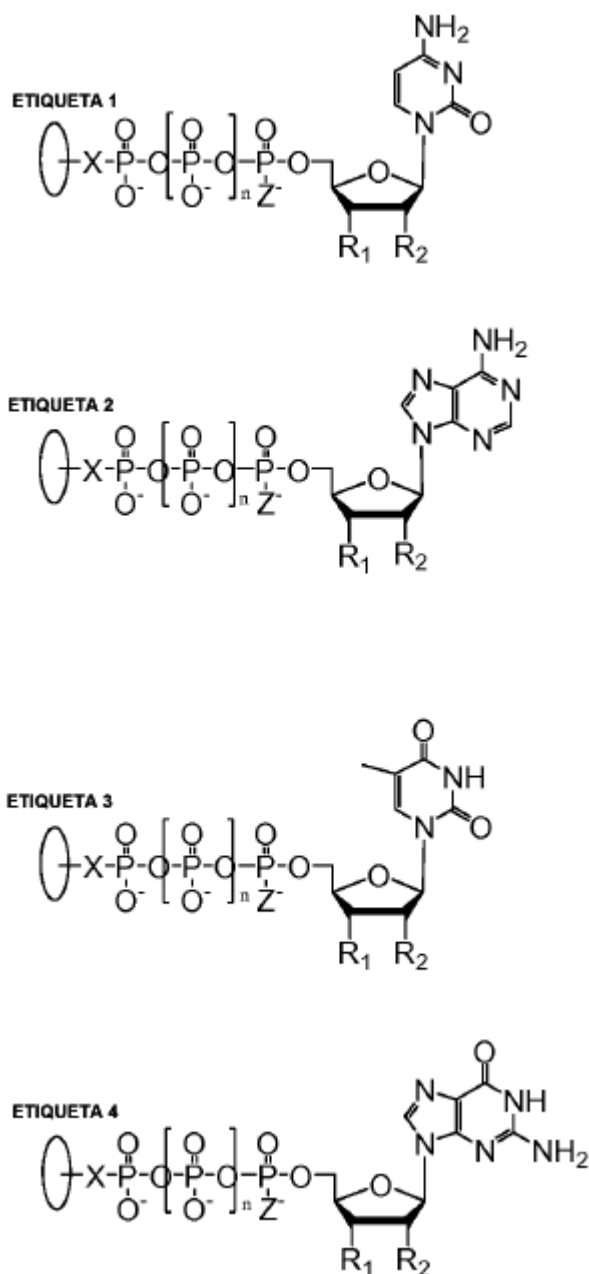
En una realización los métodos comprenden además una etapa de lavado tras cada iteración de la etapa (c) para retirar análogos de dNPP no incorporados del contacto con el ADN monocatenario.

- 5 En una realización los métodos comprenden además que el ADN monocatenario, la disolución de electrolito y el nanoporo en la membrana están ubicados dentro de un único recipiente.

En una realización de los métodos en la que R_1 es $-O-CH_2N_3$, los métodos comprenden opcionalmente además tratar el análogo de dNPP incorporado para retirar el $-CH_2N_3$ y dar como resultado un grupo OH unido a la posición 3' permitiendo así la incorporación de un análogo de dNPP adicional.

- 10 En una realización de los métodos en la que R_1 es $-O-2$ -nitrobencilo, los métodos comprenden opcionalmente además tratar el análogo de nucleótido incorporado para retirar el -2 -nitrobencilo y dar como resultado un grupo OH unido a la posición 3' permitiendo así la incorporación de un análogo de dNPP adicional.

En una realización de los métodos los análogos de dNPP tienen las siguientes estructuras:

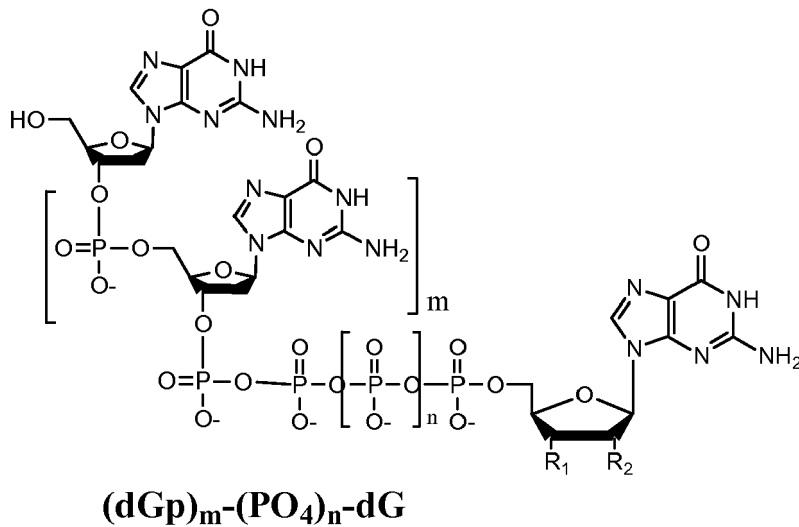
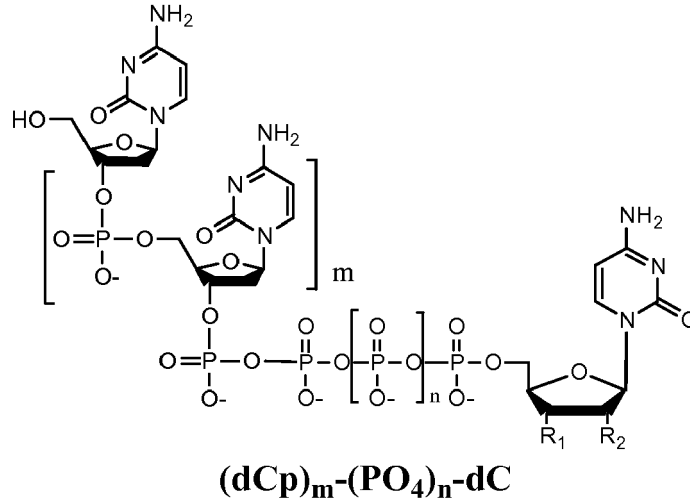


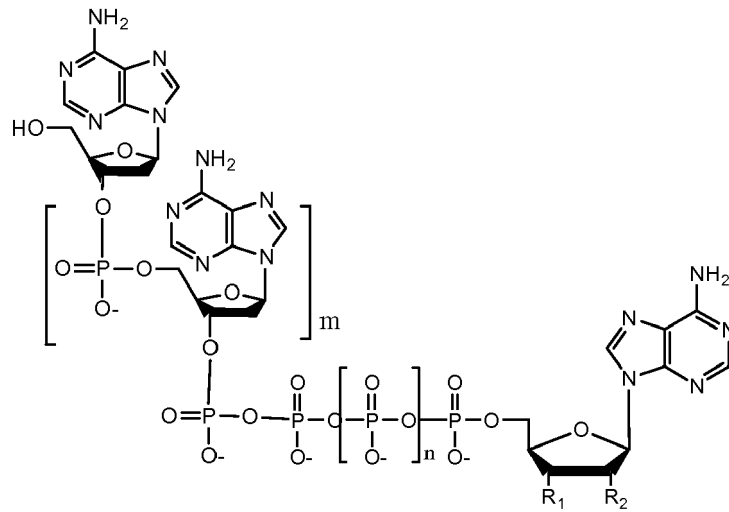
- 15 en las que R_1 es OH, en las que R_2 es H u OH, en las que Z es O, S o BH_3 , y en las que la base es adenina,

guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina.

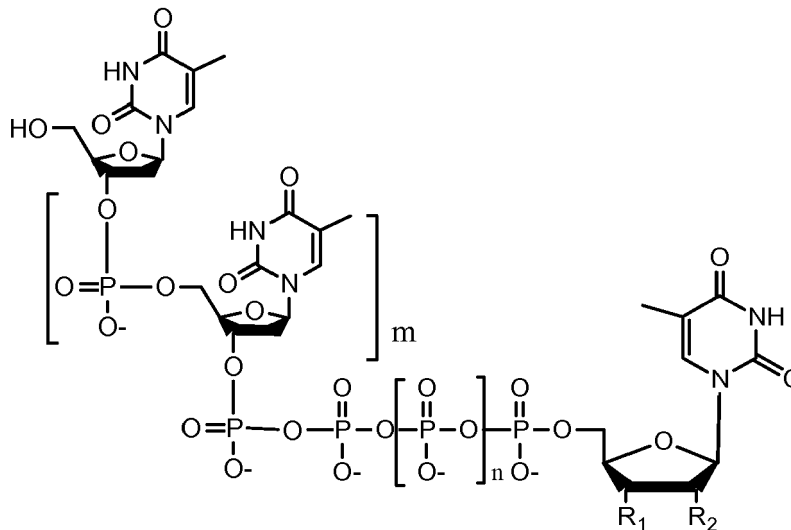
5 En una realización de los métodos la etiqueta es un mononucleótido, un dinucleótido, un trinucleótido, un tetranucleótido, un pentanucleótido o un hexanucleótido y en la que la base del mononucleótido, el dinucleótido, el trinucleótido, el tetranucleótido, el pentanucleótido o el hexanucleótido es el mismo tipo de base que la base del análogo de dNPP.

En una realización de los métodos la etiqueta se elige de las siguientes:





(dAp)_m-(PO₄)_n-dA

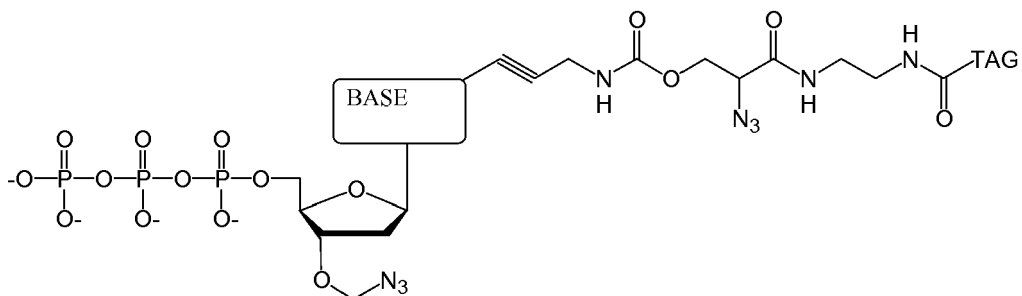


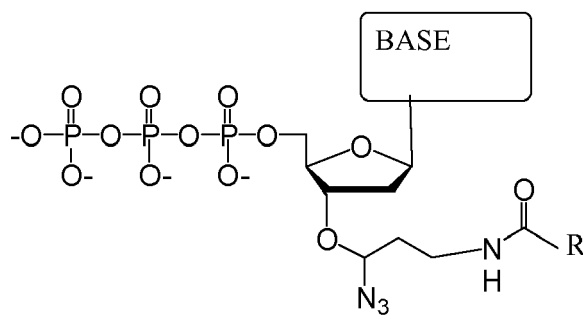
(dTp)_m-(PO₄)_n-dT

5 en las que en cada estructura n es, independientemente, 1, 2, 3 ó 4, y m es, independientemente, un número entero de desde 0 hasta 100, y en las que cuando m es 0 el fosfato terminal del dNPP está unido directamente al átomo de O en 3' del nucleósido mostrado en el lado izquierdo de la estructura, y en las que el valor de n es diferente para cada tipo de base.

En una realización de los métodos m es un número entero de desde 0 hasta 50. En una realización de los métodos, m es un número entero de desde 0 hasta 10.

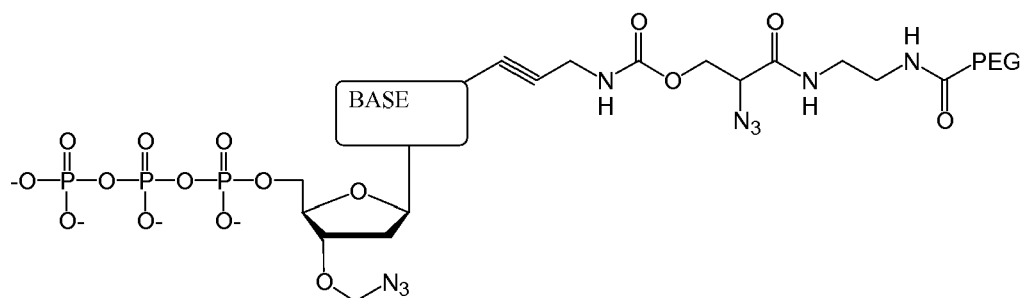
En una realización de los métodos el análogo de dNPP tiene la estructura:





en las que R es un hidrocarbilo sustituido o no sustituido, de hasta 3000 Dalton, y en las que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina.

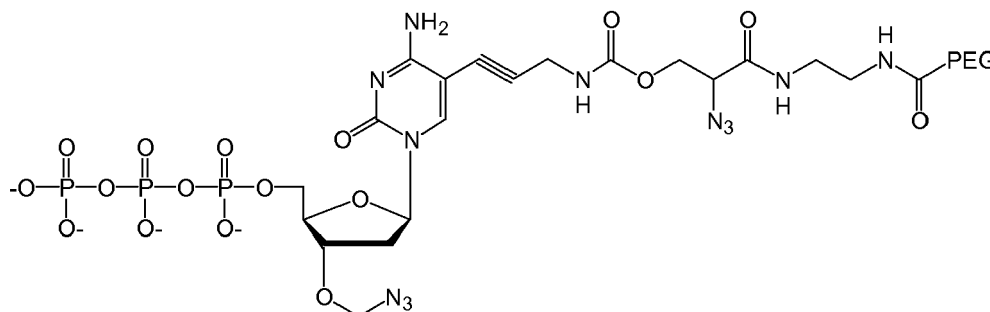
En una realización de los métodos el análogo de dNPP tiene la estructura:



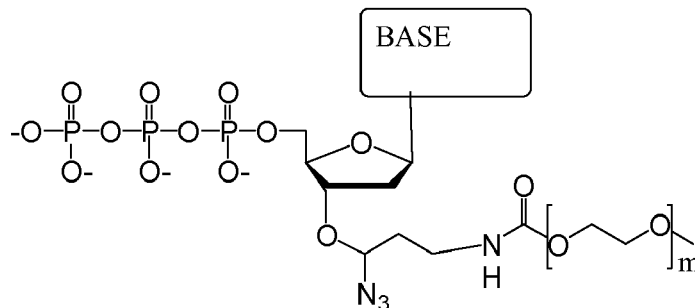
5

en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina.

En una realización de los métodos el análogo de dNPP tiene la estructura:



En una realización de los métodos el análogo de dNPP tiene la estructura:



10

en la que m es un número entero de desde 1-50, y en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina.

En una realización de los métodos el cambio electrónico es un cambio en la amplitud de corriente.

En una realización de los métodos el cambio electrónico es un cambio en la conductancia del nanoporo.

15 En una realización de los métodos el nanoporo es biológico. En una realización de los métodos, el nanoporo es

proteico. En una realización de los métodos el nanoporo comprende alfa-hemolisina. En una realización de los métodos, el nanoporo es grafeno. En una realización de los métodos, el nanoporo es un nanoporo en estado sólido. En una realización de los métodos el nanoporo está en una membrana en estado sólido.

5 En una realización de los métodos el ADN monocatenario, el cebador o la ADN polimerasa están unidos a una superficie sólida.

En otra realización de los métodos el nanoporo es parte de una matriz de nanoporos.

Procedimiento para producir un análogo de trifosfato de nucleótido, en el que el análogo de trifosfato de nucleótido difiere de un trifosfato de nucleótido por tener una etiqueta unida al fosfato terminal del mismo, que comprende:

10 a) poner en contacto un trifosfato de nucleótido con diciclohexilcarbodiimida/dimetilformamida en condiciones que permiten la producción de un trimetafosfato cíclico;

b) poner en contacto el producto resultante de la etapa a) con una etiqueta que tiene un grupo hidroxilo o amino unido a la misma en condiciones que permiten la apertura nucleófila del trimetafosfato cíclico para unir la etiqueta a un fosfato terminal formando así el análogo de trifosfato de nucleótido.

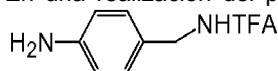
15 Procedimiento para producir un análogo de trifosfato de nucleótido, en el que el análogo de trifosfato de nucleótido difiere de un trifosfato de nucleótido por tener una etiqueta unida al fosfato terminal del mismo, que comprende:

a) poner en contacto un trifosfato de nucleótido con diciclohexilcarbodiimida/dimetilformamida en condiciones que permiten la producción de un trimetafosfato cíclico;

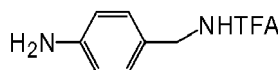
b) poner en contacto el producto resultante de la etapa a) con un nucleófilo para formar un compuesto funcionalizado con -OH o -NH₂;

20 c) hacer reaccionar el producto de la etapa b) con una etiqueta que tiene un grupo -COR unido a la misma en condiciones que permiten que la etiqueta se una indirectamente a un fosfato terminal formando así el análogo de trifosfato de nucleótido.

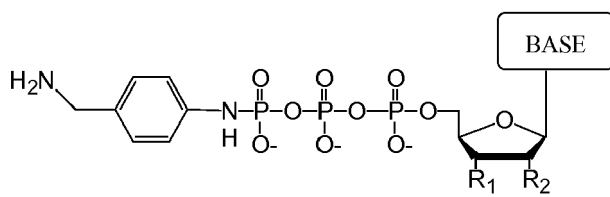
En una realización del presente procedimiento el nucleófilo es H₂N-R-OH, H₂N-R-NH₂, R'S-R-OH, R'S-R-NH₂, o



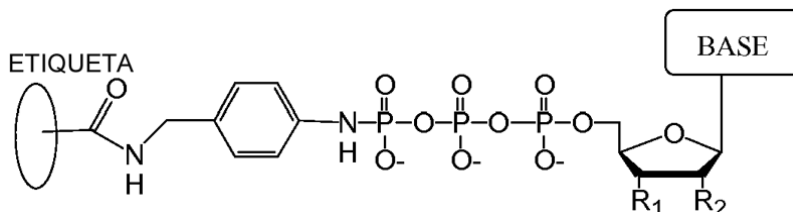
25 En una realización el presente procedimiento comprende en la etapa b) poner en contacto el producto resultante de la etapa a) con un compuesto que tiene la estructura:



y después NH₄OH para formar un compuesto que tiene la estructura:



30 y hacer reaccionar el producto de la etapa b) con una etiqueta que tiene un grupo -COR unido a la misma en condiciones que permiten que la etiqueta se una indirectamente a un fosfato terminal formando así el análogo de trifosfato de nucleótido que tiene la estructura:

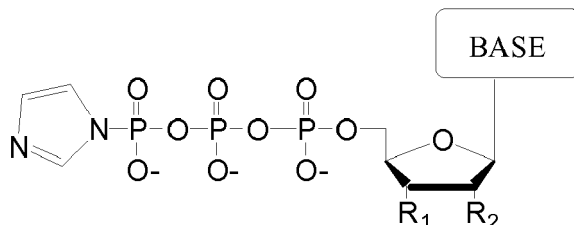


en la que R₁ es OH, en la que R₂ es H u OH, en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina.

35 Procedimiento para producir un análogo de tetrafosfato de nucleótido, en el que el análogo de tetrafosfato de nucleótido difiere de un tetrafosfato de nucleótido por tener una etiqueta unida al fosfato terminal del mismo, que

comprende:

a) poner en contacto un trifosfato de nucleótido con 1,1'-carbonildiimidazol/dimetilformamida en condiciones que permiten la formación de la siguiente estructura:

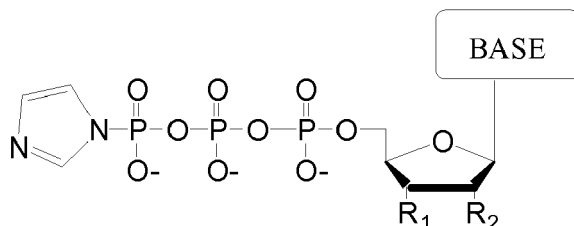


5 en la que R₁ es OH, en la que R₂ es H u OH, en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina;

b) poner en contacto el producto resultante de la etapa a) con una etiqueta que tiene un grupo monosfosfato unido a la misma en condiciones que permiten la formación del análogo de tetrafosfato de nucleótido.

10 Procedimiento para producir un análogo de tetrafosfato de nucleótido, en el que el análogo de tetrafosfato de nucleótido difiere de un tetrafosfato de nucleótido por tener una etiqueta unida al fosfato terminal del mismo, que comprende:

a) poner en contacto un trifosfato de nucleótido con 1,1'-carbonildiimidazol/dimetilformamida en condiciones que permiten la formación de la siguiente estructura:



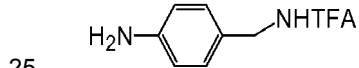
15 en la que R₁ es OH, en la que R₂ es H u OH, en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina;

b) poner en contacto el producto resultante de la etapa a) con ácido fosfórico en condiciones que permiten la formación de un tetrafosfato de nucleótido;

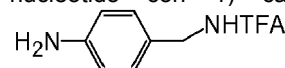
20 c) poner en contacto el tetrafosfato de nucleótido con 1) carbonildiimidazol/dimetilformamida; 2) un nucleófilo y después 3) NH₄OH para formar un compuesto funcionalizado con -OH o -NH₂;

d) poner en contacto el producto de la etapa c) con una etiqueta que tiene un grupo -COR unido a la misma en condiciones que permiten que la etiqueta se una indirectamente a un fosfato terminal formando así el análogo de tetrafosfato de nucleótido.

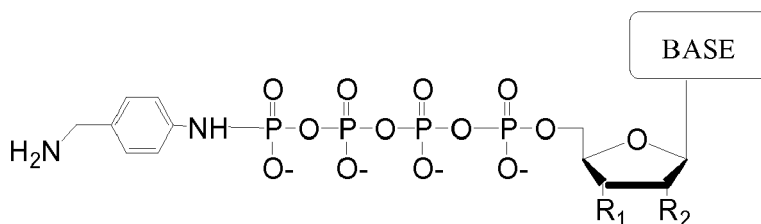
En una realización del presente procedimiento, el nucleófilo es H₂N-R-OH, H₂N-R-NH₂, R'S-R-OH, R'S-R-NH₂, o



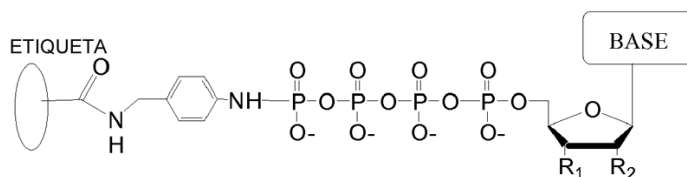
En una realización, el presente procedimiento comprende en la etapa b) poner en contacto el tetrafosfato de nucleótido con 1) carbonildiimidazol/ dimetilformamida; 2) un compuesto que tiene la estructura:



y después 3) NH₄OH para formar un compuesto que tiene la estructura:



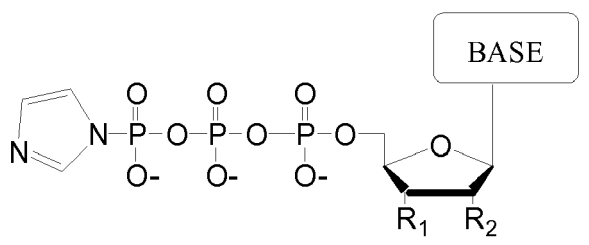
y poner en contacto el producto de la etapa b) con una etiqueta que tiene un grupo -COR unido a la misma en condiciones que permiten que la etiqueta se una indirectamente a un fosfato terminal formando así el análogo de trifosfato de nucleótido que tiene la estructura:



- 5 en la que R₁ es OH, en la que R₂ es H u OH, en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina.

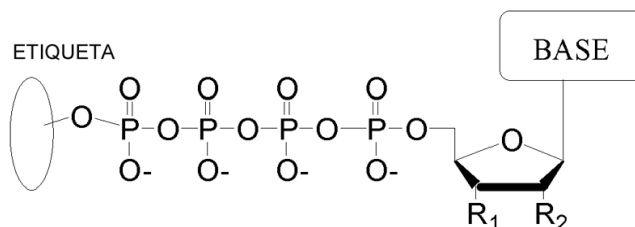
Procedimiento para producir un análogo de tetrafosfato de nucleótido, en el que el análogo de tetrafosfato de nucleótido difiere de un tetrafosfato de nucleótido por tener una etiqueta unida al fosfato terminal del mismo, que comprende:

- 10 a) poner en contacto un trifosfato de nucleótido con 1,1'-carbonildiimidazol/dimetilformamida en condiciones que permiten la formación de la siguiente estructura:



b) poner en contacto el producto resultante de la etapa a) con ácido fosfórico en condiciones que permiten la formación de un tetrafosfato de nucleótido;

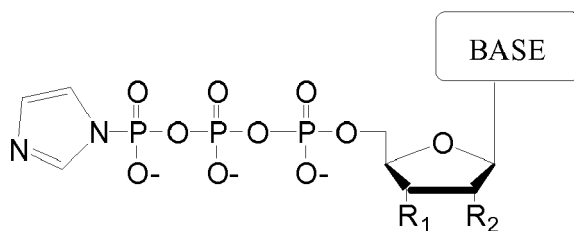
- 15 c) poner en contacto el tetrafosfato de nucleótido con carbonildiimidazol/dimetilformamida y una etiqueta que tiene un grupo hidroxilo o amino unido a la misma para formar un compuesto que tiene la estructura:



en la que R₁ es OH, en la que R₂ es H u OH, en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina.

- 20 Procedimiento para producir un análogo de pentafosfato de nucleótido, en el que el análogo de pentafosfato de nucleótido difiere de un pentafosfato de nucleótido por tener una etiqueta unida al fosfato terminal del mismo, que comprende:

a) poner en contacto un trifosfato de nucleótido con 1,1'-carbonildiimidazol/dimetilformamida en condiciones que permiten la formación de la siguiente estructura:



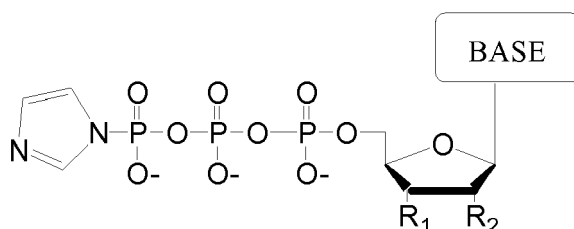
- 25 en la que R₁ es OH, en la que R₂ es H u OH, en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina;

b) poner en contacto el producto resultante de la etapa a) con una etiqueta que tiene un grupo pirofosfato unido a la misma en condiciones que permiten la formación del análogo de pentafofosfato de nucleótido.

Procedimiento para producir un análogo de pentafofosfato de nucleótido, en el que el análogo de pentafofosfato de nucleótido difiere de un pentafofosfato de nucleótido por tener una etiqueta unida al fosfato terminal del mismo, que comprende:

5

a) poner en contacto un trifosfato de nucleótido con 1,1'-carbonildiimidazol/dimetilformamida en condiciones que permiten la formación de la siguiente estructura:



10

en la que R₁ es OH, en la que R₂ es H u OH, en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina;

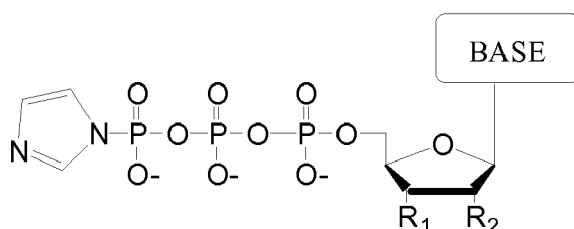
b) poner en contacto el producto resultante de la etapa a) con un grupo pirofosfato en condiciones que permiten la formación de un pentafofosfato de nucleótido;

c) poner en contacto el pentafofosfato de nucleótido con carbonildiimidazol/dimetilformamida y una etiqueta que tiene un grupo hidroxilo o amino unido a la misma para formar el análogo de pentafofosfato de nucleótido.

15

Procedimiento para producir un análogo de hexafofosfato de nucleótido, en el que el análogo de hexafofosfato de nucleótido difiere de un hexafofosfato de nucleótido por tener una etiqueta unida al fosfato terminal del mismo, que comprende:

a) poner en contacto un trifosfato de nucleótido con 1,1'-carbonildiimidazol/dimetilformamida en condiciones que permiten la formación de la siguiente estructura:



20

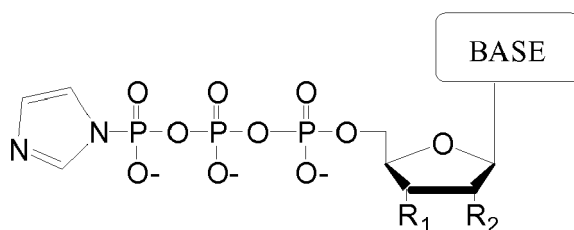
en la que R₁ es OH, en la que R₂ es H u OH, en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina;

b) poner en contacto el producto resultante de la etapa a) con una etiqueta que tiene un grupo trifosfato unido a la misma en condiciones que permiten la formación del análogo de hexafofosfato de nucleótido.

25

Procedimiento para producir un análogo de hexafofosfato de nucleótido, en el que el análogo de hexafofosfato de nucleótido difiere de un hexafofosfato de nucleótido por tener una etiqueta unida al fosfato terminal del mismo, que comprende:

a) poner en contacto un trifosfato de nucleótido con 1,1'-carbonildiimidazol/dimetilformamida en condiciones que permiten la formación de la siguiente estructura:



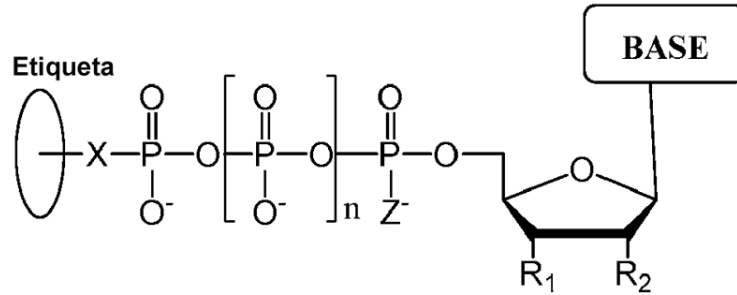
30

en la que R₁ es OH, en la que R₂ es H u OH, en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina;

b) poner en contacto el producto resultante de la etapa a) con un grupo trifosfato en condiciones que permiten la formación de un hexafosfato de nucleótido;

c) poner en contacto el hexafosfato de nucleótido con carbonildiimidazol/dimetilformamida y una etiqueta que tiene un grupo hidroxilo o amino unido a la misma para formar el análogo de hexafosfato de nucleótido.

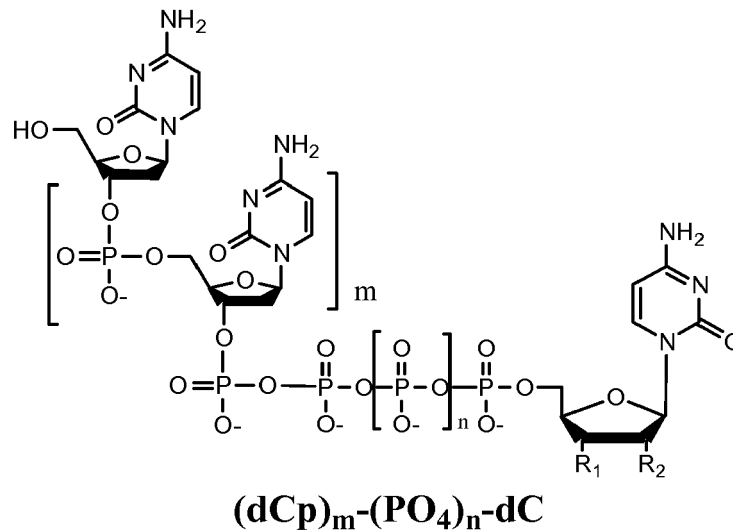
5 Compuesto que tiene la estructura:

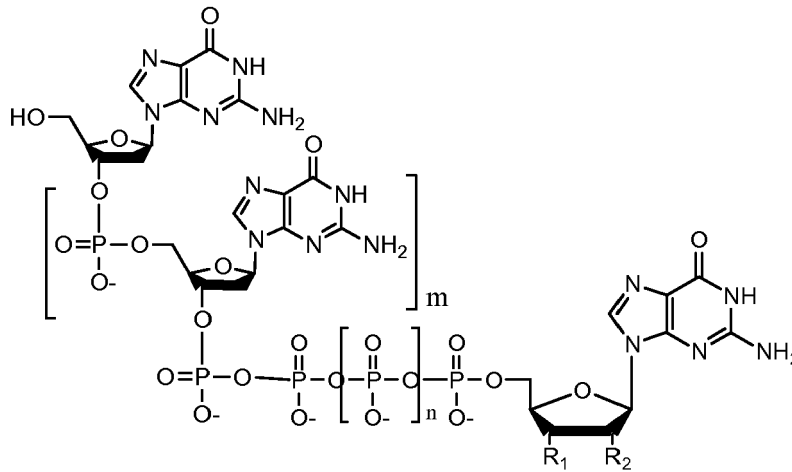


10 en la que la etiqueta es etilenglicol, un aminoácido, un hidrato de carbono, un colorante, mononucleótido, dinucleótido, trinucleótido, tetranucleótido, pentanucleótido o hexanucleótido, en la que R₁ es OH, en la que R₂ es H u OH, en la que X es O, NH, S o CH₂, en la que Z es O, S o BH₃, en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina, y en la que n es 1, 2, 3 ó 4.

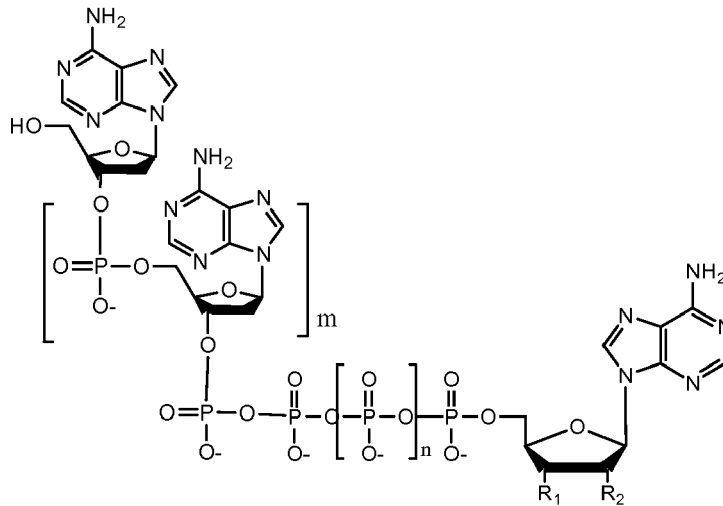
En una realización R₂ es H. En una realización R₂ es OH.

Compuesto que tiene la estructura:

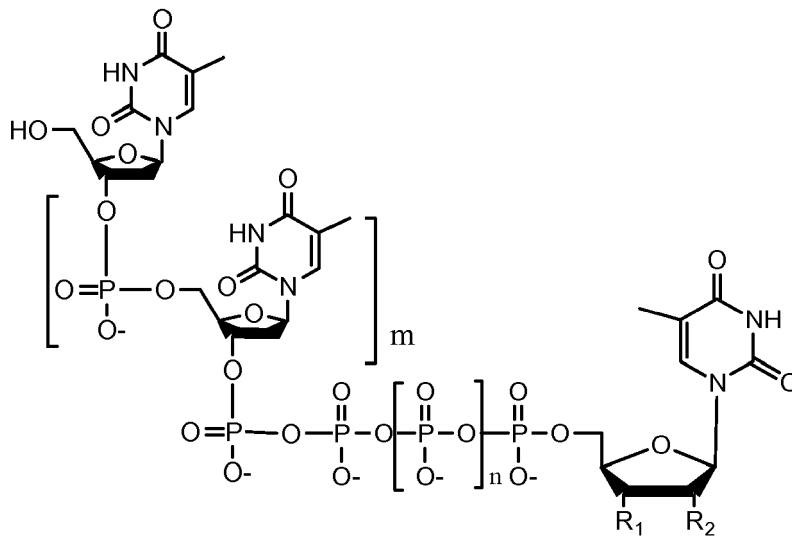




(dGp)_m-(PO₄)_n-dG



(dAp)_m-(PO₄)_n-dA



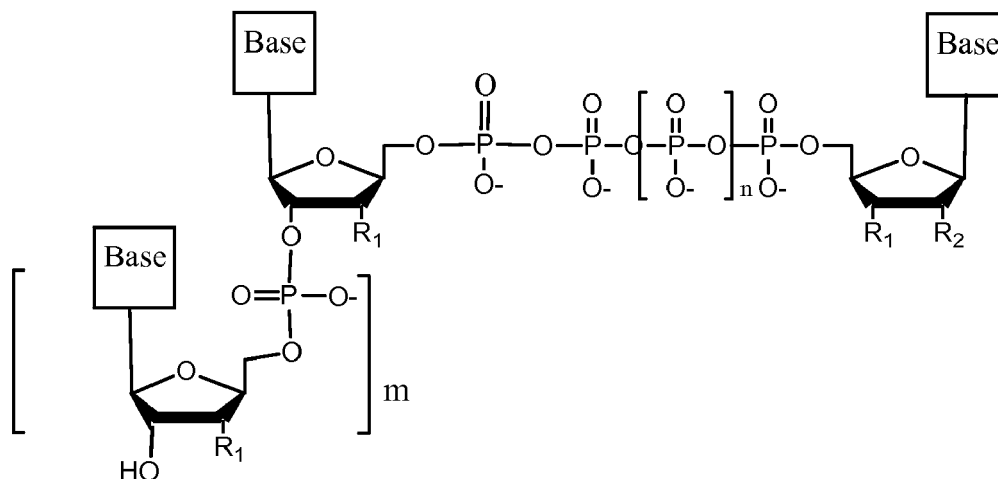
(dTp)_m-(PO₄)_n-dT

en las que en cada estructura n es, independientemente, 1, 2, 3 ó 4, y m es, independientemente, un número entero

de desde 0 hasta 100, y en las que cuando m es 0 el fosfato terminal del dNTP está unido directamente al átomo de O en 3' del nucleósido mostrado en el lado izquierdo de la estructura, en las que R₁ es -OH, o -O-CH₂N₃, y R₂ es H u OH.

5 En una realización m es de desde 0 hasta 50. En una realización m es de desde 0 hasta 10. En una realización R₁ es -OH. En una realización R₂ es -H. En una realización R₂ es -OH.

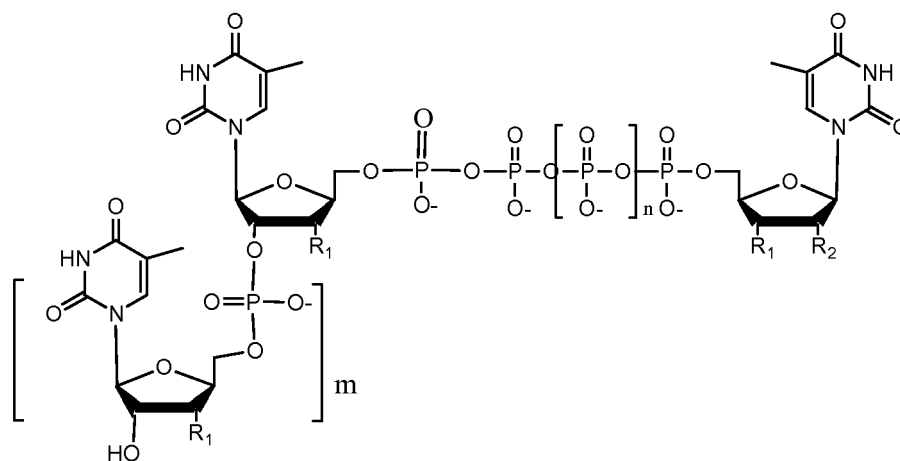
Compuesto que tiene la estructura:



en la que m un número entero de desde 0 hasta 100, y en el que el compuesto comprende un único tipo de base, y en la que la base es adenina, guanina, citosina, uracilo o timina o un derivado de cada una de las mismas.

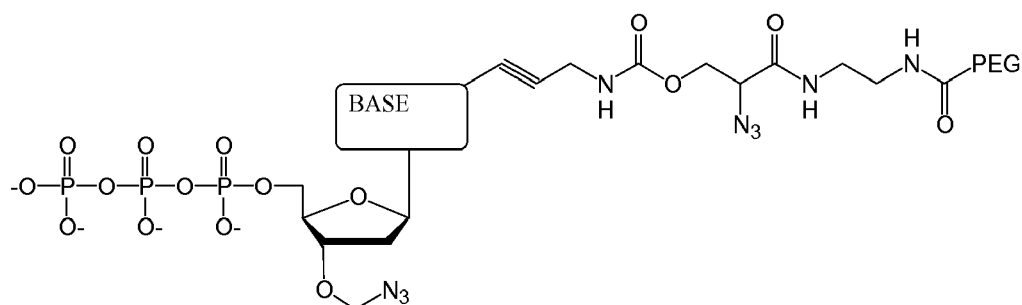
10 En una realización m es de desde 0 hasta 50. En una realización m es de desde 0 hasta 10.

En una realización el compuesto tiene la estructura:



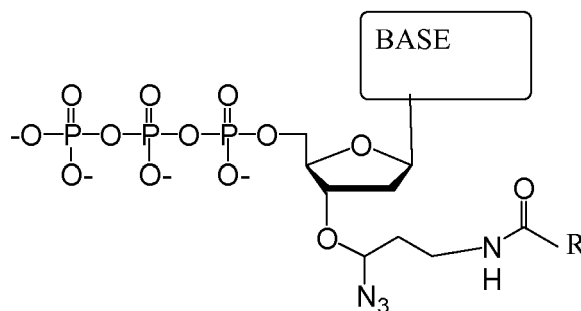
en la que m es un número entero de desde 0 hasta 100.

Compuesto que tiene la estructura:



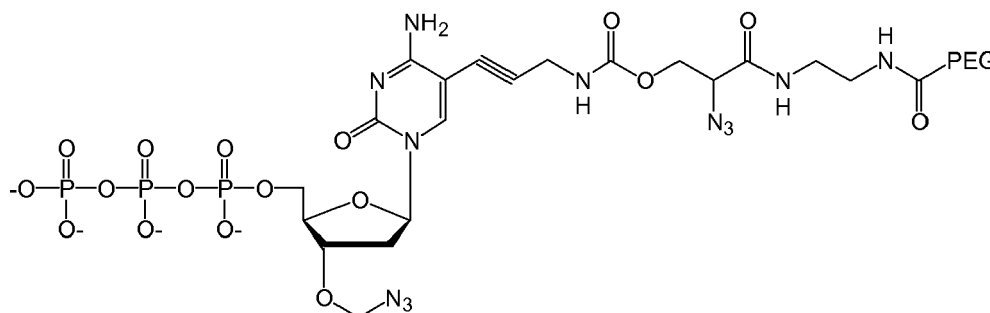
en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina.

Compuesto que tiene la estructura:

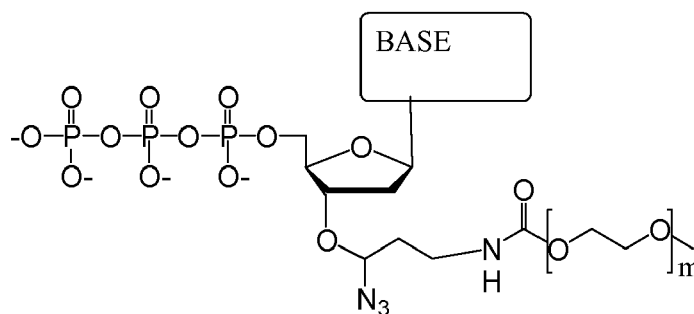


5 en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina, y R es un hidrocarbilo sustituido o no sustituido, de hasta 3000 Dalton.

Compuesto que tiene la estructura:

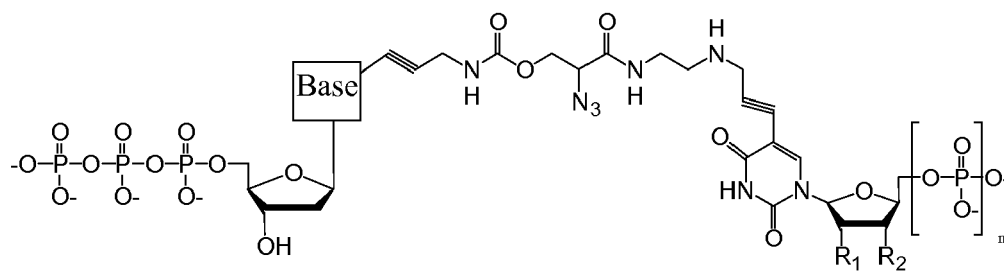


Compuesto que tiene la estructura:



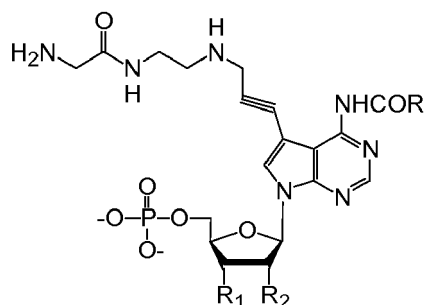
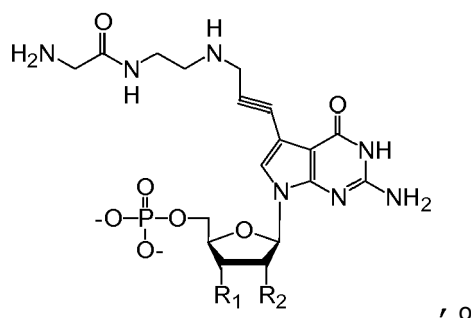
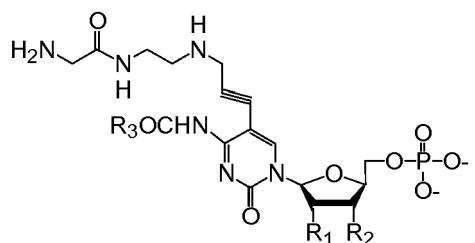
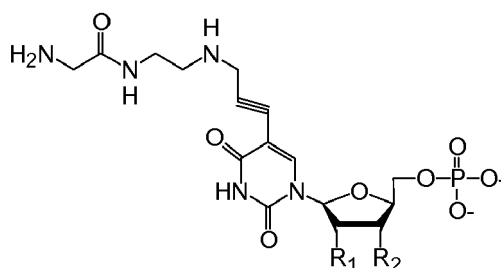
10 en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina, y m es un número entero de desde 1-50.

Compuesto que tiene la estructura:



15 en la que n es 1 ó 2 y la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina.

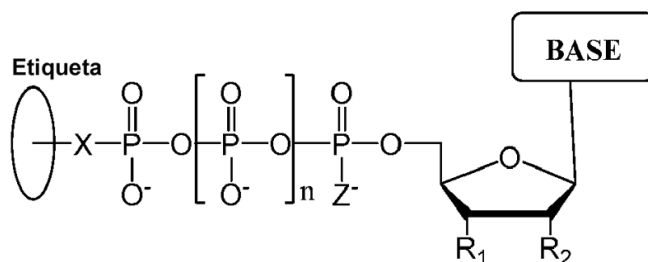
Compuesto que tiene la estructura:



5 en las que R_1 es $-OH$, o $-O-CH_2N_3$, y R_2 es H u OH .

Método para determinar la secuencia de nucleótidos de un ARN monocatenario que comprende:

10 (a) poner en contacto el ARN monocatenario, en el que el ARN monocatenario está en una disolución de electrolito en contacto con un nanoporo en una membrana, en el que el ARN monocatenario tiene un cebador hibridado con una parte del mismo, con una ARN polimerasa y al menos cuatro análogos de polifosfato de ribonucleótido (rNPP) en condiciones que permiten que la ARN polimerasa catalice la incorporación de uno de los análogos de rNPP en el cebador si es complementario al residuo de nucleótido del ARN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ARN monocatenario hibridado con el residuo de nucleótido 3'-terminal del cebador, para formar un producto de extensión de ARN, en el que cada uno de los cuatro análogos de rNPP tiene la estructura:



5 en la que la base es adenina, guanina, citosina, timina o uracilo, o un derivado de cada una de las mismas, en la que R_1 es OH, en la que R_2 es OH, en la que X es O, NH, S o CH_2 , en la que n es 1, 2, 3 ó 4, en la que Z es O, S o BH_3 , y con la condición de que (i) el tipo de base en cada análogo de rNPP es diferente del tipo de base en cada uno de los otros tres análogos de rNPP, y (ii) o bien el valor de n de cada análogo de rNPP es diferente del valor de n de cada uno de los otros tres análogos de rNPP, o bien el valor de n de cada uno de los cuatro análogos de rNPP es el mismo y el tipo de etiqueta en cada análogo de rNPP es diferente del tipo de etiqueta en cada uno de los otros tres análogos de rNPP,

10 en el que la incorporación del análogo de rNPP da como resultado la liberación de un polifosfato que tiene la etiqueta unida al mismo; y

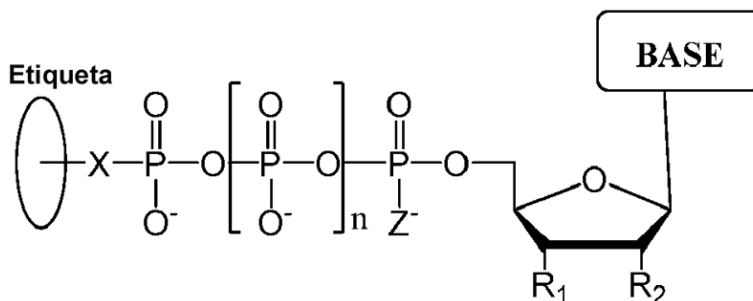
15 (b) determinar qué análogo de rNPP se ha incorporado en el cebador para formar un producto de extensión de ARN en la etapa (a) aplicando una tensión a través de la membrana y midiendo un cambio electrónico a través del nanoporo que resulta de que el polifosfato que tiene la etiqueta unida al mismo generado en la etapa (a) se transloca a través del nanoporo, en el que el cambio electrónico es diferente para cada valor de n, o para cada tipo de etiqueta diferente, según sea apropiado, identificando de ese modo el residuo de nucleótido en el ARN monocatenario complementario al análogo de rNPP incorporado; y

20 (c) realizar de manera iterativa las etapas (a) y (b) para cada residuo de nucleótido del ARN monocatenario que está secuenciándose, en el que en cada iteración de la etapa (a) el análogo de rNPP se incorpora en el producto de extensión de ARN resultante de la iteración anterior de la etapa (a) si es complementario al residuo de nucleótido del ARN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ARN monocatenario hibridado con el residuo de nucleótido 3'-terminal del producto de extensión de ARN,

determinando así la secuencia de nucleótidos del ARN monocatenario.

Método para determinar la secuencia de nucleótidos de un ARN monocatenario que comprende:

25 (a) poner en contacto el ARN monocatenario, en el que el ARN monocatenario está en una disolución de electrolito en contacto con un nanoporo en una membrana y en el que el ARN monocatenario tiene un cebador hibridado con una parte del mismo, una ARN polimerasa y un análogo de polifosfato de ribonucleótido (rNPP) en condiciones que permiten que la ARN polimerasa catalice la incorporación del análogo de rNPP en el cebador si es complementario al residuo de nucleótido del ARN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ARN monocatenario hibridado con el residuo de nucleótido 3'-terminal del cebador, para formar un
30 producto de extensión de ARN, en el que el análogo de rNPP tiene la estructura:



en la que la base es adenina, guanina, citosina, uracilo o timina, en la que R_1 es -OH, -O- CH_2N_3 o -O-2-nitrobencilo, en la que R_2 es -OH, en la que X es O, NH, S o CH_2 , en la que n es 1, 2, 3 ó 4, en la que Z es O, S o BH_3 ,

35 y en el que si no se incorpora el análogo de rNPP, repetir de manera iterativa la puesta en contacto con un análogo de rNPP diferente hasta que se incorpora un análogo de rNPP, con la condición de que (1) el tipo de base en cada análogo de rNPP es diferente del tipo de base en cada uno de los otros análogos de rNPP, y (2) o bien el valor de n de cada análogo de rNPP es diferente del valor de n de cada uno de los otros tres análogos de rNPP, o bien el valor de n de cada uno de los cuatro análogos de rNPP es el mismo y el tipo de etiqueta en cada análogo de rNPP es diferente del tipo de etiqueta en cada uno de los otros tres análogos de rNPP,

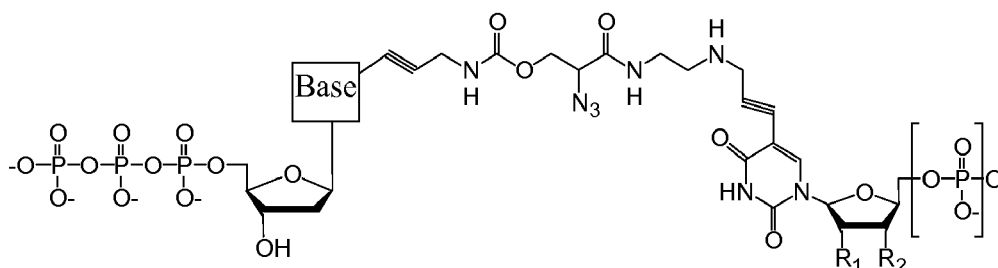
en el que la incorporación de un análogo de rNPP da como resultado la liberación de un polifosfato que tiene la etiqueta unida al mismo;

5 (b) determinar qué análogo de rNPP se ha incorporado en el cebador para formar un producto de extensión de ARN en la etapa (a) aplicando una tensión a través de la membrana y midiendo un cambio electrónico a través del nanoporo que resulta de que el polifosfato que tiene la etiqueta unida al mismo generado en la etapa (a) se transloca a través del nanoporo, en el que el cambio electrónico es diferente para cada valor de n , o diferente para cada tipo de etiqueta, según sea apropiado, identificando de ese modo el residuo de nucleótido en el ARN monocatenario complementario al análogo de dNPP incorporado;

10 (c) realizar de manera iterativa las etapas (a) y (b) para cada residuo de nucleótido del ARN monocatenario que está secuenciándose, en el que en cada iteración de la etapa (a) el análogo de rNPP se incorpora en el producto de extensión de ARN resultante de la iteración anterior de la etapa (a) si es complementario al residuo de nucleótido del ARN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ARN monocatenario hibridado con el residuo de nucleótido 3'-terminal del producto de extensión de ARN,

determinando así la secuencia de nucleótidos del ARN monocatenario.

15 En una realización el análogo de dNPP tiene la estructura:



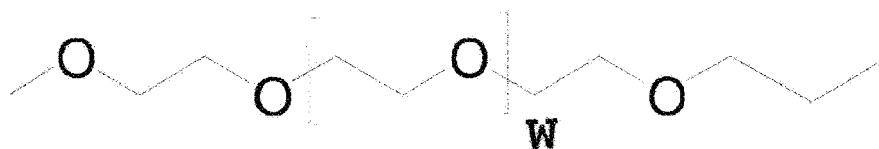
en la que n es 1 ó 2 y la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina.

20 En una realización el nanoporo biológico está integrado con electrónica de CMOS. En otra realización el nanoporo en estado sólido está integrado con electrónica de CMOS.

En una realización la unión a la superficie sólida se realiza a través de uniones de biotina-estreptavidina. En otra realización la ADN polimerasa se une a la superficie sólida a través de superficie de oro modificada con una monocapa autoensamblada de alcanotiol funcionalizada con grupos amino, en la que los grupos amino se modifican para dar ésteres NHS para la unión a grupos amino en la ADN polimerasa.

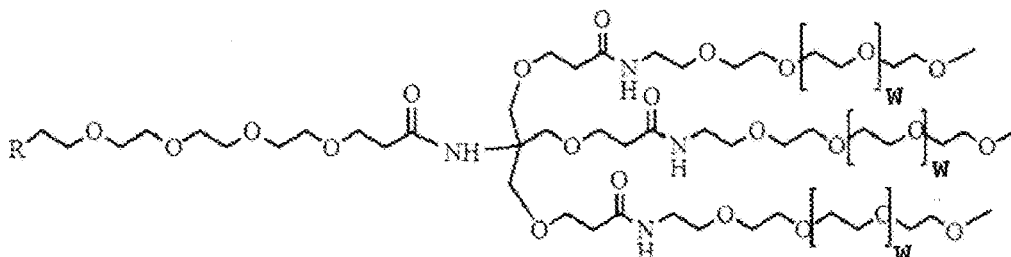
25 En una realización el análogo de dNPP es un polifosfato de nucleósido marcado con etiqueta de fosfato terminal. En una realización adicional cada tipo de análogo de dNPP tiene una etiqueta de polietilenglicol que difiere en cuanto al tamaño de las etiquetas de polietilenglicol de cada uno de los otros tres tipos de análogos de dNPP.

En una realización la etiqueta tiene la siguiente estructura:



30 en la que W es un número entero de entre 0 y 100.

En otra realización la etiqueta tiene la siguiente estructura:



en la que R es NH₂, OH, COOH, CHO, SH o N₃, y W es un número entero de desde 0 hasta 100.

Composición que comprende al menos cuatro análogos de polifosfato de desoxinucleótido (dNPP), que tienen, cada uno, una estructura seleccionada de las estructuras expuestas en las reivindicaciones 74 y 75, en la que cada uno de los cuatro análogos de dNPP comprende un tipo de base diferente del tipo de base de los otros tres análogos de dNPP.

En una realización, cada uno de los cuatro análogos de dNPP tiene una etiqueta de polietilenglicol que es diferente en cuanto al tamaño de las etiquetas de polietilenglicol de cada uno de los otros tres análogos de dNPP.

En una realización la carga neta en el polifosfato de nucleósido marcado con etiqueta es neutra. En otra realización la etiqueta liberada tiene una carga positiva.

10 En una realización, el método comprende además una etapa de tratamiento con fosfatasa alcalina tras la etapa b), en el que la fosfatasa alcalina hidroliza los grupos fosfato libres en el pirofosfato con etiqueta liberado.

En una realización se inmovilizan múltiples copias del ADN monocatenario sobre una perla.

Un "derivado" de adenina, guanina, citosina, timina o uracilo, incluye una 7-desaza-purina y una 5-metilpirimidina. Los ejemplos incluyen 7-desaza-adenina, 7-desaza-guanina y 5-metil-citosina.

15 La presente invención también proporciona un compuesto que tiene la estructura de cualquiera de los compuestos expuestos en las figuras y/o los esquemas de la presente solicitud.

La presente invención también proporciona un análogo de dNPP que comprende una etiqueta que tiene la estructura de cualquiera de las etiquetas expuestas en las figuras y/o los esquemas de la presente solicitud.

20 En una realización, la etiqueta es un hidrocarbilo, sustituido o no sustituido, tal como un alquilo, alquenilo, alquinilo, y que tiene una masa de 3000 Dalton o menos.

Tal como se usa en el presente documento, "alquilo" incluye grupos hidrocarbonados alifáticos saturados, tanto ramificados como de cadena lineal, que tienen el número especificado de átomos de carbono y pueden estar sustituidos o no sustituidos. Por tanto, se define que C1-Cn tal como en "alquilo C1-Cn" incluye grupos que tienen 1, 2, ..., n-1 o n carbonos en una disposición lineal o ramificada. Por ejemplo, se define que un "alquilo C1-C5" incluye grupos que tienen 1, 2, 3, 4 ó 5 carbonos en una disposición lineal o ramificada, e incluye específicamente metilo, etilo, n-propilo, isopropilo, n-butilo, t-butilo, y pentilo.

25 Tal como se usa en el presente documento, "alquenilo" se refiere a un radical hidrocarbonado no aromático, lineal o ramificado, que contiene al menos 1 doble enlace carbono-carbono, y puede estar presente hasta el número máximo posible de dobles enlaces carbono-carbono no aromáticos, y puede estar sustituido o no sustituido. Por ejemplo, "alquenilo C2-C5" significa un radical alquenilo que tiene 2, 3, 4 ó 5, átomos de carbono, y hasta 1, 2, 3 ó 4, dobles enlaces carbono-carbono, respectivamente. Los grupos alquenilo incluyen etenilo, propenilo y butenilo.

30 El término "alquinilo" se refiere a un radical hidrocarbonado lineal o ramificado, que contiene al menos 1 triple enlace carbono-carbono, y puede estar presente hasta el número máximo posible de triples enlaces carbono-carbono no aromáticos, y puede estar sustituido o no sustituido. Por tanto, "alquinilo C2-C5" significa un radical alquinilo que tiene 2 ó 3 átomos de carbono y 1 triple enlace carbono-carbono, o que tiene 4 ó 5 átomos de carbono y hasta 2 triples enlaces carbono-carbono. Los grupos alquinilo incluyen etinilo, propinilo y butinilo.

35 El término "sustituido" se refiere a un grupo funcional tal como se describió anteriormente tal como un alquilo, o un hidrocarbilo, en el que al menos un enlace con un átomo de hidrógeno contenido en el mismo se sustituye por un enlace con un átomo distinto de hidrógeno o distinto de carbono, siempre que se mantengan las valencias normales y que la(s) sustitución/sustituciones de(n) como resultado un compuesto estable. Los grupos sustituidos también incluyen grupos en los que uno o más enlaces con un átomo de carbono o de hidrógeno se sustituyen por uno o más enlaces, incluyendo dobles o triples enlaces, con un heteroátomo. Los ejemplos no limitativos de sustituyentes incluyen los grupos funcionales descritos anteriormente y, por ejemplo, N, por ejemplo para formar -CN.

40 Se entiende que los sustituyentes y patrones de sustitución en los compuestos de la presente invención puede seleccionarlos un experto habitual en la técnica para proporcionar compuestos que son químicamente estables y que pueden sintetizarse fácilmente mediante técnicas conocidas en la técnica, así como los métodos expuestos a continuación, a partir de materiales de partida fácilmente disponibles. Si un sustituyente está en sí mismo sustituido con más de un grupo, se entiende que estos múltiples grupos pueden estar en el mismo carbono o en carbonos diferentes, siempre que se obtenga como resultado una estructura estable.

45 Al elegir los compuestos de la presente invención, un experto habitual en la técnica reconocerá que los diversos sustituyentes, es decir R₁, R₂, etc., deben elegirse de conformidad con principios bien conocidos de conectividad de estructuras químicas.

En las estructuras de compuestos representadas en el presente documento, generalmente no se muestran los

átomos de hidrógeno, excepto en azúcares de ribosa y desoxirribosa. Sin embargo, se entiende que existen suficientes átomos de hidrógeno en los átomos de carbono representados como para satisfacer la regla del octeto.

Tal como se usa en el presente documento, y a menos que se mencione lo contrario, cada uno de los siguientes términos tendrá la definición expuesta a continuación.

5	A	-	Adenina;
	C	-	Citosina;
	ADN	-	Ácido desoxirribonucleico;
	G	-	Guanina;
	ARN	-	Ácido ribonucleico;
10	T	-	Timina; y
	U	-	Uracilo
	dNPP	-	Polifosfato de desoxirribonucleótido
	rNPP	-	Polifosfato de ribonucleótido.

15 “Ácido nucleico” significará, a menos que se especifique lo contrario, cualquier molécula de ácido nucleico, incluyendo, sin limitación, ADN, ARN e híbridos de los mismos. En una realización las bases de ácido nucleico que forman moléculas de ácido nucleico pueden ser las bases A, C, G, T y U, así como derivados de las mismas. En la técnica se conocen bien derivados de estas bases, y se muestran a modo de ejemplos en PCR Systems, Reagents and Consumables (Perkin Elmer Catalogue 1996-1997, Roche Molecular Sistemas, Inc., Branchburg, Nueva Jersey, EE.UU.).

20 Un polifosfato de nucleótido, tal como un polifosfato de desoxirribonucleótido (“dNPP”) o un polifosfato de ribonucleótido (“rNPP”), es un nucleótido que comprende múltiples, es decir tres, cuatro, cinco, seis o más, fosfatos unidos de una manera lineal a su átomo de carbono de azúcar en 5'. Un análogo de polifosfato de nucleótido es un análogo de un polifosfato de desoxirribonucleótido de este tipo o de un polifosfato de ribonucleótido de este tipo tal como se define en el presente documento, que difiere del mismo por tener una etiqueta unida al mismo en una
25 posición especificada. Tales análogos pueden incorporarse en un cebador o cadena de extensión de ácido nucleico, tal como una cadena de extensión de ADN, mediante puesta en contacto con una ácido nucleico polimerasa apropiada en las condiciones de polimerización de ácido nucleico apropiadas conocidas por los expertos en la técnica.

En una realización el dNPP es un trifosfato de desoxinucleótido.

30 Tal como se usa en el presente documento un tetranucleótido, un pentanucleótido, o un hexanucleótido, abarcan 4, 5 ó 6, respectivamente, residuos de monómero de ácido nucleico unidos mediante enlaces fosfodiéster, en los que el residuo terminal libre puede ser un nucleótido o un nucleósido. En una realización, el residuo terminal libre es un nucleósido y los demás residuos son nucleótidos.

35 “Sustrato sólido” significará cualquier medio adecuado presente en la fase sólida al que puede fijarse un ácido nucleico. Los ejemplos no limitativos incluyen chips, pocillos, perlas, estructuras de nanoporos y columnas. En una realización no limitativa el sustrato sólido puede estar presente en una disolución, incluyendo una disolución acuosa de electrolito.

40 “Hibridar” significará el apareamiento de un ácido nucleico monocatenario con otro ácido nucleico (tal como cebador) basándose en el principio bien conocido de complementariedad de secuencias. En una realización el otro ácido nucleico es un ácido nucleico monocatenario. La propensión a la hibridación entre ácidos nucleicos depende de la temperatura y la fuerza iónica de su medio, la longitud de los ácidos nucleicos y el grado de complementariedad. El efecto de estos parámetros sobre la hibridación se conoce bien en la técnica (véase Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Nueva York.). Tal como se usa en el presente documento, la hibridación de una secuencia de cebador, o de un producto de extensión
45 de ADN, con otro ácido nucleico significará el apareamiento suficiente de tal manera que el cebador, o producto de extensión de ADN, respectivamente, puede extenderse mediante creación de un enlace fosfodiéster con un nucleótido o análogo de nucleótido disponible que puede formar un enlace fosfodiéster, con el mismo.

50 Tal como se usa en el presente documento, a menos que se especifique lo contrario, una base que es “diferente de” otra base o una lista mencionada de bases significará que la base tiene una estructura diferente de la otra base o bases. Por ejemplo, una base que es “diferente de” adenina, timina y citosina incluirá una base que es guanina o una base que es uracilo.

“Cebador” tal como se usa en el presente documento (una secuencia de cebador) es un oligonucleótido corto, de manera habitual químicamente sintetizado, de longitud apropiada, por ejemplo de aproximadamente 18-24 bases, suficiente para hibridarse con un ADN diana (por ejemplo un ADN monocatenario) y permitir la adición de un residuo de nucleótido al mismo, o la síntesis de oligonucleótidos o polinucleótidos a partir del mismo, en condiciones adecuadas bien conocidas en la técnica. En una realización el cebador es un cebador de ADN, es decir un cebador que consiste, o que consiste ampliamente, en residuos de desoxirribonucleótido. Los cebadores se diseñan para tener una secuencia que es el complemento inverso de una región de ADN de molde/diana con la que se hibrida el cebador. La adición de un residuo de nucleótido al extremo 3' de un cebador mediante formación de un enlace fosfodiéster da como resultado un producto de extensión de ADN. La adición de un residuo de nucleótido al extremo 3' del producto de extensión de ADN mediante formación de un enlace fosfodiéster da como resultado un producto de extensión de ADN adicional.

En una realización el ADN monocatenario, ARN, cebador o sonda se une a un sustrato sólido mediante química de cicloadición de azida-alquino 1,3-dipolar. En una realización el ADN, ARN, cebador o sonda se une a un sustrato sólido a través de una molécula de polietilenglicol. En una realización el ADN, ARN, cebador o sonda está marcado con alquino. En una realización el ADN, ARN, cebador o sonda se une a un sustrato sólido a través de una molécula de polietilenglicol y un sustrato sólido está funcionalizado con azida. En una realización el ADN, ARN, cebador o sonda se inmoviliza sobre el sustrato sólido mediante una unión azido, una unión alquililo o interacción biotina-estreptavidina. La inmovilización de ácidos nucleicos se describe en *Immobilization of DNA on Chips II*, editado por Christine Wittmann (2005), Springer Verlag, Berlín. En una realización el ADN es ADN monocatenario. En una realización el ARN es ARN monocatenario.

En una realización el sustrato sólido está en forma de un chip, una perla, un pocillo, un tubo capilar, un portaobjetos, una oblea, un filtro, una fibra, un medio poroso, un nanotubo poroso o una columna. Esta invención también proporciona el presente método, en el que el sustrato sólido es un metal, oro, plata, cuarzo, sílice, un material de plástico, polipropileno, un vidrio o diamante. Esta invención también proporciona el presente método, en el que el sustrato sólido es una sustancia porosa no metálica a la que se une o en la que se impregna un metal o combinación de metales. La superficie sólida puede estar en diferentes formas incluyendo los ejemplos no limitativos de un chip, una perla, un tubo, una matriz, un nanotubo. La superficie sólida puede fabricarse de materiales comunes para micromatrices de ADN, incluyendo los ejemplos no limitativos de vidrio o nailon. La superficie sólida, por ejemplo perlas/microperlas, puede inmovilizarse a su vez en otra superficie sólida tal como un chip.

En una realización se separan muestras de ácido nucleico, ADN, ARN, cebador o sonda, en compartimentos, pocillos o rebajes diferenciados en una superficie o en un recipiente.

Esta invención también proporciona el presente método, en el que aproximadamente 1000 o menos copias de la muestra de ácido nucleico, ADN, ARN, cebador o sonda, se unen a la superficie sólida. Esta invención también proporciona la presente invención en la que 2×10^7 , 1×10^7 , 1×10^6 o 1×10^4 o menos copias de la muestra de ácido nucleico, ADN, ARN, cebador o sonda, se unen a la superficie sólida.

En una realización la muestra de ácido nucleico inmovilizada, ADN, ARN, cebador o sonda, se inmoviliza a una alta densidad. Esta invención también proporciona la presente invención en la que más de, o hasta, 1×10^7 , 1×10^8 , 1×10^9 copias de la muestra de ácido nucleico, ADN, ARN, cebador o sonda, se unen al sustrato sólido.

En una realización la ADN polimerasa es polimerasa 9°N o una variante de la misma, ADN polimerasa I de *E. Coli*, ADN polimerasa de bacteriófago T4, Sequenase, ADN polimerasa Taq o polimerasa 9°N (exo-)A485L/Y409V.

En una realización de los métodos o de las composiciones descritos en el presente documento, el ADN es monocatenario. En una realización de los métodos o de las composiciones descritos en el presente documento, el ARN es monocatenario, Phi29, o variantes del mismo.

En una realización de los métodos descritos para secuenciación de ARN, la polimerasa es una ARN polimerasa, transcriptasa inversa o polimerasa apropiada para polimerización de ARN tal como se conoce en la técnica.

Los grupos de unión pueden ser fotoescindibles. En una realización se usa luz UV para escindir fotoquímicamente los restos y grupos de unión fotoquímicamente escindibles. En una realización, el grupo de unión fotoescindible es un resto 2-nitrobencilo.

El grupo $-\text{CH}_2\text{N}_3$ puede tratarse con TCEP (tris(2-carboxietil)fosfina) para retirarlo del átomo de O en 3' de un análogo de dNPP, o análogo de rNPP, creando así un grupo 3'-OH.

En la patente estadounidense n.º 6.664.079 se dan a conocer métodos para la producción de análogos de nucleótidos con los grupos reactivos ocupados de manera escindible y/o unidos de manera escindible.

Un “residuo de nucleótido” es un único nucleótido en el estado en el que existe tras incorporarse en, y pasar a ser de ese modo un monómero de, un polinucleótido. Por tanto, un residuo de nucleótido es un monómero de nucleótido de un polinucleótido, por ejemplo ADN, que está unido a un monómero de nucleótido adyacente del polinucleótido a través de un enlace fosfodiéster en la posición 3' de su azúcar y está unido a un segundo monómero de nucleótido

adyacente a través de su grupo fosfato, con las excepciones de que (i) un residuo de nucleótido 3'-terminal sólo está unido a un monómero de nucleótido adyacente del polinucleótido mediante un enlace fosfodiéster a partir de su grupo fosfato, y (ii) un residuo de nucleótido 5'-terminal sólo está unido a un monómero de nucleótido adyacente del polinucleótido mediante un enlace fosfodiéster a partir de la posición 3' de su azúcar.

5 Debido a reglas de apareamiento de bases bien conocidas, la determinación de la identidad (de la base) de análogo de dNPP (o análogo de rNPP) incorporado en un cebador o producto de extensión de ADN (o producto de extensión de ARN) midiendo la señal eléctrica única de la etiqueta que se transloca a través del nanoporo, y de ese modo la identidad del análogo de dNPP (o análogo de rNPP) que se incorporó, permite la identificación del residuo de nucleótido complementario en el polinucleótido monocatenario con el que se hibrida el cebador o producto de extensión de ADN (o producto de extensión de ARN). Por tanto, si el análogo de dNPP que se incorporó comprende una adenina, una timina, una citosina o una guanina, entonces el residuo de nucleótido complementario en el ADN monocatenario se identifica como una timina, una adenina, una guanina o una citosina, respectivamente. La purina adenina (A) se empareja con la pirimidina timina (T). La pirimidina citosina (C) se empareja con la purina guanina (G). De manera similar, con respecto al ARN, si el análogo de rNPP que se incorporó comprende una adenina, un uracilo, una citosina o una guanina, entonces el residuo de nucleótido complementario en el ARN monocatenario se identifica como un uracilo, una adenina, una guanina o una citosina, respectivamente.

La incorporación en un oligonucleótido o polinucleótido (tal como un cebador o cadena de extensión de ADN) de un análogo de dNPP o de rNPP significa la formación de un enlace fosfodiéster entre el átomo de carbono en 3' del residuo de nucleótido 3'-terminal del polinucleótido y el átomo de carbono en 5' del análogo de dNPP o análogo de rNPP, respectivamente.

Tal como se usa en el presente documento, a menos que se especifique lo contrario, una base (por ejemplo de un análogo de polifosfato de nucleótido) que es diferente del tipo de base de una molécula mencionada, por ejemplo otro análogo de polifosfato de nucleótido, significa que la base tiene una estructura química diferente de la otra base o bases/base o bases de referencia. Por ejemplo, una base que es diferente de adenina incluirá una base que es guanina, una base que es uracilo, una base que es citosina y una base que es timina. Por ejemplo, una base que es diferente de adenina, timina y citosina incluirá una base que es guanina y una base que es uracilo.

Tal como se usa en el presente documento, a menos que se especifique lo contrario, una etiqueta (por ejemplo de un análogo de polifosfato de nucleótido) que es diferente del tipo de etiqueta de una molécula mencionada, por ejemplo otro análogo de polifosfato de nucleótido, significa que la etiqueta tiene una estructura química diferente de la estructura química de la otra etiqueta o etiquetas/etiqueta o etiquetas de referencia.

“Nanoporo” incluye, por ejemplo, una estructura que comprende (a) un primer y un segundo compartimento separados por una barrera física, barrera que tiene al menos un poro con un diámetro, por ejemplo, de desde aproximadamente 1 hasta 10 nm, y (b) un medio para aplicar un campo eléctrico a través de la barrera de modo que una molécula cargada tal como ADN, nucleótido, análogo de nucleótido o etiqueta, puede pasar desde el primer compartimento a través del poro hasta el segundo compartimento. De manera ideal, el nanoporo comprende además un medio para medir la firma electrónica de una molécula que pasa a través de su barrera. La barrera de nanoporo puede ser sintética o producirse parcialmente de manera natural. Las barreras pueden incluir, por ejemplo, bicapa lipídicas que tienen en las mismas α -hemolisina, canales de proteína oligomérica tales como porinas, y péptidos sintéticos y similares. Las barreras también pueden incluir placas inorgánicas que tienen uno o más orificios de un tamaño adecuado. En el presente documento “nanoporo”, “barrera de nanoporo” y el “poro” en la barrera de nanoporo se usan algunas veces de manera equivalente.

En la técnica se conocen dispositivos de nanoporos y se dan a conocer nanoporos y métodos que los emplean en las patentes estadounidenses n.^{os} 7.005.264 B2; 7.846.738; 6.617.113; 6.746.594; 6.673.615; 6.627.067; 6.464.842; 6.362.002; 6.267.872; 6.015.714; 5.795.782; y las publicaciones estadounidenses n.^{os} 2004/0121525, 2003/0104428 y 2003/0104428.

En una realización de las moléculas y los métodos dados a conocer en el presente documento la etiqueta se une al resto de la molécula mediante un grupo de unión químico que es escindible.

En una realización el nanoporo está en una membrana en estado sólido. En una realización la membrana es una membrana de nitruro de silicio. En una realización el nanoporo es un bioporo. En una realización el poro es proteico. En una realización el poro es un poro de alfa-hemolisina. En una realización el poro es un poro de grafeno.

En una realización el ADN, ARN o ácido nucleico monocatenario está ubicado en un lado de la membrana en la que está ubicado el nanoporo y la membrana está ubicada en una disolución de electrolito conductor.

Cuando se proporciona un intervalo de valores, a menos que el contexto indique claramente lo contrario, se entiende que cada número entero intermedio del valor, y cada décima de cada número entero intermedio del valor, a menos que el contexto indique claramente lo contrario, entre el límite superior e inferior de ese intervalo, y cualquier otro valor mencionado o intermedio en el intervalo mencionado, queda abarcado dentro de la invención. Los límites superior e inferior de estos intervalos menores pueden incluirse independientemente en los intervalos menores, y también quedan abarcados dentro de la invención, sujeto a cualquier límite específicamente excluido en el intervalo

mencionado. Cuando el intervalo mencionado incluye uno o ambos de los límites, intervalos que excluyen (i) cualquiera o (ii) ambos de esos límites incluidos también quedan incluidos en la invención.

5 Todas las combinaciones de los diversos elementos descritos en el presente documento quedan dentro del alcance de la invención. Todas las subcombinaciones de los diversos elementos descritos en el presente documento también quedan dentro del alcance de la invención.

Esta invención se entenderá mejor mediante referencia a los siguientes detalles experimentales, pero los expertos en la técnica apreciarán fácilmente que los experimentos específicos detallados sólo son ilustrativos de la invención tal como se describe de manera más completa en las reivindicaciones que siguen más adelante.

Detalles experimentales y comentarios

10 La invención dada a conocer en el presente documento se refiere a nucleótidos modificados para el análisis de moléculas individuales de ADN (o ARN, cambiando lo que sea necesario) usando nanoporos. Pueden realizarse modificaciones en diversas posiciones de un nucleótido, es decir el fosfato terminal, la base y/o el 2' o 3'-OH para formar un análogo de nucleótido. Tras una reacción de extensión con polimerasa en un complejo de molde-cebador, el pirofosfato unido a etiqueta liberado pasa a través de un nanoporo y se monitoriza el bloqueo de corriente
15 resultante para determinar la base de nucleótido añadida. Si la modificación o la etiqueta está en el resto de base, o el 2'/3'-OH del resto de azúcar del nucleótido, entonces tras la incorporación mediante ADN/ARN polimerasa, el grupo de unión-etiqueta se escinde de la base/azúcar mediante medios químicos o fotoquímicos y el grupo de unión-etiqueta liberado pasa a través de un nanoporo para identificar el nucleótido añadido.

20 Se diseñan y sintetizan 5'-polifosfatos de nucleósidos que portan diferentes números de grupos fosfato como grupos de unión y modificados con etiquetas unidas al fosfato terminal de los nucleótidos. Tras la incorporación mediante ADN/ARN polimerasa en una reacción de extensión de molde-cebador, el polifosfato unido a etiqueta liberado (di, tri, tetra, penta, etc.) puede detectarse usando un nanoporo para producir datos de secuencia. Opcionalmente, los polifosfatos con etiqueta liberados también pueden tratarse con fosfatasa alcalina para proporcionar etiquetas libres. Usando cuatro etiquetas diferentes que son distintas y específicas para cada base de nucleótido, puede
25 determinarse la secuencia del ADN o ARN de molde.

Se proporcionan nucleótidos que portan diferentes números de grupos fosfato o etiquetas para la síntesis de nucleótidos modificados, que son sustratos eficaces en reacciones con polimerasa. El polifosfato unido a etiqueta liberado se detecta usando un nanoporo para determinar condiciones para el diseño y la modificación de los nucleótidos para lograr señales de bloqueo diferenciadas.

30 También se proporcionan nucleótidos que portan una etiqueta con grupo de unión unida al resto de base de nucleótido, y/o el 2'/3'-OH del resto de azúcar, para reacción con ADN polimerasa para generar producto de extensión de ADN de base individual marcado con etiqueta con grupo de unión. Estos nucleótidos son buenos sustratos para ADN/ARN polimerasas habitualmente usadas. La etiqueta con grupo de unión unida al producto de ADN extendido se escinde mediante medios químicos o fotoquímicos para generar el cebador listo para su extensión
35 adicional usando los nucleótidos modificados. La etiqueta con grupo de unión liberada se hace pasar a través de nanoporo y se identifica basándose en la diferencia de tamaño, forma y carga en la etiqueta para producir datos de secuencia.

Tal como se da a conocer en el presente documento, estas herramientas moleculares facilitan la secuenciación de moléculas individuales usando nanoporo en la resolución de bases individuales.

40 En este caso se dan a conocer varias mejoras en el enfoque de nanoporo: 1) lograr una distinción precisa y evidente de las cuatro bases (A, C, G y T) que constituyen las moléculas de ácido nucleico; 2) potenciar y diferenciar la intensidad de las señales de detección; 3) desarrollar un método eficaz para distinguir y procesar las señales de bloqueo electrónicas generadas; 4) controlar la velocidad de translocación de ácidos nucleicos a través del poro, tal como ralentizar el movimiento de etiquetas para mejorar la capacidad de distinción de base a base; y 5) diseñar y
45 producir nanoporos sintéticos nuevos y más eficaces para diferenciar los cuatro nucleótidos diferentes en el ADN.

En la figura 2 se muestran las estructuras de cuatro nucleótidos. A y G son purinas, mientras que C y T son pirimidinas. El tamaño molecular global de A y G es muy similar, mientras que el tamaño de C y T es similar. Se ha mostrado que los nanoporos pueden diferenciar entre purinas y pirimidinas [Akeson *et al.* 1999 y Meller *et al.* 2000], pero no pueden distinguir entre purinas individuales, A y G, o entre pirimidinas individuales, C y T.

50 Estudios anteriores han mostrado modificaciones de 5'-trifosfatos de nucleósidos, incluyendo introducir más grupos fosfato para producir tetra, penta o hexa-fosfatos, introducir colorante directamente en el fosfato terminal, o unir un grupo de unión entre el fosfato terminal y el colorante [Kumar *et al.*, 2006 y 2008]. Los tetra y penta-fosfatos son mejores sustratos de ADN polimerasa, y se han desarrollado hexa-fosfatos de nucleótidos marcados con colorante [Kumar *et al.* 2005; Sood *et al.* 2005; Eid *et al.* 2009].

55 En el presente documento se dan a conocer análogos de nucleótidos que están diseñados para potenciar la distinción de cada nucleótido mediante modificación de los nucleótidos en el resto fosfato terminal. Se sintetizan 5'-

polifosfatos de nucleósidos y se unen diferentes etiquetas (tales como polietilenglicoles (PEG) de diferentes longitudes/masas, aminoácidos, hidratos de carbono, oligonucleótidos, colorantes o moléculas orgánicas/inorgánicas) al grupo fosfato terminal. Tras la extensión con reacciones con polimerasa, se generan restos de polifosfato unido a etiqueta (figura 3) y se producen señales diferentes específicas para cada base cuando los restos de polifosfato unido a etiqueta pasan a través del nanoporo. Estas modificaciones amplían la distinción de las bases mediante nanoporos debido al aumento de diferencias de tamaño, masa o carga de unidades de polifosfato marcado con etiqueta liberadas entre los cuatro nucleótidos (A, G, C y T).

La velocidad de translocación de ADN a través del nanoporo se reduce debido a la voluminosidad de los polifosfatos unidos a etiqueta liberados, aunque no se necesita reducir la velocidad de translocación de las etiquetas a través del nanoporo siempre que puedan diferenciarse las etiquetas. Por tanto, pueden lograrse la precisión y fiabilidad requeridas para la secuenciación de base a base. Otros parámetros analíticos en la secuenciación en nanoporos, tales como la concentración del polinucleótido, magnitud de la tensión aplicada, temperatura y valor de pH de la disolución, se optimizan con el fin de obtener los resultados más precisos y fiables para la detección y el análisis de la cadena de ADN.

Los enfoques de moléculas individuales para la secuenciación permiten la posibilidad de derivar haplotipos para estudios genéticos y permitir la secuenciación directa de ARNm. Entre los posibles enfoques de moléculas individuales para descodificar la secuencia de moléculas de ADN o ARN se encuentra el uso de nanoporos biológicos o sintéticos como detectores de las bases de ADN individuales.

El enfoque existente de secuenciación por síntesis (SBS) usa terminadores reversibles de nucleótidos fluorescentes y escindibles (CF-NRT) [Guo *et al.* 2010]. El método de SBS se basa en la capacidad de detenerse tras cada adición de nucleótido durante la reacción con polimerasa y el uso de fluoróforos específicos para distinguir entre las 4 bases. Sin embargo, una limitación principal de SBS para la secuenciación de moléculas individuales es el requisito de detectores de fluorescencia caros y software rápido de obtención de imágenes. El método y procedimiento dados a conocer en el presente documento aprovechan las ventajas de SBS, especialmente su alta precisión, con la velocidad y sensibilidad del nanoporo como detector de impedancia de corriente iónica.

Aunque se ha investigado mucho sobre la introducción de ADN a través de nanoporos, con la esperanza de distinguir cada base a medida que pasa a través debido a su efecto variable sobre la corriente iónica, ha resultado muy difícil lograr esto, debido tanto a la velocidad de transmisión como al efecto de bases circundantes que pueden contribuir con sus propios efectos sobre iones y contraiones que pasan a través de los poros [Timp *et al.* 2010]. El uso de ciclodextrinas u otras estructuras en forma de anillo en la luz de poros proteicos ayuda a proporcionar un mecanismo de trinquete para ralentizar el tiempo de tránsito [Astier *et al.* 2006], pero la capacidad para reconocer de manera absoluta cada base para su secuenciación a medida que pasa sigue suponiendo un desafío. Una estrategia alternativa que usa exonucleasa para permitir que un nucleótido atraviese el poro cada vez ha conducido a una distinción de bases individuales [Clarke *et al.* 2009]. Sin embargo, todavía existen dificultades en el control del tiempo de reacción de la exonucleasa para diferentes longitudes de ADN y nucleótido y la velocidad a la que los iones liberados llegan al poro con este enfoque.

La propia reacción con polimerasa presenta alta capacidad de procesamiento y velocidades estables de incorporación de bases. De hecho, se han usado reacciones con polimerasa para controlar el movimiento de cadenas de ADN a través de nanoporos para la distinción directa de bases [Benner *et al.* 2007, Cockroft *et al.* 2008, Hurt *et al.* 2009]. Durante la reacción con polimerasa, existe una liberación de un resto pirofosfato (PPi). Por tanto, si se une una etiqueta diferente al trifosfato para cada uno de los cuatro nucleótidos, estos pueden distinguirse a medida que se liberan y pasan a través de un nanoporo apropiado para la determinación de la secuencia de ADN. Estos análogos de pirofosfato relativamente pequeños, o moléculas equivalentes con grupos adicionales con carga positiva, pueden alcanzar el poro de manera extremadamente rápida. La velocidad de incorporación de nucleótidos mediante polimerasas es de aproximadamente 1000 nucleótidos por segundo, es decir un milisegundo por adición de base, mientras que la velocidad de transporte a través del nanoporo es de 1 molécula por microsegundo. Por tanto, con sistemas fluidicos e ingeniería apropiados, no hay problemas de desajuste de fase para secuenciar ADN con el presente enfoque, ni tampoco hay dificultades con la descodificación de tramos de homopolímeros. Se ha mostrado que puede distinguirse entre un amplio intervalo de tamaño de polietilenglicoles que se diferencian en tan sólo una o dos unidades de carbono mediante el efecto que tienen sobre las corrientes de bloqueo en nanoporos [Reiner *et al.* 2010, Robertson *et al.* 2007], una resolución esencialmente equivalente a la obtenida mediante un espectrómetro de masas. Por tanto, tal como se describe a continuación, se unen cadenas de PEG de diferente longitud al fosfato terminal de dATP, dCTP, dGTP y dTTP. A medida que se incorpora cada nucleótido durante la reacción con polimerasa, se libera un grupo fosfato marcado con etiqueta de manera específica en el nanoporo, proporcionando una señal de bloqueo de corriente distinta para indicar qué nucleótido se incorpora. La velocidad de secuenciación es extremadamente rápida, limitada sólo por la velocidad de la reacción con polimerasa. Como enfoque alternativo para marcar con etiqueta los nucleótidos, también se usan diferentes longitudes de cadena de fosfato (por ejemplo, tri, tetra y penta-fosfatos).

Adicionalmente, también se usan nanoporos en estado sólido que tienen ventajas en cuanto a un mejor control y flexibilidad de fabricación, por tanto garantizan un rápido transporte vectorial de polifosfatos marcados con etiqueta pero no de los precursores de nucleótido o el ADN hacia y a través de los nanoporos o nanocanales. Para lograr

esto se incorporan dos características de diseño importantes. En primer lugar, los precursores (polifosfatos de nucleótidos marcados con etiqueta) se sintetizan con una carga global neutra, mientras que los fosfatos marcados con etiqueta escindidos tienen una carga global positiva. Usando una corriente que atrae los iones positivos, los nanoporos sólo necesitan distinguir las cuatro moléculas marcadas con etiqueta liberadas alternativas. La carga diferencial en precursores y productos se logra incorporando en las etiquetas un número de lisinas o argininas (con carga positiva) que equilibran exactamente el número de fosfatos (con carga negativa). Tras la incorporación del α -fosfato en el cebador en crecimiento, hay una lisina más que fosfatos en el producto liberado. Opcionalmente, puede usarse fosfatasa alcalina para escindir todos los fosfatos para producir una etiqueta de PEG con una carga positiva más fuerte. En segundo lugar, para garantizar que los fosfatos liberados se mueven inmediatamente a través del poro más cercano, se inmoviliza la ADN polimerasa en la entrada del poro, por ejemplo a través de una unión biotina-estreptavidina. A medida que la cadena de ADN se inserta a través de la polimerasa, los productos marcados con etiqueta liberados sólo tienen que difundirse la misma distancia corta para alcanzar el nanoporo.

También es importante reconocer las ventajas del mecanismo de transducción bioelectrónico con respecto a enfoques ópticos. Para técnicas de transducción óptica de moléculas individuales, la señal de un único fluoróforo es normalmente de < 2500 fotones/s (correspondiente a niveles de corriente detectados del orden de 50 fA) a niveles de ruido corto altos, que requieren óptica compleja para recopilar cada fotón emitido, haciendo que el ajuste a escala de las plataformas a densidades superiores sea difícil. Las reacciones de síntesis debe ralentizarse hasta 1 Hz para permitir tiempos de integración suficientes para estas señales ópticas débiles con ruido. Los desafíos para las técnicas ópticas han abierto la posibilidad de enfoques de detección bioelectrónica, que tienen niveles de señal significativamente superiores (normalmente superiores en más de tres órdenes de magnitud), permitiendo la posibilidad de detección con banda ancha alta con el diseño conjunto apropiado de transductor, detector y amplificador. Los niveles de señal para nanoporos pueden ser de hasta 100 pA a partir de alfa-hemolisina [Kasianowicz *et al.* 1996], 300 pA para MspA [Derrington *et al.* 2010], y más de 4 nA a partir de nanoporos en estado sólido [Wanunu *et al.* 2010].

Se ha dirigido un esfuerzo significativo hacia el desarrollo de tecnología de nanoporos como mecanismo de transducción bioelectrónico [Benner *et al.* 2007, Deamer *et al.* 2002, Kasianowicz *et al.* 1996, Branton 2008, Branton *et al.* 2008, Chen 2004, Gershow *et al.* 2007, Nealy 2007, Matysiak *et al.* 2006]. Dos atributos esenciales de este sensor electrónico le proporcionan sensibilidad de moléculas individuales. El primero es la geometría muy localizada (nanoescala) de sensibilidad de carga en el propio poro. El diámetro de un poro puede ser de 2-3 nm, y debido a la carga de electrolito la exploración de la corriente medida resulta altamente insensible frente a fuentes de carga a más de unos pocos nanómetros desde el poro. En segundo lugar, el sensor de nanoporo proporciona una ganancia a través del efecto de la carga de movimiento comparativamente lento que tiene un biopolímero sobre una concentración cercana de iones de sal de movilidad superior. Sin embargo, los nanoporos están extremadamente limitados por el tiempo relativamente corto que pasan las biomoléculas en la región sensible a carga del poro. Esto se aborda directamente mediante el uso de etiquetas, que pueden optimizarse para producir niveles de señal altos y acontecimientos de translocación más prolongados. Al mismo tiempo, la cointegración de CMOS de estos poros se aprovecha para mejorar drásticamente los anchos de banda limitados por ruido para la detección en un dispositivo de nanoporos. Esta plataforma soporta poros tanto en estado sólido como biológicos. Esta integración de estado sólido, junto con sistemas microfluídicos asociados, también permite de manera única aumentar a escala este diseño para dar grandes matrices con electrónica integrada para detección.

EJEMPLO 1

I. Diseño y síntesis de nucleótidos modificados

El efecto de la voluminosidad del polifosfato con etiqueta sobre señales de bloqueo electrónico generadas por un nanoporo se determina usando diversos nucleótidos unidos a fosfato con etiquetas o grupos de diferente tamaño unidos al fosfato terminal del nucleótido. En la figura 4 se muestran estructuras de cuatro 5'-polifosfatos de nucleósidos marcados con etiqueta de fosfato. En primer lugar, se sintetiza una serie de 5'-tri, tetra, penta y hexa-fosfatos de nucleósidos. En estos nucleótidos, el fosfato terminal está unido con un grupo de unión a través del cual se unen diferentes etiquetas, por ejemplo etilenglicoles de diferente longitud y masa u otras moléculas que aumentan la voluminosidad o carga del polifosfato liberado. Estos nucleótidos se someten a prueba con nanoporo para determinar qué etiquetas o grupos voluminosos unidos al fosfato terminal se correlacionan con una diferencia más drástica en la señal de bloqueo electrónico entre las diferentes bases.

1) Polifosfatos de nucleósidos modificados con fosfato terminal

a. 5'-Trifosfatos de nucleósidos marcados con etiqueta de fosfato terminal

Tal como se muestra en la figura 5, pueden sintetizarse 5'-trifosfatos de nucleósidos marcados con etiqueta de fosfato terminal haciendo reaccionar el dNTP correspondiente con DCC/DMF para dar trimetafosfato cíclico que puede abrirse con nucleófilos apropiados para dar 5'-trifosfato de nucleósido unido a etiqueta o grupo de unión. Esto puede usarse en una reacción de extensión de molde-cebador y el pirofosfato unido a etiqueta liberado puede leerse usando un nanoporo. Alternativamente, el grupo de unión unido al fosfato puede hacerse reaccionar con etiqueta-éster NHS para proporcionar un 5'-trifosfato de nucleósido unido a etiqueta alternativo.

b. 5'-Tetrafosfatos de nucleósidos marcados con etiqueta de fosfato terminal

5 Para la síntesis de 5'-tetrafosfatos de nucleósidos marcados con etiqueta de fosfato terminal, en primer lugar se hace reaccionar el trifosfato correspondiente con CDI en DMF para activar el grupo fosfato terminal que después se hace reaccionar con ácido fosfórico o etiqueta-monofosfato para dar el tetrafosfato (figura 6). El fosfato terminal en el tetrafosfato puede activarse adicionalmente con CDI seguido por reacción con nucleófilos apropiados para proporcionar un tetrafosfato unido a grupo de unión que puede usarse adicionalmente para unir etiquetas de diferente masa, longitud o volumen, tales como m-dPEG-éster NHS, también mostrado en la figura 6.

c. 5'-Penta y hexafosfatos de nucleósidos marcados con etiqueta de fosfato terminal

10 La síntesis de 5'-penta y hexafosfatos de nucleósidos marcados con etiqueta de fosfato terminal sigue el mismo principio tal como se muestra en la figura 7. Pueden prepararse a partir de o bien trifosfatos o bien tetrafosfatos activados haciendo reaccionar con ácido fosfórico, pirofosfato o fosfatos unidos a etiqueta. Alternativamente, puede unirse un grupo de unión a penta o hexafosfato seguido por reacción con ésteres NHS activados.

d. Polifosfatos de nucleósidos unidos a etiqueta de oligonucleótido

15 Hay varios problemas con el enfoque actual a la secuenciación con nanoporos tales como el reconocimiento de las bases a medida que pasan a través del nanoporo y la velocidad o tasa de transporte para permitir registrar el reconocimiento de la nucleobase. El ADN pasa a través de un nanoporo de α -hemolisina a una velocidad de 1-5 μ s, lo cual es demasiado rápido como para registrar para los experimentos de secuenciación de moléculas individuales. Se ha producido cierto progreso para superar estos problemas mediante una variedad de estrategias de ingeniería de proteínas incluyendo el uso de frenos moleculares (oligonucleótidos cortos unidos de manera covalente) [Bailey, H. 2006].

25 Tal como se da a conocer en el presente documento, pueden unirse oligonucleótidos cortos al fosfato terminal de un polifosfato de nucleósido mediante reacción del fosfato terminal activado con el 3'-OH o el 5'-OH del oligonucleótido. Alternativamente, puede activarse el 3'- o 5'-fosfato del oligonucleótido con CDI o imidazol/DCC y hacerse reaccionar con 5'-polifosfatos de nucleósidos. En las figuras 8(a) y 8(b) se muestran estructuras de fosfatos de nucleósidos unidos a oligonucleótido (oligo-3' a 5'-fosfato; oligo-5' a 5'-fosfato), respectivamente. En la figura 8(c) se muestra el subproducto de reacción con polimerasa que se monitoriza mediante paso a través del nanoporo.

30 La velocidad de migración a través del nanoporo del subproducto de reacción con polimerasa puede controlarse uniendo oligonucleótidos de diferente longitud a diferentes 5'-polifosfatos de nucleósidos. Por ejemplo, si el nucleósido dA tiene 1 ó 2 unidades de oligo-dA unidas, dT puede tener 3 unidades de oligo-dT, dC puede tener 4 unidades de oligo-dC, y dG puede tener 5 unidades de oligo-dG. Pueden usarse diferentes combinaciones del número de oligonucleótidos para cada nucleótido para controlar el transporte y el tiempo de retención en un nanoporo.

El transporte y el tiempo de retención en un nanoporo también pueden controlarse añadiendo diferentes números de grupos fosfato a los nucleótidos. Por tanto, la carga y la masa pueden variar para cada polifosfato de nucleótido.

35 Ejemplos de estructura de etiqueta con grupo de unión

En la tabla 1 se proporcionan ejemplos específicos de grupos reactivos en los fosfatos terminales o el resto de base de nucleósido y grupos con los que pueden reaccionar los grupos. Los grupos reactivos con los que pueden reaccionar pueden estar presentes o bien en el grupo de unión o bien en la etiqueta.

TABLA 1

Posibles sustituyentes reactivos y grupos funcionales que reaccionan con los mismos	
Grupos reactivos	Grupos funcionales
Ésteres de succinimidilo	Amino primario, amino secundario
Anhídridos, haluros de ácido	Grupos amino e hidroxilo
Carboxilo	Amino, hidroxilo, tioles
Aldehído, isotiocianato e isocianatos	Grupos amino
Vinilsulfona y diclorotriazina	Grupos amino
Haloacetamidas	Tioles, imidazoles
Maleimidas	Tioles, hidroxilo, amino
Tioles	Tioles, maleimida, haloacetamida
Fosforamiditos, P. activado	Grupos hidroxilo, amino, tiol
Azido	Alquino

40 En el presente documento se incluyen etiquetas que pueden detectarse por el nanoporo, pero no se limitan de ningún modo a estos grupos de compuestos. Un experto en la técnica puede cambiar el/los grupo(s) funcional(es) para obtener una etiqueta adecuada.

Las etiquetas incluyen compuestos alifáticos, aromáticos, arilo, heteroarilo con uno o más anillos de 4-8 miembros y pueden estar opcionalmente sustituidos con halo, hidroxilo, amino, nitro, alcoxilo, ciano, alquilo, arilo, heteroarilo, ácido, aldehído, azido, alquenilo, alquinilo u otros grupos. Estos incluyen polietilenglicoles (PEG), hidratos de carbono, aminoácidos, péptidos, colorantes fluorescentes, fluorogénicos (no fluorescentes pero que se vuelven fluorescentes tras la eliminación de un grupo protector), cromogénicos (incolores pero que se vuelven de color tras la eliminación de un grupo protector), compuestos quimioluminiscentes, nucleósidos, mono, di o polifosfatos de nucleósidos, oligonucleótidos, arilo, heteroarilo o compuestos alifáticos. En la figura 16 se facilitan algunos ejemplos.

En la figura 17 se muestra a modo de ejemplo la estructura de nucleótidos marcados con etiqueta de PEG-fosfato y algunos ejemplos de posibles PEG con diferentes grupos reactivos para reaccionar con grupos funcionales.

10 En el presente documento se proporcionan algunos otros ejemplos de los colorantes o compuestos que pueden usarse para unirse al fosfato terminal o al resto de base de los nucleótidos. Estos no son de ningún modo los únicos compuestos que pueden usarse. Se indican en el presente documento como ejemplos y un experto en la técnica puede obtener fácilmente una etiqueta con grupo de unión adecuada que puede unirse al nucleótido y detectarse mediante un nanoporo.

15 Otros ejemplos de etiquetas adecuadas son:

Colorantes fluorescentes: colorantes de xantina, colorantes de bodipy, colorantes de cianina.

Compuestos quimioluminiscentes: compuestos de 1,2-dioxetano (Tropix Inc., Bedford, MA).

Aminoácidos y péptidos: aminoácidos que se producen de manera natural o modificados y polímeros de los mismos.

Hidratos de carbono: glucosa, fructosa, galactosa, manosa, etc.

20 *NMP y NDP:* monofosfatos de nucleósidos, difosfatos de nucleósidos. Ácidos alifáticos o aromáticos, alcoholes, tioles, sustituidos con halógenos, ciano, nitro, alquilo, alquenilo, alquinilo, azido u otros grupos de este tipo.

2) 5'-trifosfatos de nucleósidos con base modificada

Se sintetiza una variedad de terminadores reversibles de nucleótidos (NRT) para la secuenciación por síntesis (SBS) de ADN en los que un grupo de unión escindible une un colorante fluorescente a la base de nucleótido y el 3'-OH del nucleótido se bloquea con un grupo de terminación reversible pequeño [Ju *et al.* 2006, Guo *et al.* 2008 & 2010]. Usando estos NRT, la síntesis de ADN se detiene de manera reversible en cada posición. Tras registrar la señal fluorescente a partir de la base incorporada, se retiran los restos escindibles de los nucleótidos incorporados y se repite el ciclo.

También puede usarse el mismo tipo de nucleótidos para la secuenciación de ADN con nanoporos. Tal como se muestra en la figura 9(A), pueden sintetizarse un pequeño grupo de bloqueo en 3'-OH y una etiqueta unida a la base unida a través de un grupo de unión escindible. Tras la reacción de extensión con polimerasa, se escinden tanto el grupo de bloqueo en 3'-O como la etiqueta de la base y puede usarse la etiqueta liberada para pasar a través del nanoporo y monitorizarse la señal de bloqueo. Pueden usarse cuatro etiquetas diferentes (por ejemplo polietilenglicoles (PEG) de diferente longitud y peso molecular, tal como se muestra en la figura 9(A)), uno para cada una de las cuatro bases, diferenciando así las señales de bloqueo.

Alternativamente, no se usa el grupo de bloqueo en 3'-O porque se ha mostrado que un grupo o base de nucleótido voluminoso puede prevenir que la ADN polimerasa añada más de un nucleótido cada vez [Harris *et al.* 2008]. Tal como se muestra en la figura 9(B), se introduce un dNMP voluminoso a través de un grupo de unión escindible. Por tanto, se introducen diferentes dNMP a través de un grupo de unión según el dNTP original. Por ejemplo, con el nucleótido dTTP, se introduce un dTMP (para dATP, se introduce un dAMP; para dGTP, un dGMP y para dCTP, un dCMP). Tras la incorporación con polimerasa y escisión con TCEP, se generan dNMP modificados que se hacen pasar a través del canal de nanoporo y se detectan mediante métodos apropiados.

3) 5'-Trifosfatos de nucleósidos modificados en 2'- o 3'-OH

Puede llevarse a cabo la síntesis de los cuatro 5'-trifosfatos de nucleósidos modificados en 3' [Guo *et al.* 2008, Li *et al.* 2003, Seo *et al.* 2004]. Los dNTP unidos a 3'-O-2-nitrobencilo y 3'-O-azidometilo (figuras 10A y 10B, respectivamente) son buenos sustratos para ADN polimerasas. Tras la incorporación mediante ADN/ARN polimerasa en una reacción de secuenciación, estos nucleótidos marcados con etiqueta en 3'-O terminan la síntesis tras una extensión de una única base debido al grupo de bloqueo en 3'-OH. Una extensión adicional sólo es posible tras la escisión del grupo de bloqueo a partir de la posición 3'-O. El grupo 3'-O-2-nitrobencilo puede escindirse de manera eficaz mediante luz UV y 2'-O-azidometilo mediante tratamiento con TCEP para generar el grupo OH libre para una extensión adicional. Se monitoriza el producto escindido de la reacción (figuras 10C o 10D) para determinar el bloqueo electrónico mediante paso a través del nanoporo y registro de la señal. Se sintetizan cuatro dNTP protegidos con nitrobencilo sustituidos diferentes y cuatro dNTP sustituidos con azidometilo diferentes, uno para cada una de las cuatro bases de ADN.

II. Extensión de ADN usando nucleótidos modificados

1) Nucleótidos marcados con etiqueta de fosfato

Se usan polifosfatos de nucleósidos marcados con etiqueta de fosfato terminal descritos anteriormente en reacciones con polimerasa para generar productos de extensión. Tal como se muestra en la figura 11, tras una reacción con polimerasa, se obtiene el subproducto liberado del nucleótido marcado con etiqueta de fosfato, polifosfato con etiqueta, y el ADN extendido está libre de cualquier modificación. Entonces se usa el polifosfato con etiqueta liberado en un nanoporo modificado por ingeniería mediante técnicas de grabación de canales individuales para análisis de secuenciación. Los polifosfatos con etiqueta liberados también pueden tratarse con fosfatasa alcalina para proporcionar etiquetas libres que también pueden detectarse. Usando cuatro etiquetas diferentes para los cuatro nucleótidos (A, T, G y C) para generar cuatro polifosfatos marcados con etiqueta diferentes que se diferencian por la masa, carga o volumen, puede determinarse la secuencia del ADN.

2) Nucleótidos marcados con etiqueta de base con grupos de unión escindibles

Se sintetizan trifosfatos de nucleótidos marcados con etiqueta de base para la secuenciación por síntesis (SBS) de ADN y secuenciación de moléculas individuales [Guo *et al.* 2008 y 2010]. La adición de grandes grupos voluminosos en la posición 5 de pirimidinas (C y T) y la posición 7 de 7-desazapurinas (G y A) puede bloquear la adición de más de un nucleótido en una reacción con ADN polimerasa. Se sintetizan nucleótidos modificados con un grupo de unión escindible, un grupo voluminoso y diferentes cargas unidas a la base de nucleótido. Los nucleótidos modificados también pueden tener un pequeño grupo de bloqueo en el 3'-OH de los nucleótidos. Estos nucleótidos modificados se usan en una reacción de extensión con polimerasa. Tal como se muestra en la figura 12, tras la extensión con el nucleótido apropiado, el grupo de unión y la etiqueta de la base de nucleótido y del 3'-O de azúcar, si está bloqueado, se escinden por medios químicos o fotoquímicos y se usa la etiqueta con grupo de unión liberada en un nanoporo modificado por ingeniería mediante técnicas de grabación de canales individuales para análisis de secuenciación.

3) Nucleótidos marcados con etiqueta en 2' ó 3' con grupos de unión escindibles

También pueden unirse un grupo de unión y una etiqueta al 2'- o 3'-OH de nucleótidos. Tras una reacción de extensión con polimerasa, se escinde la etiqueta con grupo de unión del producto extendido mediante reacción química, fotoquímica o enzimática para liberar el 3'-OH libre para una extensión adicional. Tal como se muestra en la figura 13, después se usa la etiqueta con grupo de unión liberada en un nanoporo modificado por ingeniería mediante técnicas de grabación de canales individuales para análisis de secuenciación.

III. Estudio de secuenciación de ADN usando nanoporo

Se evalúa la distinción de diferentes nucleótidos en la secuenciación de ADN usando nanoporo siguiendo la estrategia mostrada en las figuras 11-13. Para validar la capacidad de un nanoporo para distinguir las cuatro etiquetas con grupo de unión diferentes en ADN, se realiza una serie de experimentos tal como se muestra en la figura 14. La ADN/ARN polimerasa puede unirse al nanoporo y se añade un molde que va a secuenciarse junto con el cebador. También puede inmovilizarse o bien el molde de ADN o bien el cebador encima del nanoporo y después formar posteriormente un complejo de molde-cebador tras la adición de una ADN polimerasa. A este complejo de molde-cebador se le añaden cuatro nucleótidos marcados con etiquetas diferentes en conjunto o de manera secuencial. Tras la incorporación catalizada por polimerasa del nucleótido correcto, el nucleótido añadido libera el polifosfato unido a etiqueta (en el caso de nucleótidos marcados con fosfato terminal) que entonces pasa a través del nanoporo para generar la señal eléctrica que va a registrarse y usarse para identificar la base añadida. Opcionalmente, también puede tratarse el polifosfato con etiqueta liberado con fosfatasa alcalina para proporcionar una etiqueta libre que también puede detectarse mediante paso a través del nanoporo. Cada etiqueta genera una firma de bloqueo electrónico diferente debido a la diferencia de tamaño, masa o carga. En el caso de nucleótidos modificados en la base o modificados en 2'/3', tras la extensión con ADN/ARN polimerasa, se escinde la etiqueta del cebador extendido mediante medios químicos, fotoquímicos o enzimáticos y se monitoriza la firma electrónica de la etiqueta liberada. La forma, tamaño, masa, carga u otras propiedades de la etiqueta pueden ajustarse según los requisitos.

Tal como se da a conocer en el presente documento, se distinguen y caracterizan señales de cada uno de los nucleótidos (figura 15) y las transiciones entre nucleótidos de diferentes identidades. Se analizan la magnitud y duración de las firmas de bloqueo sobre el esquema de acontecimientos y se comparan con esquemas conocidos. Por tanto, con estos diseños y modificaciones químicos racionales de los elementos estructurales del ADN, se optimiza el uso de nanoporo para descifrar la secuencia de ADN a nivel de moléculas individuales con resolución de bases individuales.

Para implementar esta estrategia novedosa para la secuenciación de ADN, puede construirse una matriz de nanoporos sobre una superficie plana para realizar una secuenciación de ADN en paralelo masiva tal como se muestra en la figura 18. La matriz de nanoporos también puede construirse sobre un chip de silicio u otras superficies de este tipo. El nanoporo puede construirse a partir de la proteína con bicapas lipídicas u otras capas de este tipo (poro de α -hemolisina, porina A de *Mycobacterium smegmatis*, MspA) [Derrington *et al.* 2010] o pueden ser

nanoporos en estado sólido sintéticos fabricados en nitruro de silicio, óxido de silicio u óxidos de metales [Storm *et al.* 2005; Wanunu *et al.* 2008] o un híbrido entre poro en estado sólido y α -hemolisina [Hall *et al.* 2010].

La figura 18 muestra un esquema de matriz de nanoporos para secuenciación por síntesis de ADN en paralelo masiva. Los nanoporos pueden detectar cada subproducto de adición de nucleótido catalizado por ADN/ARN polimerasa (etiqueta unida al fosfato o la base y/o 2',3'-OH del resto de azúcar) a medida que pasa a través del nanoporo. Las propiedades eléctricas de diferentes etiquetas distinguirán las bases basándose en su propiedad de bloqueo en los nanoporos. La matriz de nanoporos mostrada en la figura 18 pueden leer, cada uno, la misma secuencia o secuencias diferentes. Aumentar el número de veces que se lee cada secuencia dará como resultado una mejor calidad de los datos de secuencia resultantes.

10 EJEMPLO 2

I. Síntesis de desoxiguanosina 5'-tetrafosfatos marcados con PEG (dG4P-PEG):

Se sintetizan desoxiguanosina 5'-tetrafosfatos marcados con PEG (dG4P-PEG) según la figura 19. En primer lugar, 2'-desoxiguanosina trifosfato (dGTP) reacciona con CDI en DMF para activar el grupo fosfato terminal que después se hace reaccionar con fosfato de dibutilamonio para dar el tetrafosfato. Se activa adicionalmente el fosfato terminal en este tetrafosfato con EDAC en tampón de imidazol 0,1 M seguido por reacción con diaminoheptano para proporcionar un tetrafosfato unido a amino que se hace reaccionar adicionalmente con mPEG-ésteres NHS para proporcionar los cuatro PEG-dG4P requeridos. Tras la incorporación con polimerasa, la carga neta en el PEG liberado es de -3 (PEG-NH-trifosfato).

II. Pruebas de nucleótidos modificados en reacciones de extensión de bases individuales.

20 Se caracterizan los dG4P-PEG mediante espectroscopía de masas de MALDI-TOF tal como se muestra en la tabla II.

Tabla II. Resultados de MS de MALDI-TOF para dG4P-PEG		
	P.M. calculado	P.M. medido
dG4P-PEG24	1798	1798
dG4P-PEG37	2371	2374

Los dG4P-PEG son sustratos excelentes para ADN polimerasa en la extensión de cebador. En la figura 20 se muestran los espectros de masas de MALDI-TOF de los productos de extensión de ADN.

25 EJEMPLO 3 – Detección de moléculas individuales mediante nanoporo de los PEG usados para marcar los nucleótidos

Polietilenglicol es un polímero no electrolito que se une débilmente a cationes (por ejemplo, se une a iones K^+ con una $K_d \sim 2$ M). Por tanto, la carga neta en el polímero depende de la concentración de catión móvil y de la presencia de otros restos que se unen químicamente al mismo. Se ha demostrado que un único nanoporo de α -hemolisina puede distinguir fácilmente entre polímeros de PEG de diferentes tamaños con una resolución mejor que monomérica, es decir, mejor que 44 g/mol [Reiner *et al.* 2010; Robertson *et al.* 2007]. Ese nivel de distinción se vuelve posible porque el polímero reduce la conductancia del poro debido a la exclusión de volumen (la conductancia de poro disminuye al aumentar el tamaño de polímero) y mediante la unión a cationes móviles que de lo contrario fluirían libremente a través del poro [Reiner *et al.* 2010]. Además, el tiempo de residencia del polímero en el poro es altamente sensible a la carga del polímero, que para PEG aumenta a escala en proporción a la longitud del polímero. Un nanoporo debe poder distinguir entre PEG de diferentes tamaños que están químicamente unidos a otros restos. Se someten a prueba en nanoporo PEG (PEG 16, 24, 37 y 49) para marcar nucleótidos y generan firmas de bloqueo electrónico diferenciadas a nivel de moléculas individuales tal como se muestra en la figura 21.

Para investigar el efecto de la voluminosidad de los polifosfatos marcados con diversas etiquetas sobre señales de bloqueo electrónico generadas en el nanoporo, se sintetizan diversos nucleótidos unidos a fosfato con etiquetas de polietilenglicol (PEG) de diferente tamaño unidas al fosfato terminal del nucleótido. En primer lugar, tal como se muestra en la figura 4, se sintetiza una serie de 5'-tri, tetra y penta-fosfatos de nucleósidos con el fosfato terminal unido a través de un grupo de unión al que pueden unirse diferentes etiquetas, por ejemplo PEG u otras moléculas de diferente longitud y masa para aumentar el tamaño molecular o modificar la carga del polifosfato liberado. Después se someten a prueba estos nucleótidos en reacciones con polimerasa acopladas con detección mediante nanoporo para ver qué etiquetas o grupos voluminosos unidos al fosfato terminal producen diferencias más drásticas en señales de bloqueo electrónico entre diferentes bases.

I. Examen y selección de 4 etiquetas de PEG con señales de bloqueo de nanoporos diferenciadas

Recientemente se ha mostrado que cuando una molécula de polietilenglicol (PEG) entra en un único poro de α -hemolisina, provoca estados de conductancia dependientes de la masa diferenciados con tiempos de residencia medios característicos [Robertson *et al.* 2007]. La figura 22A muestra que el espectro de masas basado en conductancia resuelve claramente las unidades de repetición de etilenglicol, y el tiempo de residencia aumenta con

la masa de PEG.

1.a Pruebas de PEG para determinar firmas de bloqueo de nanoporos.

Se seleccionan PEG de diferente longitud y peso molecular (disponibles comercialmente de Quanta Biodesign Ltd u otros proveedores) y se monitorizan las señales de bloqueo de nanoporos, tal como se describe en el ejemplo 2. Tal como se muestra en la figura 22A, se distinguen claramente PEG de 28-48 unidades de etilenglicol mediante nanoporo. Por tanto, se seleccionan PEG con un amplio intervalo de unidades de etilenglicol que presentan señales de bloqueo de nanoporos muy diferenciadas como etiquetas para marcar los nucleótidos A, C, G y T. En la figura 22B se muestran ejemplos. También se evalúan PEG ramificados como etiquetas ya que pueden modificarse con cargas positivas de una manera más directa. En la parte inferior de la figura 22 se muestran estructuras de algunos PEG lineales y ramificados.

1.b Diseño y síntesis de PEG marcados con fosfato seleccionados en 1.a

En la secuenciación con nanoporos, las señales de bloqueo de corriente en el nanoporo se generan mediante los PEG-fosfatos liberados durante la reacción con polimerasa. Por tanto, se diseñan y sintetizan PEG marcados con fosfato con grupos de unión con carga positiva, y se someten a prueba estas moléculas con nanoporos orgánicos (por ejemplo, α -hemolisina) y sintéticos (fase sólida) para evaluar sus señales de bloqueo de corriente. Se convierten los PEG seleccionados en sus trifosfatos tal como se muestra en la figura 23. Por ejemplo, puede convertirse amino-butanol protegido con Fmoc en el trifosfato correspondiente haciéndolo reaccionar en primer lugar con oxiclورو de fósforo seguido por reacción con pirofosfato de tributilamonio en una reacción en una sola etapa. Tras la purificación, se activa el trifosfato con DCC/DMF o CDI/DMF para proporcionar trifosfato activado que reacciona con el grupo OH de los PEG para generar PEG-trifosfatos. El mismo esquema puede aplicarse para PEG tanto lineales como también ramificados. Estos PEG-fosfatos se someten a prueba en nanoporos para optimizar las condiciones para generar señales de bloqueo de corriente diferenciadas.

Se sintetizan los grupos de unión de poliaminoácido (polilisina, poliarginina, polilisina interrumpida) mediante estrategias de síntesis peptídica convencionales; si se incorpora una unión éster en la cadena de polifosfato, debe poder usarse fosfatasa alcalina para escindirlos, dando como resultado etiquetas más fuertemente positivas para el análisis con nanoporos. También pueden incorporarse cargas positivas en las cadenas de PEG.

1.c Diseño y síntesis de una biblioteca de 5'-trifosfatos de nucleósidos marcados con etiqueta de fosfato terminal.

Se diseñan y se sintetizan 5'-tri, tetra y penta-fosfatos de nucleósidos marcados con etiqueta de fosfato terminal. Estas moléculas se someten a prueba en la reacción con polimerasa y se seleccionan las óptimas para la detección con nanoporos. Se ha mostrado que ADN polimerasas aceptan 5'-tri, tetra, y penta-fosfatos de nucleósidos marcados con fosfato terminal con una variedad de etiquetas, incluyendo polilisininas pequeñas o grandes, aminoácidos, una variedad de colorantes con carga negativa o positiva, tales como colorantes de transferencia de energía, y unidades de etilenglicol, como sustratos excelentes para la extensión de cebador [Kumar *et al.* 2006 y 2008; Sood *et al.* 2005; y Eid *et al.* 2009].

1.c.1 Diseño y síntesis de 5'-trifosfatos de nucleósidos marcados con etiqueta de fosfato terminal.

Tal como se muestra en la figura 24, se sintetizan 5'-trifosfatos de nucleósidos marcados con etiqueta de fosfato terminal haciendo reaccionar el dNTP correspondiente con DCC/DMF para proporcionar un trimetafosfato cíclico que puede abrirse con nucleófilos para generar un 5'-trifosfato de nucleósido unido a etiqueta o grupo de unión. Además, el grupo de unión unido al fosfato puede hacerse reaccionar con PEG-ésteres NHS para proporcionar 5'-trifosfatos de nucleósidos unidos a PEG alternativos. El 5'-trifosfato de nucleósido marcado con etiqueta de fosfato terminal resultante se usa en la reacción de extensión de molde-cebador y se detecta el pirofosfato unido a etiqueta liberado y se diferencia mediante sus parámetros de bloqueo de corriente de nanoporo específicos.

1.c.2 Diseño y síntesis de 5'-tetrafosfatos de nucleósidos marcados con etiqueta de fosfato terminal.

Para la síntesis de 5'-tetrafosfatos de nucleósidos marcados con etiqueta de fosfato terminal, en primer lugar se hace reaccionar el trifosfato correspondiente con CDI en DMF para activar el grupo fosfato terminal que después se hace reaccionar con ácido fosfórico o etiqueta-monofosfato para dar el tetrafosfato tal como se muestra en la figura 25. El fosfato terminal en el tetrafosfato puede activarse adicionalmente con EDAC en tampón de imidazol 0,1 M seguido por reacción con un nucleófilo apropiado para proporcionar un tetrafosfato unido a grupo de unión que puede usarse para unirse a etiquetas de diferente masa, longitud o carga, tales como m-PEG-ésteres NHS. En este caso, se usan cuatro trimetil-lisinas para neutralizar la carga de cuatro fosfatos. Tras la incorporación con polimerasa, la carga neta en el PEG liberado es de +1 o, si se trata con fosfatasa alcalina, de +4, lo cual puede detectarse por el nanoporo.

1.c.3 Diseño y síntesis de 5'-penta-fosfatos de nucleósidos marcados con etiqueta de fosfato terminal.

La síntesis de 5'-penta-fosfatos de nucleósidos marcados con etiqueta de fosfato terminal sigue el mismo principio tal como se muestra en la figura 26. Pueden prepararse a partir de o bien trifosfatos o bien tetrafosfatos activados

haciéndolos reaccionar con ácido fosfórico, pirofosfato o fosfatos unidos a etiqueta. Alternativamente, puede unirse un grupo de unión al pentafosfato seguido por reacción con ésteres NHS activados.

5 Los polifosfatos de nucleósidos marcados con etiqueta de fosfato terminal descritos anteriormente se usan en la reacción con polimerasa para generar productos de extensión. Siguiendo el esquema mostrado en la figura 11, se evalúa el rendimiento de los polifosfatos de nucleósidos marcados con etiqueta de fosfato terminal en la extensión con polimerasa. En primer lugar se realiza una reacción de extensión de bases individuales y se caracteriza el producto de extensión de ADN mediante espectroscopía de masas de MALDI-TOF para evaluar la eficacia de incorporación. Tras establecer las condiciones de reacción optimizadas, se inmoviliza el molde sobre perlas magnéticas y se repite la reacción de extensión de bases individuales, tras lo cual se aíslan las etiquetas con polifosfato liberadas de la disolución para la detección usando un único nanoporo. Esta reacción se realiza de manera continua para evaluar los 4 nucleótidos (A, C, G y T) y sus etiquetas liberadas correspondientes detectadas por el nanoporo. La reacción con polimerasa continua con los nucleótidos con etiqueta de polifosfato y la clara distinción de la etiqueta con polifosfato liberada mediante nanoporo establecen la viabilidad del enfoque.

15 Tal como se muestra en la figura 11, tras la reacción con polimerasa, se obtiene el subproducto liberado del nucleótido marcado con etiqueta de fosfato (polifosfato con etiqueta) y la cadena de ADN extendido está libre de cualquier modificación. Esto es ventajoso porque cualquier cicatriz que permanezca en las cadenas de ADN en crecimiento puede afectar a su capacidad para reconocerse por polimerasa con adiciones de nucleótidos crecientes, terminando eventualmente la síntesis de ADN adicional. El polifosfato unido a etiqueta liberado se somete a ensayo en el nanoporo para evaluar la sensibilidad y precisión de secuenciación. En experimentos iniciales se someten a prueba las etiquetas para determinar sus señales de bloqueo antes de llevar a cabo reacciones de SBS. La secuencia de ADN puede determinarse si se usan etiquetas diferentes para los cuatro nucleótidos para generar cuatro polifosfatos con etiqueta diferentes que difieren en cuanto a la masa, carga o volumen, y se proporcionan 4 señales de bloqueo diferenciadas.

II. Detección de los fosfatos marcados con etiqueta liberados mediante nanoporos proteicos

25 Se usa un único nanoporo de α -hemolisina para detectar PEG que están unidos a nucleótidos unidos a través de un grupo de unión de múltiples fosfatos y el mismo polímero tras haberse escindido el resto de nucleótido/ribosa mediante la reacción con ADN polimerasa. Cada uno de las cuatro bases de ADN diferentes se une a un polímero de PEG con una longitud única. Por tanto, se identifica cada base que se retira del PEG mediante la polimerasa. Dado que los nucleótidos sin reaccionar no pueden separarse de los polifosfatos marcados con etiqueta liberados, especialmente en situaciones en tiempo real, se aprovecha la extrema sensibilidad del método frente a la carga molecular para distinguir entre el producto de reacción liberado y el material de partida. Se mide la conductancia de α -hemolisina individual usando soportes de vidrio cónicos [White *et al.*, 2006 y 2007] que permiten la recogida de datos a 100 kHz y ruido de pA_{RMS} de ~ 4 pA. Se miden las distribuciones de profundidad de bloqueo y tiempo de residencia tanto de los nucleótidos marcados con etiqueta como de productos marcados con etiqueta a lo largo de un amplio intervalo de potenciales transmembrana para determinar condiciones óptimas para la distinción de nucleótidos y para extender la actual comprensión teórica de interacciones PEG-nanoporo [Robertson *et al.* 2007] frente a moléculas con cargas fijas. La caracterización y el entendimiento teórico permiten la identificación inequívoca de los nucleótidos incorporados en polinucleótidos mediante polimerasa. Por tanto, con estos diseños y modificaciones químicas racionales de los elementos estructurales del ADN, se optimiza el uso de nanoporos para descifrar el ADN a nivel de moléculas individuales con una resolución de bases individuales en nanoporos proteicos o sintéticos.

EJEMPLO 4 - Fabricación de un único nanoporo en estado sólido para secuenciación de moléculas individuales

45 La transición de un nanoporo proteico a un nanoporo en estado sólido hace que la fabricación de matrices de nanoporos de alta densidad sea posible, una etapa clave para proporcionar un secuenciador de ADN electrónico de moléculas individuales de alto rendimiento. En este caso, se desarrolla una plataforma de nanoporo en estado sólido individual integrada para caracterizar los nucleótidos marcados con etiqueta en la reacción con polimerasa basándose en el conocimiento obtenido a partir del nanoporo proteico.

Plataforma de nanoporo integrada.

50 Se desarrolló electrónica de CMOS de poco ruido integrada especializada, que cuando se integra con nanoporos en estado sólido, proporciona ventajas de rendimiento significativas con respecto a técnicas de medición "convencionales" que emplean amplificadores electrofisiológicos externos, tales como el instrumento Axopatch 200B. Estas ventajas proceden de aprovechar la realimentación capacitiva (en vez de resistiva) en un diseño de amplificador de integración personalizado. La corriente CC, que es característica de esta y otras interfaces bioelectrónicas, se retira con una fuente de corriente de poco ruido que funciona en un bucle de servocontrol de CC. 55 Las capacitancias de entrada de amplificador reducidas y las capacitancias parásitas reducidas asociadas con la cointegración mejoran el rendimiento de ruido a altas frecuencias, permitiendo anchos de banda que se aproximan a 1 MHz para poros en estado sólido. Tal alta resolución temporal, en combinación con las etiquetas desarrolladas, proporcionará una alta flexibilidad para realizar ajustes precisos de esta plataforma para un rendimiento en tiempo real y de alta sensibilidad.

El uso de este circuito integrado de nanoporo integrado con CMOS (CNP) en una configuración o bien de dos chips o bien de un chip tal como se muestra en la figura 27. En el primer caso, el poro se empaqueta junto con el CNP tal como se muestra en la figura 27B. En el segundo, el poro se fabrica directamente en el CNP tal como se muestra en la figura 27C con sistemas fluidicos a ambos lados del chip. En ambos casos, el electrodo en cis, que se conecta con la entrada del amplificador, está integrado directamente en la superficie del CNP. La configuración de un chip tiene la ventaja de poder ajustarse a escala fácilmente para dar una plataforma multiplexada a costa de complejidad de fabricación adicional. La capacidad de someter a procedimiento posterior dados de CMOS fabricados (que no tienen más de 5 mm en un lado) es una capacidad única establecida a lo largo de los últimos cinco años [Huang *et al.* 2011, Lei *et al.* 2008, y Levine *et al.* 2009]. Este enfoque hace uso completamente de flujos de procedimiento de fundición existentes en vez de requerir el desarrollo de nuevos procedimientos.

El enfoque de fabricación de un chip avanza adaptando técnicas de fabricación de nanoporos en estado sólido convencionales [Rosenstein *et al.* 2011]. En zonas del dado reservadas para los sensores, todos los metales se han bloqueado, dejando un apila gruesa de capas alternantes de ocupación de nitruro de silicio y relleno de vidrio. La mayor parte de la pila dieléctrica se graba usando un plasma de CHF₃ acoplado de manera inductiva. Tras depositar y producir un patrón de una máscara de grabado de Si₃N₄ de PECVD en el lado posterior del dado, se realizan aberturas localizadas en el sustrato de silicio usando una grabación con hidróxido de potasio anisotrópico. Después se usa una inmersión breve en ácido fluorhídrico tamponado para aislar una única capa de 50 nm de nitruro de silicio a partir de la pila dieléctrica original como membrana en suspensión. Finalmente, se perforan nanoporos a través de estas membranas de nitruro con un microscopio electrónico de transmisión de alta resolución.

En la figura 28A se muestra el ruido medido de este sistema, junto con una medición del ruido de línea base para una configuración del instrumento Axopatch 200B. Para el ancho de banda más alto soportado por el instrumento Axopatch (B=100 kHz), el amplificador integrado tiene un umbral mínimo de ruido de 3,2 pA_{RMS}, en comparación con 9 pA_{RMS} para el instrumento Axopatch. Al ancho de banda más alto caracterizado para el amplificador integrado (B=1 MHz), el nivel de ruido es de 27 pA_{RMS}, en contraste con 247 pA_{RMS} modelado extrapolando la respuesta del instrumento Axopatch más allá de su intervalo soportado (aproximadamente un factor de diez veces el ruido inferior). Como punto de comparación, para una señal de 1 nA, sólo se transportan aproximadamente 6250 iones a través del poro en 1 us. Un nivel de ruido relacionado con la entrada de 27 pA_{RMS} para el amplificador integrado permite una resolución de tan sólo 150 iones en este intervalo.

También es importante observar que este rendimiento eléctrico superior se obtiene con un amplificador integrado que ocupa un área de tan sólo 0,2 mm² en un chip de CMOS en comparación con un amplificador Axopatch montado en rack, lo que demuestra la importancia de la electrónica innovadora. Cuando se conecta un nanoporo a la entrada de amplificador, la introducción de ruido de 1/f y capacitancia de membrana aumenta el espectro de ruido por encima de la línea base de cabezal abierto. La figura 28b muestra un espectro de ruido típico en este caso, que demuestra umbrales mínimos de ruido de tan sólo 10 pA_{RMS} y 163 pA_{RMS} para anchos de banda de 100 kHz y 1 MHz, respectivamente. Se muestran comparaciones medidas con el instrumento Axopatch hasta 100 kHz para el mismo nanoporo. A 100 kHz, hay una reducción de más de un factor de dos en la potencia de ruido relacionado con la entrada para el CNP. Si el instrumento Axopatch pudiera medirse a anchos de banda mayores, habría una diferencia de potencia de ruido de un factor de seis a 1 MHz.

Esta plataforma también permite la integración de nanoporos biológicos, proporcionando incluso más flexibilidad. Se crean nanoporos biológicos en membranas lipídicas (normalmente 1,2-dioleoil-sn-glicero-3-fosfolina (DOPC)) formadas sobre un agujero en una membrana de teflón entre dos células de fluido. La superficie debe ser suficientemente hidrófila para que se forme la membrana a partir de vesículas unilaminares. La conductancia entre las dos cámaras de la célula se monitoriza mientras se añade la proteína de membrana a una de las células, que se purga inmediatamente una vez detectada la incorporación. Las membranas usadas para fabricar los nanoporos también pueden usarse como soportes sólidos para bicapas lipídicas con la perforación de agujeros más grandes en las membranas, sobre los que se forma la bicapa lipídica [Clarke *et al.* 2009; Benner *et al.*, 2007; Hou *et al.*, 2009; y Wang *et al.* 2011]. Se han diseñado por ingeniería membranas lipídicas bicapa (BLM) planas con diferentes canales proteicos sobre soportes sólidos con patrones con agujeros en nanopatrones (~100 nm de diámetro), así como anclándolas directamente sobre oro a través de un montaje de monocapa autoensamblada [Axelrod *et al.*, 1976, Bultmann *et al.* 1991, Dutta *et al.* 2010, Jenkins *et al.* 2001, Nam *et al.* 2006, Palegrosdemange *et al.* 1991, Shen *et al.* 2009, Srinivasan *et al.* 2001, Yang *et al.* 2003, Yin *et al.* 2005]. Además, se ha mostrado la formación de BLM contiguas con un coeficiente de difusión de 4 μm²/s sobre sustratos con nanopatrones; BLM formadas sobre montajes de SAM-oro proporcionaron un coeficiente de 0,8 μm²/s. Ambos se encuentran dentro del intervalo de difusión ideal de 0,1-10 μm²/s representativo de BLM bien formadas [Axelrod *et al.* 1976, Bultmann *et al.* 1991]. Las caracterizaciones eléctricas de estas BLM indican una membrana con alta impedancia con una resistencia de 1,4 GW-mm², haciendo que sea apropiada para un análisis eléctrico adicional de nanoporos biológicos formados en la membrana [Oliver *et al.* 1994, Shi *et al.* 2000, Wieheman 1988].

Inmovilización de polimerasa a superficies que portan nanoporos

El tamaño de la polimerasa es de aproximadamente 5 nm x 5 nm. Una polimerasa se sitúa cerca de la entrada de cada nanoporo. Para lograr esto para los nanoporos en estado sólido, es necesario que (1) se modifique una única posición sobre la superficie con grupos funcionales durante la fabricación de CMOS para unirse a la polimerasa; (2)

que los sitios sean lo suficientemente pequeños como para que sólo una molécula de polimerasa pueda unirse; (3) que estén lo suficientemente separados como para que haya pocas posibilidades de difusión de los polifosfatos marcados con etiqueta liberados a un canal cercano; y (4) que el agente de reticulación sea lo suficientemente flexible como para que la enzima esté funcionalmente intacta. El anclaje de polimerasa se logra combinando un punto de unión en patrones con el uso de una concentración apropiada de disolución de polimerasa durante la incubación de manera que como máximo se une una molécula de enzima.

El establecimiento de un punto de anclaje apropiado para la polimerasa se logra aprovechando enfoques de fabricación existentes para nanoporos en estado sólido. Normalmente, para maximizar las señales de transducción, estos poros se crean reduciendo el grosor de una membrana de Si₃N₄ soportada usando litografía de haz de electrones para definir una ventana cuyo grosor se reduce posteriormente con un grabado con plasma (por ejemplo SF₆). Después se perfora el nanoporo en la región de grosor reducido usando ablación con haz de electrones. El pocillo creado mediante esta ventana (figura 29) crea un lugar natural para anclar la polimerasa, garantizando una estrecha proximidad a la entrada del nanoporo. Antes de grabar la ventana de grosor reducido, puede aumentarse la membrana original con una capa epitaxial enterrada de material de unión. Una vez grabada la ventana, puede convertirse en una región de pared lateral selectiva para la unión de polimerasa. Los materiales de unión incluyen dióxido de silicio u oro. Puede haber una selectividad limitada con dióxido de silicio, sin embargo, porque también puede formarse un óxido sobre la superficie de nitruro de silicio en las condiciones apropiadas.

En principio, con superficies de dióxido de silicio, pueden usarse uniones de biotina-estreptavidina [Korlach *et al.* 2008 y 2010], usando moléculas de PEG biotiniladas sobre los parches de sílice e incubando polimerasa marcada en el extremo con biotina en presencia de estreptavidina. El resto de la superficie se pasiva con poli(ácido vinilfosfónico). Debido a las cuestiones planteadas anteriormente, en vez de eso es preferible modificar la superficie de oro con una monocapa autoensamblada de alcanotiol (SAM) funcionalizada con grupos amino [Love *et al.* 2005]. Estos pueden modificarse fácilmente para dar ésteres NHS para la unión a grupos amino en la polimerasa. El grosor y la homogeneidad de la capa se determinan mediante elipsometría o microscopía de fuerza atómica.

Desarrollo de nucleótidos modificados en 5'-con grupos de unión con carga positiva

Se genera un sistema para la difusión rápida de las etiquetas liberadas hacia los poros mientras se repelen los nucleótidos precusores y el ADN por los poros. Se modifican por ingeniería los nucleótidos marcados con etiqueta de modo que tras la incorporación en el ADN, la etiqueta liberada del nucleósido tiene una carga positiva acumulativa mientras que los nucleótidos con etiqueta intacta permanecen neutros. Esto permite dirigir de manera activa la etiqueta liberada específicamente a través del canal de detección, si el canal tiene carga negativa según métodos conocidos [Vanunu *et al.* 2007]. Dado que todas las demás moléculas libres presentes en la mezcla de reacción (cebadores, nucleótidos sin reaccionar, molde), aparte de la etiqueta, tienen carga negativa, sólo la etiqueta liberada que porta carga positiva se ve atraída al interior del canal, aumentando la especificidad de detección y reduciendo el ruido. Puede usarse un número diferente de grupos cargados en diferentes etiquetas, dependiendo de la base de nucleótido específica. Por tanto, puede usarse la carga acumulativa de la etiqueta junto con su tamaño para la distinción de bases. Tras la incorporación y liberación de la etiqueta, si se considera que el polifosfato enmascara la carga positiva, puede retirarse usando reacciones secundarias (por ejemplo, fosfatasa alcalina inmovilizada en un segundo sitio aguas abajo en el poro). La etiqueta con carga positiva puede dirigirse al interior del canal con carga negativa para su detección y reconocimiento.

Difusión y deriva

Un aspecto crítico de este sistema de secuenciación es la captura fiable y a tiempo de la etiqueta liberada de cada nucleótido mediante el nanoporo adyacente. Deben diseñarse por ingeniería condiciones tales que las etiquetas se capturen rápidamente y en el orden correcto. Adicionalmente, debe minimizarse la tasa de captura de etiquetas no incorporadas, y la interferencia de canales adyacentes debe ser despreciable. Crear el pocillo a la entrada del poro (tal como se muestra en la figura 29) ayuda a este procedimiento, que también depende de la estrecha proximidad de la polimerasa a la abertura de nanoporo. El análisis de procedimientos de captura de nanoporo considera generalmente un procedimiento radialmente simétrico que rodea al poro. La geometría indica que, en ausencia de un campo eléctrico, una molécula tiende a difundir más lejos desde un poro, oponiéndose a la atracción electrostática. Con un gradiente de tensión, existe una distancia crítica *L* a la que el movimiento molecular debido a la difusión y la electroforesis son iguales [Gershow *et al.* 2007]. Esta distancia crítica es función de la corriente iónica (*I*) y la conductividad de electrolito (σ), así como la constante de difusión (*D*) y la movilidad (μ) de la molécula de

$$L = \frac{I \mu}{4 \sigma D}$$

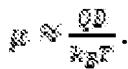
analito. La captura es un procedimiento estadístico, pero aproximadamente el 50% de las moléculas a una distancia *L* se capturan. Esta probabilidad aumenta para distancias más cortas, y supera el 90% para $d < L/3$. Durante este procedimiento, las moléculas se capturan normalmente en una escala de tiempo del orden

$$t_{\text{captura}} = \frac{L^2}{D}$$

de Colocando la polimerasa dentro de *L*/3 del nanoporo se capturan casi todas las moléculas. También garantiza que t_{captura} es significativamente más rápido que la velocidad de incorporación de polimerasa,

para capturar bases en el orden correcto.

Un valor aproximado para el coeficiente de difusión de moléculas de PEG de 25 unidades en agua es de $D=3e-10 \text{ m}^2/\text{s}$ [Shimada *et al.* 2005], que es del mismo orden de magnitud que un fragmento de ADNmc de longitud similar [Nkodo *et al.* 2001]. Suponiendo la validez de la relación de Nernst-Einstein (aunque esto no siempre es válido para polímeros), la movilidad puede estimarse como función de la constante de difusión y la carga neta (Q),



Entonces, para estas estimaciones, con $I=5 \text{ nA}$ en KCl 1 M, véase la siguiente tabla.

	+1e	+4e
50% de captura	2,1 nm	5,8 nm
90% de captura	0,7 nm	1,9 nm
t_{captura}	7,1 ns	114 ns

EJEMPLO 5 – Fabricación de una matriz de nanoporos en estado sólido

Además de un rendimiento mejorado, sólo con la electrónica integrada es posible producir matrices de nanoporos paralelos de manera masiva. Esto implica la topología de un chip mostrada en la figura 27C en la que se integran nanoporos directamente en el dado de CMOS con sistemas fluidicos a ambos lados del chip. En la figura 27C también se muestra el enfoque para integrar múltiples poros. En este caso, se usan pocillos de resina fotosensible SU-8 para aislar nanoporos individuales unos de otros. Esto es un enfoque similar al de Rothberg *et al.* 2011. Sin embargo, en Rothberg *et al.* los pocillos todavía pueden permanecer “conectados” por el depósito de disolución por encima del chip. En el presente caso, dado que se necesita aislamiento eléctrico entre los depósitos en cis, se usa una tapa de PDMS para sellar los pocillos para la medición tras la introducción de reactivos tal como se muestra en la figura 27C. Se integran 64 nanoporos en estado sólido en el mismo dado de 5 mm por 5 mm. El diseño de amplificador que integra corriente, que tendría que duplicarse en cada sitio de poro, sólo mide 250 um por 150 um, pero tiene que dejarse espacio adicional para la fabricación del propio poro. A medida que se desarrollan adicionalmente técnicas de fabricación para reducir el área de chip, esto puede ajustarse fácilmente a escala a una matriz de 16 por 16 electrodos.

EJEMPLO 6 – Pirosecuenciación usando nucleótido marcado con etiqueta de fosfato y detección con nanoporos

La pirosecuenciación es un método de secuenciación por síntesis (SBS) que se basa en la detección de pirofosfato que se libera cuando se incorpora un nucleótido en la cadena de ADN en crecimiento en la reacción con polimerasa [Ronaghi *et al.* 1998]. En este enfoque, a cada uno de los cuatro dNTP se le añade de manera secuencial un cóctel de enzimas, sustratos, y los componentes habituales de reacción con polimerasa. Si el nucleótido añadido es complementario a la primera base disponible en el molde, se incorporará el nucleótido y se liberará un pirofosfato. A través de una cascada enzimática, el pirofosfato liberado se convierte en ATP, y después pasa a ser una señal de luz visible mediante luciferasa de las luciérnagas. Por otro lado, si el nucleótido añadido no se incorpora, no se producirá luz y el nucleótido simplemente se degradará por la enzima apirasa. La pirosecuenciación se ha aplicado satisfactoriamente a la detección de polimorfismos de un solo nucleótido (SNP) y secuenciación de ADN. Se desarrolló una plataforma de secuenciación comercial combinando pirosecuenciación y amplificación de molde de ADN en micropelotas individuales para la secuenciación de ADN de alto rendimiento [Margulies *et al.* 2005]. Sin embargo, hay dificultades inherentes en la pirosecuenciación para determinar el número de nucleótidos incorporados en regiones homopoliméricas (por ejemplo, un tramo de varias T seguidas) del molde. Además de esto, hay otros aspectos de pirosecuenciación que todavía necesitan mejorarse. Por ejemplo, cada uno de los cuatro nucleótidos tiene que añadirse y detectarse por separado. La acumulación de nucleótidos no degradados y otros componentes también pueden reducir la precisión del método cuando se secuencia un molde de ADN largo.

Esto es un enfoque de pirosecuenciación modificado que se basa en la detección de etiqueta o fosfatos con etiqueta liberados durante la reacción con polimerasa. En este enfoque se usan nucleótidos marcados con etiqueta de fosfato en reacción catalizada con polimerasa en un complejo de molde-cebador. Tras la incorporación de los nucleótidos marcados con etiqueta, se libera el resto de etiqueta de fosfato, que puede detectarse mediante paso a través de un nanoporo. Puede usarse la misma etiqueta en cada nucleótido o puede usarse una etiqueta de peso molecular y longitud diferentes (tal como PEG). Se ha mostrado que polietilenglicoles (PEG) de diferente longitud y masa pueden resolverse a una sensibilidad de moléculas individuales cuando pasan a través de un nanoporo de hemolisina [Robertson *et al.* 2009].

Puede usarse un canal de α -hemolisina para detectar ácidos nucleicos a nivel de moléculas individuales [Kasianowicz *et al.* 1996]. El polipéptido monomérico se autoensambla para dar una bicapa lipídica para formar un poro heptamérico, con una abertura limitante de 1,5 nm de diámetro. En una disolución acuosa de sal iónica, el poro formado por el canal de α -hemolisina transporta una corriente iónica intensa y estable cuando se aplica una tensión apropiada a través de la membrana. La abertura limitante del nanoporo permite que moléculas de ácido nucleico (diámetro $\sim 2,0 \text{ nm}$) monocatenarias lineales, pero no bicatenarias, pasen a través. Los ácidos nucleicos polianiónicos se impulsan a través del poro mediante el campo eléctrico aplicado, que bloquea o reduce la corriente iónica. Este paso genera una firma electrónica única. Por tanto, se obtendrá un diagrama de acontecimientos

específicos, que es la representación gráfica del tiempo de translocación frente a la corriente de bloqueo, y se usará para distinguir la longitud y la composición de polinucleótidos mediante técnicas de grabación de canales individuales basándose en parámetros característicos tales como corriente de translocación, duración de translocación, y sus correspondientes dispersiones en el diagrama. Se seleccionan cuatro etiquetas de PEG, que se ha mostrado que proporcionan señales de bloqueo de corriente diferenciadas en nanoporos, para acoplarse con cuatro nucleótidos (A, C, G, T) en el fosfato terminal. Estos análogos de nucleótido novedosos se usan en una reacción con polimerasa y usan nanoporos para detectar las etiquetas liberadas para descodificar las bases incorporadas tal como se muestra en la figura 14.

Hay varias ventajas en este enfoque:

- 1) Evitar el uso de muchas enzimas diferentes (ahorra coste y complejidad).
- 2) Adición de polifosfato de nucleósido unido a etiqueta individual de manera secuencial o de los cuatro nucleótidos con diferentes etiquetas unidas a cada nucleótido.
- 3) Uso de PEG como etiquetas que pueden detectarse mediante nanoporo a una resolución de unidades individuales.
- 4) Secuenciación de detección de moléculas individuales en tiempo real a medida que la etiqueta pasa a través del nanoporo.
- 5) Secuenciación en paralelo de manera masiva, bajo coste y alto rendimiento.

Tal como se muestra en la figura 14, se inmoviliza ADN polimerasa en el nanoporo y se añade el molde-cebador junto con los nucleótidos marcados con etiqueta de PEG. Al incorporar el nucleótido marcado con etiqueta de PEG correcto, los fosfatos con PEG liberados pasan a través del nanoporo y se mide la señal de bloqueo electrónico. PEG de diferente longitud tienen diferentes señales de bloqueo, por tanto, pueden usarse 4 PEG diferentes para 4 nucleótidos diferentes.

Los nucleótidos pueden añadirse uno cada vez, si se añade el nucleótido correcto proporciona una señal de bloqueo diferenciada. Sin embargo, si el nucleótido no es complementario a la base de ácido nucleico de molde, no se incorporará y por tanto no se detectará ninguna señal. En un modo en paralelo, masivo, puede construirse una matriz de alta densidad de micro/nanopocillos para realizar el procedimiento bioquímico. Cada micro/nanopocillo contiene un molde de ADN y dispositivo de nanoporo diferentes. Se detectan los PEG liberados con una sensibilidad de moléculas individuales.

En la figura 30 se muestran métodos generales para la síntesis de 5'-polifosfato nucleósido marcado con etiqueta. Pueden sintetizarse 5'-tri, tetra, penta o hexa-fosfatos de nucleósidos marcados con fosfato terminal a partir de los 5'-trifosfatos de nucleósidos (NTP) correspondientes. Por tanto, en primer lugar se activa trifosfato con DCC/DMF que puede hacerse reaccionar directamente con el nucleófilo con etiqueta para dar NTP unido a etiqueta o puede hacerse reaccionar con un nucleófilo con grupo de unión con el que puede hacerse reaccionar una etiqueta-NHS o etiqueta activada de manera apropiada para proporcionar NTP unido a grupo de unión con etiqueta. Para la síntesis de tetrafosfatos (N4P) o penta-fosfatos (N5P) de nucleósidos unidos con etiqueta, en primer lugar se hace reaccionar el trifosfato activado con ácido fosfórico o pirofosfato para dar tetra y penta-fosfato, respectivamente, que puede hacerse reaccionar con nucleófilo con grupo de unión seguido por la reacción con etiquetas activadas apropiadas.

La síntesis de nucleótidos marcados con PEG se comentó anteriormente en los ejemplos 2 y 3. Los nucleótidos marcados con PEG tienen cargas -3, -4, -5 o -6 basándose en el uso de tri, tetra, penta o hexa-fosfatos. Tras la reacción de extensión de cebador catalizada con polimerasa, la carga neta en las etiquetas con PEG liberadas será de uno menos (-1) que el PEG-nucleótido de partida, lo cual es suficiente para distinguir las mediante la señal de bloqueo iónico de nanoporo (el PEG-nucleótido sin reaccionar también es más voluminoso que los fosfatos con PEG liberados, de ahí una señal de bloqueo iónico diferente). Alternativamente, si está presente fosfatasa alcalina en la mezcla de reacción, el PEG liberado será neutro (los grupos fosfato libres se hidrolizan por la fosfatasa alcalina). También puede hacerse que las etiquetas con PEG liberadas tengan carga positiva tal como se muestra a continuación de modo que pueden detectarse fácilmente por nanoporos. De manera similar, también puede hacerse que tengan carga altamente negativa.

Síntesis de polifosfatos de nucleósidos unidos a etiqueta con carga positiva:

Los polifosfatos de nucleósidos unidos a etiquetas con carga positiva se sintetizan tal como se muestra en la figura 25. En primer lugar, se hace reaccionar un aminoácido trimetil-(lisina)_n-glicina (K[(Me)₃]_n-Gly) con carga positiva con el PEG-éster NHS y después se activa para formar el PEG-K[(Me)₃]_n-Gly-éster NHS. Se hace reaccionar este éster activado con el polifosfato de nucleótido terminado en amino tal como se muestra en las figuras 19 y 25. La carga neta en el tetrafosfato de nucleósido es neutra, pero tras la incorporación con polimerasa el PEG liberado tiene una carga positiva +1, y si se añade fosfatasa alcalina al cóctel de reacción la carga neta en el PEG liberado es de +4. Por tanto, la etiqueta liberada puede separarse e identificarse fácilmente mediante paso a través del nanoporo.

Síntesis de polifosfatos de nucleósidos unidos a PEG bloqueados en 3' para secuenciación por síntesis con detección con nanoporos.

La síntesis de polifosfatos de nucleósidos bloqueados en 3' sigue esencialmente la misma ruta que la mostrada para polifosfatos de nucleósidos unidos a etiqueta, excepto porque el 5'-trifosfato de nucleósido de partida es dNTP bloqueado en 3'-O. Tal como se muestra en la figura 31, en primer lugar se hace reaccionar 3'-O-azidometil-dNTP (6) con CDI o DCC/DMF seguido por reacción con ácido fosfórico (tetrafosfato) o pirofosfato (pentafosfatos). Esto se hace reaccionar, tras la purificación, con el nucleófilo apropiado para proporcionar fosfato terminado en amino que después se hace reaccionar con el PEG-éster NHS apropiado (neutro, con carga positiva o con carga negativa) para proporcionar el polifosfato de nucleósido unido a PEG bloqueado en 3'-O requerido.

- 5
- 10 Esquema de secuenciación con nucleótidos con PEG y detección con nanoporos (muchas copias de una molécula de ADN se inmovilizan sobre una perla y adición secuencial de un PEG-nucleótido cada vez).

Tal como se muestra en la figura 32, las moléculas de ADN se inmovilizan sobre una perla. Por tanto, cada perla tiene muchas copias de la misma molécula de ADN. Se añade la perla se añade a un micro/nanopocillo que está unido a un nanoporo. El ADN forma el complejo con la ADN polimerasa que o bien está unida al nanoporo o bien se añade al micro/nanopocillo junto con el nucleótido unido a PEG. Los nucleótidos pueden añadirse uno cada vez, si se añade el nucleótido correcto se incorpora y libera una etiqueta con PEG que proporciona una señal de bloqueo diferenciada cuando pasa a través de un nanoporo. Sin embargo, si el nucleótido no es complementario a la base de ácido nucleico de molde, no se incorporará y por tanto no se detectará ninguna señal. En este caso, puede usarse un PEG de igual longitud y peso molecular en los cuatro nucleótidos, o, si se desea, también pueden usarse cuatro PEG diferentes. Por tanto, la adición de base de ácido nucleico puede detectarse fácilmente mediante la señal de bloqueo de nanoporo con una sensibilidad de moléculas individuales.

- 15
- 20

Secuenciación por síntesis con PEG-nucleótidos bloqueados en 3'-O y detección con nanoporos (muchas copias de una molécula de ADN se inmovilizan sobre una perla y adición simultánea de los cuatro PEG-nucleótidos bloqueados en 3'-O).

- 25 Las regiones homopoliméricas del ADN pueden secuenciarse de manera correcta usando este enfoque. Por tanto, si el grupo 3'-OH del nucleótido está bloqueado mediante un resto reversible, la síntesis de ADN se detendrá tras la adición de un único nucleótido. La síntesis puede continuarse tras la retirada del grupo de bloqueo para generar un grupo 3'-OH libre. Tal como se muestra en la figura 33, los cuatro 3'-O-azidometil-nucleótidos unidos a PEG de diferente tamaño pueden añadirse al micro/nanopocillo de reacción y siempre que se incorpora un nucleótido correcto, se lee la etiqueta de PEG liberada mediante paso a través del nanoporo y se detecta la señal iónica. Dado que el grupo 3'-OH está bloqueado, sólo se añade un nucleótido cada vez. Este grupo bloqueado en 3'-O puede escindirse mediante tratamiento con TECP y por tanto el grupo OH libre está listo para una incorporación de nucleótido adicional. Mediante adición y escisión de nucleótidos repetidas, puede secuenciarse una región homopolimérica de manera correcta y sencilla.

- 35 Pirosecuenciación paralela de manera masiva usando nanoporos:

Tal como se muestra en la figura 34, en un modo en paralelo, masivo, puede construirse una matriz de alta densidad de micropocillos para realizar el procedimiento bioquímico. Cada micro/nanopocillo contiene un molde de ADN y dispositivo de nanoporo diferentes. Los PEG liberados se detectan con una sensibilidad de moléculas individuales.

Sumario del experimento:

- 40
- 1) Cualquier etiqueta de diferente tamaño, longitud, peso molecular, carga, unida al fosfato terminal del nucleótido que puede detectarse mediante nanoporo tras la incorporación con polimerasa.
 - 2) Etiqueta unida a los tri, tetra, penta, hexa-fosfatos.
 - 3) Detección electrónica.
 - 4) Grupo de moléculas de ADN unidas a la perla o superficie sólida y sensibilidad de detección de moléculas individuales (alta densidad y alta sensibilidad).
 - 5) Región homopolimérica secuenciada fácilmente usando nucleótidos bloqueados en 3'-O unidos a etiqueta.
 - 6) Adición de un nucleótido con etiqueta por ciclo.
 - 7) Adición de los cuatro nucleótidos marcados con etiqueta de manera reversible juntos para secuenciación de regiones homopoliméricas.
- 50
- 8) Alta sensibilidad, precisión y velocidad.
 - 9) Secuenciación en paralelo masiva.

Bibliografía:

1. Akesson, M., Branton, D., Kasianowicz, J.J., Brandin, E. y Deamer, D.W. Microsecond time-scale discrimination between polycytidylic acid and polyadenylic acid segments within single RNA molecules. *Biophys. J.* 1999, 77, 3227-3233.
- 5 2. Aksimentiev, A. *et al.*, Microscopic Kinetics of DNA Translocation through Synthetic Nanopores. *Biophysical Journal* 2004 87, 2086-2097.
3. Astier, Y., Braha, O. y Bayley, H. Toward single molecule DNA sequencing: direct identification of ribonucleoside and deoxyribonucleoside 5'-monophosphates by using an engineered protein nanopore equipped with a molecular adapter. *J Am Chem Soc* 128, 1705-10 (2006).
- 10 4. Axelrod, D., Koppel, D.E., Schlessinger, J., Elson, E. y Webb, W.W. Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. *Biophysical Journal* 16, 1055-1069 (1976).
5. Bailey, H. Sequencing single molecules of DNA. *Curr. Opinion Chem Biol.* 2006, 10, 628-637.
6. Benner, S. *et al.* Sequence-specific detection of individual DNA polymerase complexes in real time using a nanopore. *Nat Nanotechnol* 2, 718-24 (2007).
- 15 7. Bezrukov, S.M., y Kasianowicz, J.J. Neutral polymers in the nanopore of alamethicin and alpha-hemolysin. *Biologicheskie Membrany* 2001, 18, 453-457.
8. Bokhari, S.H. y Sauer, J.R., A Parallel Graph Decomposition Algorithm for DNA Sequencing with Nanopores. *Bioinformatics* 2005 21(7), 889-896.
9. Branton, D. Nanopore sequencing. *Nat. Biotechnol.* 26, 1146-1153 (2008).
- 20 10. Branton, D. *et al.* The potential and challenges of nanopore sequencing. *Nat Biotechnol* 26, 1146-53 (2008).
11. Bultmann, T., Vaz, W.L., Melo, E.C., Sisk, R.B. y Thompson, T.E. Fluid-phase connectivity and translational diffusion in a eutectic, two-component, two-phase phosphatidylcholine bilayer. *Biochemistry* 30, 5573-9 (1991).
12. Chandler, E.L., Smith, A.L., Burden, L.M., Kasianowicz y Burden, D.L. Membrane Surface Dynamics of DNAThreaded Nanopores Revealed by Simultaneous Single-Molecule Optical and Ensemble Electrical Recording. *Langmuir* 2004, 20, 898-905.
- 25 13. Chen, P. Probing single DNA molecule transport using fabricated nanopores. *Nano Lett.* 4, 2293-2298 (2004).
14. Clarke, J., Wu, H., Jayasinghe, L., Patel, A., Reid, S. y Bayley, H. Continuous base identification for singlemolecule nanopore DNA sequencing. *Nat. Biotech.* 2009, 1-6.
15. Cockroft, S.L., Chu, J., Amorin, M. y Ghadiri, M.R. A single-molecule nanopore device detects DNA polymerase activity with single-nucleotide resolution. *J Am Chem Soc* 130, 818-20 (2008).
- 30 16. Deamer, D.W. y Branton, D. Characterization of nucleic acids by nanopore analysis. *Acc. Chem. Res.* 2002, 35(10), 817-825.
17. Deamer, D.W. Nanopore Analysis of Nucleic Acids Bound to Exonucleases and Polymerases, *Annual Review of Biophysics*, Vol. 39: 79-90 (fecha de publicación del volumen, junio de 2010).
- 35 18. Derrington, I.M. *et al.* Nanopore DNA sequencing with MspA. *Proc Natl Acad Sci U S A* 107, 16060-5 (2010).
19. Dutta, D., Pulsipher, A. y Yousaf, M.N. Selective Tethering of Ligands and Proteins to a Microfluidically Patterned Electroactive Fluid Lipid Bilayer Array. *Langmuir* 26, 9835-9841 (2010).
20. Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G. *et al.* Real-time DNA sequencing from single polymerase molecules. *Science* 2009, 323, 133-138.
- 40 21. Fologea, D. *et al.*, Slowing DNA Translocation in a Solid State Nanopore. *Nano Letters* 2005 5(9), 1734-1737.
22. Fologea, D. *et al.*, Detecting Single Stranded DNA with a Solid State Nanopore. *Nano Letters* 2005 5(10), 1905-1909.
23. Garaj, S. *et al.* Graphene as a subnanometre trans-electrode membrane. *Nature* 467, 190-3 (2010).
- 45 24. Gershow, M. y Golovchenko, J.A. Recapturing and trapping single molecules with a solid state nanopore. *Nature Nanotechnology* 2, 775-779 (2007).

25. Guo, J.; Xu, N., Li, Z., Zhang, S.; Wu, J., Kim, D. H., Marma, M. S., Meng, Q., Cao, H., Li, X., Shi, S., Yu, L., Kalachikov, S., Russo, J. J., Turro, N. J., Ju, J. Four-color DNA sequencing with 3'-O-modified nucleotide reversible terminators and chemically cleavable fluorescent dideoxynucleotides. *Proc. Natl. Acad. Sci. USA* 2008, 105(27), 9145-9150
- 5 26. Guo, J., Yu, L., Turro, N. J., y Ju, J. An integrated system for DNA sequencing by synthesis using novel nucleotide analogues. *Accounts of Chemical Research* 2010, 43(4), 551-563.
27. Hall, A.R. *et al.* Hybrid pore formation by directed insertion of alpha-haemolysin into solid-state nanopores. *Nat Nanotechnol* 5, 874-7 (2010).
28. Harris, T. D., Buzby, P. J. *et al.* Single-molecule DNA sequencing of a viral genome. *Science* 2008, 320, 106-109.
- 10 29. Healy, K. Nanopore-based single-molecule DNA analysis. *Nanomedicine (Lond)* 2, 459-81 (2007).
30. Heng, J.B. *et al.*, Stretching DNA Using the Electric Field in a Synthetic Nanopore. *Nano Letters* 2005 5(10), 1883-1888.
31. Heng, J.B. *et al.*, The Electromechanics of DNA in a synthetic Nanopore. *Biophysical Journal* 2006, 90, 1098-1106.
- 15 32. Henrickson, S.E., Misakian, M., Robertson, B. y Kasianowicz, J.J. Driven asymmetric DNA transport in a nanometer-scale pore. *Physical Review Letters* 2000, 85, 3057-3060.
33. Hou, X. *et al.* A biomimetic potassium responsive nanochannel: G-quadruplex DNA conformational switching in a synthetic nanopore. *J Am Chem Soc* 131, 7800-5 (2009).
- 20 34. Huang, T.C. *et al.* Gene expression analysis with an integrated CMOS microarray by time-resolved fluorescence detection. *Biosens Bioelectron* 26, 2660-5 (2011).
35. Hurt, N., Wang, H., Akesson, M. y Lieberman, K.R. Specific nucleotide binding and rebinding to individual DNA polymerase complexes captured on a nanopore. *J Am Chem Soc* 131, 3772-8 (2009).
36. Jenkins, A.T.A., Neumann, T. y Offenhausser, A. Surface plasmon microscopy measurements of lipid vesicle adsorption on a micropatterned self-assembled monolayer. *Langmuir* 17, 265-267 (2001).
- 25 37. Ju, J., Kim, D. H., Bi, L., Meng, Q., Bi, X., Li, Z., Li, X., Marma, M. S., Shundi, S., Wu, J., Edwards, J. R., Romu, A., y Turro, N. J. Four-color DNA sequencing by synthesis using cleavable fluorescent nucleotide reversible terminators. *Proc. Natl. Acad. Sci. USA* 2006, 103(52), 19635-19640
38. Kasianowicz, J.J., Brandin, E., Branton, D. y Deamer, D.W. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl. Acad. Sci. USA* 1996, 93, 13770-13773.
- 30 39. Kasianowicz, J.J. Nanometer-scale pores: potential applications for DNA characterization and analyte detection. *Disease Markers* 2003, 18, 185-191.
40. Kasianowicz, J.J. Nanopore. Flossing with DNA. *Nature Materials* 2004, 3, 355-356.
41. Korlach, J. *et al.* Selective aluminum passivation for targeted immobilization of single DNA polymerase molecules in zero-mode waveguide nanostructures. *Proc Natl Acad Sci U S A* 105, 1176-81 (2008).
- 35 42. Korlach, J. *et al.* Real-time DNA sequencing from single polymerase molecules. *Methods Enzymol* 472, 431-55 (2010).
43. Kumar, S, y Sood, A. Labeled Nucleoside Polyphosphates. Patente estadounidense de 2006, 7.041.812
44. Kumar, S., McDougall, M., Sood, A., Nelson, J., Fuller, C., Macklin, J. y Mitsis, P. Terminal-Phosphate-Labeled Nucleotides with New Linkers. US Patent 2008, 7,393,640
- 40 45. Kumar, S., Sood, A., Wegener, J., Finn, P., Nampalli, S., Nelson, J., Sekher, A., Mitsis, P., Macklin, J. y Fuller, C. W. Terminal Phosphate Labeled Nucleotides: Synthesis, Applications and Linker effect on incorporation by DNA Polymerases. *Nucleosides, Nucleotides & Nucleic Acids* (2005) 24 (5-7), 401-408
46. Lee S.E., Sidorov A., Gourlain T., Mignet N., Thorpe S.J., Brazier J.A., Dickman M.J., Hornby D.P., Grasby, J.A. y Williams, D.M. Enhancing the catalytic repertoire of nucleic acids: a systematic study of linker length and rigidity. *Nucleic Acids Research* 2001, 29(7), 1565-1573.
- 45 47. Lei, N., Watson, B.O., MacLean, J.N., Yuste, R. y Shepard, K.L. A 256-by-56 CMOS Microelectrode Array for Extracellular Neural Stimulation of Acute Brain Slices. in *Solid-State Circuits Conference, 2008. ISSCC 2008. Digest of Technical Papers. IEEE International* 148-603 (2008).

48. Levine, P.M., Gong, P., Levicky, R. y Shepard, K.L. Real-time, multiplexed electrochemical DNA detection using an active complementary metal-oxide-semiconductor biosensor array with integrated sensor electronics. *Biosens Bioelectron* 24, 1995-2001 (2009).
- 5 49. Li, Z., Bai, X., Ruparel, H., Kim, S., Turro, N.J. y Ju, J. A photocleavable fluorescent nucleotide for DNA sequencing and analysis. *Proc. Natl. Acad. Sci. USA* 2003, 100, 414-419.
50. Love, J.C., Estroff, L.A., Kriebel, J.K., Nuzzo, R.G. y Whitesides, G.M. Self-assembled monolayers of thiolates on metals as a form of nanotechnology. *Chem Rev* 105, 1103-69 (2005).
51. Margulies, M. *et al.* Genome Sequencing in Open Microfabricated High Density Picoliter Reactors. *Nature* 437: 376-380 (2005).
- 10 52. Mathe, J. *et al.*, Nanopore Unzipping of Individual Hairpin Molecules. *Biophysical Journal* 2004 87, 3205-3212.
53. Matysiak, S., Montesi, A., Pasquali, M., Kolomeisky, A.B. y Clementi, C. Dynamics of polymer translocation through nanopores: theory meets experiment. *Phys Rev Lett* 96, 118103 (2006).
54. McNally, B. *et al.* Optical recognition of converted DNA nucleotides for single-molecule DNA sequencing using nanopore arrays. *Nano Lett* 10, 2237-44 (2010).
- 15 55. Meller, A., Nivon, L., Brandin, E., Golovchenko, J. y Branton, D. Rapid nanopore discrimination between single polynucleotide molecules. *Proc. Natl. Acad. Sci. USA* 2000, 97, 1079-1084.
56. Meller, A. *et al.*, Single Molecule Measurements of DNA Transport Through a Nanopore. *Electrophoresis* 2002 23, 2583-2591.
57. Merchant, C.A. *et al.* DNA translocation through graphene nanopores. *Nano Lett* 10, 2915-21 (2010).
- 20 58. Nam, J.M., Nair, P.M., Neve, R.M., Gray, J.W. y Groves, J.T. A fluid membrane-based soluble ligand-display system for live-cell assays. *ChemBioChem* 7, 436-440 (2006).
59. Nkodo, A.E. *et al.* Diffusion coefficient of DNA molecules during free solution electrophoresis. *ELECTROPHORESIS* 22, 2424-2432 (2001).
- 25 60. Oliver, A.E. y Deamer, D.W. Alpha-Helical Hydrophobic Polypeptides Form Proton-Selective Channels in Lipid Bilayers. *Biophysical Journal* 66, 1364-1379 (1994).
61. Palegrosdemange, C., Simon, E.S., Prime, K.L. y Whitesides, G.M. Formation of Self-Assembled Monolayers by Chemisorption of Derivatives of Oligo(Ethylene Glycol) of Structure Hs(Ch₂)₁₁(Och₂ch₂)₂Meta-Oh on Gold. *Journal of the American Chemical Society* 113, 12-20 (1991).
- 30 62. Perkins, T.T., Quake, S.R., Smith, D.E. y Chu, S. Relaxation of a single DNA molecule observed by optical microscopy. *Science* 1994, 264, 822-826.
63. Reiner, J.E., Kasianowicz, J.J., Nablo, B.J. y Robertson, J.W. Theory for polymer analysis using nanopore-based single-molecule mass spectrometry. *Proc Natl Acad Sci U S A* 107, 12080-5 (2010).
64. Rief, M., Clausen-Schaumann, H. y Gaub, H.E. Sequence-dependent mechanics of single DNA molecules. *Nat. Struct. Biol.* 1999, 6, 346-349.
- 35 65. Robertson, J. W. F., Rodrigues, C. W., Stanford, V. M., Rubinson, K. A., Krasilnikov, O. V. y Kasianowicz, J. J. Single-molecule mass spectrometry in solution using a solitary nanopore. *Proc. Natl. Acad. Sci. USA* 2007, 104, 8207-8211.
66. Ronaghi M., Uhlen, M. y Nyren, P. A sequencing method based on real-time pyrophosphate. *Science* 281, 363-365 (1998)
- 40 67. Rosenstein, J., V. Ray, M. Drndic, y K.L. Shepard. Solid-state nanopores integrated with low-noise preamplifiers for high-bandwidth DNA analysis. in *Life Science Systems and Applications Workshop (LISSA)*, 2011 IEEE/NIH (2011).
- 45 68. Rosenstein, J., Ray, V., Drndic, M. y Shepard, K.L. Nanopore DNA sensors in CMOS with on-chip low-noise preamplifiers. in *Solid-State Sensors, Actuators and Microsystems Conference (TRANSDUCERS)*, 2011 16th International 874-877 (2011).
69. Rothberg, J.M. *et al.* An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 475, 348-352 (2011).
70. Sauer-Budge, A.F. *et al.*, Unzipping Kinetics of Double Stranded DNA in a Nanopore. *Physical Review Letters*

2003 90(23), 238101-1 - 238101-4.

71. Schneider, G.F. *et al.* DNA translocation through graphene nanopores. *Nano Lett* 10, 3163-7 (2010).
72. Seo, T.S., Bai, X., Ruparel, H., Li, Z., Turro, N.J. y Ju, J. Photocleavable fluorescent nucleotides for DNA sequencing on a chip constructed by site-specific coupling chemistry. *Proc. Natl. Acad. Sci. USA* 2004, 101, 5488-5493.
73. Shapovalov, G. y Lester, H. A. Getting transitions in bacterial ion channels measured at 3 microseconds resolution. *J. Gen. Physiol.* 2004, 124, 151-161.
74. Shen, K., Tsai, J., Shi, P. y Kam, L.C. Self-aligned supported lipid bilayers for patterning the cell-substrate interface. *J Am Chem Soc* 131, 13204-5 (2009).
75. Shi, H. y Ratner, B.D. Template recognition of protein-imprinted polymer surfaces. *Journal of Biomedical Materials Research* 49, 1-11 (2000).
76. Shimada, K., Kato, H., Saito, T., Matsuyama, S. y Kinugasa, S. Precise measurement of the self-diffusion coefficient for poly(ethylene glycol) in aqueous solution using uniform oligomers. *Journal of Chemical Physics* 122 (2005) .
77. Smith, S.B., Cui, Y. y Bustamante, C. Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science* 1996, 271, 795-799.
78. Sood, A., Kumar, S., Wegener, J., Nampalli, S., Nelson, J., Macklin, J. y Fuller, C. W. Terminal Phosphate Labeled Nucleotides with improved substrate properties for Nucleic Acid Assays. *J. Am. Chem. Soc.* 2005, 127(8), 2394-2395.
79. Srinivasan, M.P., Ratto, T.V., Stroeve, P. y Longo, M.L. Patterned supported bilayers on self-assembled monolayers: Confinement of adjacent mobile bilayers. *Langmuir* 17, 7951-7954 (2001).
80. Storm, A.J. *et al.* Fast DNA translocation through a solid-state nanopore. *Nano Lett* 5, 1193-1197 (2005).
81. Timp, W. *et al.* Nanopore Sequencing: Electrical Measurements of the Code of Life. *IEEE Trans Nanotechnol* 9, 281-294 (2010).
82. Wang, H. *et al.*, DNA heterogeneity and Phosphorylation unveiled by Single-Molecule Electrophoresis. *PNAS* 2004 101(37), 13472-13477.
83. Wang, Y., Zheng, D., Tan, Q., Wang, M.X. y Gu, L.Q. Nanopore-based detection of circulating microRNAs in lung cancer patients. *Nat Nanotechnol* 6, 668-74 (2011).
84. Wanunu, M. y Meller, A. Chemically modified solid-state nanopores. *Nano Lett* 7, 1580-5 (2007).
85. Wanunu, M., Sutin, J., McNally, B., Chow, A. y Meller, A. DNA translocation governed by interactions with solidstate nanopores. *Biophys J* 95, 4716-25 (2008).
86. Wanunu, M. *et al.* Rapid electronic detection of probe-specific microRNAs using thin nanopore sensors. *Nat Nanotechnol* 5, 807-14 (2010).
87. White, R.J. *et al.* Ionic conductivity of the aqueous layer separating a lipid bilayer membrane and a glass support. *Langmuir* 22, 10777-83 (2006).
88. White, R.J. *et al.* Single ion-channel recordings using glass nanopore membranes. *J Am Chem Soc* 129, 11766-75 (2007).
89. Wiehelman, K. Investigation of the bicinchoninic acid protein assay: identification of the groups responsible for color formation. *Analytical Biochemistry* 175(1988).
90. Yang, T.L., Baryshnikova, O.K., Mao, H.B., Holden, M.A. y Cremer, P.S. Investigations of bivalent antibody binding on fluid-supported phospholipid membranes: The effect of hapten density. *Journal of the American Chemical Society* 125, 4779-4784 (2003).
91. Yin, P. Tethered Bilayer Membrane Sensors with Small Transmembrane Peptide Ion Channels - Recent Developments, Future Research and Potential Applications. in *Advances in Planar Lipid Bilayers and Liposomes*, Vol. Volume 2 (ed. Ottova-Leitmannova, A.) 49-76 (Academic Press, 2005).
92. Vercoetere, W., Winters-Hilt, S., Olsen, H., Deamer, D., Haussler, D. y Akeson, M. Rapid discrimination among individual DNA hairpin molecules at single-nucleotide resolution using an ion channel. *Nat. Biotech* 2001, 19, 248-252.

93. Vercoutere, W.A. *et al.*, Discrimination Among Individual Watson-Crick Base Pairs at the Terminus of Single DNA Hairpin Molecules. *Nucleic Acids Research* 2003 31(4), 1311-1318.

REIVINDICACIONES

1. Método para determinar la secuencia de nucleótidos de una molécula de ácido nucleico monocatenario que comprende:

5 (a) poner en contacto la molécula de ácido nucleico monocatenario, en el que la molécula de ácido nucleico monocatenario está en una disolución de electrolito en contacto con un nanoporo en una membrana y en el que la molécula de ácido nucleico monocatenario tiene un cebador hibridado con una parte de la misma, con una ácido nucleico polimerasa y al menos cuatro análogos de polifosfato de nucleótido (NPP) marcados con etiqueta de fosfato en condiciones que permiten que la ácido nucleico polimerasa catalice la incorporación de uno de los análogos de NPP en el cebador si es complementario al residuo de nucleótido de la molécula de ácido nucleico monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido de la molécula de ácido nucleico monocatenario hibridado con el residuo de nucleótido 3'-terminal del cebador, para formar un producto de extensión de ácido nucleico

15 y con la condición de que (i) el tipo de base en cada análogo de NPP es diferente del tipo de base en cada uno de los otros tres análogos de NPP, y (ii) o bien el número de fosfatos en la parte de polifosfato de cada análogo de NPP es diferente del número de fosfatos en la parte de polifosfato de cada uno de los otros tres análogos de NPP, o bien el número de fosfatos en la parte de polifosfato de cada uno de los cuatro análogos de NPP es el mismo y el tipo de etiqueta en cada análogo de NPP es diferente del tipo de etiqueta en cada uno de los otros tres análogos de NPP,

20 en el que la incorporación del análogo de NPP da como resultado la liberación de un polifosfato que tiene la etiqueta unida al mismo; y

(b) determinar qué análogo de NPP se ha incorporado en el cebador para formar un producto de extensión de ácido nucleico en la etapa (a) aplicando una tensión a través de la membrana y midiendo un cambio electrónico a través del nanoporo que resulta de que el polifosfato que tiene la etiqueta unida al mismo generado en la etapa (a) se transloca a través del nanoporo, en el que el cambio electrónico es diferente para cada número de fosfatos en la parte de polifosfato, o para cada tipo de etiqueta diferente, según sea apropiado, identificando de ese modo el residuo de nucleótido en el ácido nucleico monocatenario complementario al análogo de NPP incorporado; y

(c) realizar de manera iterativa las etapas (a) y (b) para cada residuo de nucleótido de la molécula de ácido nucleico monocatenario que está secuenciándose, en el que en cada iteración de la etapa (a) se incorpora el análogo de NPP en el producto de extensión de ácido nucleico resultante de la iteración anterior de la etapa (a) si es complementario al residuo de nucleótido de la molécula de ácido nucleico monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido de la molécula de ácido nucleico monocatenario hibridado con el residuo de nucleótido 3'-terminal del producto de extensión de ácido nucleico,

35 determinando así la secuencia de nucleótidos de la molécula de ácido nucleico monocatenario.

2. Método para determinar la secuencia de nucleótidos de una molécula de ácido nucleico monocatenario que comprende:

40 (a) poner en contacto la molécula de ácido nucleico monocatenario, en el que la molécula de ácido nucleico monocatenario está en una disolución de electrolito en contacto con un nanoporo en una membrana y en el que la molécula de ácido nucleico monocatenario tiene un cebador hibridado con una parte de la misma, una ácido nucleico polimerasa y un análogo de polifosfato de nucleótido (NPP) marcado con etiqueta de fosfato en condiciones que permiten que la ácido nucleico polimerasa catalice la incorporación del análogo de NPP en el cebador si es complementario al residuo de nucleótido de la molécula de ácido nucleico monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido de la molécula de ácido nucleico monocatenario hibridado con el residuo de nucleótido 3'-terminal del cebador, para formar un producto de extensión de ácido nucleico;

45 y en el que si no se incorpora el análogo de NPP, se repite de manera iterativa la puesta en contacto con un análogo de NPP diferente hasta que se incorpora un análogo de NPP, con la condición de que (1) el tipo de base en cada análogo de NPP es diferente del tipo de base en cada uno de los otros análogos de NPP, y (2) o bien el número de fosfatos en la parte de polifosfato de cada análogo de NPP es diferente del número de fosfatos en la parte de polifosfato de cada uno de los otros tres análogos de NPP, o bien el número de fosfatos en la parte de polifosfato de cada uno de los cuatro análogos de NPP es el mismo y el tipo de etiqueta en cada análogo de NPP es diferente del tipo de etiqueta en cada uno de los otros tres análogos de NPP,

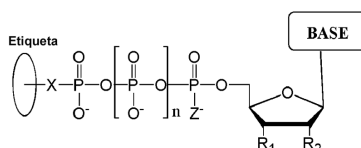
55 en el que la incorporación de un análogo de NPP da como resultado la liberación de un polifosfato que tiene la etiqueta unida al mismo;

(b) determinar qué análogo de NPP se ha incorporado en el cebador para formar un producto de extensión de ácido nucleico en la etapa (a) aplicando una tensión a través de la membrana y midiendo un cambio electrónico a través del nanoporo que resulta de que el polifosfato que tiene la etiqueta unida al mismo generado en la etapa (a) se transloca a través del nanoporo, en el que el cambio electrónico es diferente para cada número de fosfatos en la parte de polifosfato, o para cada tipo de etiqueta diferente, según sea apropiado, identificando de ese modo el residuo de nucleótido en la molécula de ácido nucleico monocatenario complementario al análogo de NPP incorporado; y

(c) realizar de manera iterativa las etapas (a) y (b) para cada residuo de nucleótido de la molécula de ácido nucleico monocatenario que está secuenciándose, en el que en cada iteración de la etapa (a) se incorpora el análogo de NPP en el producto de extensión de ácido nucleico resultante de la iteración anterior de la etapa (a) si es complementario al residuo de nucleótido de la molécula de ácido nucleico monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido de la molécula de ácido nucleico monocatenario hibridado con el residuo de nucleótido 3'-terminal del producto de extensión de ácido nucleico,

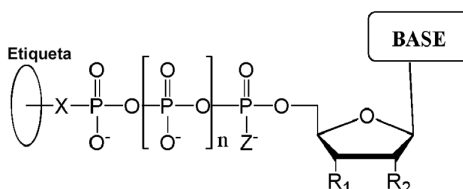
determinando así la secuencia de nucleótidos de la molécula de ácido nucleico monocatenario.

3. Método según la reivindicación 1 ó 2, en el que (i) si se pone en contacto la molécula de ácido nucleico con al menos cuatro análogos de NPP, cada análogo de NPP tiene la estructura:



en la que la base es adenina, guanina, citosina, timina o uracilo, o un derivado de cada una de las mismas, en la que R_1 es OH, en la que R_2 es H u OH, en la que X es O, NH, S o CH_2 o un grupo de unión, en la que n es 1, 2, 3, 4 ó 5, y en la que Z es O, S o BH_3 ; o

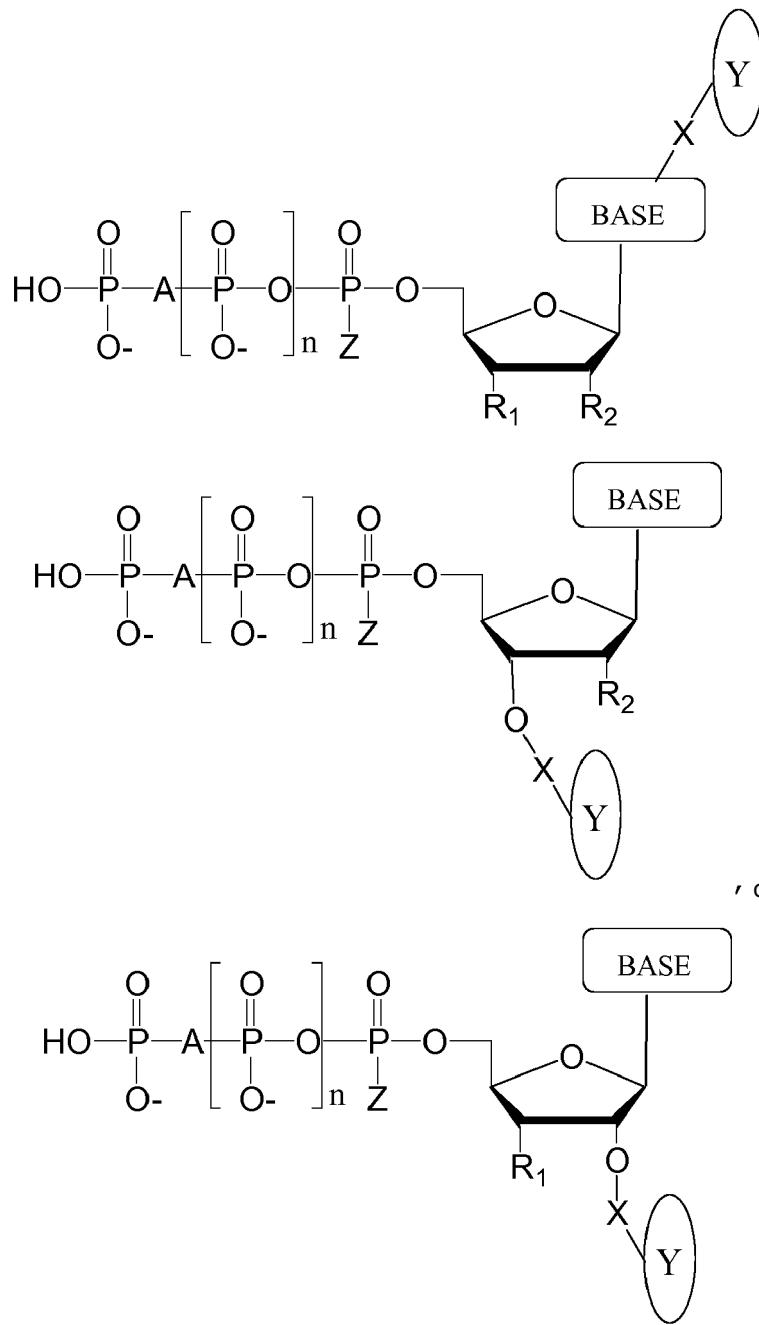
en el que (ii) si se pone en contacto la molécula de ácido nucleico con un análogo de NPP, en el que se repite la puesta en contacto con un análogo de NPP diferente hasta que se incorpora un análogo de NPP, cada análogo de NPP tiene la estructura:



en la que la base es adenina, guanina, citosina, timina o uracilo, o un derivado de cada una de las mismas, en la que R_1 es OH, $-O-CH_2N_3$ o $-O-2$ -nitrobenzilo, en la que R_2 es H u OH, en la que X es O, NH, S o CH_2 o un grupo de unión, en la que n es 1, 2, 3, 4 ó 5, y en la que Z es O, S o BH_3 .

4. Método para determinar la secuencia de nucleótidos de un ADN monocatenario que comprende:

(a) poner en contacto el ADN monocatenario, en el que el ADN monocatenario está en una disolución de electrolito en contacto con un nanoporo en una membrana y en el que el ADN monocatenario tiene un cebador hibridado con una parte del mismo, con una ADN polimerasa y al menos cuatro análogos de polifosfato de desoxirribonucleótido (dNPP) en condiciones que permiten que la ADN polimerasa catalice la incorporación de uno de los análogos de dNPP en el cebador si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-terminal del cebador, para formar un producto de extensión de ADN, en el que cada uno de los cuatro análogos de dNPP tiene una estructura elegida de las siguientes:



5

en las que la base es adenina, guanina, citosina, timina o uracilo, o un derivado de cada una de las mismas, en las que Y es una etiqueta, en las que R₁, si está presente, es OH, en las que R₂, si está presente, es H, en las que X es un grupo de unión escindible, en las que Z es O, S o BH₃, en las que n es 1, 2, 3, 4 ó 5, en las que A es O, S, CH₂, CHF, CFF o NH, y con la condición de que (i) el tipo de base en cada análogo de dNPP es diferente del tipo de base en cada uno de los otros tres análogos de dNPP, y (ii) el tipo de etiqueta en cada análogo de dNPP es diferente del tipo de etiqueta en cada uno de los otros tres análogos de dNPP;

10

(b) escindir la etiqueta del análogo de dNPP incorporado en la etapa (a);

15

(c) determinar qué análogo de dNPP se incorporó en la etapa (a) aplicando una tensión a través de la membrana y midiendo un cambio electrónico a través del nanoporo que resulta de que la etiqueta del análogo de dNPP incorporado en la etapa (a) se transloca a través del nanoporo, en el que el cambio electrónico es diferente para cada tipo diferente de etiqueta, identificando de ese modo el residuo de nucleótido en el ADN monocatenario complementario al análogo de dNPP incorporado; y

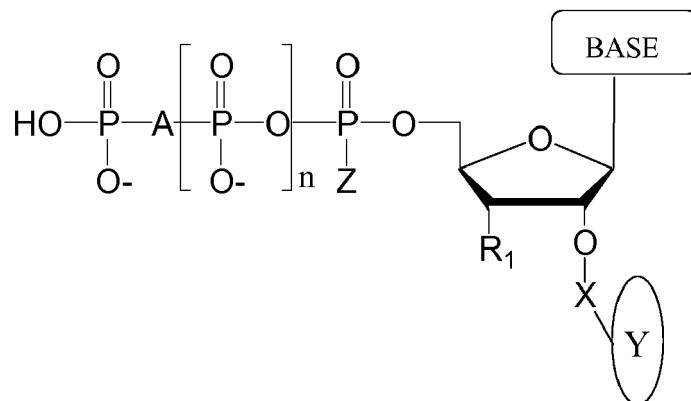
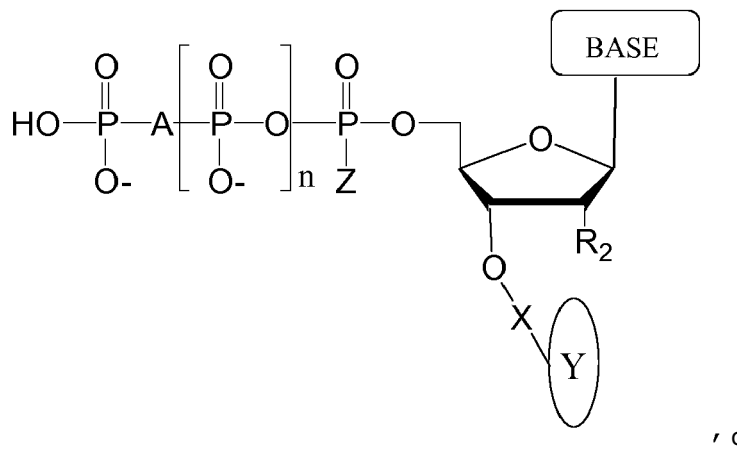
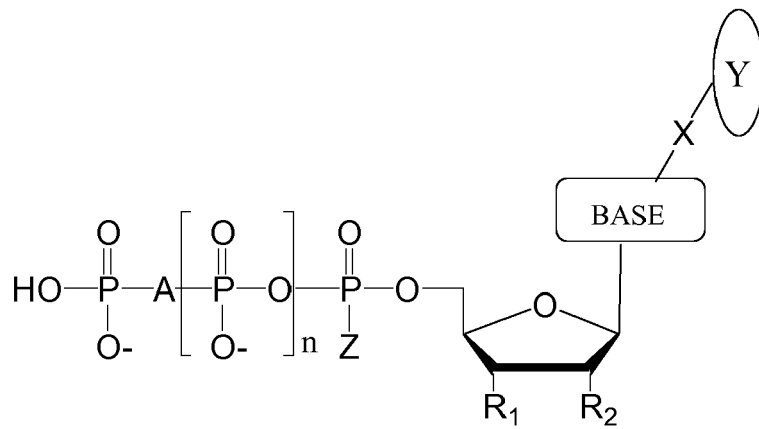
(d) realizar de manera iterativa las etapas (a), (b) y (c) para cada residuo de nucleótido del ADN monocatenario que está secuenciándose, en el que en cada iteración de la etapa (a) se incorpora el

análogo de dNPP en el producto de extensión de ADN resultante de la iteración anterior de la etapa (a) si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-terminal del producto de extensión de ADN,

5 determinando así la secuencia de nucleótidos del ADN monocatenario.

5. Método para determinar la secuencia de nucleótidos de un ADN monocatenario que comprende:

10 (a) poner en contacto el ADN monocatenario, en el que el ADN monocatenario está en una disolución de electrolito en contacto con un nanoporo en una membrana, en el que el ADN monocatenario tiene un cebador hibridado con una parte del mismo, una ADN polimerasa y un análogo de polifosfato de desoxirribonucleótido (dNPP) en condiciones que permiten que la ADN polimerasa catalice la incorporación del análogo de dNPP en el cebador si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-terminal del cebador, para formar un producto de extensión de ADN, en el que el análogo de dNPP tiene la estructura:



en las que la base es adenina, guanina, citosina, uracilo o timina, o un derivado de cada una de las mismas, en las que Y es una etiqueta, y en las que R₁ si está presente es OH, -OCH₂N₃ o -O-2-nitrobencilo, R₂ si está presente es H, en las que X es un grupo de unión escindible, en las que Z es O, S o BH₃, en las que n es 1, 2, 3, 4 ó 5, en las que A es O, S, CH₂, CHF, CFF o NH,

5 y si no se incorpora el análogo de dNPP, repetir de manera iterativa la puesta en contacto con un análogo de dNPP diferente hasta que se incorpora un análogo de dNPP, con la condición de que (1) el tipo de base en cada análogo de dNPP es diferente del tipo de base en cada uno de los otros análogos de dNPP, y (2) el tipo de etiqueta en cada análogo de dNPP es diferente del tipo de etiqueta en cada uno de los otros análogos de dNPP,

10 en el que la incorporación de un análogo de dNPP da como resultado la liberación de un polifosfato que tiene la etiqueta unida al mismo;

(b) escindir la etiqueta del análogo de dNPP incorporado en la etapa (a);

15 (c) determinar qué análogo de dNPP se incorporó en la etapa (a) para formar un producto de extensión de ADN aplicando una tensión a través de la membrana y midiendo un cambio electrónico a través del nanoporo que resulta de que la etiqueta del análogo de dNPP incorporado en la etapa (a) se transloca a través del nanoporo, en el que el cambio electrónico es diferente para cada tipo de etiqueta, identificando de ese modo el residuo de nucleótido en el ADN monocatenario complementario al análogo de dNPP incorporado; y

20 (d) realizar de manera iterativa las etapas (a), (b) y (c) para cada residuo de nucleótido del ADN monocatenario que está secuenciándose, en el que en cada iteración de la etapa (a) se incorpora el análogo de dNPP en el producto de extensión de ADN resultante de la iteración anterior de la etapa (a) si es complementario al residuo de nucleótido del ADN monocatenario que está inmediatamente en 5' con respecto a un residuo de nucleótido del ADN monocatenario hibridado con el residuo de nucleótido 3'-terminal del producto de extensión de ADN,

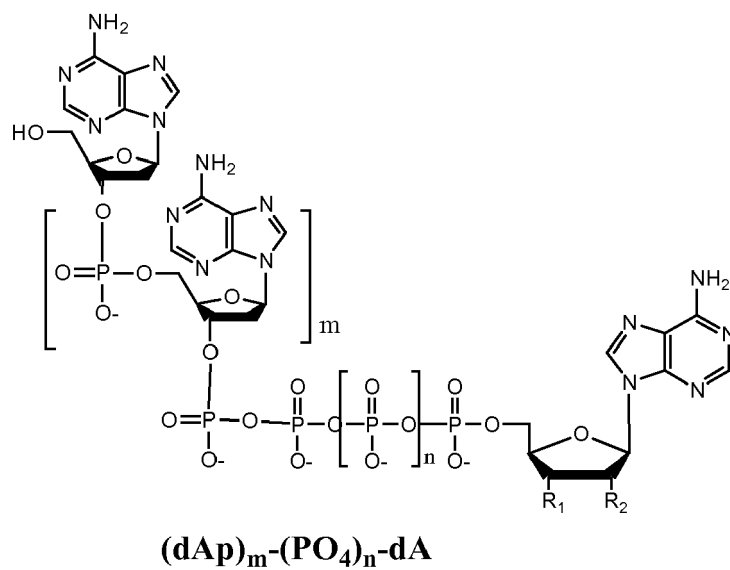
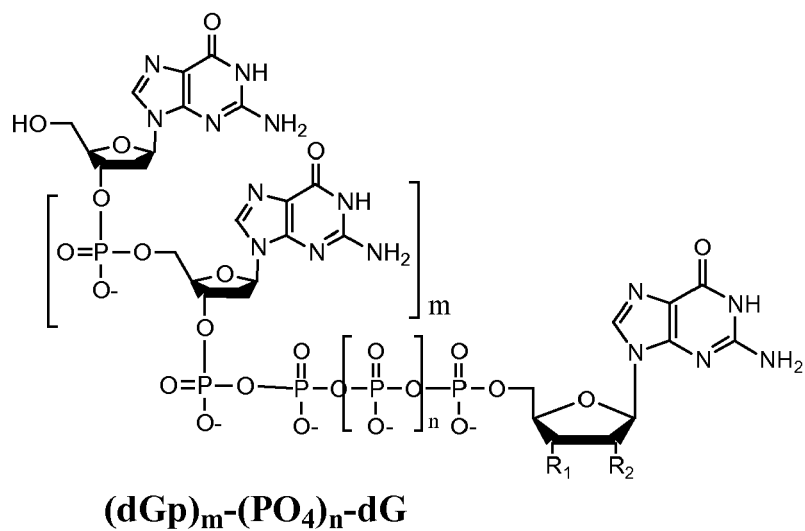
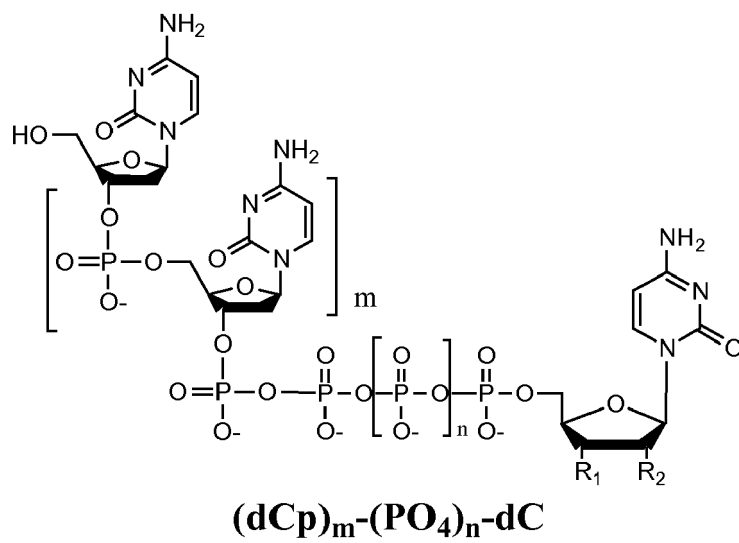
25 determinando así la secuencia de nucleótidos del ADN monocatenario.

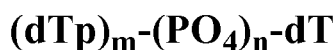
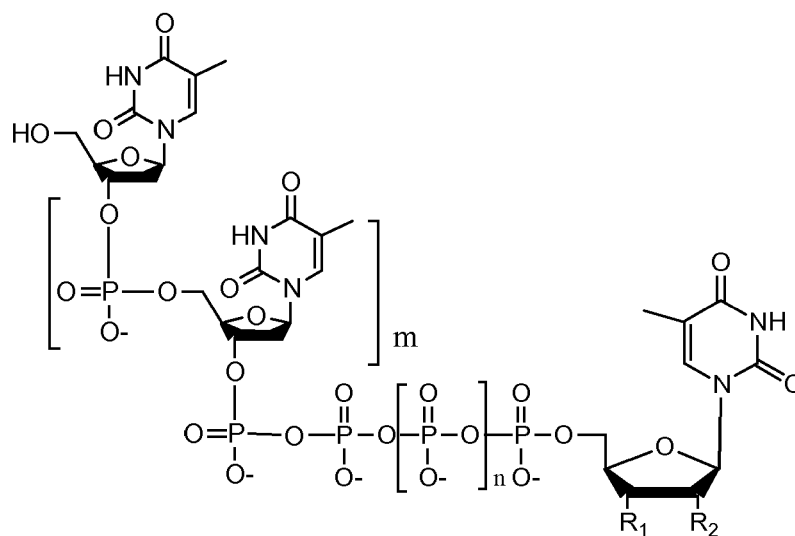
6. Método según la reivindicación 4 ó 5, en el que la etapa (b) se realiza antes que la etapa (c).

7. Método según una cualquiera de las reivindicaciones 1-6, en el que la etiqueta es polietilenglicol, un aminoácido, un hidrato de carbono, un colorante, un mononucleótido, un dinucleótido, un trinucleótido, un tetranucleótido, un pentanucleótido o un hexanucleótido, un colorante fluorescente, un compuesto quimioluminiscente, un aminoácido, un péptido, un hidrato de carbono, un monofosfato de nucleótido, un difosfato de nucleótido, un ácido alifático o un ácido aromático o un alcohol o un tiol no sustituido o sustituido con uno o más halógenos, un grupo ciano, un grupo nitro, un grupo alquilo, un grupo alquenilo, un grupo alquinilo o un grupo azido, en el que la base se selecciona del grupo que consiste en adenina, guanina, citosina, timina, 7-desazaguanina, 7-desazaadenina o 5-metilcitosina.

8. Método según una cualquiera de las reivindicaciones 1-7, en el que el tipo de etiqueta en cada análogo de NPP o dNPP es diferente del tipo de etiqueta en cada uno de los otros análogos de NPP o dNPP; en el que cada análogo de NPP o dNPP comprende una etiqueta que tiene una carga diferente de cualquier otra etiqueta; o en el que cada análogo de NPP o dNPP comprende un grupo de unión que tiene una longitud diferente de la de cualquier otro grupo de unión.

9. Método según una cualquiera de las reivindicaciones 1-8, en el que el análogo de NPP es un análogo de dNPP y la etiqueta es un mononucleótido, un dinucleótido, un trinucleótido, un tetranucleótido, un pentanucleótido o un hexanucleótido y en el que la base del mononucleótido, el dinucleótido, el trinucleótido, el tetranucleótido, el pentanucleótido o el hexanucleótido es el mismo tipo de base que la base del análogo de dNPP, o el análogo de dNPP se elige de los siguientes:





en el que en cada estructura n es, independientemente, 1, 2, 3 ó 4, R₁ es -OH, R₂ es H, y m es, independientemente, un número entero de desde 0 hasta 100, un número entero de desde 0 hasta 50, o un número entero de desde 0 hasta 10, y en el que cuando m es 0 el fosfato terminal del dNPP está unido directamente al átomo de O en 3' del nucleósido mostrado en el lado izquierdo de la estructura, y en el que el valor de n es diferente para cada tipo de base.

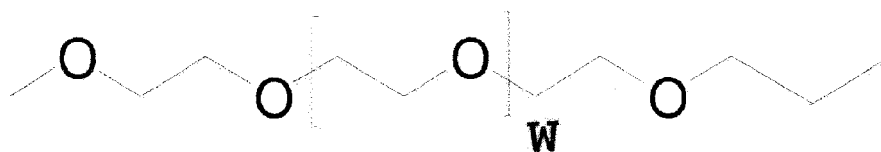
5

10. Método según cualquiera de las reivindicaciones 1-9, en el que el cambio electrónico es un cambio en la amplitud de corriente o la conductancia del nanoporo, en el que el nanoporo es un nanoporo biológico, preferiblemente proteico tal como una alfa-hemolisina, o un nanoporo en estado sólido, preferiblemente grafeno, opcionalmente en el que el nanoporo está en una membrana en estado sólido, opcionalmente en el que el nanoporo está integrado con electrónica de CMOS, opcionalmente en el que el ADN monocatenario, el ARN monocatenario, el cebador, la ADN polimerasa o la ARN polimerasa está unido a una superficie sólida, en el que la unión se realiza a través de uniones de biotina-estreptavidina, en el que la ADN o ARN polimerasa está unida a través de superficie de oro modificada con una monocapa autoensamblada de alcanotiol funcionalizada con grupos amino modificados para dar ésteres NHS, u opcionalmente en el que el nanoporo es parte de una matriz de nanoporos.

10

15

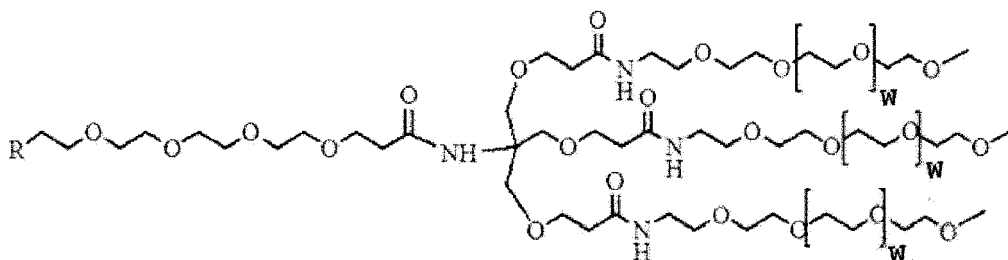
11. Método según una cualquiera de las reivindicaciones 1-10, en el que la etiqueta tiene la siguiente estructura:



20

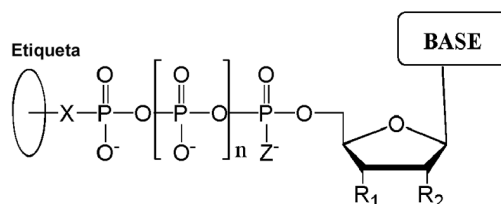
en la que W es un número entero de entre 0 y 100, o

en el que la etiqueta tiene la siguiente estructura:



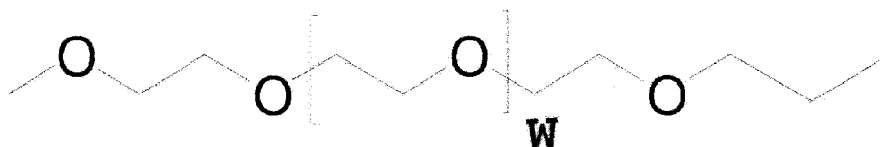
en la que R es NH₂, OH, COOH, CHO, SH o N₃, y W es un número entero de desde 0 hasta 100.

12. Compuesto que tiene la estructura:



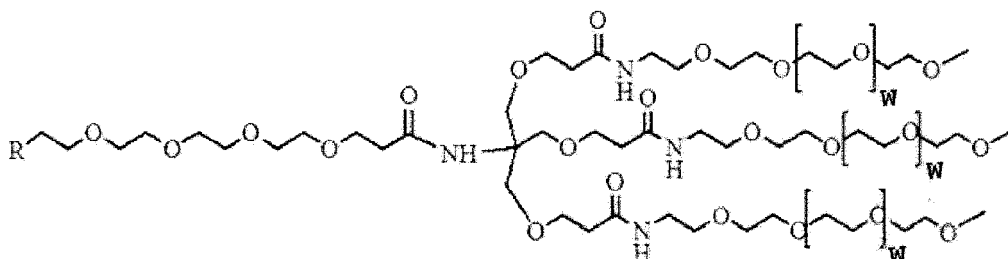
5 en la que la etiqueta es polietilenglicol, un aminoácido, un hidrato de carbono, un mononucleótido, dinucleótido, trinucleótido, tetranucleótido, pentanucleótido o hexanucleótido, en la que R₁ es OH, en la que R₂ es H u OH, en la que X es O, NH, S o CH₂, en la que Z es O, S o BH₃, en la que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina, y en la que n es 1, 2, 3, 4 ó 5;

en el que si la etiqueta es polietilenglicol, la etiqueta tiene la siguiente estructura:



en la que W es un número entero de entre 0 y 100, o

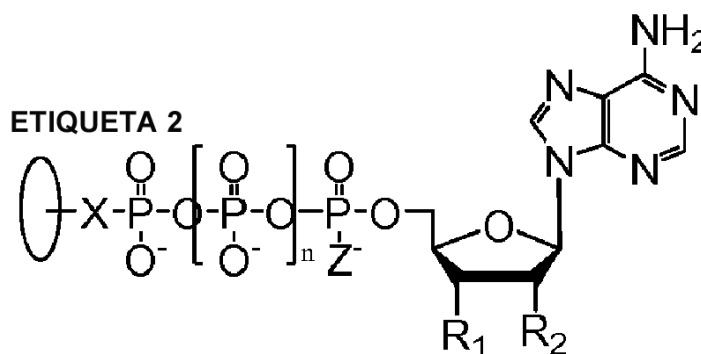
la etiqueta tiene la siguiente estructura:



10

en la que R es NH₂, OH, COOH, CHO, SH o N₃, y W es un número entero de desde 0 hasta 100,

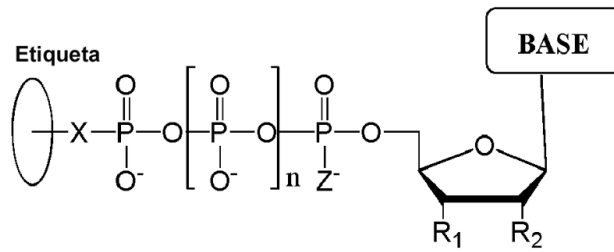
y en el que el compuesto no es:



15

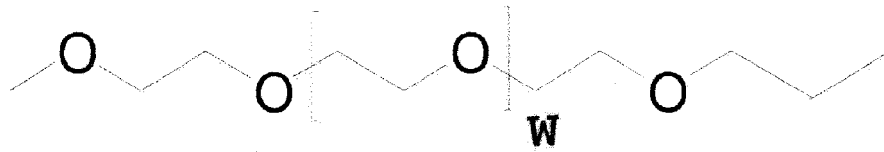
en el que ETIQUETA2 es polietilenglicol, un aminoácido, un hidrato de carbono, un mononucleótido, dinucleótido, trinucleótido, tetranucleótido, pentanucleótido o hexanucleótido, en el que R₁ es OH, en el que R₂ es H u OH, en el que X es O, NH, S o CH₂, en el que Z es O, S o BH₃, y en el que n es 1, 2, 3, 4 ó 5.

13. Compuesto que tiene la estructura:



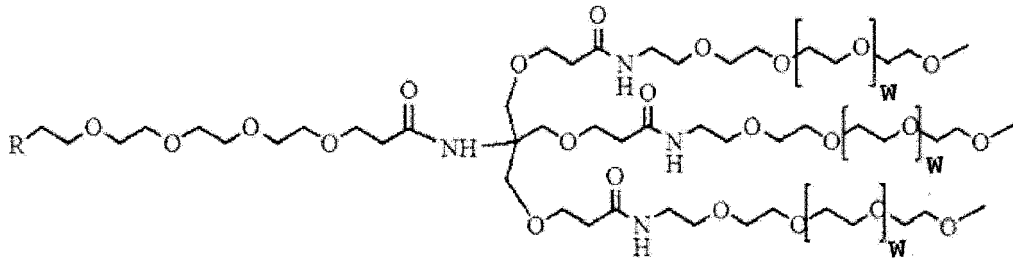
en la que la etiqueta es polietilenglicol, un aminoácido, o un hidrato de carbono, en la que R_1 es OH, en la que R_2 es H u OH, en la que X es O, NH, S o CH_2 , en la que Z es O, S o BH_3 , en la que la base es adenina, y en la que n es 1, 2, 3, 4 ó 5;

5 en el que si la etiqueta es polietilenglicol, la etiqueta tiene la siguiente estructura:



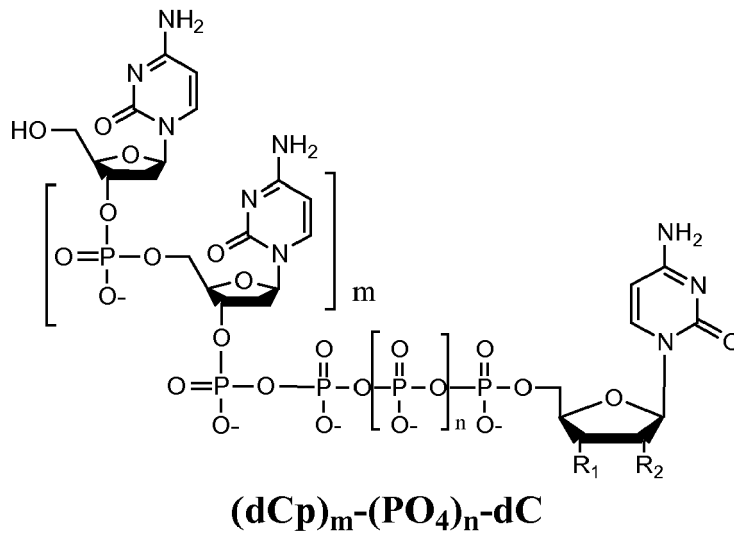
en la que W es un número entero de entre 0 y 100, o

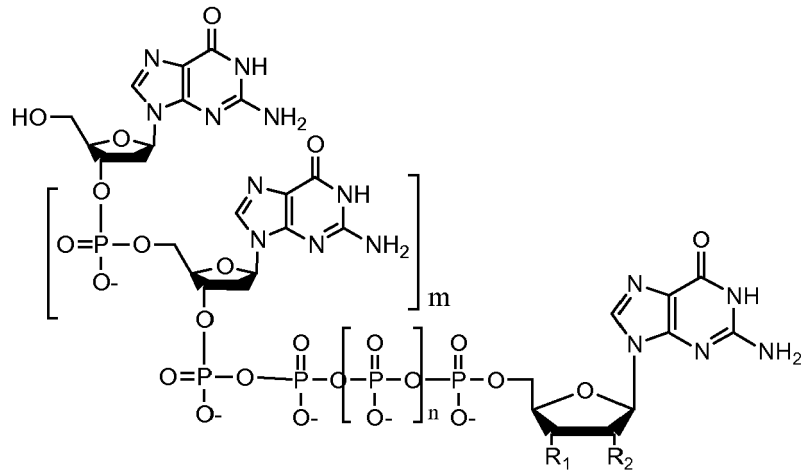
la etiqueta tiene la siguiente estructura:



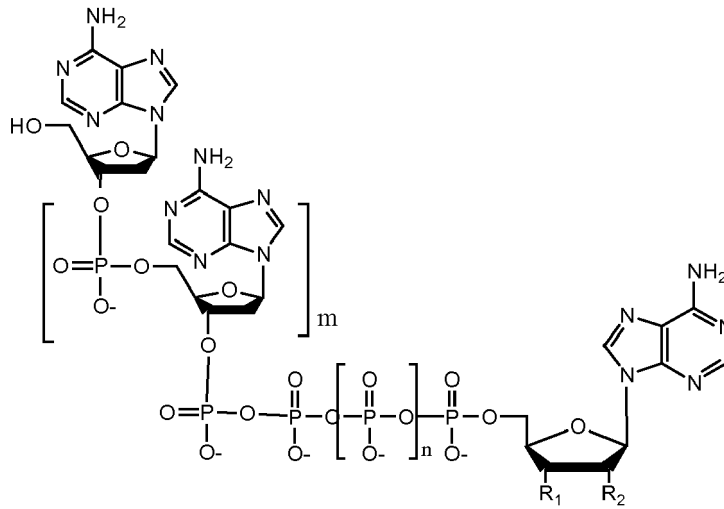
10 en la que R es NH_2 , OH, $COOH$, CHO, SH o N_3 , y W es un número entero de desde 0 hasta 100.

14. Compuesto que tiene la estructura:

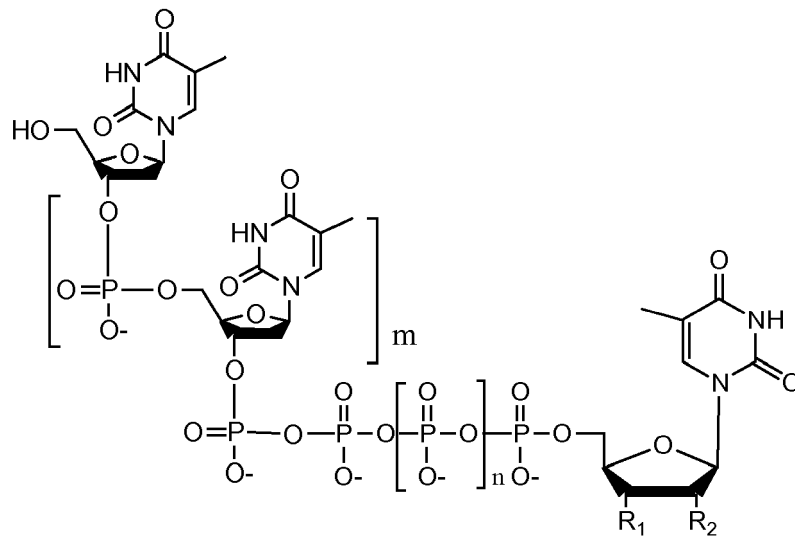




(dGp)_m-(PO₄)_n-dG



(dAp)_m-(PO₄)_n-dA

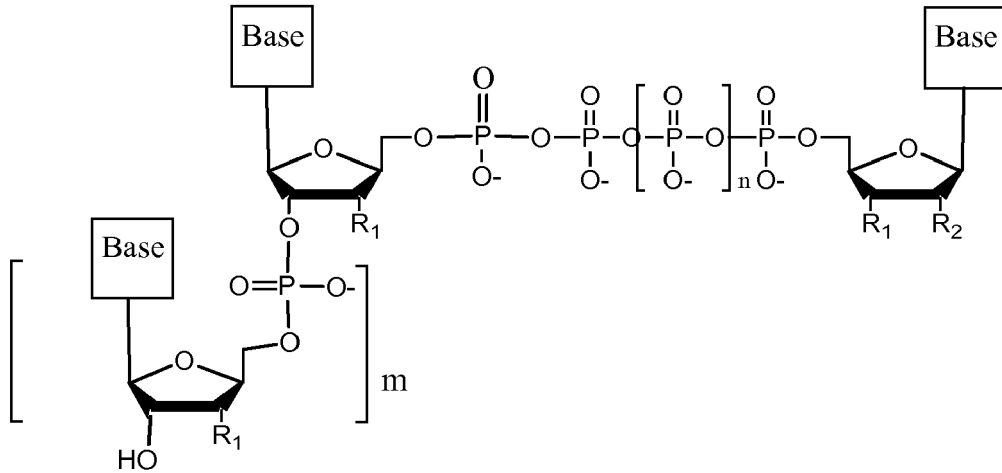


(dTp)_m-(PO₄)_n-dT

en la que en cada estructura n es, independientemente, 1, 2, 3 ó 4, y m es, independientemente, un número

entero de desde 0 hasta 100, un número entero de desde 0 hasta 50, o un número entero de desde 0 hasta 10, y en la que cuando m es 0 el fosfato terminal del dNPP está unido directamente al átomo de O en 3' del nucleósido mostrado en el lado izquierdo de la estructura, en la que R₁ es -OH, o -O-CH₂N₃, y R₂ es H u OH.

5 15. Compuesto que tiene la estructura:



en la que m es un número entero de desde 0 hasta 100, desde 0 hasta 50, o desde 0 hasta 10, en la que R₁ es OH, en la que R₂ es H u OH, en la que n es 1, 2, 3 ó 4, y en la que el compuesto comprende un único tipo de base, y en la que la base es guanina, citosina, uracilo o timina o un derivado de cada una de las mismas.

10

Figura 1

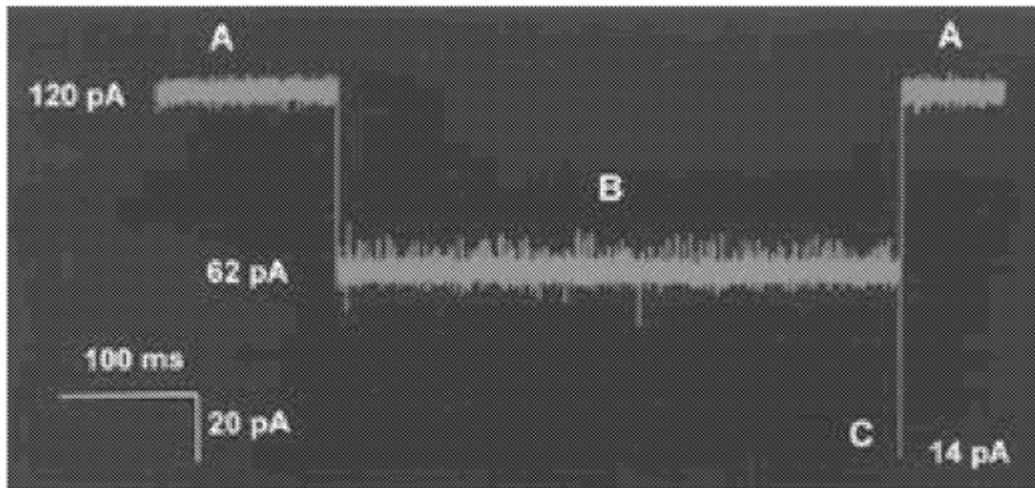
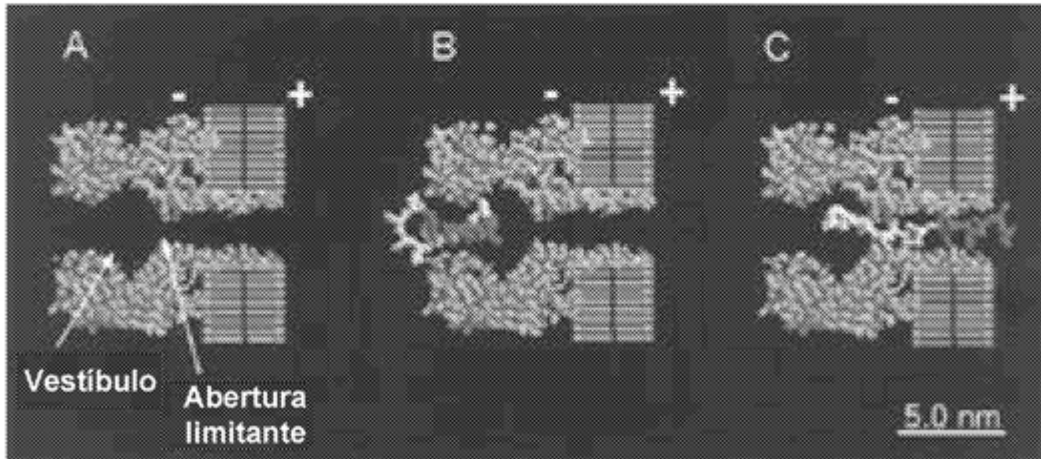


Figure 2

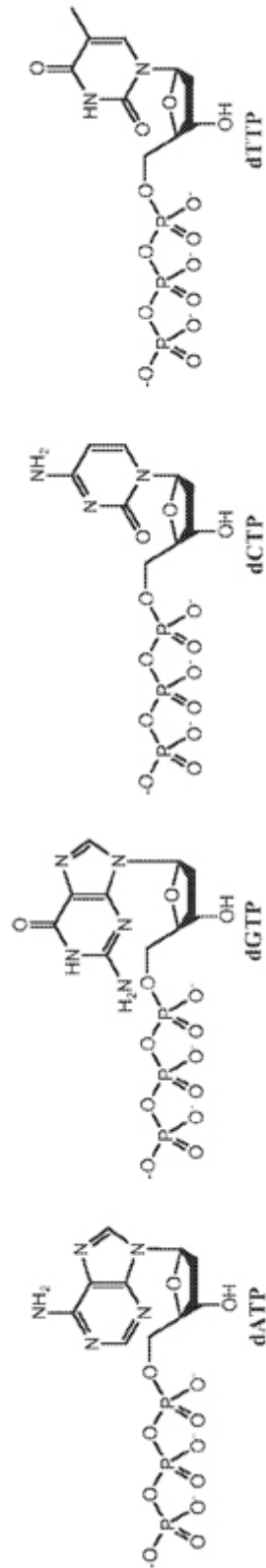


Figura 3

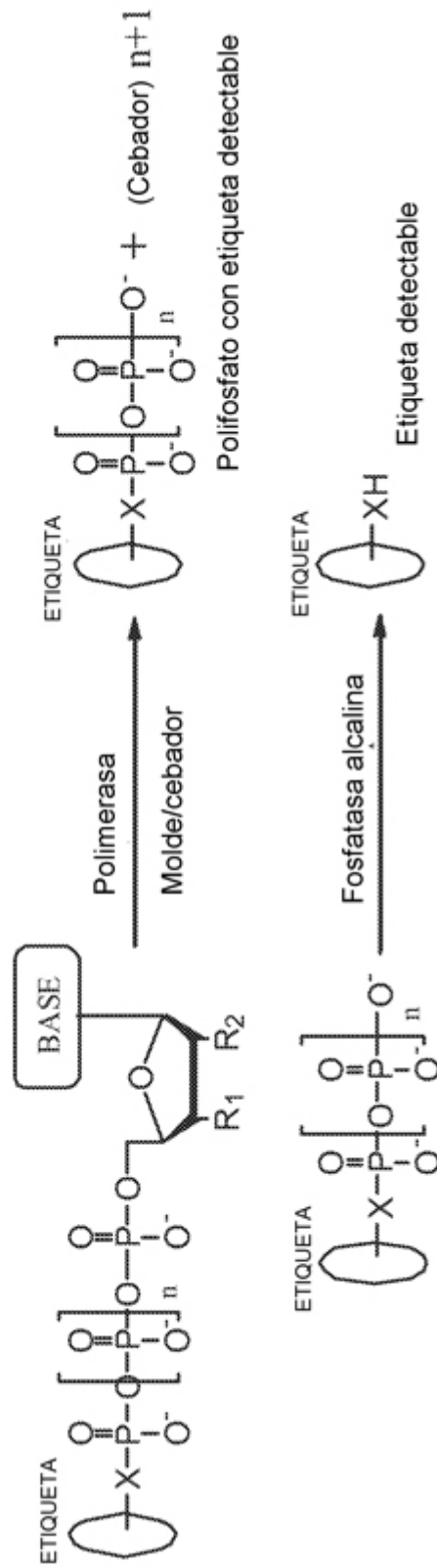


Figura 4

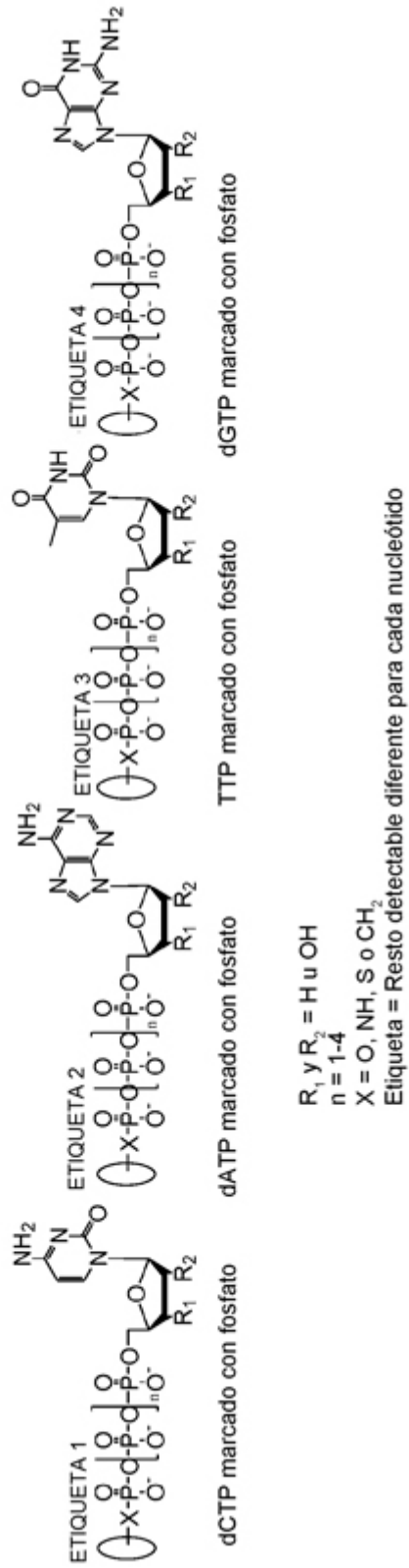
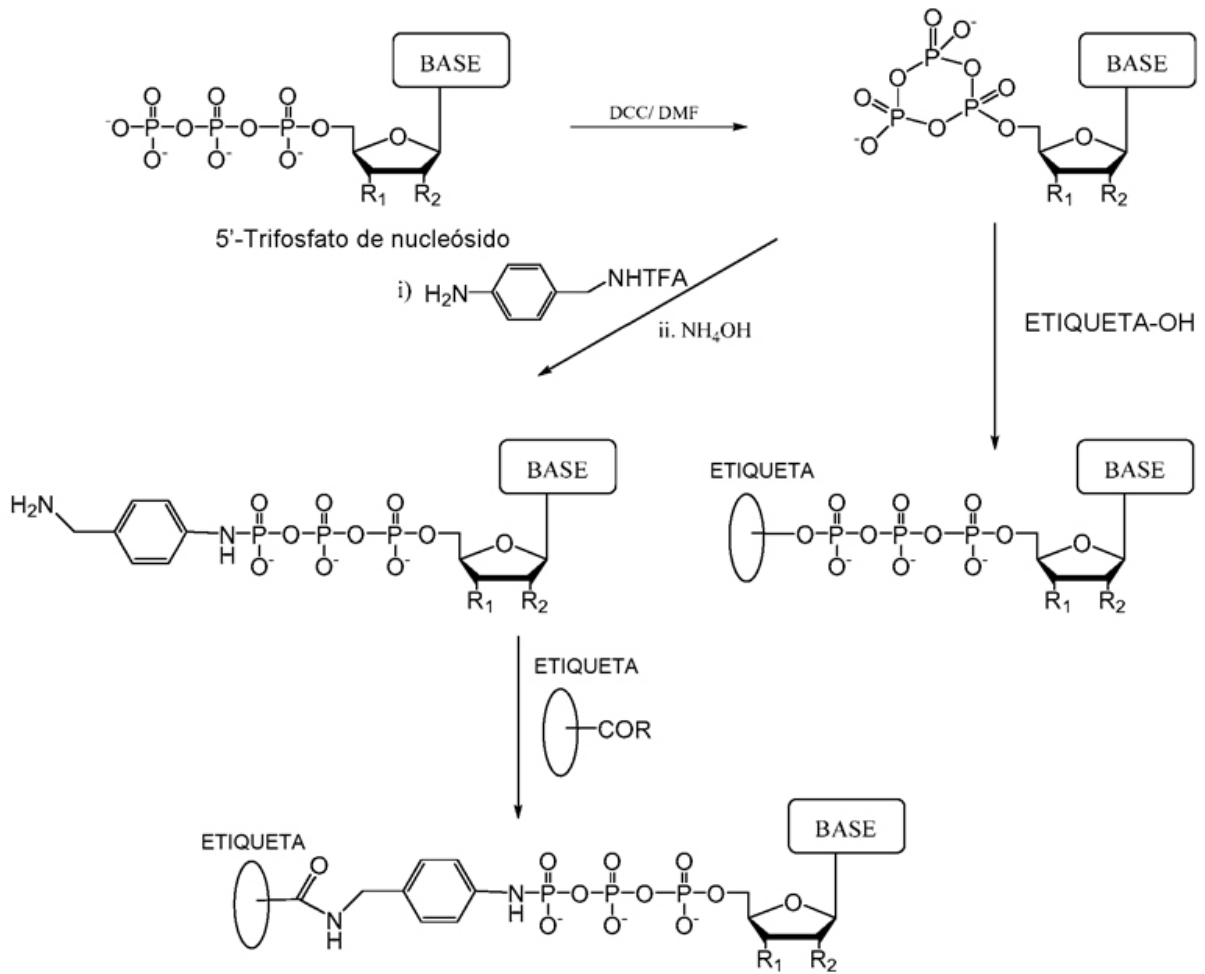


Figura 5

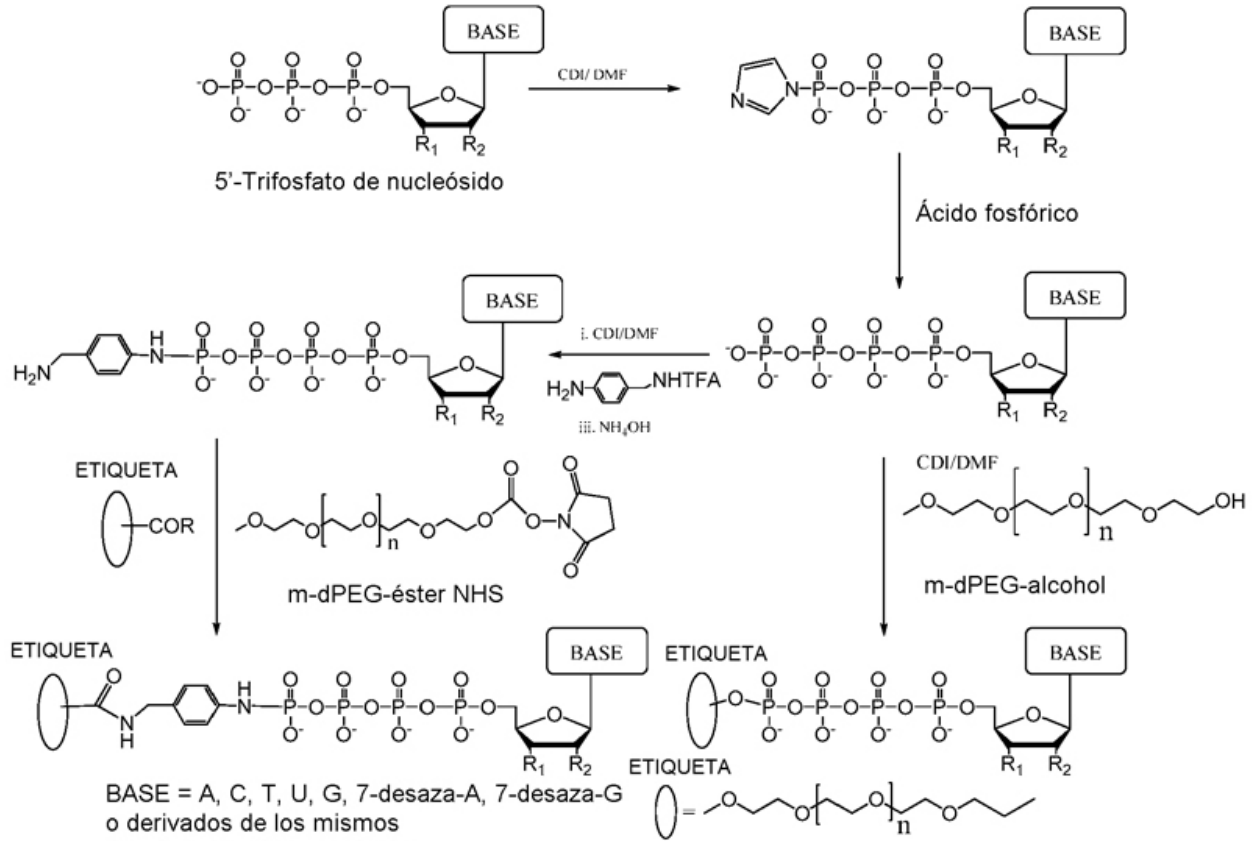


BASE = A, C, T, U, G, 7-desaza-A, 7-desaza-G o derivados de los mismos
 R_1 y R_2 = H u OH

ETIQUETA-OH = resto detectable en forma hidroxilo, tal como, alcoholes alifáticos o aromáticos

ETIQUETA-COR = resto detectable para reaccionar con el grupo amino, tal como ésteres NHS, cloruro de ácido etc

Figura 6



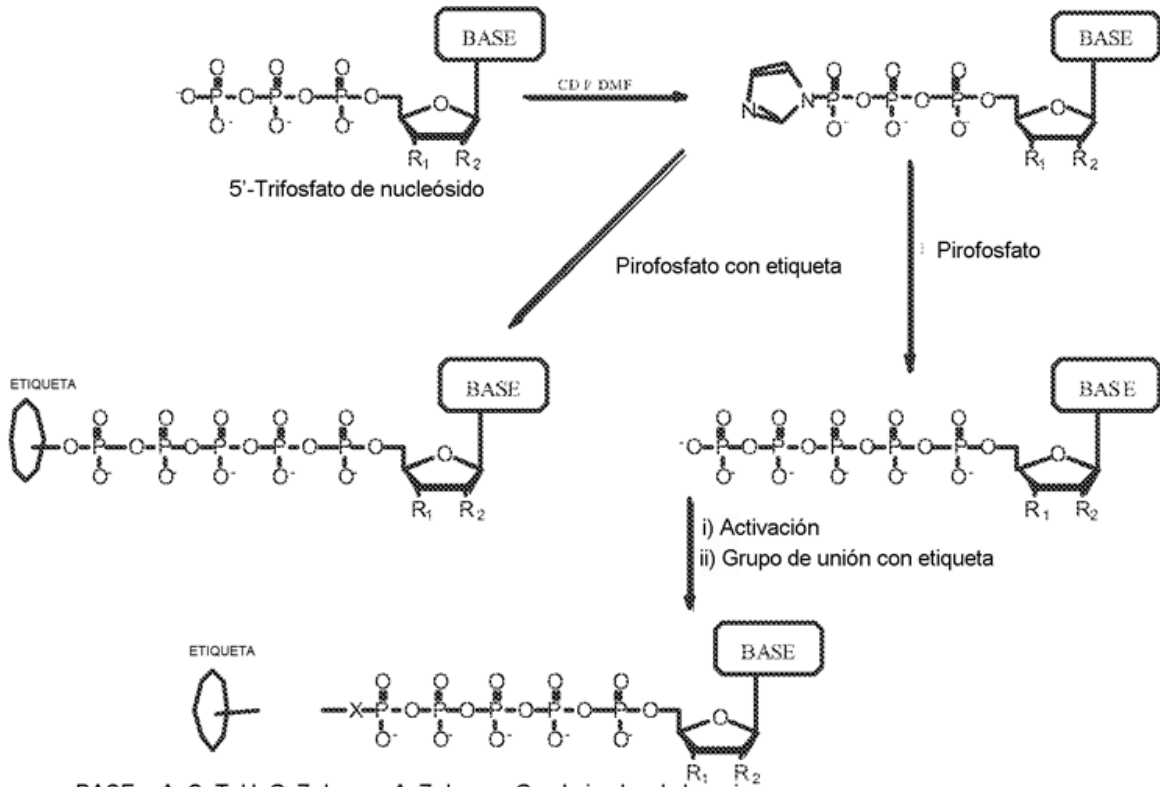
BASE = A, C, T, U, G, 7-desaza-A, 7-desaza-G
o derivados de los mismos

R₁ y R₂ = H u OH; n = 1-100

ETIQUETA = resto detectable , tal como, alcoholes alifáticos o aromáticos, aminoácidos, colorante, hidratos de carbono, oligos etc

ETIQUETA-COR = resto detectable para reaccionar con el grupo amino, tal como ésteres NHS, cloruro de ácido etc

Figura 7



BASE = A, C, T, U, G, 7-desaza-A, 7-desaza-G o derivados de los mismos

R₁ y R₂ = H u OH

X = O, NH, S o CH₂

ETIQUETA = resto detectable, tal como, alcoholes alifáticos o aromáticos, aminoácidos, colorante, hidratos de carbono, oligos etc

ETIQUETA-COR = resto detectable para reaccionar con el grupo amino, tal como ésteres NHS, cloruro de ácido etc

Figura 8

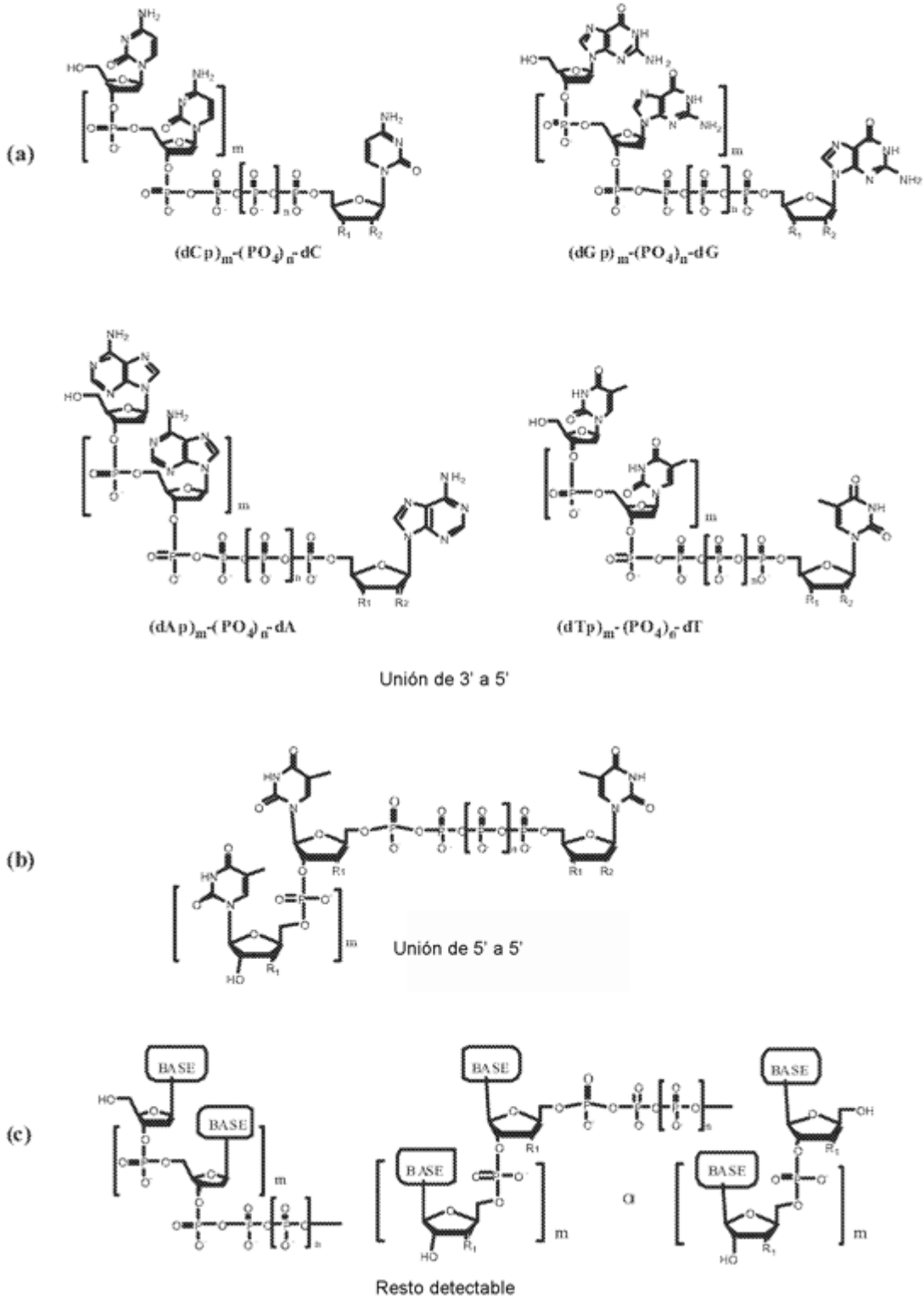
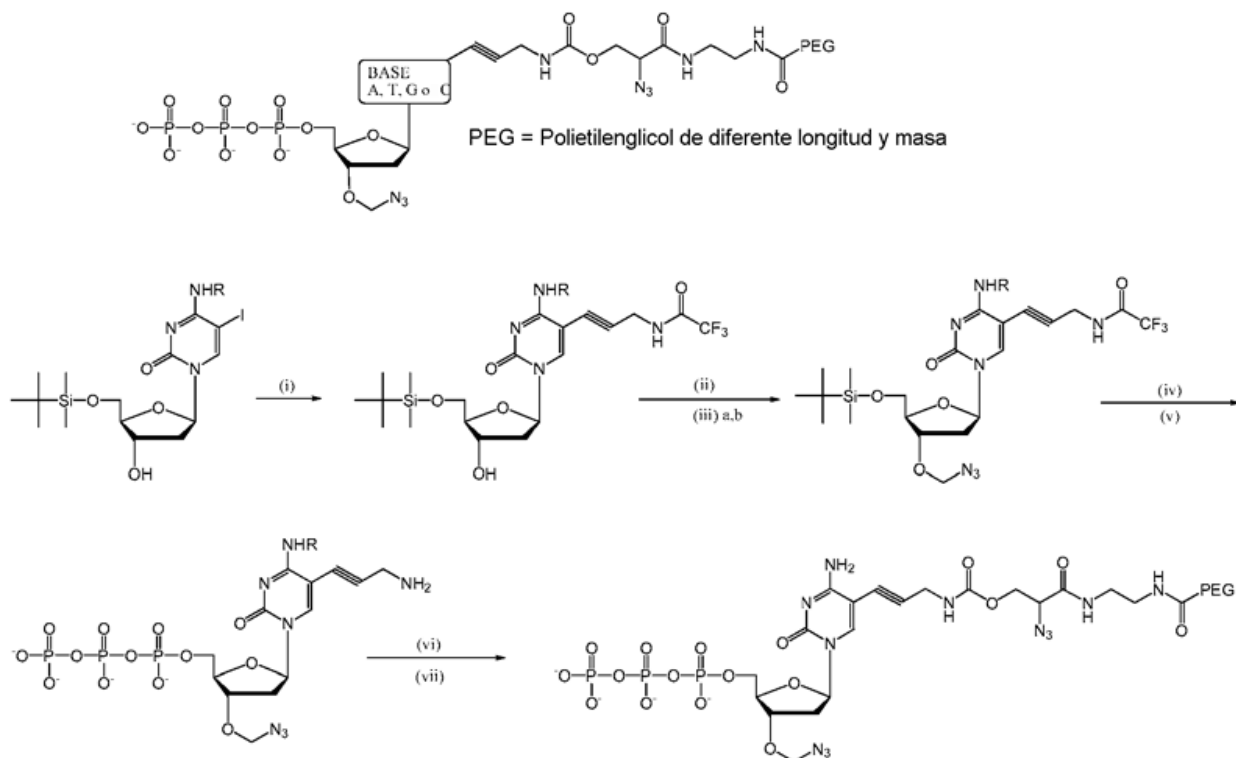


Figura 9(A)



i) N-propargiltrifluoroacetamida, $\text{Pd}(\text{PPh}_3)_4$, CuI , Et_3N , DMF , t.a. 10 h; ii) DMSO , Ac_2O , AcOH , t.a. 48 h; iii) ciclohexeno, SO_2Cl_2 , CH_2Cl_2 , 0°C , t.a. 1 h; iiib) NaN_3 , DNF , t.a. 4 h; iv) TBAF , v) POCl_3 , $(\text{MeO})_3\text{PO}$, $(\text{Bu}_3\text{NH})_2\text{H}_2\text{P}_2\text{O}_7$, TEAB , Bu_3N , NH_4OH ; vi) $\text{ClC}(\text{O})\text{OCH}_2\text{CHN}_3\text{CONHCH}_2\text{CH}_2\text{-NHCO-PEG}$; viii) $\text{NH}_4\text{OH-MeOH}$

Figura 9(B)

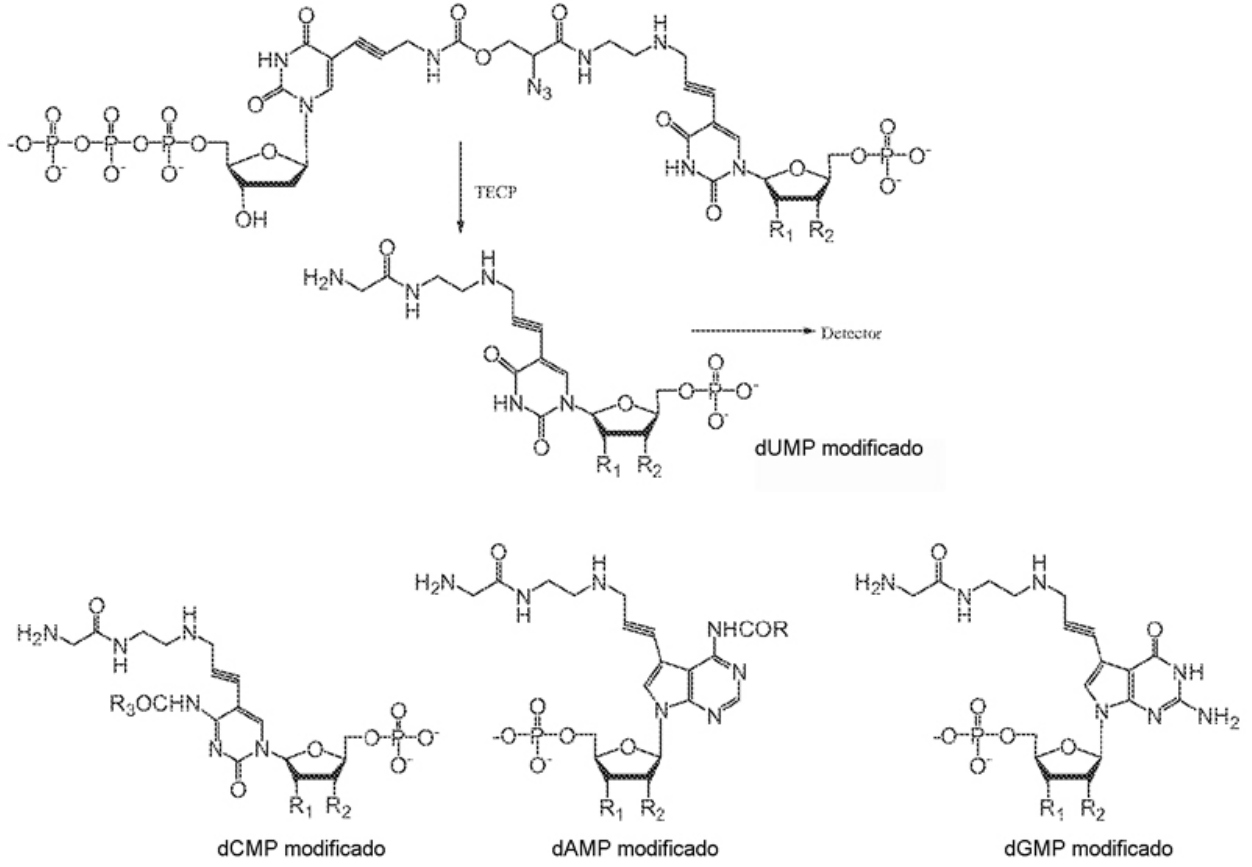


Figura 10

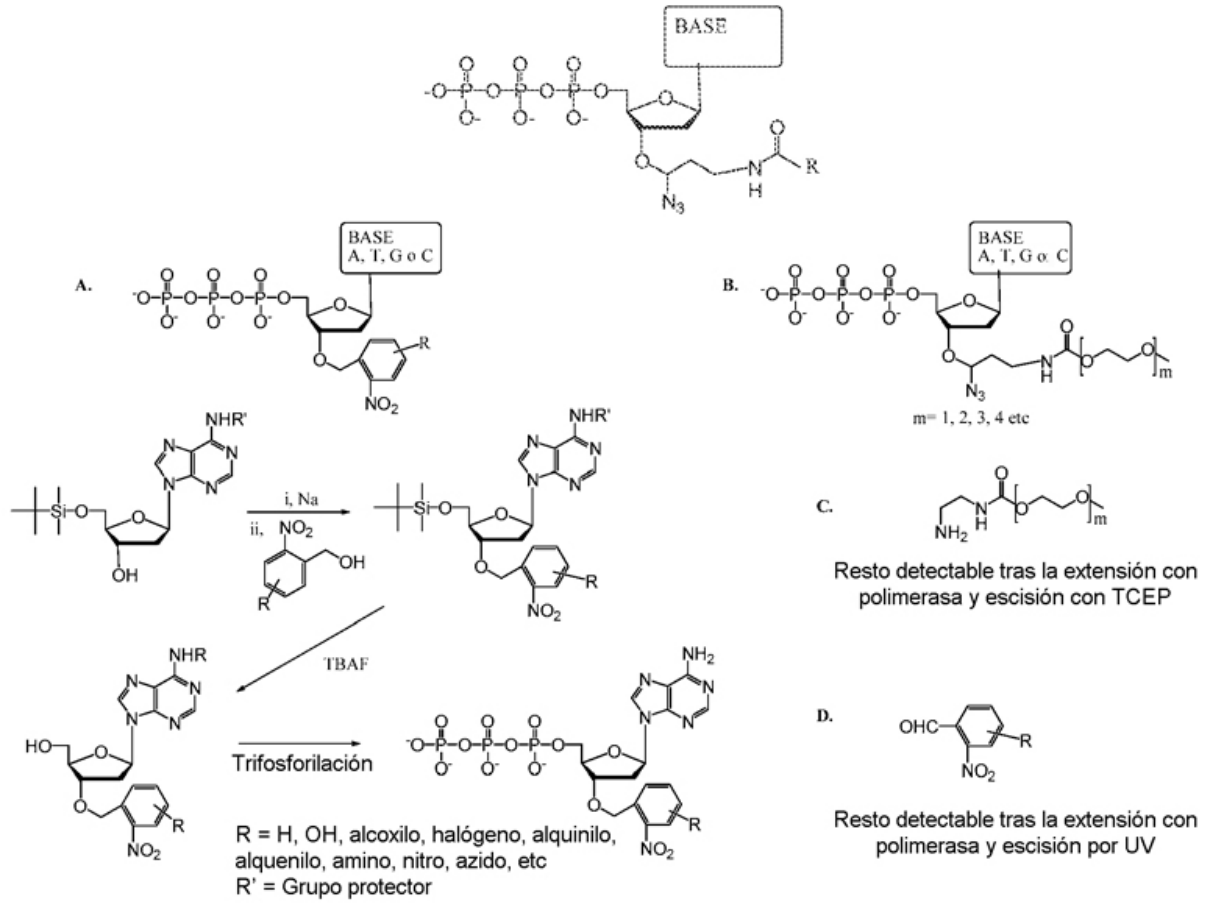


Figura 11

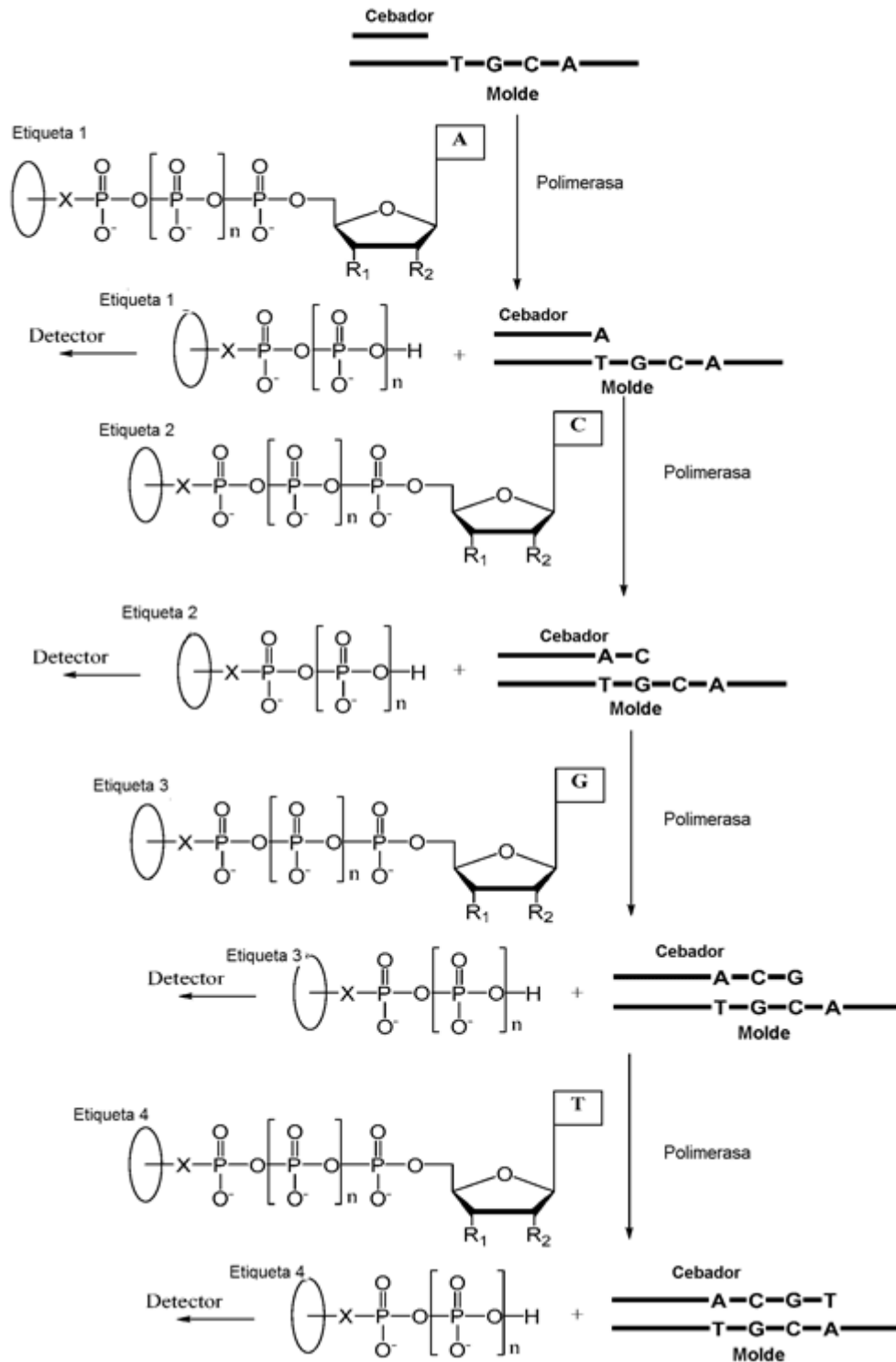


Figura 12

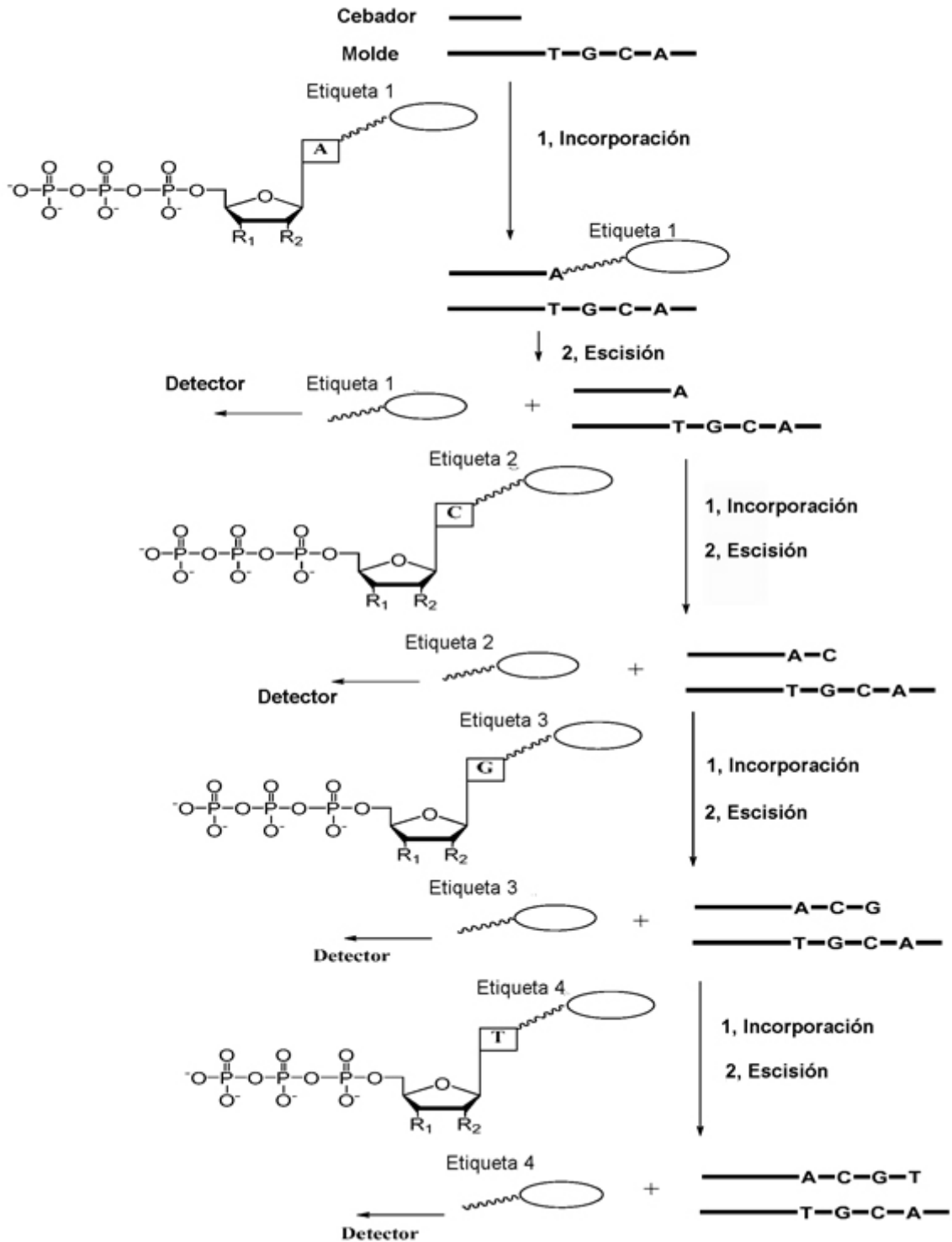


Figura 13

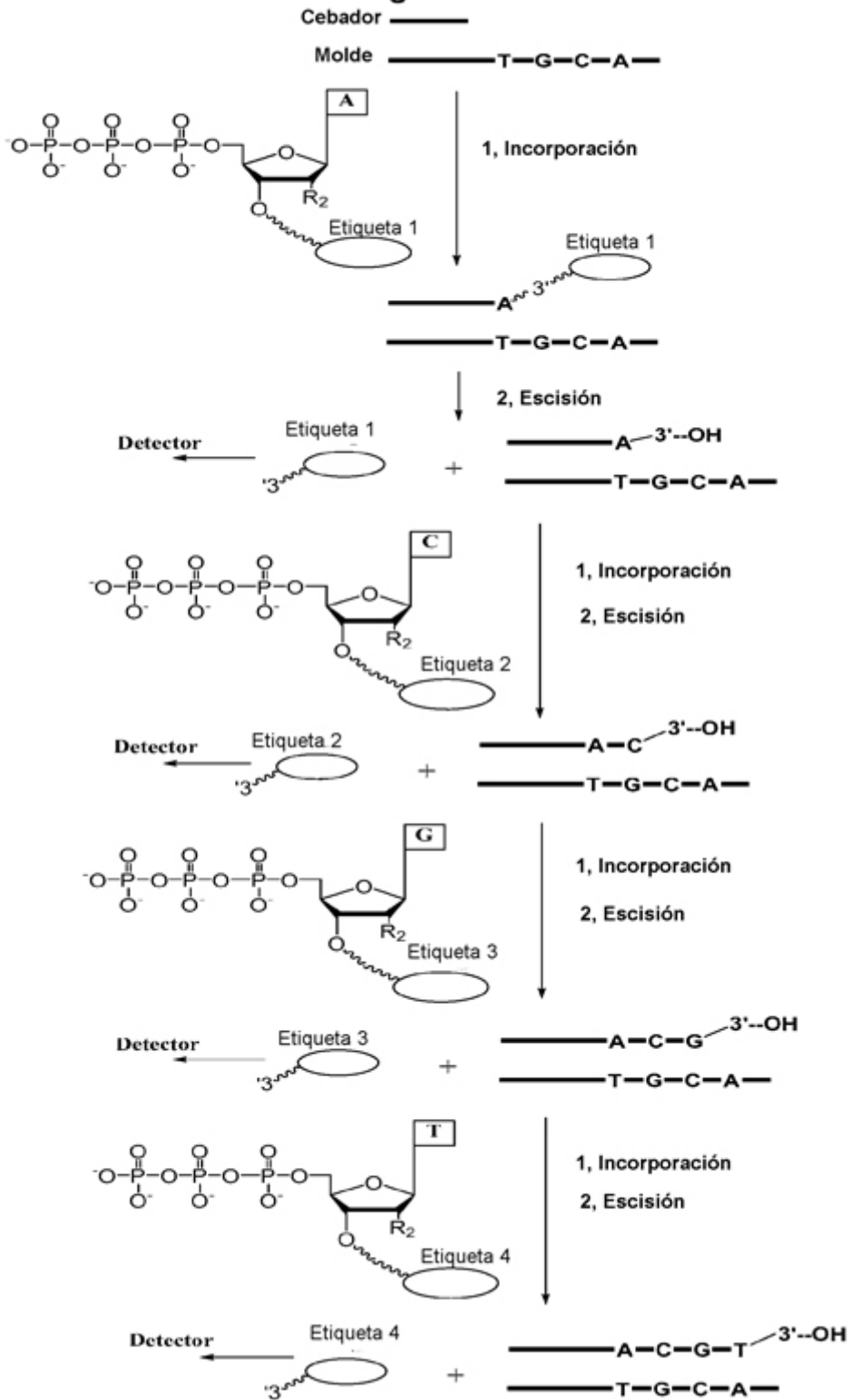


Figura 14

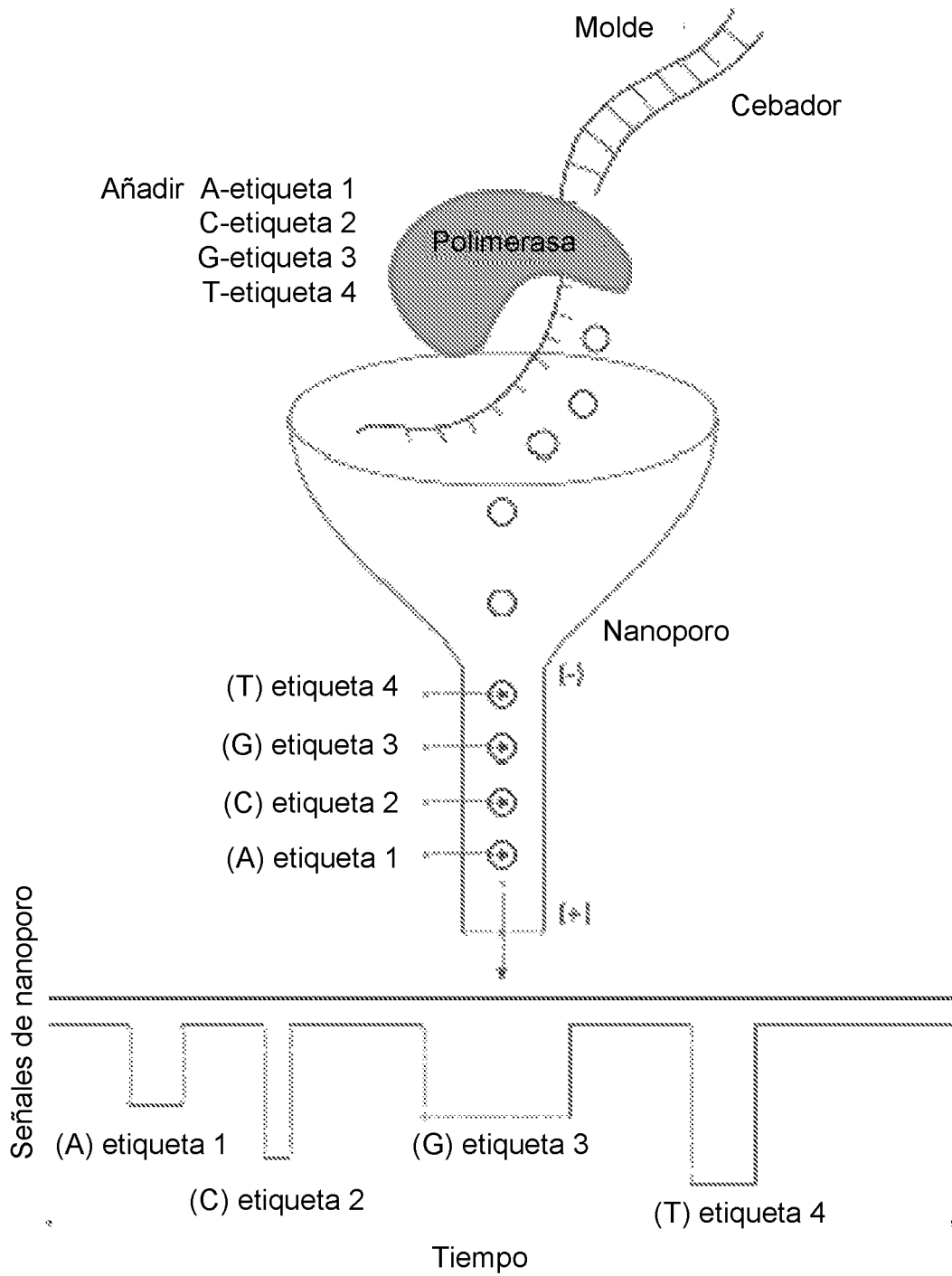


Figura 15

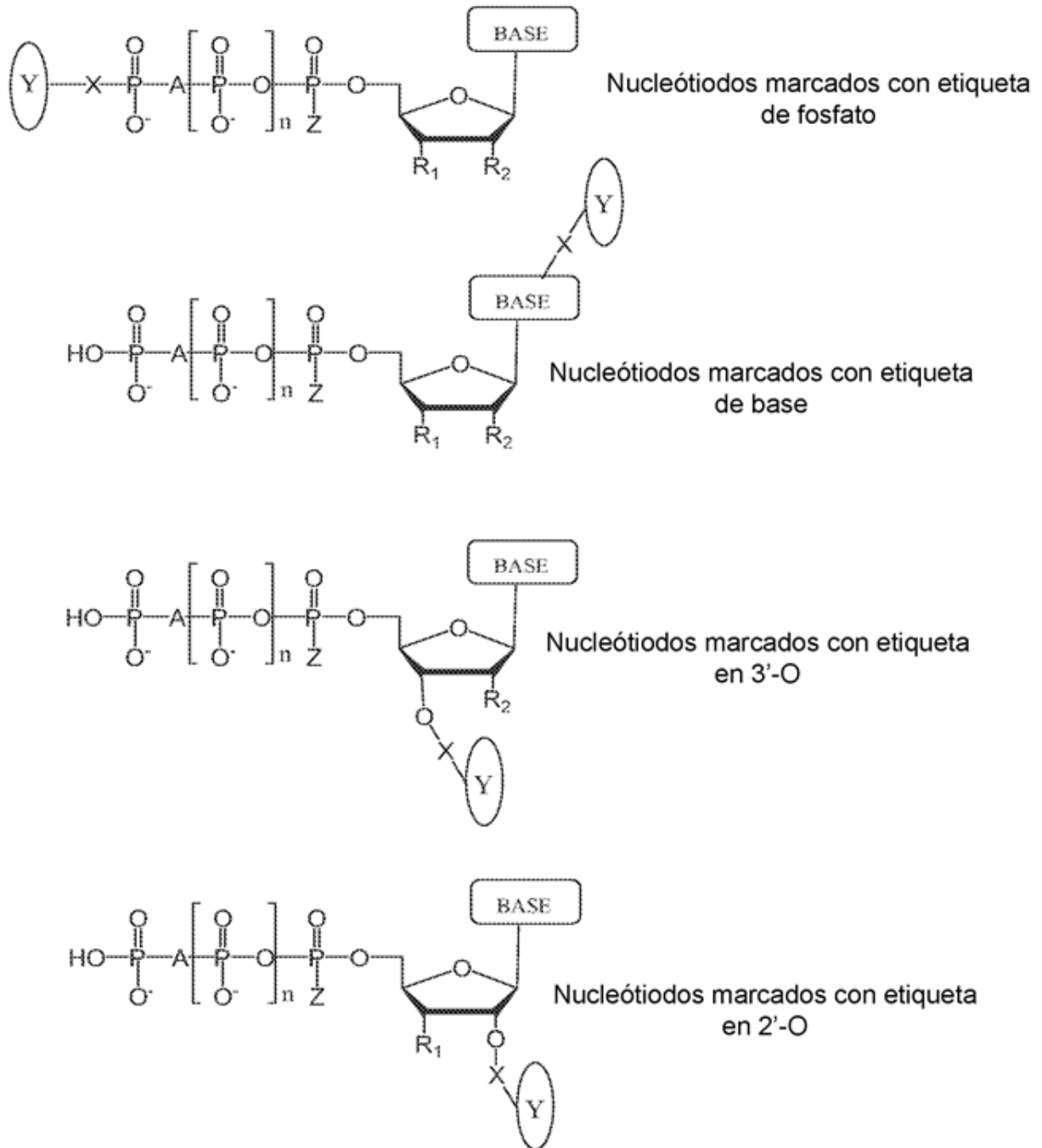
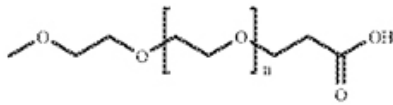
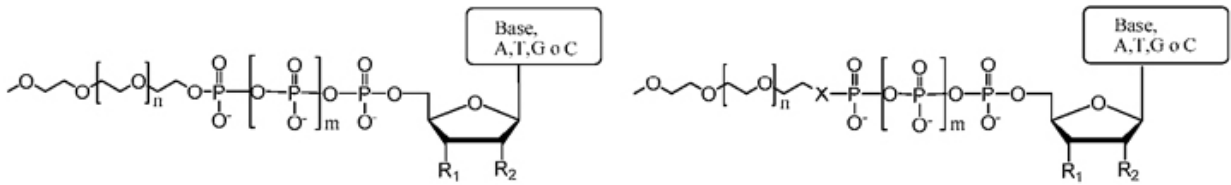
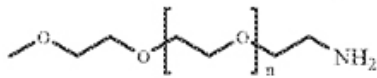


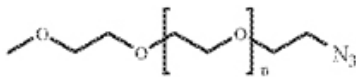
Figura 16



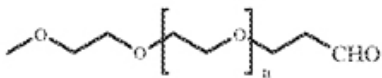
Hacer reaccionar con grupo amino



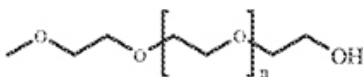
Hacer reaccionar con aldehídos o ácidos activados



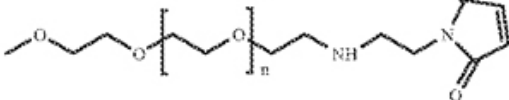
Hacer reaccionar con alquinos



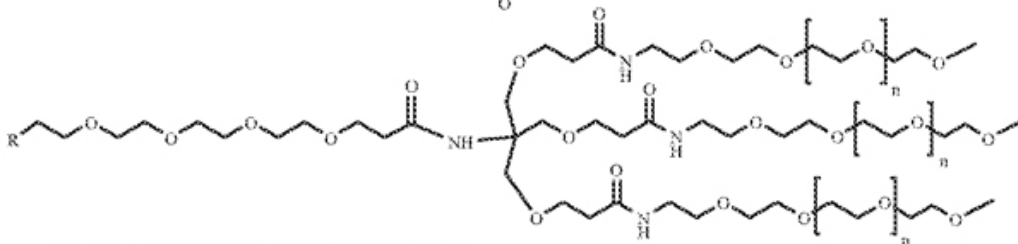
Hacer reaccionar con grupo amino



Hacer reaccionar con fosfatos activados



Hacer reaccionar con tioles



R= NH₂, OH, COOH, COOR', CHO, SH, N₃

Figura 17

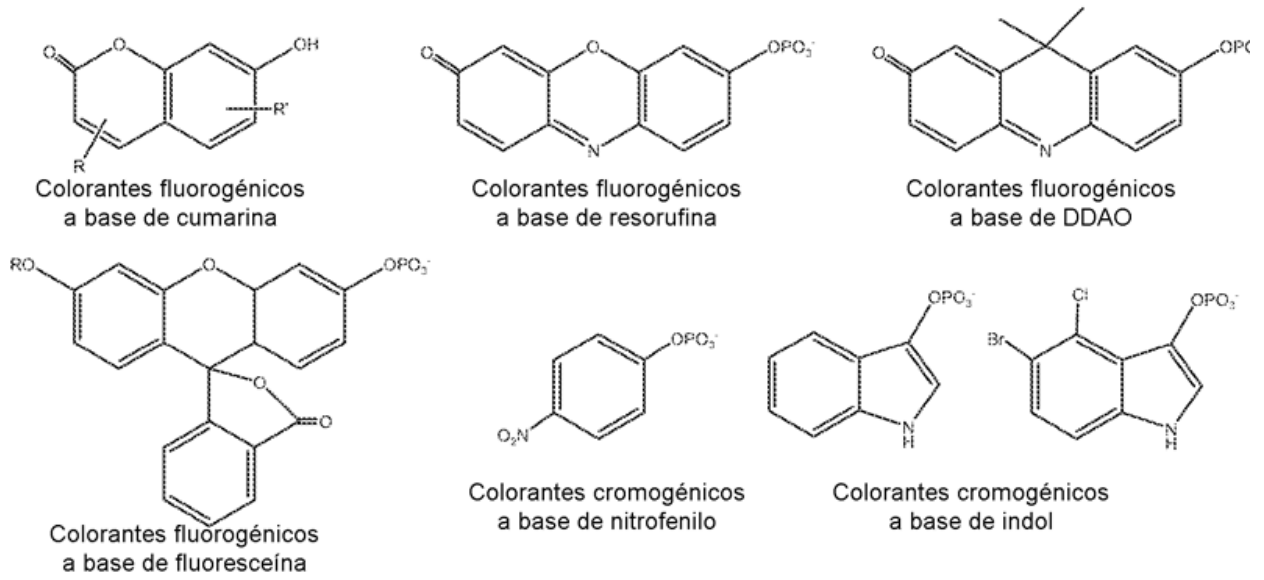


Figura 18

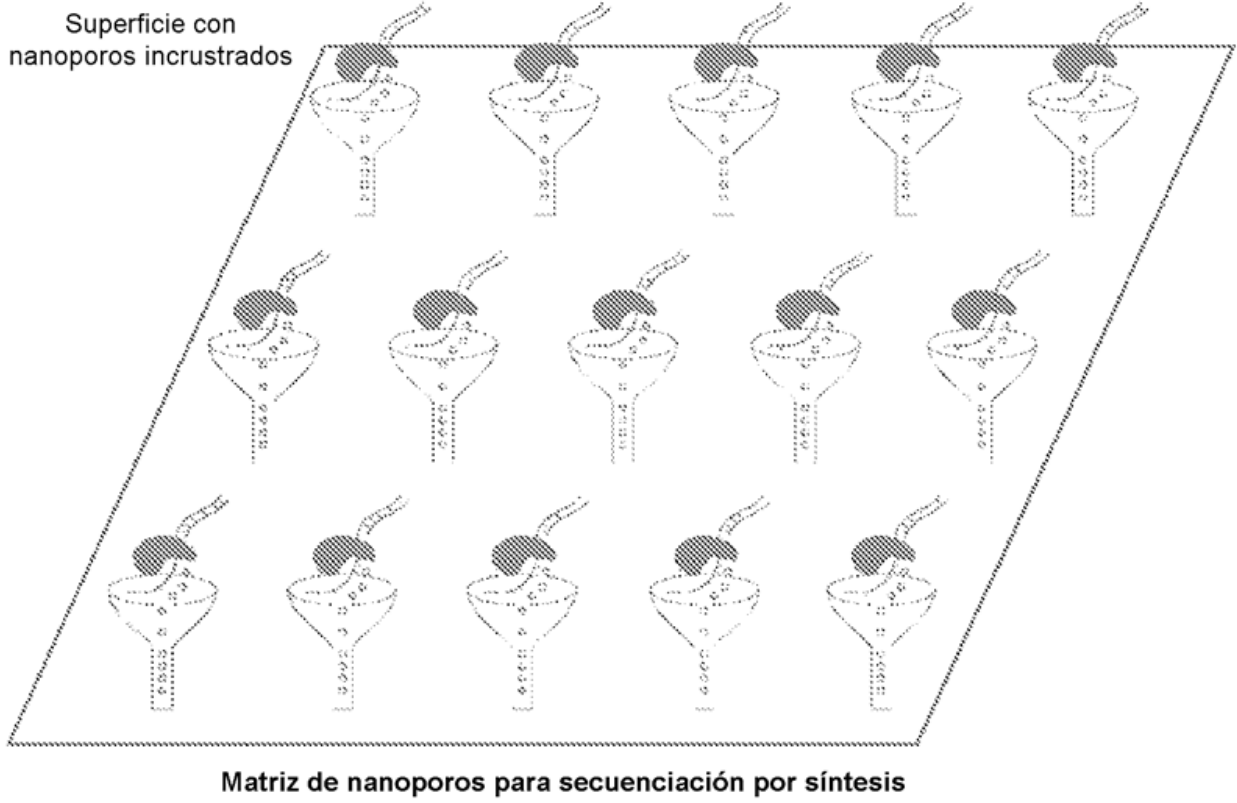
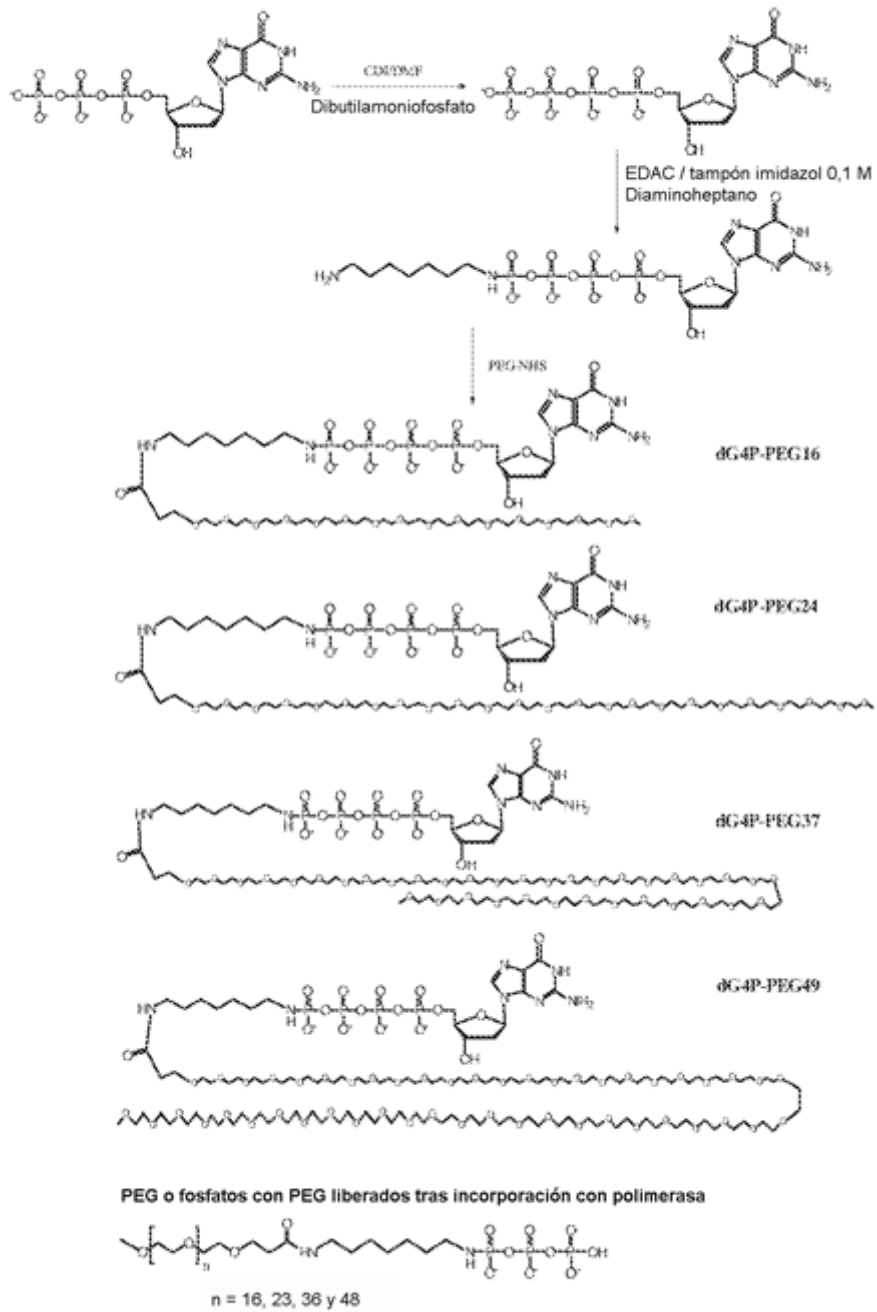
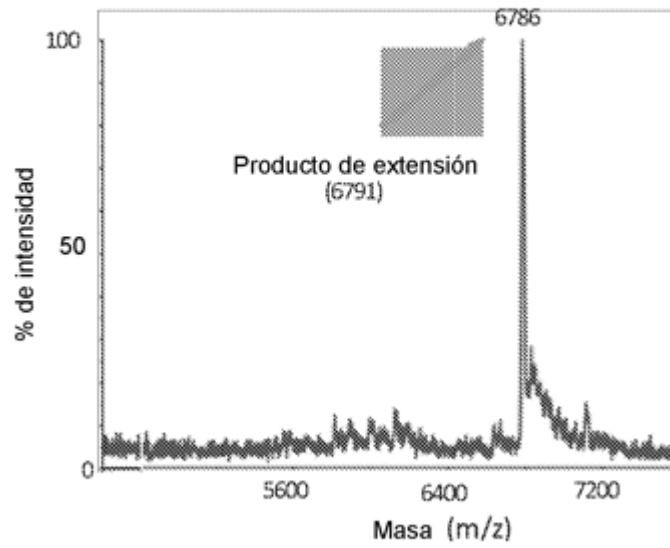


Figura 19

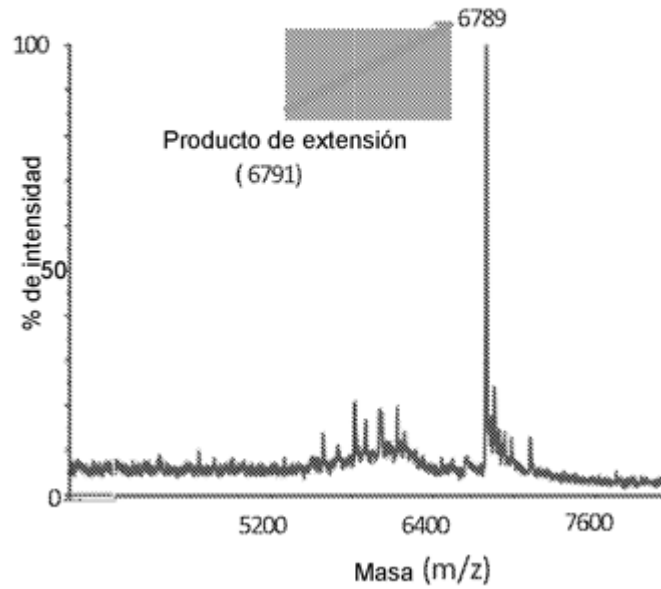


Síntesis de nucleótidos marcados con etiqueta de PEG de primera generación

Figura 20



Reacción de extensión con dG4P-PEG24



Reacción de extensión con dG4P - PEG37

Figura 21

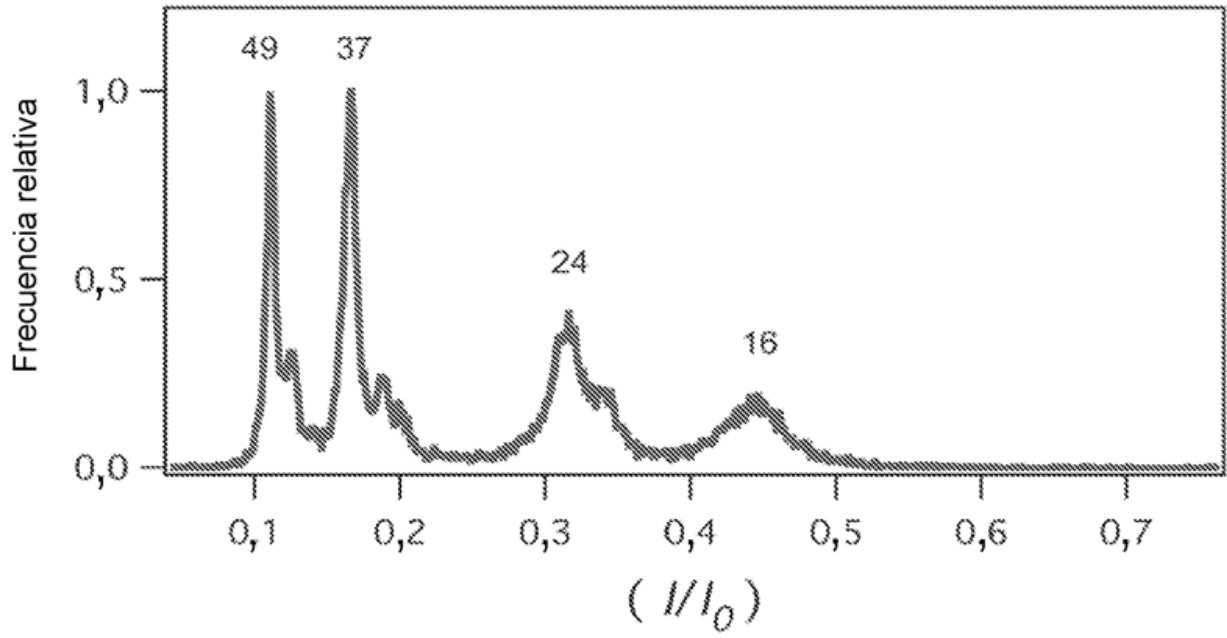
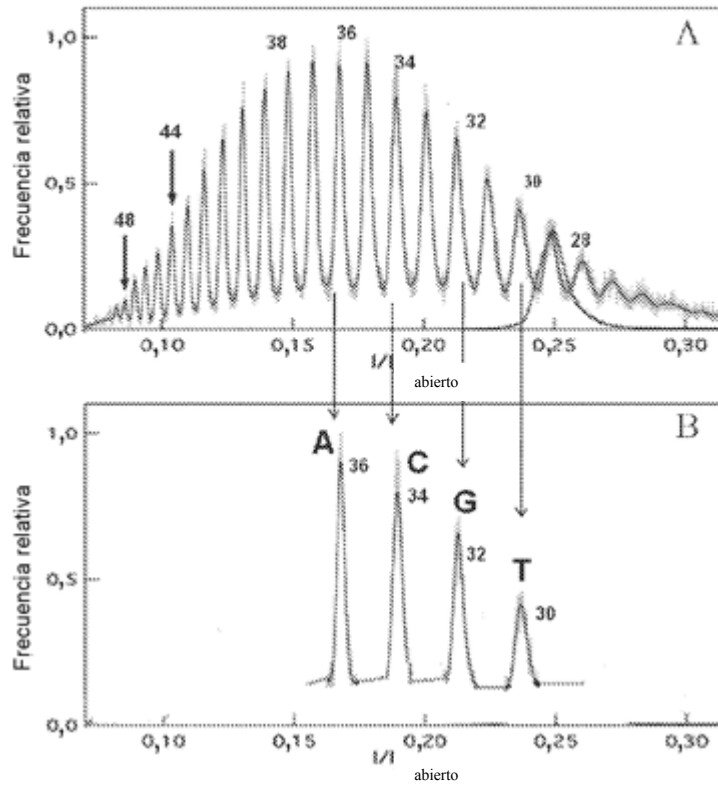
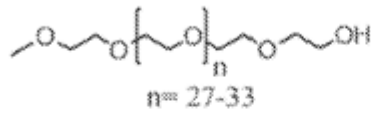


Figura 22

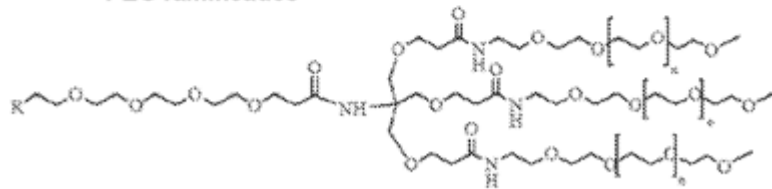


PEG lineales



- $n=27$; dPEG₃₀: 30 unidades de etilenglicol **Para T**
- $n=29$; dPEG₃₂: 32 unidades de etilenglicol **Para G**
- $n=31$; dPEG₃₄: 34 unidades de etilenglicol **Para C**
- $n=33$; dPEG₃₆: 36 unidades de etilenglicol **Para A**

PEG ramificados



R = NH₂, OH, COOH, COOR', CHO, SH, N₃

Figura 23

Síntesis de trifosfatos con PEG con carga

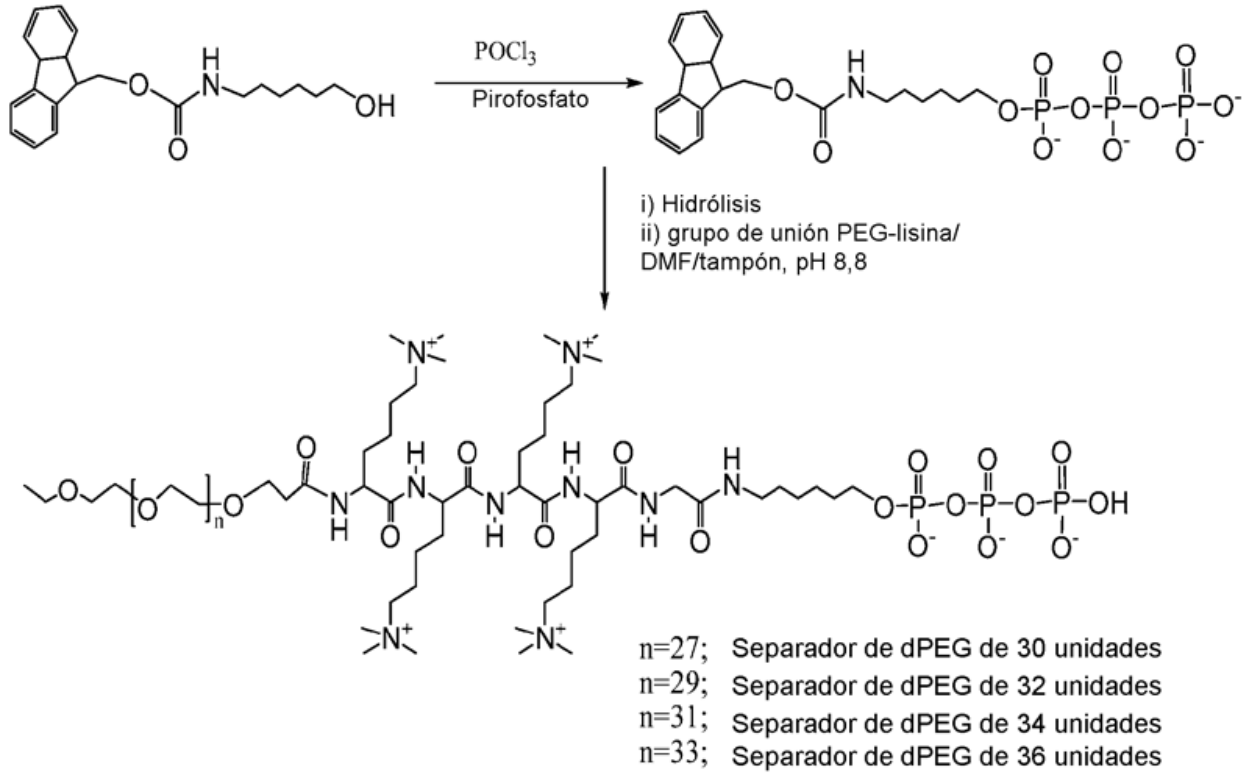


Figura 24

Síntesis de 5'-trifosfatos de nucleósidos marcados con etiqueta de fosfato

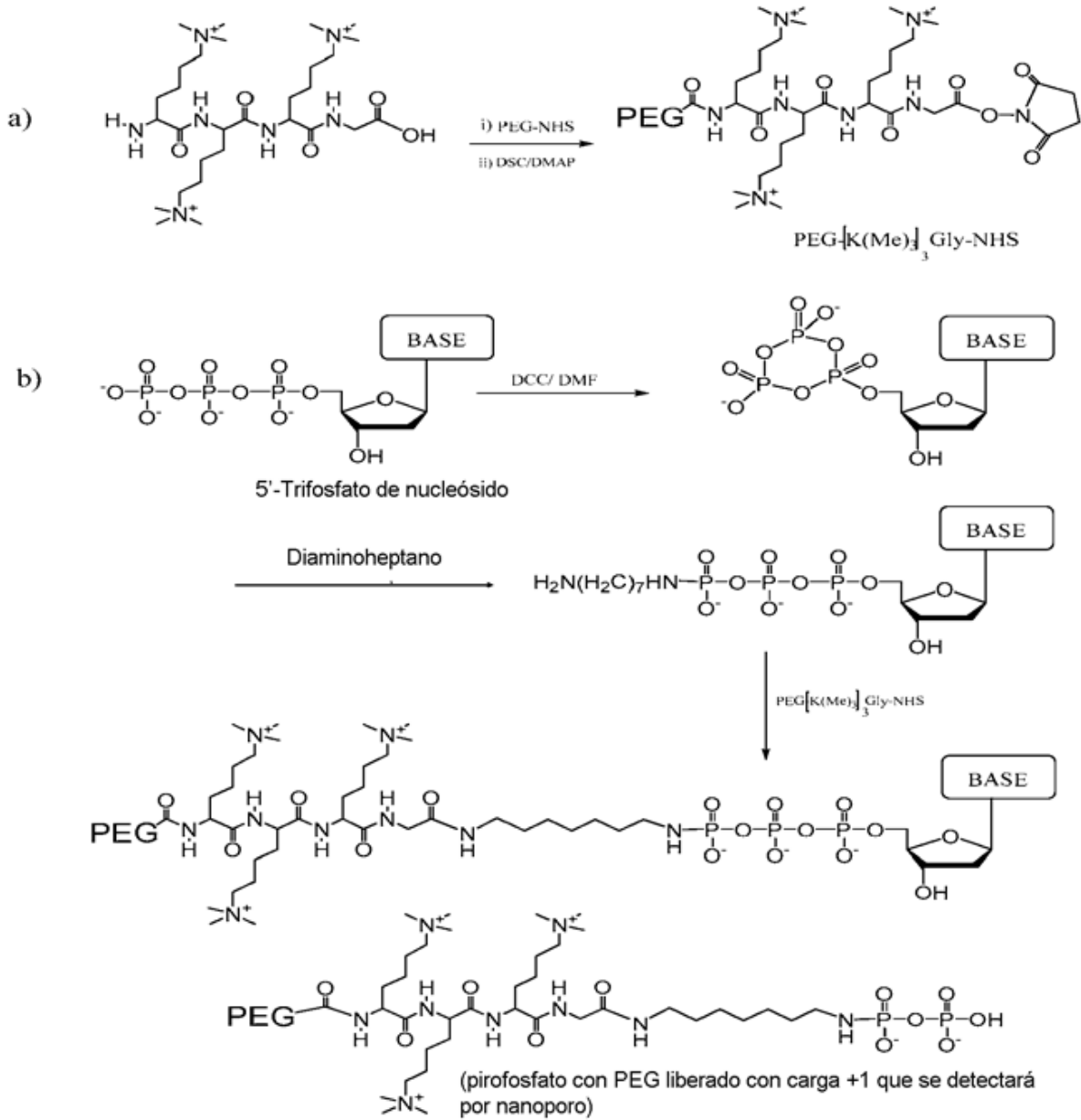


Figura 25

Síntesis de 5'-tetrafosfatos de nucleósidos marcados con etiqueta de fosfato

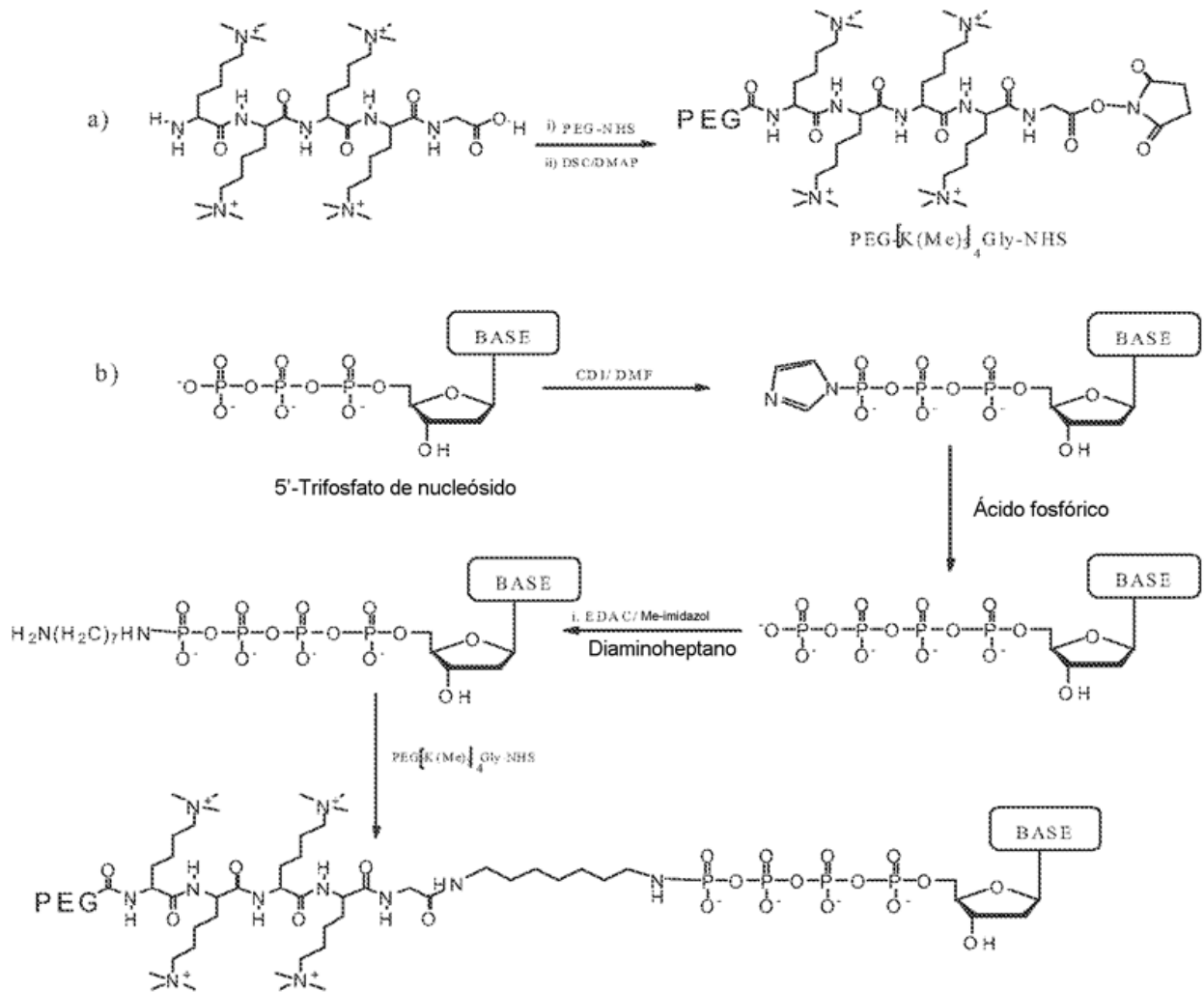


Figura 26

Síntesis de 5'-pentafosfatos de nucleósidos marcados con etiqueta de fosfato terminal

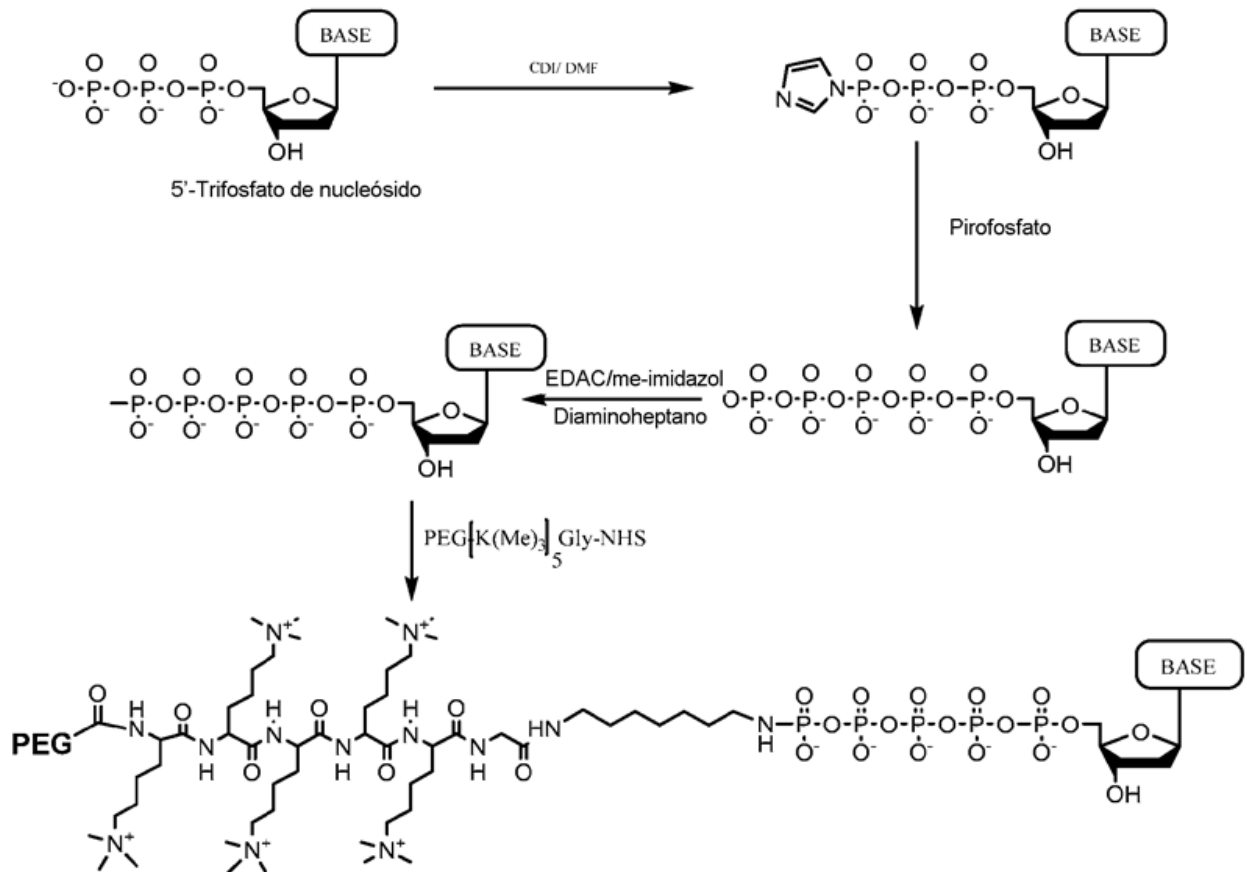


Figura 28

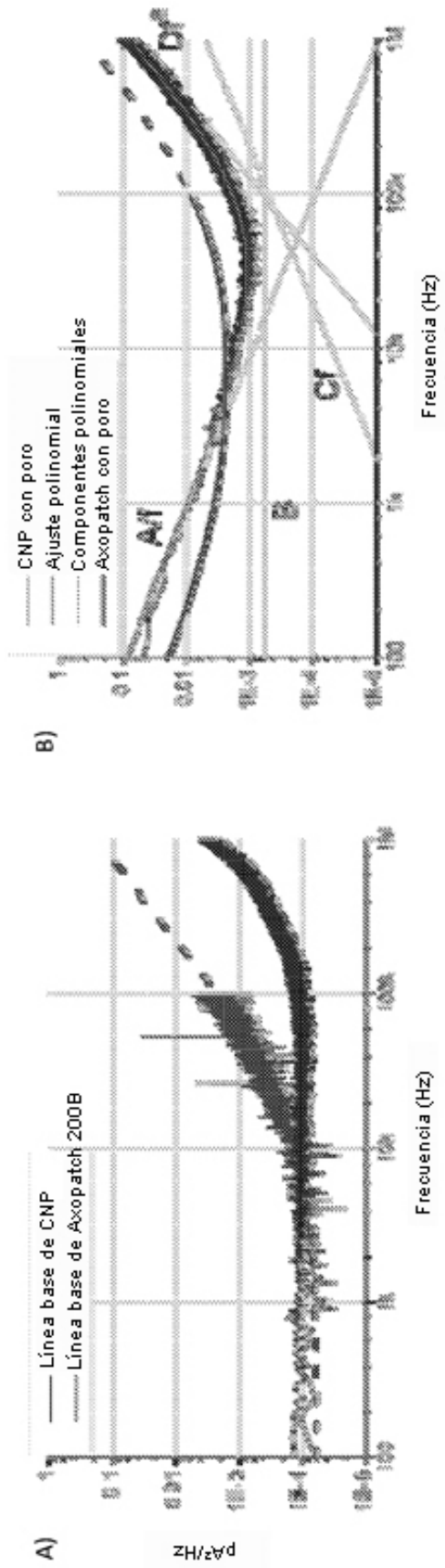


Figura 29

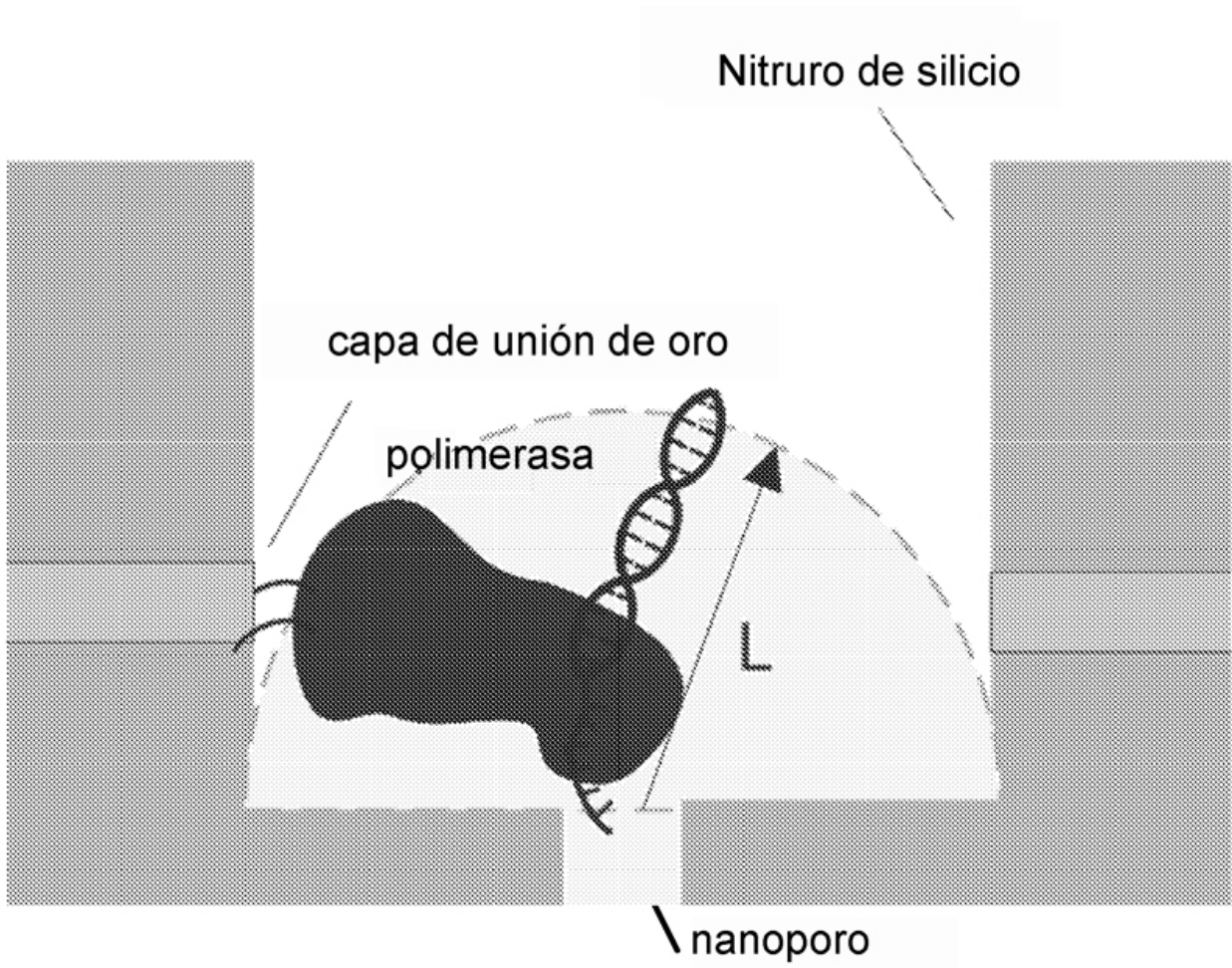


Figura 30

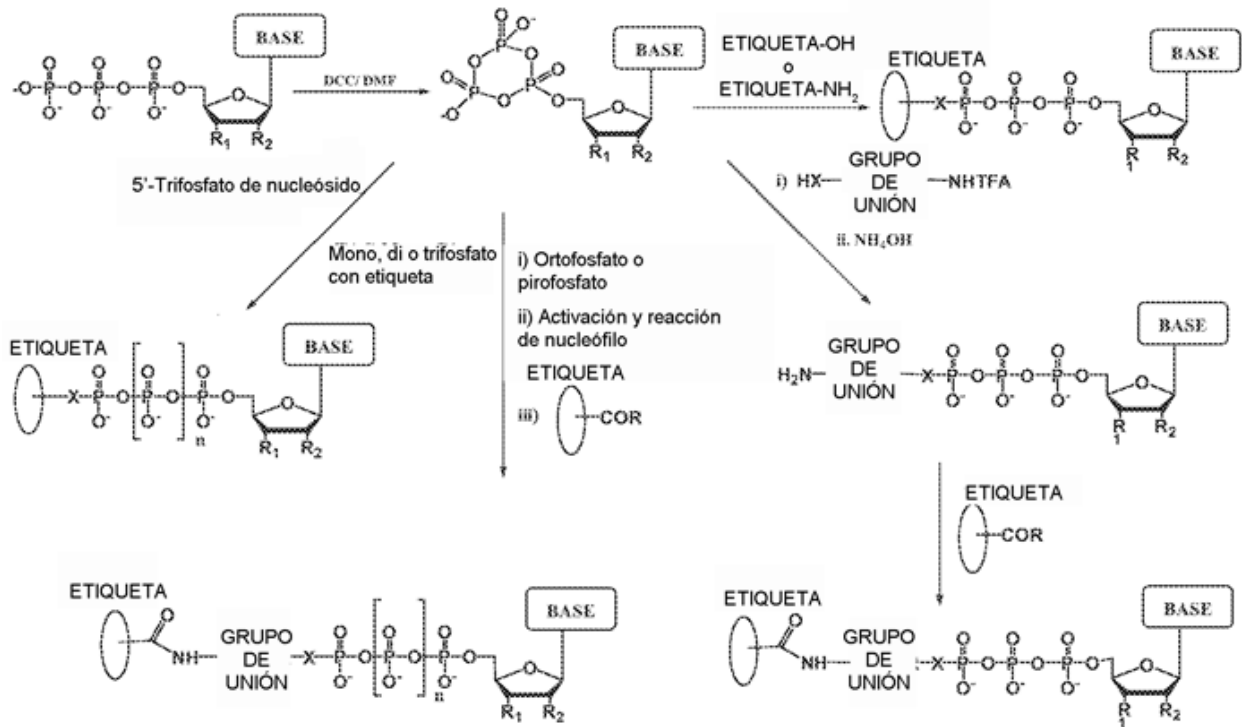
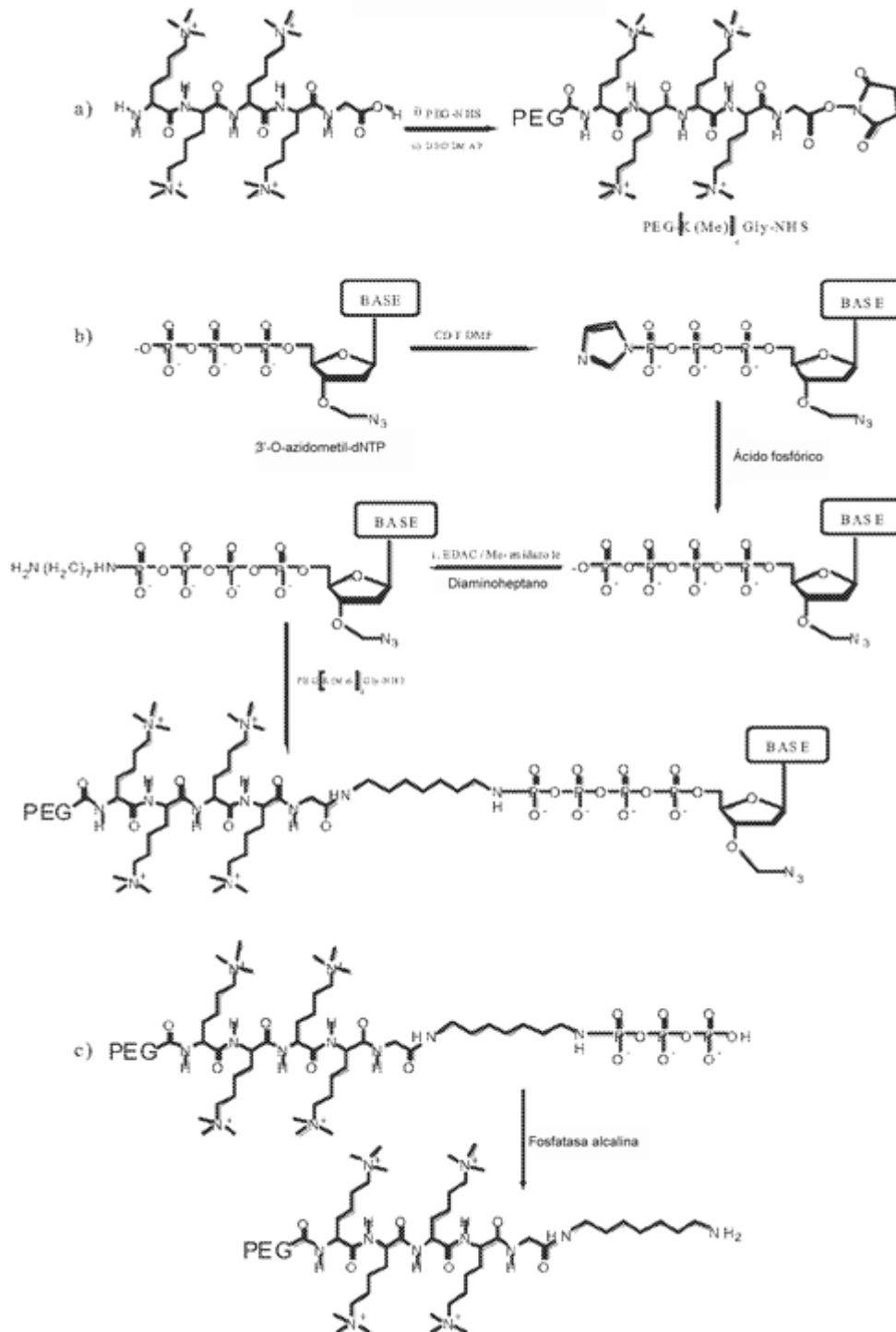


Figura 31



Síntesis de PEG-nucleótidos bloqueados en 3'-O

Figura 32

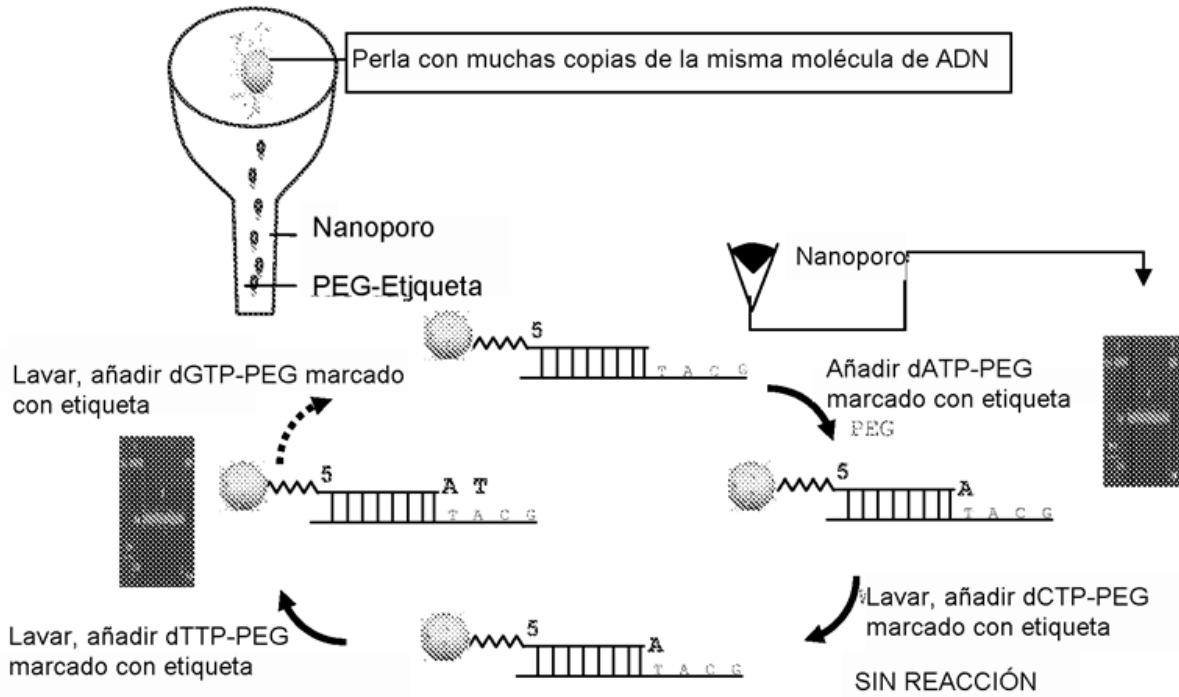


Figura 33

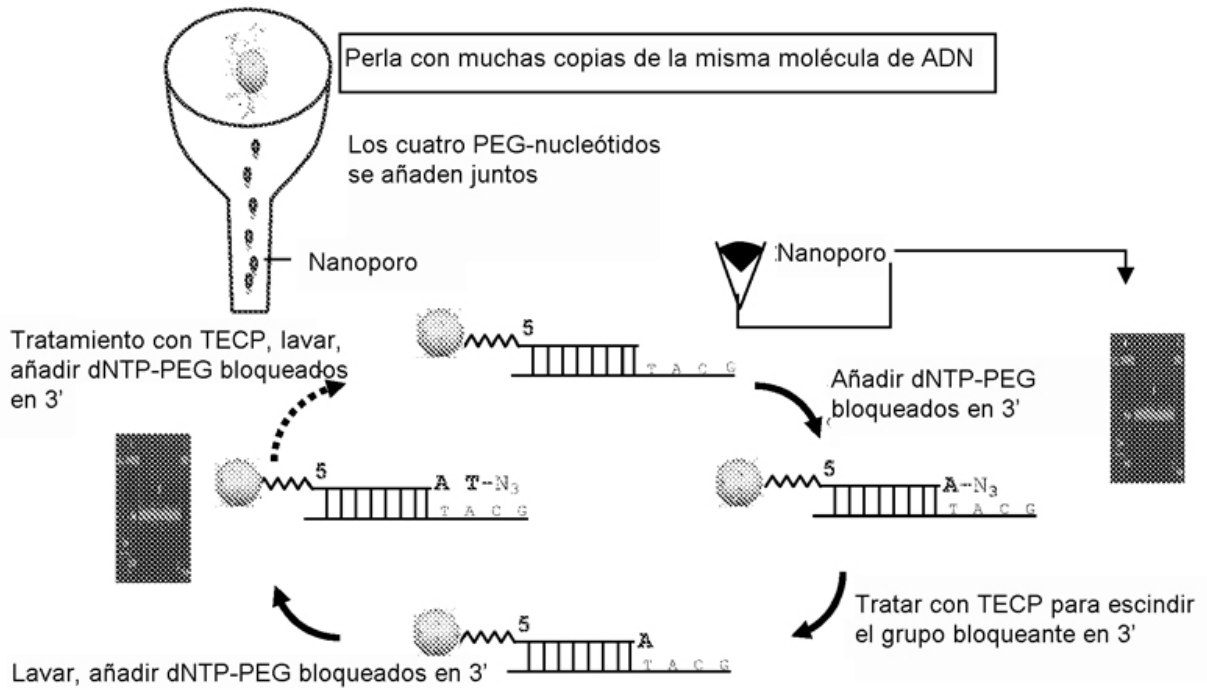


Figura 34

