

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 539 254**

51 Int. Cl.:

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **26.05.2011 E 11722796 (7)**

97 Fecha y número de publicación de la concesión europea: **18.03.2015 EP 2576822**

54 Título: **Método de secuenciación de ADN mediante polimerización**

30 Prioridad:

27.08.2010 US 377621 P
27.05.2010 EP 10305563

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
29.06.2015

73 Titular/es:

**CENTRE NATIONAL DE LA RECHERCHE
SCIENTIFIQUE (CNRS) (33.3%)**
3, rue Michel-Ange
75016 Paris, FR;
ECOLE NORMALE SUPÉRIEURE (33.3%) y
UNIVERSITÉ PIERRE ET MARIE CURIE (PARIS 6)
(33.3%)

72 Inventor/es:

BENSIMON, DAVID;
CROQUETTE, VINCENT;
ALLEMAND, JEAN-FRANÇOIS;
MANOSAS, MARIA y
DING, FANG-YUAN

74 Agente/Representante:

CURELL AGUILÁ, Mireia

ES 2 539 254 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

Método de secuenciación de ADN mediante polimerización.

5 La presente invención se refiere a un método rápido para la determinación de una secuencia de ácido nucleico, ADN o ARN, que es útil, en particular, para la secuenciación de un ácido nucleico desconocido, o alternativamente, para la detección de una secuencia de ácido nucleico específica para diagnóstico.

10 Actualmente, la determinación de secuencia de ácido nucleico está en el corazón de la biología molecular. Por ejemplo, un amplio abanico de fenómenos biológicos se pueden evaluar mediante secuenciación de ADN de alto rendimiento, por ejemplo variación genética, expresión de ARN, interacciones proteína-ADN, y conformación cromosómica (véanse, para unos pocos ejemplos, Mitreva & Mardis, *Methods Mol Biol.*, 533: 153-87, 2009; Mardis, *Genome Med.*, 1(4): 40, 2009; Cloonan et al., *Nat Methods*, 5(7): 613-619, 2008; Valouev et al., *Genome Res.*, 18(7): 1051-63, 2008, Valouev et al., *Nat Methods.*, 5(9):829-34, 2008; Orscheln et al., *Clin Infect Dis.*, 49(4): 536-42, 2009; Walter et al., *Proc Natl Acad Sci USA.*, 106(31): 12950-5, 2009; Mardis et al., *N Engl J Med.*, 361(11): 1058-66, 2009, Hutchinson, *Nucl. Acids Res.*, 35(18): 6227-6237, 2007).

20 Además, la demostración de la presencia de una secuencia de ADN específica en una muestra fisiológica constituye, actualmente, la línea principal de desarrollo de métodos de diagnóstico, por ejemplo para identificar la probabilidad de que las bacterias desarrollen resistencia antibiótica, anomalías genéticas, los riesgos de cáncer asociado con modificaciones genéticas e infecciones víricas, por ejemplo infecciones asociadas con VIH o con virus de la hepatitis (véanse, por ejemplo, Zhang et al., *Nature*, 358: 591-593, 1992; Turner et al., *J Bacteriol.*, 176(12): 3708-3722, 1994; Weston et al., *Infection and Impunity*, 77(7): 2840-2848, 2009).

25 La secuenciación de ácido nucleico se lleva a cabo actualmente principalmente con implementaciones semiautomatizadas, a base de capilares, de la bioquímica de Sanger. El método clásico comprende una etapa de amplificación del ADN de interés, seguido de una etapa de "secuenciación en ciclo", en la que cada ronda de extensión del cebador se termina estocásticamente mediante la incorporación de didesoxinucleótidos (ddNTPs) marcados fluorescentemente. La secuencia se determina mediante separación electroforética de alta resolución de los productos de extensión monocatenarios, marcados en los extremos, en un gel polimérico a base de capilar. La electroforesis simultánea en 96 o 384 capilares independientes proporciona un nivel limitado de paralelización.

35 La elevada demanda de secuenciación de bajo coste ha conducido al desarrollo de tecnologías de secuenciación de alto rendimiento que paralelizan el proceso de secuenciación, produciendo miles o millones de secuencias a la vez (Shendure y Ji, *Nat Biotechnol.*, 26(10): 1135-45, 2008). Las tecnologías de secuenciación de alto rendimiento están destinadas a reducir el coste de la secuenciación del ADN más allá de lo que es posible con métodos estándar de terminador marcado con colorante. Actualmente, este rendimiento tan elevado se logra con sacrificios sustanciales en longitud y exactitud de las lecturas individuales cuando se compara con la secuenciación de Sanger. Los ejemplos de tales métodos nuevos incluyen las tecnologías de 454 y Solexa. Estas tecnologías permiten la secuenciación aleatoria de genomas completos sin clonar en *E. coli* o cualquier célula hospedante. Las bibliotecas de fragmentos de ADN cortos, flanqueados por adaptadores, capturados en la superficie de perlas se amplifican mediante PCR de emulsión. La secuenciación se lleva a cabo usando síntesis cebada mediante ADN polimerasa. En el método de 454 (también conocido como "pirosecuenciación"), la matriz se presenta con cada uno de los cuatro dNTPs, secuencialmente, y la cantidad de incorporación se monitoriza mediante detección luminométrica del pirofosfato liberado. Una diferencia clave entre este método y el Solexa es que este último usa nucleótidos que terminan la cadena. El marcador fluorescente en la base terminadora se puede eliminar para dejar un término 3' no bloqueado, haciendo la terminación de la cadena un proceso reversible. La tecnología SOLiD se basa en la ligación de sondas dibásicas marcadas fluorescentemente a un cebador secuenciador hibridado a una secuencia adaptadora en el molde de la biblioteca amplificado de forma clonal. La especificidad de la sonda dibásica se logra interrogando cada 1ª y 2ª base en cada reacción de ligación. Se realizan múltiples ciclos de ligación, detección y escisión, determinando el número de ciclos la longitud de lectura eventual. En contraste con las tres tecnologías previas, que requieren todas ellas una primera etapa de amplificación, la plataforma Helicos permite la secuenciación de moléculas de ADN individuales. Esta tecnología se basa en el uso de un sistema de detección muy sensible de la incorporación de nucleótidos fluorescentes para interrogar directamente moléculas de ADN individuales vía secuenciación mediante síntesis.

55 Tales métodos se describen, por ejemplo, en la patente US nº 4.882.127, patente US nº 4.849.077; patente US nº 7.556.922; patente US nº 6.723.513; solicitud de patente PCT nº WO 03/066896; solicitud de patente US nº US 2008/0020392; solicitud de patente PCT nº WO 2006/084132; solicitud de patente US nº US 2009/0186349; solicitud de patente US nº US 2009/0181860; solicitud de patente US nº US 2009/0181385; solicitud de patente US nº US 2006/0275782; patente europea EP-B1-1141399; Shendure y Ji, *Nat Biotechnol.*, 26(10): 1135-45, 2008; Pihlak et al., *Nat Biotechnol.*, 26(6): 676-684, 2008; Fuller et al., *Nature Biotechnol.*, 27(11): 1013-1023, 2009; Mardis, *Genome Med.*, 1(4): 40, 2009; Metzker, *Nature Rev. Genet.*, 11(1): 31-46, 2010. En particular, la solicitud de patente PCT nº WO 2007/111924 describe un método para la determinación de una secuencia de ácido nucleico, en el que el movimiento de una polimerasa procesiva selectiva de nucleótido se registra a lo largo de un molde polinucleotídico.

Sin embargo, todos los métodos desarrollados hasta el momento presentan serios inconvenientes. En particular, todos ellos usan nucleótidos marcados (por ejemplo, fluorescentes), contribuyendo así a incrementar seriamente los costes globales. Además, todos estos nuevos métodos salvo uno (la plataforma Helicos) requieren amplificación de la secuencia diana antes de la secuenciación, lo que por un lado consume tiempo, por otro lado incrementa la probabilidad de errores, y tiene mucha tendencia a la contaminación. Además, los métodos que implican técnicas mecánicas en lugar de bioquímicas carecen de sensibilidad (Maier et al., Proc. Natl., Acad. Sci. U.S.A., 97(22): 12002-12007, 2000; Wuite et al., Nature, 404(6773): 103-106, 2000; documento US 2010/0035252). De este modo, todavía existe la necesidad de un nuevo método muy sensible que permita la secuenciación de moléculas individuales.

Descripción detallada de la invención

La presente invención se refiere a un método para la determinación de una secuencia de ácido nucleico mediante manipulación física. En particular, dicho método comprende las etapas de determinar la localización física del sitio en el que se produce una pausa de la replicación, y deducir a partir de ello información sobre la secuencia del ácido nucleico.

El método según la presente invención, basado en técnicas físicas y tratamientos electrónicos, difiere de los enfoques actuales, que son químicos o bioquímicos. Sus ventajas son numerosas:

- 1) Permite la secuenciación de una molécula individual, y de este modo no requiere una etapa de amplificación previa (por ejemplo mediante PCR).
- 2) Es bastante más barato que los métodos de la técnica, puesto que se usan moléculas de ácido nucleico estándar, que son mucho menos caras que los nucleótidos marcados (ya sea con fluoróforos o con algunos otros grupos).
- 3) Permite determinar la localización (en pb) de una hebra complementaria recientemente sintetizada a lo largo de un ácido nucleico bicatenario midiendo la distancia entre los dos extremos de la mencionada molécula de ácido nucleico bicatenario.
- 4) En una forma de realización, permite determinar, en un experimento de polimerización, la posición de un nucleótido dado a lo largo de la hebra.
- 5) La medida se puede repetir periódicamente en una segunda escala de tiempo, conduciendo así a la eliminación de falsos positivos (detención espúrea de la polimerasa), estadística mejorada y una reducción significativa en desviaciones instrumentales.
- 6) El experimento se puede repetir muchas veces en la misma molécula, mejorando así la estadística y la fiabilidad de la medida, puesto que el ácido nucleico monocatenario recientemente sintetizado se puede eliminar (por ejemplo reduciendo la fuerza o la fuerza iónica, o usando una actividad de helicasa o exonucleasa) tras la etapa de replicación. Permite la secuenciación paralela de diversas moléculas de ácido nucleico bicatenario, puesto que cada molécula se puede manipular independientemente de las otras.
- 7) Se puede usar una enzima no modificada para sintetizar la nueva hebra, lo que reduce los costes asociados y mejora la tasa de error, en comparación con la secuenciación de moléculas individuales de tercera generación, que introduce mutación específica del sitio en la polimerasa para escindir el ligador de colorante, o une la enzima a un lado, o la modifica con un punto cuántico.

La presente invención se refiere a un método para la determinación de una secuencia de ácido nucleico basado en la localización física en la molécula de ácido nucleico secuenciada de los sitios en los que la replicación está pausada o bloqueada.

Por "determinación de una secuencia de ácido nucleico", se quiere hacer referencia en la presente memoria no solo al descifrado de la sucesión real de bases en un ácido nucleico, sino también a todas las actividades que conducen directa o indirectamente a la obtención de alguna información sobre la secuencia de ácido nucleico, tal como la detección de una secuencia particular en una molécula de ácido nucleico o la detección de una diferencia entre las secuencias de dos moléculas de ácido nucleico diferentes.

La mayoría de los métodos para determinar una secuencia de ácido nucleico se basa en la síntesis cebada de una nueva hebra mediante una polimerasa procesiva. En estos métodos, un cebador se hibrida a una de las hebras del molde de ácido nucleico bicatenario; se sintetiza una nueva hebra a partir del cebador mediante una polimerasa; la síntesis se detiene o se bloquea en sitios específicos; y la detección de estas pausas o bloqueos en la polimerización da información sobre la secuencia del mencionado ácido nucleico.

Se ha descubierto según la invención que es posible explotar los parámetros físicos asociados con este bloqueo para obtener información sobre la secuencia del ácido nucleico bicatenario. Más precisamente, se ha encontrado que es posible localizar físicamente en la mencionada molécula de ácido nucleico bicatenario el sitio en el que se produce la pausa o bloqueo de la replicación; la posición física específica de la pausa o bloqueo proporciona entonces información sobre la secuencia del mencionado ácido nucleico bicatenario.

La presente invención surge de la observación de que es posible medir la distancia física entre los dos extremos de una molécula de ácido nucleico bicatenario parcialmente desnaturalizada cuando la mencionada molécula está bajo tensión. En un procedimiento de secuenciación mediante síntesis, la progresión de una horquilla de replicación está asociada con el desenrollamiento de la molécula de ácido nucleico bicatenario, dejando atrás dos extremos libres que se unen en la horquilla. Cuando la replicación se bloquea en un sitio específico, la molécula de ácido nucleico bicatenario se bloquea en una conformación en la que las dos hebras enfrente de la horquilla de replicación todavía están hibridadas, mientras que las dos hebras parentales detrás de la horquilla están separadas. Se ha descubierto que es posible medir la distancia física entre los dos extremos separados de la mencionada molécula de ácido nucleico bicatenario, cuando la mencionada molécula de ácido nucleico bicatenario está bajo tensión. La posición física en la mencionada molécula de ácido nucleico bicatenario del sitio en el que se produce la pausa o bloqueo de la replicación se puede deducir entonces a partir de la mencionada distancia, dando como resultado cierta información sobre la secuencia de la mencionada molécula de ácido nucleico bicatenario.

De este modo, el método de la invención se refiere a un método para la determinación de una secuencia de ácido nucleico, comprendiendo dicho método las etapas de:

- a) desnaturalizar una molécula de ácido nucleico bicatenario correspondiente a la mencionada secuencia de ácido nucleico;
- b) hibridar una molécula de ácido nucleico monocatenario ("el cebador") con la mencionada molécula de ácido nucleico bicatenario desnaturalizada;
- c) aplicar una tensión al cebador hibridado/molécula de ácido nucleico bicatenario obtenido en b);
- d) incubar el cebador hibridado/molécula de ácido nucleico bicatenario obtenido en la etapa b) con una polimerasa en condiciones que conducirán a al menos una pausa en la replicación; y
- e) determinar la posición de la mencionada pausa en la replicación con respecto a un extremo del ácido nucleico bicatenario.

Por "desnaturalización", se quiere hacer referencia en la presente memoria al proceso de separación de hebras de una molécula de ácido nucleico bicatenario que se produce cuando se rompe la mayoría de los enlaces de hidrógeno entre las mencionadas hebras. El proceso de desnaturalización produce una molécula de ácido nucleico desnaturalizada, mediante lo cual se quiere hacer referencia en la presente memoria a las dos hebras complementarias separadas que resultan de la desnaturalización de una molécula de ácido nucleico bicatenario. Por "renaturalización", se hace referencia en la presente memoria al proceso mediante el cual dos hebras complementarias separadas se vuelven a formar a través de hibridación en una doble hélice. Como se usa en la presente memoria, "hibridación" es el proceso de establecer una interacción no covalente, específica de la secuencia, entre dos o más hebras complementarias de ácidos nucleicos en un único híbrido.

Existen varias posibilidades conocidas por el experto en la materia para desnaturalizar el ácido nucleico. En una manera muy preferida, las dos hebras se separan sometiéndolas a una fuerza física. Por ejemplo, los extremos libres del mencionado ácido nucleico bicatenario se pueden desarmar, rompiendo así todos los enlaces entre las bases emparejadas, y abriendo el ácido nucleico bicatenario.

En este tipo de método de determinación de la secuencia, puede ser ventajoso, a fin de facilitar el reemparejamiento, colocar los extremos libres del ADN bicatenario (es decir, los extremos que no están unidos a soportes) para que se unan entre sí covalentemente o casi covalentemente antes de desmontarlos. En una forma de realización preferida, la molécula de ácido nucleico bicatenario es una horquilla. En otra forma de realización preferida, el extremo 5' de una hebra se une directamente de forma covalente al extremo 3' de la otra hebra. Si se desea que el ácido nucleico bicatenario se represente gráficamente en el contexto de la presente invención, es posible asemejarlo a una "cremallera", que se abre (o se cierra): la desnaturalización del ácido nucleico bicatenario es la apertura de la cremallera, la renaturalización es el cierre de nuevo de la cremallera.

El ácido nucleico monocatenario de la invención puede ser en particular una molécula de ADN o de ARN, ya sea natural o modificada. Las expresiones "ácido desoxirribonucleico" y "ADN", como se usan aquí, significan un polímero compuesto de desoxirribonucleótidos. Las expresiones "ácido ribonucleico" y "ARN", como se usan aquí, significan un polímero compuesto de ribonucleótidos. El mencionado ácido nucleico monocatenario también puede estar formado de nucleótidos modificados, tales como ácido nucleico bloqueado (LNA), que son nucleótidos en los que el resto de ribosa está modificado con un puente extra que conecta el oxígeno de 2' y el carbono de 4', o ácido

nucleico peptídico (PNA), en el que la cadena principal está compuesta de unidades que se repiten de N-(2-aminoetil)-glicina enlazadas mediante enlaces peptídicos.

La invención se aplica a cualquier tipo de ácido nucleico bicatenario. Más a menudo, el ácido nucleico bicatenario será ADN, pero se entiende que la invención también se aplica a dúplex de ADN monocatenario-ADN monocatenario, perfectamente emparejados o no emparejados perfectamente, o alternativamente a dúplex de ADN monocatenario-ARN monocatenario, perfectamente emparejados o no emparejados perfectamente, o alternativamente a dúplex de ARN monocatenario-ARN monocatenario, perfectamente emparejados o no emparejados perfectamente. Además, el dúplex puede consistir en el reemparejamiento al menos parcial de dos hebras individuales obtenidas de muestras de orígenes diferentes. Finalmente, la invención también se aplica a las estructuras secundarias de un solo ADN monocatenario o de un solo ARN monocatenario.

En una configuración típica, las moléculas de ácido nucleico bicatenario se pueden anclar específicamente sobre dos sustratos sólidos (por ejemplo, portaobjetos de microscopio, micropipeta, micropartícula). Uno de los extremos se puede unir directa o indirectamente a una superficie, mientras que el otro extremo se une directa o indirectamente a una superficie móvil. En esta forma de realización, se aplica una tensión en ambos extremos del ácido nucleico bicatenario cuando los soportes se separan. Cuando la tensión es mayor que un valor umbral, las dos hebras se separan, y la molécula de ácido nucleico está desnaturalizada. La tensión aplicada está preferentemente por encima o es igual a 15 pN; está más preferentemente por encima o es igual a 16 pN; incluso está más preferentemente por encima o es igual a 17 pN; en un aspecto mucho más preferido, está por encima o es igual a 18 pN. Esta fuerza puede variar con la temperatura, con el tipo de nucleótido y tampón, pero el experto en la materia adaptará fácilmente la mencionada fuerza con respecto a estos parámetros a fin de obtener la separación de las dos hebras.

En una forma de realización preferida de la invención, el ácido nucleico bicatenario está desnaturalizado aplicando una tensión mayor que un valor umbral. La incubación del ácido nucleico bicatenario desnaturalizado con un ácido nucleico monocatenario (el "cebador") conduce a la hibridación del mencionado cebador monocatenario. Preferentemente, la secuencia de la molécula de ácido nucleico monocatenario es complementaria a al menos parte de la secuencia de la molécula de ácido nucleico bicatenario.

Cuando la tensión se disminuye hasta alrededor de un valor intermedio, las dos hebras del ácido nucleico bicatenario desnaturalizado se pueden volver a hibridar. Para obtener la rehibridación de las dos hebras, se aplica una tensión de entre 10 y 12 pN; más preferentemente es 12 pN; incluso más preferentemente, es 11 pN; todavía más preferentemente, es 10 pN. Según la invención, la actividad de polimerasa es activa en estas condiciones de tensión, dando como resultado la extensión del cebador mediante incorporación nucleotídica en una nueva hebra. Todavía más preferentemente, el ácido nucleico bicatenario es una horquilla. Como se usa en la presente memoria, "horquilla" significa una doble hélice en la que el extremo 5' de una hebra está enlazado físicamente al extremo 3' de la otra hebra a través de un bucle no emparejado. El mencionado enlace físico puede ser covalente o no covalente. Preferentemente, el mencionado enlace físico es un enlace covalente. De este modo, una horquilla consiste en un tallo bicatenario y un bucle monocatenario no emparejado. En la horquilla, los extremos de las dos hebras que no están acopladas en el bucle están libres, y de este modo se pueden desarmar. Esto da como resultado el desemparejamiento del ácido nucleico bicatenario, produciendo así una molécula de ácido nucleico bicatenario desnaturalizada. Es posible abrir completamente una molécula de ácido nucleico bicatenario de horquilla al tirar en cada extremo de la mencionada molécula de ácido nucleico con una fuerza mayor que un valor umbral. Cuando la tensión aplicada a la molécula se disminuye hasta un valor intermedio, la molécula de ácido nucleico se vuelve a autohibridar para volver a formar una horquilla. Bajo esta tensión intermedia, se produce una nueva hebra mediante la actividad de polimerasa de la enzima, hasta que se encuentra un bloqueo. Según la invención, la determinación de la posición del mencionado bloqueo en la replicación con respecto a un extremo del ácido nucleico bicatenario da información sobre la secuencia del mencionado ácido nucleico bicatenario.

El uso de una horquilla hace posible, en particular, llevar a cabo ciclos de emparejamiento y desemparejamiento, y de este modo mejorar la relación de señal/ruido.

En el contexto de la invención, el bucle puede ser de cualquier longitud comprendida entre 0 y 60 nucleótidos. Se cree que para formar una horquilla estable, es necesaria una región de bucle de al menos alrededor de 4 ó 5 nucleótidos. Sin embargo, también es posible llevar a cabo la invención con bucles de una longitud mucho más corta. De hecho, se ha descubierto que en algunas formas de realización de la invención puede ser ventajoso usar una horquilla cuyo bucle consiste en 0 nucleótidos. En este caso, el extremo 3' de una hebra está enlazado directa y físicamente al extremo 5' de la otra hebra. Las técnicas que permiten que los extremos libres del ácido nucleico bicatenario se unan juntos son conocidas, y algunas se describirán con mayor detalle en adelante.

Mediante la determinación del bloqueo, se quiere hacer referencia en la presente memoria a la determinación de los parámetros físicos asociados con el bloqueo. El más útil de estos parámetros es la posición del bloqueo en la molécula de ácido nucleico bicatenario, correspondiendo dicha posición a la posición del último nucleótido incorporado en la hebra individual recientemente sintetizada. De hecho, se ha descubierto que se puede determinar de forma precisa la posición en el ácido nucleico bicatenario estirado en la que se produce la pausa en la renaturalización: el uso de horquilla permite al experto en la materia determinar la distancia física entre los dos

extremos libres de la horquilla en cualquier momento durante el proceso de desnaturalización/renaturalización.

Por "extremo libre" se quiere hacer referencia en la presente memoria al extremo de una hebra que no está enlazado covalentemente a una extremidad de la otra hebra; como se explica anteriormente, estos extremos libres se pueden unir cada uno a una superficie diferente. Por ejemplo, una de estas superficies puede ser móvil, mientras que la otra puede ser inmóvil. El experto en la materia apreciará así fácilmente que, a fin de medir la distancia entre los extremos libres del ácido nucleico bicatenario de horquilla, es posible medir simplemente la distancia entre las dos superficies.

Esta distancia es máxima (z_{high} a una fuerza (F_{open}), que es mayor que el valor umbral mencionado anteriormente) cuando la molécula de horquilla está desnaturalizada completamente, puesto que el ácido nucleico de horquilla está entonces completamente extendido. Es mínima (z_{low} a una fuerza (F_{test}) que corresponde al valor intermedio explicado anteriormente) cuando la mencionada molécula de horquilla está completamente renaturalizada. Es ventajoso llevar a cabo todas las comparaciones de longitudes a la misma fuerza F_{test} , de manera que el ácido nucleico monocatenario tenga las mismas propiedades elásticas. Usando el retraso en el cierre del bucle, el usuario experto puede medir z_{high} (F_{test}). Cuando la replicación se bloquea en un sitio específico, la molécula de ácido nucleico bicatenario se bloquea en una conformación en la que las dos hebras en frente de la horquilla de replicación todavía están hibridadas, mientras que las dos hebras parentales detrás de la horquilla están separadas. Se puede medir la distancia entre los dos extremos libres cuando el proceso de replicación se detiene temporal o permanentemente: como se esperaba, la distancia z está comprendida entre z_{high} y z_{low} (midiéndose todas las z con $F = F_{\text{test}}$). Es inmediatamente claro que la distancia z varía con la localización en la molécula de horquilla del punto en el que la horquilla de replicación se detiene o se bloquea. Si la mencionada horquilla de replicación se detiene en una secuencia que está situada próxima a los extremos libres de la horquilla, la distancia z_{pause} será mínima. Por otro lado, si la mencionada horquilla de replicación se bloquea en una secuencia que corresponde a una parte de la horquilla que está próxima al bucle desemparejado, la distancia z será máxima (Fig. 1). La pausa se puede observar durante el desenrollamiento, inducido por polimerasa, de una horquilla bajo una tensión F_{test} . Si la replicación transcurre a una fuerza F_{open} (para la que la horquilla se abre) hasta que se bloquea, el bloqueo de la replicación se puede observar al reducir la fuerza hasta F_{test} , que permite que las hebras se vuelvan a unir en forma de cremallera hasta el punto de bloqueo.

Es posible correlacionar de forma precisa una distancia física en una molécula de ácido nucleico bicatenario con un número de bases. Por ejemplo, una distancia de 1 nm corresponde a la distancia abarcada por dos nucleótidos (1 pb) en un ácido nucleico bajo una fuerza de 10 pN. La calibración exacta frente a la fuerza se da mediante la elasticidad del ácido nucleico monocatenario. Por lo tanto, midiendo simplemente la distancia entre los dos extremos libres de la molécula de ácido nucleico bicatenario bajo tensión, es posible determinar de forma precisa dónde se bloquea la renaturalización.

De este modo, en una forma de realización, el método de la invención se refiere a un método para la determinación de una secuencia de ácido nucleico, comprendiendo dicho método las etapas de:

- a) desnaturalizar una molécula de ácido nucleico bicatenario correspondiente a la mencionada secuencia de ácido nucleico;
- b) hibridar una molécula de ácido nucleico monocatenario ("el cebador") con la mencionada molécula de ácido nucleico bicatenario desnaturalizada;
- c) aplicar una tensión al cebador hibridado/molécula de ácido nucleico bicatenario obtenido en b);
- d) incubar el cebador hibridado/molécula de ácido nucleico bicatenario obtenido en la etapa b) con una polimerasa en condiciones que conducirán a al menos una pausa en la replicación; y
- e) determinar la posición de la mencionada pausa en la replicación con respecto a un extremo del ácido nucleico bicatenario

en el que la distancia entre los dos extremos de la molécula bicatenaria se determina cuando el proceso de replicación está bloqueado. Preferentemente, la distancia entre los dos extremos de la mencionada molécula se determina cuando la molécula está completamente desnaturalizada. Incluso más preferentemente, las dos distancias se comparan y se determina la posición del bloqueo.

Como se usa en la presente memoria, "polimerasa" se refiere a una enzima que cataliza la polimerización de nucleótidos (es decir, la actividad de polimerasa). Generalmente, la enzima iniciará la síntesis en el extremo 3' del cebador hibridado a una secuencia molde polinucleotídica, y transcurrirá hacia el extremo 5' de la hebra molde. La "ADN polimerasa" cataliza la polimerización de desoxinucleótidos, mientras que la "ARN polimerasa" cataliza la polimerización de ribonucleótidos. La polimerasa según la invención es una polimerasa procesiva o una polimerasa no procesiva. Una enzima procesiva cataliza múltiples rondas de una reacción en un molde de ácido nucleico bicatenario desnaturalizado, mientras que la enzima permanece unida al mencionado molde. Como se entiende en

la presente memoria, una polimerasa será procesiva, es decir, permanecerá unida al molde de ácido nucleico bicatenario desnaturalizado, para al menos 25 nucleótidos, al menos 50 nucleótidos, al menos 100 nucleótidos, habitualmente al menos 500 nucleótidos, y puede ser procesiva para al menos 1000 nucleótidos o más. Las polimerasas según la invención incluyen ARN polimerasas dependientes de ARN, ARN polimerasas dependientes de ADN, ADN polimerasas dependientes de ADN, ADN polimerasas dependientes de ARN (transcriptasa inversa), y similares. Muchas de tales enzimas son conocidas en la técnica. Según la invención, la mencionada polimerasa es capaz de sintetizar ácidos nucleicos cuando la fuerza aplicada al molde de ácido nucleico bicatenario está en un valor intermedio, es decir, comprendido entre 10 y 12 pN, o a fuerzas elevadas para las cuales la horquilla está completamente desenrollada.

Preferiblemente, en el método de la invención se usa una polimerasa con una actividad de exonucleasa 3'-5'. Como se usa en la presente memoria, "actividad de exonucleasa 3'-5'" se refiere a la capacidad de una enzima para eliminar nucleótidos incorporados desde el extremo 3' de un polímero de ADN. Los ejemplos de tales enzimas incluyen, por ejemplo, T4 ADN polimerasa, T7 ADN polimerasa, DEEP VENT ADN polimerasa, E. coli polimerasa III, Phi29 ADN polimerasa, E. coli ADN polimerasa I, E. coli ADN polimerasa I, fragmento de Klenow, Phusion® High Fidelity ADN polimerasa, Phusion® Hot Start High Fidelity ADN polimerasa, Phire® Hot Start ADN polimerasa, 9^oNm ADN polimerasa, ADN polimerasa del virus del herpes simple tipo 1. En una forma de realización preferida, la polimerasa de la invención se puede intercambiar desde un modo activo de polimerasa a un modo activo de exonucleasa 3'-5' disminuyendo la fuerza aplicada a la molécula de ácido nucleico bicatenario por debajo de un valor mínimo. Preferiblemente, el mencionado valor mínimo es 7 pN; más preferiblemente, el mencionado valor mínimo es 6 pN; incluso más preferiblemente, el mencionado valor mínimo es 5 pN.

Incluso más preferiblemente, la mencionada polimerasa tiene, además, una actividad de desplazamiento de hebra bajo una tensión intermedia, por ejemplo cuando se aplica una fuerza entre 10 y 12 pN a la horquilla bicatenaria. Por "desplazamiento de hebra", se quiere hacer referencia en la presente memoria a la capacidad de la polimerasa para desplazar el ácido nucleico en dirección 3' durante la síntesis. Se ha descubierto en particular que la T4 ADN polimerasa y la T7 ADN polimerasa, que no es conocido que tengan ninguna actividad de desplazamiento de hebra en condiciones de tubo de ensayo, es decir, en condiciones en las que no se aplica tensión al molde bicatenario, son capaces de eliminar el ácido nucleico en dirección 3' durante la polimerización cuando la horquilla bicatenaria está bajo una fuerza ≥ 10 pN. La T4 ADN polimerasa y la T7 ADN polimerasa son así particularmente adecuadas para llevar a cabo el método de la invención. "T4 ADN polimerasa" y "T7 ADN polimerasa" se refieren en la presente memoria tanto a la enzima monomérica como a la holoenzima.

El método según la invención comprende una etapa de replicación que se lleva a cabo en condiciones que conducirán a al menos una pausa en el proceso de replicación. Preferentemente, la molécula de ácido nucleico bicatenario se somete a una tensión durante la etapa de replicación. Más preferiblemente, la mencionada tensión está alrededor de un valor intermedio, es decir, la mencionada tensión está comprendida entre 10 y 12 pN. La mencionada pausa se puede provocar por cualquiera de los medios conocidos por el experto en la materia. El método de secuenciación mediante síntesis, en el que se bloquea la síntesis de la nueva hebra, se ha usado ampliamente en la técnica. En el contexto de la presente invención, se puede adaptar cualquiera de tales métodos. Por ejemplo, la polimerasa puede ser una enzima procesiva, sensible a nucleótido, que entonces se pone en contacto con el molde de ácido nucleico bicatenario desnaturalizado en una mezcla de reacción que altera la velocidad para el movimiento procesivo de la enzima para un nucleótido específico. Por ejemplo, la mencionada mezcla de reacción comprende un conjunto de desoxinucleótidos (dNTP) en el que una de las bases está presente en una concentración muy baja. En ese caso, cada vez que la polimerasa encuentra el complemento del mencionado nucleótido, se detiene hasta que el nucleótido de baja concentración se difunde en la posición. Las posiciones de las pausas a lo largo de la molécula revelan así las posiciones del nucleótido escaso en la hebra sintetizada, y permite al experto en la materia identificar la posición de la base correspondiente en la secuencia (Greenleaf y Block, Science, 313: 801, 200; patente US nº 7.556.922). Como alternativa, es posible usar didesoxinucleótidos (ddNTPs) además de los desoxinucleótidos normales (dNTPs) encontrados en ADN. Los didesoxinucleótidos son esencialmente los mismos que los nucleótidos, excepto que contienen un grupo hidrógeno en el carbono 3' en lugar de un grupo hidroxilo (OH). Estos nucleótidos modificados, cuando se integran en una secuencia, evitan la adición de más nucleótidos. Esto ocurre debido a que no se puede formar un enlace de fosfodiéster entre el didesoxinucleótidos y el siguiente nucleótido entrante, y de este modo se termina la cadena de ADN. Por lo tanto, la incorporación de un ddNTP provocará que se detenga la reacción de polimerasa, puesto que no se puede añadir ningún nucleótido después del mencionado ddNTP. La posición del bloqueo a lo largo de la molécula revela así la posición de la incorporación del ddNTP en la hebra sintetizada, y permite al experto en la materia identificar la posición de la base correspondiente en la secuencia. La posición de cada pausa o bloqueo se puede determinar entonces mediante el método de la invención, es decir, midiendo la distancia física entre los dos extremos libres de la molécula. Además, también es posible usar trifosfatos de nucleósidos (NTPs) en lugar de ddNTPs. La diferencia está en que la incorporación de un NTP solamente detendrá transitoriamente el proceso de la polimerasa, generando un patrón similar al primer ejemplo propuesto mencionado aquí.

El método de la invención se puede usar para la secuenciación directa de un ácido nucleico desconocido. En una forma de realización preferida, se usa una enzima procesiva para sintetizar a partir de un ácido nucleico monocatenario conocido (un cebador) una secuencia de extensión creciente complementaria a una de las hebras de

horquilla, desenrollando de ese modo eficazmente la horquilla bicatenaria mantenida bajo una tensión moderada (por ejemplo, en el intervalo de 5 a 13 pN). En esta forma de realización, la polimerización se inicia abriendo la horquilla bicatenaria al incrementar transitoriamente la fuerza hasta F_{open} en presencia de un cebador monocatenario. F_{open} es una tensión mayor que el valor umbral requerido para abrir completamente la horquilla bicatenaria. Esto da como resultado la hibridación del cebador a la horquilla bicatenaria. En una etapa siguiente, la fuerza se ajusta a $F_{elongation}$ ($\leq F_{open}$), a fin de permitir que la polimerasa sintetice una nueva hebra (en un modo de desplazamiento de hebra o no). $F_{elongation}$ se ajusta preferiblemente a una tensión intermedia menor que F_{open} (el valor umbral requerido para abrir completamente la horquilla bicatenaria), permitiendo dicha tensión intermedia la replicación mediante la polimerasa. Preferiblemente, el mencionado valor intermedio está comprendido entre 10 y 12 pN. Bajo una tensión igual a $F_{elongation}$, la polimerasa sintetiza una nueva hebra a una velocidad sostenida hasta que se produce una pausa o un bloqueo. La actividad enzimática conduce así a la producción de una molécula de ácido nucleico monocatenario complementaria extendida.

En la configuración de desplazamiento de hebra, la mencionada polimerasa se puede intercambiar entre sus dos modos de operación, es decir, actividad de exonucleasa y alargamiento, ajustando la fuerza aplicada sobre la horquilla.

En una forma de realización preferida, el proceso de alargamiento se detiene antes de que la polimerasa alcance el bucle. Esto se puede lograr por diversos medios. Por ejemplo, la inserción de bases no convencionales unos pocos nucleótidos delante del bucle detendrá a la polimerasa. El mismo efecto se puede lograr con un dominio de unión bicatenario de una proteína situada justo antes del bucle.

En una forma de realización preferida adicional, la fuerza (F_{exo}) aplicada a la molécula de horquilla se disminuye por debajo de un valor mínimo, por ejemplo 5 pN, después de que se bloquea el proceso de alargamiento. Esto permite que la polimerasa intercambie su actividad de exonucleasa y desensamble la hebra que se ha sintetizado recientemente. Este proceso se detiene cuando toda la hebra se desensambla completamente. Alternativamente, el proceso de desensamblaje se detiene cuando la enzima se atasca, por ejemplo cuando encuentra un obstáculo, tal como un nucleótido modificado en el cebador. En tal caso, puede ser necesario usar una enzima para expulsar la hebra recientemente sintetizada. Como ejemplos de enzimas adecuadas, se pueden citar, por ejemplo, helicasas, incluyendo una UvrD helicasa, una recBCD helicasa, E. coli UvrD helicasa, Tte-UvrD helicasa, T7 Gp4 helicasa, RecBCD helicasa, DnaB helicasa, MCM helicasa, Rep helicasa, RecQ helicasa, PcrA helicasa, T4 UvsW helicasa, helicasa del antígeno T grande de SV40, helicasa del virus del herpes, Sgs1 helicasa de levadura, helicasas dependientes de DEAH-ATP y proteína de helicasa del virus del papiloma E1 y sus homólogos, y exonucleasas, incluyendo fosfodiesterasa de veneno de serpiente, fosfodiesterasa de bazo, Bal-31 nucleasa, E. coli exonucleasa I, E. coli exonucleasa VII, nucleasa de judía mungo, S1 nucleasa, y actividad de exonucleasa de E. coli ADN polimerasa 1, una actividad de exonucleasa de un fragmento de Klenow de ADN polimerasa 1, una actividad de exonucleasa de T4 ADN polimerasa, una actividad de exonucleasa de T7 ADN polimerasa, una actividad de exonucleasa de Taq ADN polimerasa, una actividad de exonucleasa de DEEP VENT ADN polimerasa, E. coli exonucleasa III, λ exonucleasa y una actividad de exonucleasa de VENTR ADN polimerasa.

El desensamblaje de la hebra que se acaba de sintetizar proporciona la oportunidad de repetir todo el proceso, es decir, la síntesis de una hebra bajo una tensión superior a un umbral, por ejemplo 10 pN, deteniéndose la polimerasa por ejemplo cada vez que encuentra el complemento del nucleótido raro, o deteniéndose si incorpora un ddNTP o un NTP. Si el cebador ha sido expulsado durante la etapa de desensamblaje, la síntesis irá precedida de una etapa de apertura de la horquilla y de cierre nuevamente, de manera que se pueda hibridar un cebador. El incremento de la fuerza por encima de 10 pN intercambiará nuevamente la polimerasa al modo de alargamiento, y se puede registrar un nuevo patrón de pausa. La repetición del ciclo de síntesis/desensamblaje hace posible así registrar varios patrones de pausa con el mismo nucleótido raro. Esto conduce a una relación mejorada de señal/ruido, permitiendo la obtención de una secuencia de una mayor calidad con menores errores. Como alternativa, la etapa de replicación se puede realizar en presencia de ddNTP a una fuerza elevada, en la que la horquilla está completamente abierta. En ese caso, los bloqueos que resultan de la incorporación de ddNTP se pueden detectar al reducir la fuerza hasta F_{test} .

Una vez que se han registrado suficientes estadísticas, el procedimiento se repite con una escasez de otro nucleótido u otro ddNTP. Después de que el mencionado procedimiento se ha repetido con cada nucleótido, las posiciones de todos los nucleótidos en la hebra se compilan juntas, produciendo así la secuencia completa de la molécula de ácido nucleico bicatenario original.

La implementación del método de la invención se ha hecho posible, en particular, por la existencia de dispositivos diseñados para investigar la interacción de ácido nucleico en tiempo real al nivel de moléculas individuales. Tal dispositivo se describe, por ejemplo, en las patentes US n^{os} 7.052.650 y 7.244.391. El aparato descrito en la presente memoria usa trampas magnéticas para aplicar una fuerza de escala de picoNewton sobre una perla superparamagnética de tamaño micrométrico. De forma breve, el mencionado aparato comprende un microscopio óptico, imanes y un PC. Las moléculas de ácido nucleico bicatenario se anclan en múltiples puntos en un extremo a un elemento inmóvil, por ejemplo una superficie, y en el otro extremo a una superficie móvil, en este caso una perla magnética. Se proporcionan imanes para actuar sobre la perla. En particular, los imanes se pueden usar para tirar

de la perla lejos de la superficie. Sin embargo, la implementación del método de la invención no está restringida al aparato anterior. Para implementar el método de la invención, se puede usar cualquier dispositivo que permita extender completamente y después volver a enrollar una molécula de ácido nucleico bicatenario, mientras se monitoriza al mismo tiempo la extensión de la mencionada molécula. Por ejemplo, se pueden usar pinzas ópticas; sin embargo, requieren una calibración previa de la fuerza y no se paralelizan fácilmente para medidas de alto rendimiento. Otros inconvenientes son la falta de control torsional total del ácido nucleico y el posible calentamiento local de la disolución por el láser enfocado, que puede alterar las condiciones de hibridación.

El ácido nucleico bicatenario se incuba durante unos pocos minutos en una disolución de perlas adecuadas (por ejemplo las revestidas con estreptavidina) a las que se une mediante uno de sus extremos marcados (por ejemplo biotina). Las perlas pueden ser transparentes si se usan posteriormente pinzas ópticas para la manipulación, o magnéticas si se usan trampas magnéticas o pinzas para la manipulación.

El montaje de perla-ácido nucleico se inyecta en una cámara fluidica cuya superficie se ha tratado de manera que se una al otro extremo marcado de la molécula (por ejemplo, una superficie revestida con anti-Dig para que se una al extremo marcado con Dig del ácido nucleico). Las perlas se anclan entonces a la superficie vía una horquilla del ácido nucleico; véase la figura 1a. Entonces, la distancia de la perla a la superficie se monitoriza por diversos medios conocidos por el experto en la materia: por ejemplo, se pueden usar los anillos de difracción de su imagen en una cámara para deducir su distancia, o para medir su distancia se puede usar la intensidad de luz que dispersan (o emiten por fluorescencia) cuando se iluminan en un modo evanescente. Alternativamente, se puede medir el campo magnético que generan (usando un sensor magnético tal como los sensores GMR o Hall) para deducir su distancia a un sensor en la superficie de anclaje.

Para tirar de la molécula de ácido nucleico que ancla las perlas a la superficie, se han descrito diversas técnicas. Se puede usar la luz de un haz de láser enfocado para atrapar una perla transparente cerca del punto focal. Mediante el desplazamiento relativo del haz con respecto a la superficie de anclaje, se puede aplicar una fuerza sobre la molécula sujetadora (un ensayo de pinzas ópticas típico). La fuerza ejercida que es proporcional al desplazamiento de la perla desde su posición de equilibrio, para ejercer una fuerza constante sobre la molécula sujetadora requiere un bucle de retroalimentación sobre el haz atraparador.

Para ejercer una fuerza constante sobre la perla, se ha descrito el uso del arrastre hidrodinámico generado por un flujo alrededor de la perla, pero habitualmente produce una exactitud espacial pequeña (> 100 nm). La forma de realización preferida usa una trampa magnética para tirar de las perlas superparamagnéticas ancladas a la superficie mediante una horquilla de ácido nucleico como se describe anteriormente. En esta configuración, se usan pequeños imanes colocados encima de la muestra para aplicar una fuerza constante sobre la perla anclada, cuya posición se puede determinar con una exactitud de < 1 nm (dependiendo de la fuerza de tracción y la disipación debido al arrastre hidrodinámico). En cada caso, se observa que la cremallera de la horquilla sujetadora se puede abrir completamente de forma mecánica tirando de las perlas con una fuerza mayor que alrededor de 16 pN. La reducción de la tensión sobre la molécula por debajo de alrededor de 11 pN permite que la cremallera de la horquilla se vuelva a unir espontáneamente (la transición de la apertura de la cremallera es reversible aunque histerética). Si, durante la fase de cremallera abierta, algunas moléculas en disolución (tales como las proteínas u oligonucleótidos complementarios de ADN, ARN, LNA o PNA) se han unido al ácido nucleico monocatenario estirado, estas moléculas bloquearán el cierre nuevamente de la cremallera de la horquilla cuando la fuerza se reduce por debajo de 11 pN. El principio del ensayo es así el intercambio entre dos fuerzas: una grande F_{open} para abrir la horquilla y una más pequeña F_{test} usada para permitir el cierre nuevamente de la cremallera, y medir la extensión de la molécula en bloqueos transitorios. La posición de bloqueo está relacionada con la secuencia mediante una relación lineal entre la extensión total y el bloqueo. Para la mejor exactitud, la extensión total se mide preferiblemente a la fuerza de ensayo F_{test} . Esto se logra diseñando el bucle de horquilla de manera que requiere una fracción de un segundo para volverse a enrollar una vez que la fuerza se reduce desde F_{open} a F_{test} .

A fin de unir ácidos nucleicos a superficies o soportes, se puede hacer uso de una cualquiera de las técnicas conocidas en el campo. Esencialmente, el ácido nucleico se ancla directamente al soporte, por ejemplo la microperla, que implica una funcionalización de esta superficie, por ejemplo revistiéndola con estreptavidina, un grupo COOH, y similar, capaz de reaccionar con el extremo funcionalizado del ácido nucleico.

Tales métodos necesitan, en general, funcionalizar el ácido nucleico, especialmente los extremos 3' y 5', es decir, injertar grupos químicos apropiados en ellos. Además, resulta preferido unir los otros dos extremos libres de la molécula mediante un bucle a fin de evitar que las hebras se disocien en el extremo de la operación, de manera que esta última se pueda repetir si es apropiado. Para este fin, se pueden adoptar procedimientos diferentes.

Lo más simple es funcionalizar, usando oligonucleótidos sintéticos, uno de los extremos de un ácido nucleico bicatenario con dos funciones diferentes (por ejemplo, biotina y amina), que permitan el anclaje a dos superficies pretratadas diferentes. Las dos hebras en el otro extremo se pueden unir usando un nucleótido sintético parcialmente emparejado en forma de un bucle. De esta manera, se produce un ácido nucleico monocatenario, emparejado, es decir, una horquilla, a partir de un ácido nucleico bicatenario. La ventaja de este método se basa en su capacidad para funcionalizar una población heterogénea de fragmentos de ácido nucleico grandes (según se

obtienen mediante fraccionamiento de un gen o cromosoma), que entonces se pueden analizar simultáneamente. En este caso, la muestra de ácido nucleico se fracciona usando dos (o más) enzimas de restricción, que permiten que se obtenga una subpoblación con dos sitios de restricción diferentes en sus extremos que son similares a lo largo de todos los fragmentos. Esto permite que los dos extremos se traten de manera diferente (por ejemplo, uniendo un extremo a un oligonucleótido en forma de un bucle que posee el sitio de restricción apropiado en su extremo). El inconveniente de este método reside en la interferencia estérica entre los dos grupos funcionales adyacentes, que pueden hacer difícil el acoplamiento a las superficies. Para resolver este problema, puede ser ventajoso añadir a cada extremo libre de la molécula de horquilla una secuencia "espaciadora" de bases, a cuyo extremo se añade entonces un grupo funcional; las dos secuencias espaciadoras no son complementarias, proporcionando a cada grupo funcional espacio suficiente para unirse a su superficie dedicada. Más ventajosamente, la secuencia de cada secuencia espaciadora se diseña a fin de usar cebadores secuenciadores monocatenarios de secuencia conocida en el método de secuenciación de la invención. La adición de un bucle y/o espaciadores a las moléculas de ácido nucleico bicatenario se puede llevar a cabo con cualquiera de los métodos usados habitualmente en biología molecular. Estos métodos son bien conocidos por el experto en la materia, y de este modo no hay necesidad de detallarlos en la presente memoria.

Con respecto a las técnicas actuales de anclaje, existen muchas de ellas, y derivan de las técnicas para anclar macromoléculas (proteínas, ADN, y similares) a superficies pretratadas comercialmente disponibles. La mayoría de estas técnicas se han desarrollado para ensayos inmunológicos, y para ligar proteínas (inmunoglobulinas) a superficies que portan grupos (--COOH, --NH₂, --OH, y similares) capaces de reaccionar con los extremos carboxilo (--COOH) o amina (--NH₂) de las proteínas.

El anclaje covalente del ácido nucleico se puede lograr directamente, vía el fosfato libre del extremo 5' de la molécula, que reacciona con una amina secundaria (Covalink --NH surface comercializado por PolyLabo en Estrasburgo) para formar un enlace covalente. También es posible funcionalizar ADN con un grupo amina, y después proceder de igual forma que con una proteína.

También existen superficies revestidas con estreptavidina (perlas Dynal, y similares), que permiten el anclaje casi covalente entre la estreptavidina y una molécula de ADN biotinilada. Finalmente, injertando un anticuerpo dirigido contra digoxigenina sobre una superficie (mediante los métodos mencionados anteriormente), se puede anclar un ácido nucleico funcionalizado con digoxigenina. Esto representa meramente una muestra de las muchas técnicas de anclaje posibles.

Entre las técnicas de unión y anclaje, también se deberían de mencionar, por ejemplo, las técnicas descritas en la patente EP 152.886, que usa un acoplamiento enzimático para la unión de ADN a un soporte sólido tal como celulosa.

La patente EP 146.815 también describe diversos métodos de unión de ADN a un soporte. De forma similar, la solicitud de patente WO 92/16659 propone un método que usa un polímero para unir ADN.

Naturalmente, el ácido nucleico se puede unir directamente al soporte pero, cuando sea necesario, especialmente con vistas a limitar la influencia de las superficies, el ácido nucleico se puede unir al extremo de un brazo inerte de péptido u otra naturaleza, como se describe, por ejemplo, en la patente EP 329.198.

La práctica de la invención emplea, excepto que se indique de otro modo, técnicas convencionales o química proteica, virología molecular, microbiología, tecnología de ADN recombinante, y farmacología, que están dentro de la pericia de la técnica. Tales técnicas se explican completamente en la bibliografía. (Véanse Ausubel et al., *Current Protocols in Molecular Biology*, Eds., John Wiley & Sons, Inc. Nueva York, 1995; Remington's *Pharmaceutical Sciences*, 17ª ed., Mack Publishing Co., Easton, Pa., 1985; y Sambrook et al., *Molecular cloning: A laboratory manual* 2ª edición, Cold Spring Harbor Laboratory Press - Cold Spring Harbor, NY, USA, 1989).

Los ejemplos a continuación permitirán poner de manifiesto otras características y ventajas de la presente invención.

Leyendas de las figuras

Figura 1 Principio de detección de la hibridación de oligonucleótidos a su secuencia complementaria en un ADN horquillado. El ADN horquillado que ancla la perla a la superficie (a) se abre momentáneamente la incrementar la fuerza que tira de la perla hasta un valor por encima de 16 pN. En esa fase, el fragmento complementario en disolución se hibrida a su diana en la horquilla de ADN abierta, evitando así el cierre nuevamente de la cremallera de la horquilla (b) cuando la fuerza se reduce nuevamente a su valor inicial. El enrollamiento de nuevo de la horquilla presenta cuatro mesetas que se producen en extensiones bien definidas pero con duración variable. La meseta superior a 73,71 nm está asociada con la horquilla completamente abierta de 83 pb a F_{test} , mientras que la de la parte inferior corresponde a la horquilla completamente enrollada de nuevo. Las dos mesetas intermedias a 25,47 nm y 35,17 nm se producen debido a que se han colocado dos oligos en la disolución.

A partir de este cambio en la extensión (Z_{high-Z}), es posible deducir dónde, a lo largo de la horquilla, se ha emparejado la secuencia complementaria. Aquí, según sus posiciones, los bloques coinciden con la localización 28,66 pb y 39,60 pb con una coincidencia muy buena con sus posiciones esperadas en 29 pb y 40 pb. Las posiciones de las mesetas se estiman mejor ajustando la representación gaussiana al histograma obtenido de varios ciclos de apertura/cierre (aquí, ~ 20 ciclos).

Figura 2: Ilustración de secuenciación de Sanger de una sola molécula. Registro en tiempo real de la síntesis en una horquilla de una sola molécula de 1,2 kbps con T4 ADN polimerasa (dTNP = dGNP = dCNP = 500 μ M, dATP = 5 μ M, ddATP = 400 μ M, abrazadera y cargador de abrazadera) en un tampón comercial (tampón de reacción 5X: Tris-HCl 335 mM (pH 8,8 a 25°C), MgCl₂ 33 mM, DTT 5 mM, (NH₄)₂SO₄ 84 mM, Fermentas; <http://www.fermentas.com/en/products/all/modifying-enzymes/mesophilic-polymerases/ep006-t4-dna-polymerase>). A fuerza elevada, el pico de la curva muestra la existencia de la muestra de horquilla. A fuerza media, la curva en ascenso muestra la síntesis procesiva de una nueva hebra complementaria mediante la T4 ADN polimerasa, mientras que la meseta representa la pausa que se produce cuando se incorpora un ddNTP. A fuerza baja, el borde que cae de la curva muestra la actividad de exonucleasa que elimina la hebra que se acaba de sintetizar. Esta síntesis y las fases de exonucleasa se pueden repetir en ciclos.

Figura 3: Ilustración de secuenciación basado en la concentración de dNTP no balanceada. Registro en tiempo real de la síntesis en una horquilla de una sola molécula de 1,2 kbps con T4 ADN polimerasa (dTNP = dGNP = dCNP = 33 μ M, dATP = 20 nM, abrazadera y cargador de abrazadera) en un tampón comercial (tampón de reacción 5X: Tris-HCl 335 mM (pH 8,8 a 25°C), MgCl₂ 33 mM, DTT 5 mM, (NH₄)₂SO₄ 84 mM, Fermentas; <http://www.fermentas.com/en/products/all/modifying-enzymes/mesophilic-polymerases/ep006-t4-dna-polymerase>). A fuerza elevada (~ 21 pN), el pico de la curva muestra la existencia de una muestra de horquilla, y se podría usar para hibridar un cebador, si es necesario. A fuerza media (~ 11,7 pN), el borde ascendente de la curva muestra la síntesis procesiva de una nueva hebra complementaria mediante la T4 ADN polimerasa, mientras que la pausa transitoria corresponde a las presentes cuando se necesita un dATP. A fuerza baja (~ 1,6 pN), la caída muestra la escisión de la hebra recientemente sintetizada, debido a la activación de la actividad de exonucleasa de la T4 ADN polimerasa. Esta síntesis y el ciclo de escisión son repetibles.

Figura 4: Detección de la secuenciación sobre la base de la concentración de dNTP no balanceada a su secuencia complementaria en una horquilla. Figura 4A: Registro en tiempo real de síntesis en una horquilla de una sola molécula de 83 pb con T4 ADN polimerasa (dTNP = dGNP = dCNP = 33 μ M, dATP = 20 nM, abrazadera y cargador de abrazadera) en un tampón comercial (tampón de reacción 5X: Tris-HCl 335 mM (pH 8,8 a 25°C), MgCl₂ 33 mM, DTT 5 mM, (NH₄)₂SO₄ 84 mM, Fermentas; <http://www.fermentas.com/en/products/all/modifying-enzymes/mesophilic-polymerases/ep006-t4-dna-polymerase>). A fuerza elevada (~ 21 pN), el pico de la curva muestra la existencia de una muestra de horquilla, y se podría usar para hibridar un cebador, si es necesario. A fuerza media (~ 11,7 pN), el borde ascendente de la curva muestra la síntesis procesiva de una nueva hebra complementaria mediante la T4 ADN polimerasa, mientras que la pausa transitoria corresponde a las presentes cuando se necesita un dATP. A fuerza baja (~ 1,6 pN), la caída muestra la escisión de la hebra recientemente sintetizada, debido a la activación de la actividad de exonucleasa de la T4 ADN polimerasa. Esta síntesis y el ciclo de escisión son repetibles. El histograma de posiciones de estas pausas transitorias se deduce de forma similar a como se describe en la figura 1, y se indica la asignación de bases a los picos del cromatograma (flechas). La secuencia verdadera del molde se muestra en el lado derecho de la figura. Figura 4B: Existen dos maneras para medir el cambio de extensión. En primer lugar, se puede usar como referencia la Zzip; como alternativa, también podemos usar como referencia la posición del bucle de la horquilla.

Figura 5. Ilustración del diseño de la horquilla. El fragmento de ADN de interés se liga con un bucle de ADN (abásico y con bases de LNA) y dos fragmentos de ADN parcialmente complementarios (un ADN monocatenario con biotina en el extremo, un ADN bicatenario con Dig en el extremo), que eventualmente forman una horquilla que se puede unir a una perla superparamagnética cubierta con estreptavidina (DYNAL) mientras que el otro extremo se puede unir a un cubreobjetos de vidrio tratado con anti-dig.

Ejemplos experimentales

Preparación de ADN

Un fragmento de ADN bicatenario (bc) de secuencia desconocida y de un tamaño comprendido entre unas pocas decenas y unos pocos miles de pares de bases, se liga en una de sus extremidades a un bucle de ADN. Su otra extremidad se liga a un fragmento de ADNbc que permite la unión de sus dos hebras a superficies revestidas diferentemente. Por ejemplo, el extremo 3' libre de una hebra se puede marcar con biotina que permite la unión a perlas revestidas con estreptavidina, mientras que el extremo 5' en la hebra opuesta se puede marcar con digoxigenina que permite su unión a superficies revestidas con un anticuerpo anti-Dig. Este marcaje de los extremos se puede realizar de diversas formas conocidas por el experto en la materia, tal como el uso de transferasa terminal para añadir nucleótidos modificados con biotina (o dig) o la hibridación con oligonucleótidos marcados

adecuadamente.

Aparato de estiramiento por fuerza

- 5 Este constructo de ADN se incuba durante unos pocos minutos en una disolución de perlas adecuadas (por ejemplo, las revestidas con estreptavidina) a las que se une mediante uno de sus extremos marcados (por ejemplo biotina). Las perlas pueden ser transparentes si se usan más tarde pinzas ópticas para la manipulación, o magnéticas si se usan trampas magnéticas o pinzas para la manipulación.
- 10 El ensamblaje de perla-ADN se inyecta en una cámara fluidica cuya superficie se ha tratado para que se una al otro extremo marcado de la molécula (por ejemplo, una superficie revestida con anti-Dig para unirse al extremo marcado con Dig del ADN). Las perlas se anclan de este modo a la superficie vía una horquilla de ADN; véase la figura 1a. La distancia de la perla a la superficie se monitoriza entonces por diversos medios conocidos por el experto en la materia: por ejemplo, se pueden usar los anillos de difracción de su imagen en una cámara para deducir su distancia, o se puede usar la intensidad de la luz que dispersan (o emiten mediante fluorescencia) cuando se iluminan en un modo evanescente, para medir su distancia. Alternativamente, se puede medir el campo magnético que generan (usando un sensor magnético tal como los sensores GMR o Hall) para deducir su distancia a un sensor en la superficie de anclaje.
- 15
- 20 Para ejercer tracción sobre la molécula de ADN que ancla las perlas a la superficie, se han descrito diversas técnicas. La forma de realización preferida usa una trampa magnética para ejercer tracción sobre perlas superparamagnéticas ancladas a una superficie mediante una horquilla de ADN como se describe anteriormente. En esta configuración, se usan pequeños imanes colocados encima de la muestra para aplicar una fuerza constante sobre la perla anclada, cuya posición se puede determinar con una exactitud de < 1 nm (dependiendo de la fuerza de tracción y la disipación debido al arrastre hidrodinámico). En esta serie de experimentos, se usó el aparato descrito en las patentes US n^{os} 7.052.650 y 7.244.391. Además, excepto que se indicase de otro modo, los experimentos dados a conocer en la presente memoria se llevaron a cabo en Tris 25 mM pH 7,5, KAc 150 mM, MgCl₂ 10 mM, 0,2% de BSA. En cada caso, la horquilla sujetadora se puede abrir completamente de forma mecánica tirando de las perlas con una fuerza mayor que alrededor de 16 pN. La reducción de la tensión sobre la molécula por debajo de alrededor de 11 pN permite que la horquilla se vuelva a cerrar espontáneamente en forma de cremallera (la transición de la apertura de la cremallera es reversible aunque histerética). Si, durante la fase de cremallera abierta, se produce la unión de una molécula en disolución (tal como una proteína u oligonucleótidos complementarios de ADN, ARN, LNA o PNA) al ADN monocatenario (mc) estirado, esta molécula bloqueará transitoriamente el cierre nuevamente de la cremallera de la horquilla cuando la fuerza se reduce por debajo de 11 pN. El principio del ensayo es el intercambio entre dos fuerzas: una grande F_{open} para abrir la horquilla y una más pequeña F_{test} usada para permitir el cierre nuevamente de la cremallera y medir la extensión de la molécula en bloques transitorios. La posición de bloqueo está relacionada con la secuencia mediante una relación lineal entre la extensión total y el bloqueo. Para la mejor exactitud, la extensión total se mide preferiblemente a la fuerza de ensayo F_{test} . Esto se logra diseñando el bucle de horquilla de manera que requiere una fracción de un segundo para volverse a enrollar una vez que la fuerza se reduce desde F_{open} a F_{test} .
- 30
- 35
- 40

La posición de hibridación de un oligonucleótido se puede medir con una resolución de pares de bases

45 Midiendo la extensión de la molécula de ADN (la distancia de la perla a la superficie) durante una de estas pausas de cierre nuevamente de la cremallera, es posible determinar la posición del bloqueo con una precisión nanométrica (1 nm corresponde a la distancia abarcada por dos nucleótidos (1 pb) en un ADNmc bajo una fuerza de 10 pN). La configuración de apertura de la cremallera presenta la relación más grande de extensión a par de bases (en ADNbc, la relación es solamente 0,34 nm por pb).

50 La exactitud de esta medida está limitada por dos contribuciones de ruido:

- La exactitud del método de medida,
- El movimiento browniano de la perla.

55 Se han usado diferentes técnicas para medir la posición vertical de la perla. Una de las más simples se basa en microscopía de video (patentes U.S. n^{os} 7.052.650 y 7.244.391). Los resultados en la figura 1 se obtuvieron con este método; la resolución típica alcanza 1 nm para un promedio de 1 segundo. Se han demostrado otros métodos con mejor resolución, tales como iluminación mediante láser con sensores PSD, que alcanza 0,1 nm de resolución (Greenleaf y Block, Science, 313: 801, 2006) e iluminación mediante onda evanescente (Singh-Zocchi et al., Proc Natl Acad Sci U S A., 100(13): 7605-7610, 2003, Liu et al., Biophys J., 96(9): 3810-3821, 2009).

60

La limitación intrínseca en la resolución está dada por las fluctuaciones brownianas de la perla que ejerce tracción en una molécula de ADNmc. $\langle x^2 \rangle = 4k_B T \Delta f (6 \pi \eta r) / k_{ADNmc}^2(F)$, en la que $k_{ADNmc}(F)$ es la rigidez de una molécula de ADNmc, k_B es la constante de Boltzman, T la temperatura absoluta, η la viscosidad del agua, r el radio de la perla y Δf es el intervalo de frecuencia de la medida. $k_{ADNmc}(F = 10 \text{ pN}) = 0,05/Nb$ (N/m), en el que Nb es el número de

65

bases del ADNmc. Para la horquilla de 84 pb, esto conduce a 0,04 nm de ruido a lo largo de un promedio de 1 segundo ($\Delta f = 1$ Hz). El ruido más grande en Fig. 1 ($\sigma \sim 1$ nm) es esencialmente debido al dispositivo de medida, no a las fluctuaciones intrínsecas. El ruido browniano intrínseco aumenta con el tamaño de la horquilla: una horquilla de 1200 pb conduce a un ruido de 0,6 nm cuando se promedia durante 1 segundo.

5 Diagnóstico y secuenciación mediante detección mecánica de la polimerización

Se ha mostrado que la T4 ADN polimerasa puede replicar una horquilla de ADN cuando la fuerza es suficientemente elevada para desestabilizar suficientemente la horquilla (F_{test}). Al igual que en la secuenciación de Sanger clásica, la incorporación de un ddNTP específico evitará el alargamiento adicional de la hebra naciente mediante la T4 ADN polimerasa. En nuestro método, este bloqueo se puede identificar fácilmente, como se muestra en la Fig. 6. Al disminuir la fuerza, se activa la actividad de exonucleasa de la T4 AND polimerasa, y la enzima corta la hebra nuevamente sintetizada. De este modo, mediante ciclos repetidos de síntesis y escisión, la molécula (horquilla) se puede secuenciar identificando las posiciones de bloqueo primero en presencia de ddATP, y después en presencia de cada uno de los otros ddNTPs, es decir ddTTP, ddCTP, y ddGTP.

De forma similar, la molécula de horquilla bicatenaria se puede secuenciar en un tampón que comprende un déficit de uno de los cuatro dNTPs en comparación con los otros, es decir, este dNTP está presente en una concentración muy baja en comparación con los otros. De este modo, siempre que la T4 ADN polimerasa, durante la polimerización alcance una posición que requiera la adición del nucleótido limitante, se produce una pausa transitoria, como se ejemplifica en la figura 7. Como se describe en la presente memoria anteriormente, la hebra nuevamente sintetizada se puede cortar disminuyendo la fuerza, lo que da como resultado la activación de la actividad de exonucleasa de la enzima. De este modo, mediante ciclos repetidos de síntesis y escisión, la molécula (horquilla) se puede secuenciar identificando las posiciones de pausa en primer lugar en presencia de concentraciones bajas de dATP (por ejemplo), y después por turnos de cada uno de los otros dNTPs, es decir, dTTP, dCTP, y dGTP.

Listado de secuencias

30 <110> CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) ECOLE NORMALE SUPERIEURE

<120> MÉTODO DE SECUENCIACIÓN DE ADN MEDIANTE POLIMERIZACIÓN

35 <130> 358481D28618

<150> EP 10305563.8

<151> 2010-05-27

40 <150> US 61/377,621

<151> 27-08-2010

<160> 4

45 <170> PatentIn version 3.3

<210> 1

<211> 21

<212> ADN

<213> artificial

50 <220>
<223> Oligonucleótido

<400> 1

55 aactgaccaa acgtcgggtg g 21

<210> 2

<211> 16

<212> ADN

60 <213> artificial

<220>

<223> Secuencia de tallo de horquilla

65 <400> 2

ggattcggg gtctct 16

- <210> 3
 <211> 42
 <212> ADN
 5 <213> artificial

 <220>
 <223> Secuencia de bucle de horquilla

 10 <220>
 <221> característica diversa
 <222> (20)..(20)
 <223> n representa una base abásica

 15 <220>
 <221> característica diversa
 <222> (22)..(22)
 <223> n representa una base abásica

 20 <220>
 <221> característica diversa
 <222> (24)..(24)
 <223> n representa una base abásica

 25 <220>
 <221> característica diversa
 <222> (26)..(26)
 <223> n representa una base abásica

 30 <400> 3
 tcgcgctga tcgtccactn tntntntagt ggacgatcag gc 42

 <210> 4
 <211> 41
 35 <212> ADN
 <213> artificial

 <220>
 <223> Secuencia de tallo de horquilla

 40 <400> 4
 aggaagagac ccgcgaatcc cccaccgacg tttggtcagt t 41

REIVINDICACIONES

1. Método para la determinación de una secuencia de ácido nucleico, comprendiendo dicho método las etapas de:
 - 5 a) desnaturalizar una molécula de ácido nucleico bicatenario correspondiente a dicha secuencia de ácido nucleico, en el que al menos una de las bases de una de las hebras del ácido nucleico bicatenario está unida directa o indirectamente a un soporte, y en el que al menos una de las bases de la otra hebra del ácido nucleico bicatenario está unida a un soporte móvil;
 - 10 b) hibridar una molécula de ácido nucleico monocatenario, "el cebador", con dicha molécula de ácido nucleico bicatenario desnaturalizada;
 - c) aplicar una tensión al cebador hibridado/molécula de ácido nucleico bicatenario obtenido en b);
 - 15 d) incubar el cebador hibridado/molécula de ácido nucleico bicatenario obtenido en la etapa b) con una polimerasa en condiciones que conducirán a por lo menos una pausa en la replicación; y
 - e) determinar la posición de dicha pausa en la replicación con respecto a un extremo del ácido nucleico bicatenario, comprendiendo dicha determinación las etapas de:
 - 20 • medir la distancia (z) entre los dos extremos de la molécula de ácido nucleico bicatenario que están unidos al soporte;
 - 25 • medir la distancia (z_{high}) entre los dos extremos de la molécula de ácido nucleico bicatenario que están unidos al soporte, cuando dicha molécula de ácido nucleico bicatenario está desnaturalizada; y
 - comparar z y z_{high} , y
 - 30 • determinar la posición de la pausa,
en el que dicha molécula de ácido nucleico bicatenario es una horquilla.
2. Método según la reivindicación 1, en el que el ácido nucleico bicatenario está desnaturalizado en la etapa a) alejando los soportes.
- 35 3. Método según la reivindicación 2, en el que se aplica una fuerza física superior o igual a 15 pN, preferiblemente superior o igual a 17 pN, más preferentemente superior o igual a 18 pN, a la molécula bicatenaria alejando los soportes.
- 40 4. Método según cualquiera de las reivindicaciones anteriores, en el que dicha tensión de la etapa c) está comprendida entre 12 pN y 10 pN.
5. Método según cualquiera de las reivindicaciones anteriores, en el que las etapas a) a e) se repiten varias veces para acumular mediciones e incrementar la relación señal/ruido.
- 45 6. Método según cualquiera de las reivindicaciones anteriores, en el que el cebador se extiende mediante la actividad de dicha polimerasa, y el proceso de extensión se detiene antes de que dicha polimerasa alcance el bucle de la molécula de ácido nucleico bicatenario.
- 50 7. Método según cualquiera de las reivindicaciones anteriores, que comprende una etapa adicional de reducir la fuerza física aplicada a la molécula de ácido nucleico bicatenario a menos de o igual a 5 pN.
8. Método según la reivindicación 7, en el que se activa la actividad de exonucleasa de dicha polimerasa.
- 55 9. Método según cualquiera de las reivindicaciones anteriores, que comprende una etapa adicional de desensamblar la hebra sintetizada recientemente.
10. Método según cualquiera de las reivindicaciones anteriores, en el que se repiten las etapas a) a e).
- 60 11. Método según cualquiera de las reivindicaciones anteriores, en el que la polimerasa es una enzima selectiva de nucleótido, y las condiciones de la etapa d) incluyen utilizar una mezcla de reacción que altera la velocidad para el movimiento procesivo de la enzima para un nucleótido específico.
- 65 12. Método según cualquiera de las reivindicaciones 1-10, en el que las condiciones de la etapa d) incluyen utilizar una mezcla de reacción que comprende didesoxinucleótidos además de desoxinucleótidos.

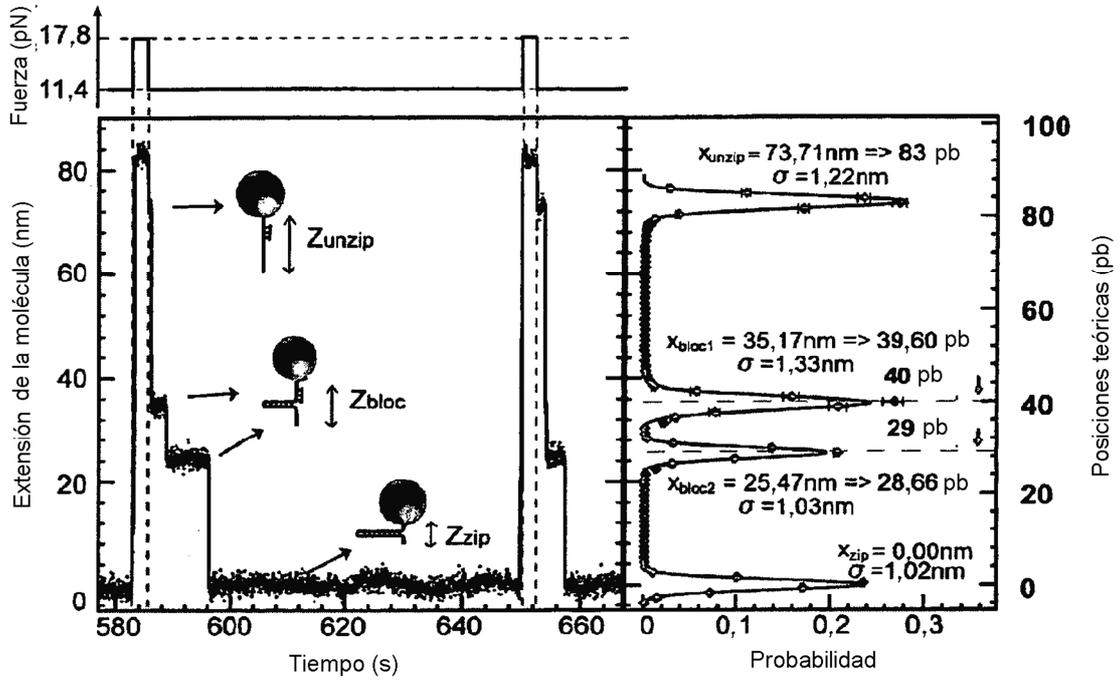


Figura 1

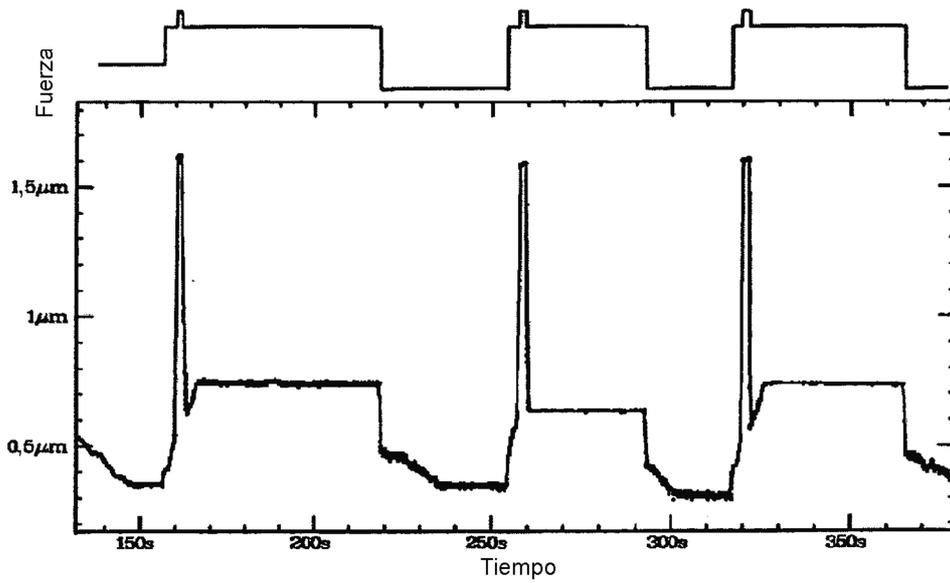


Figura 2

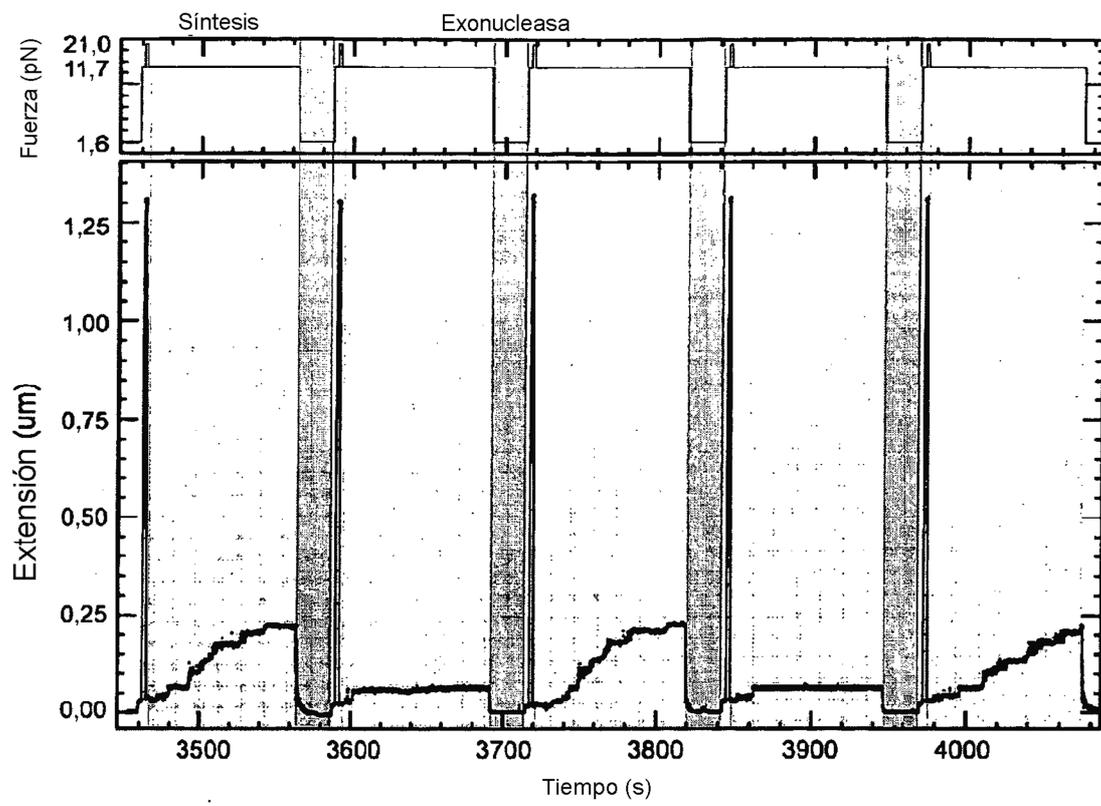
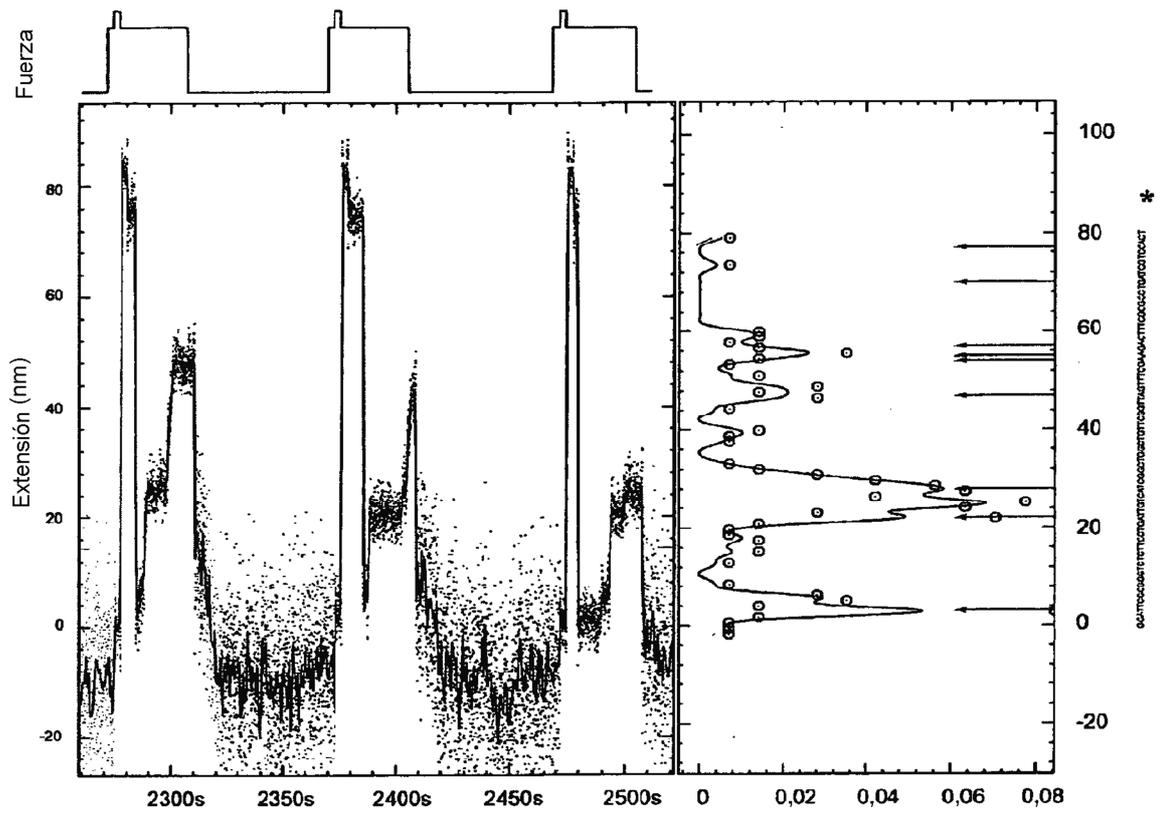
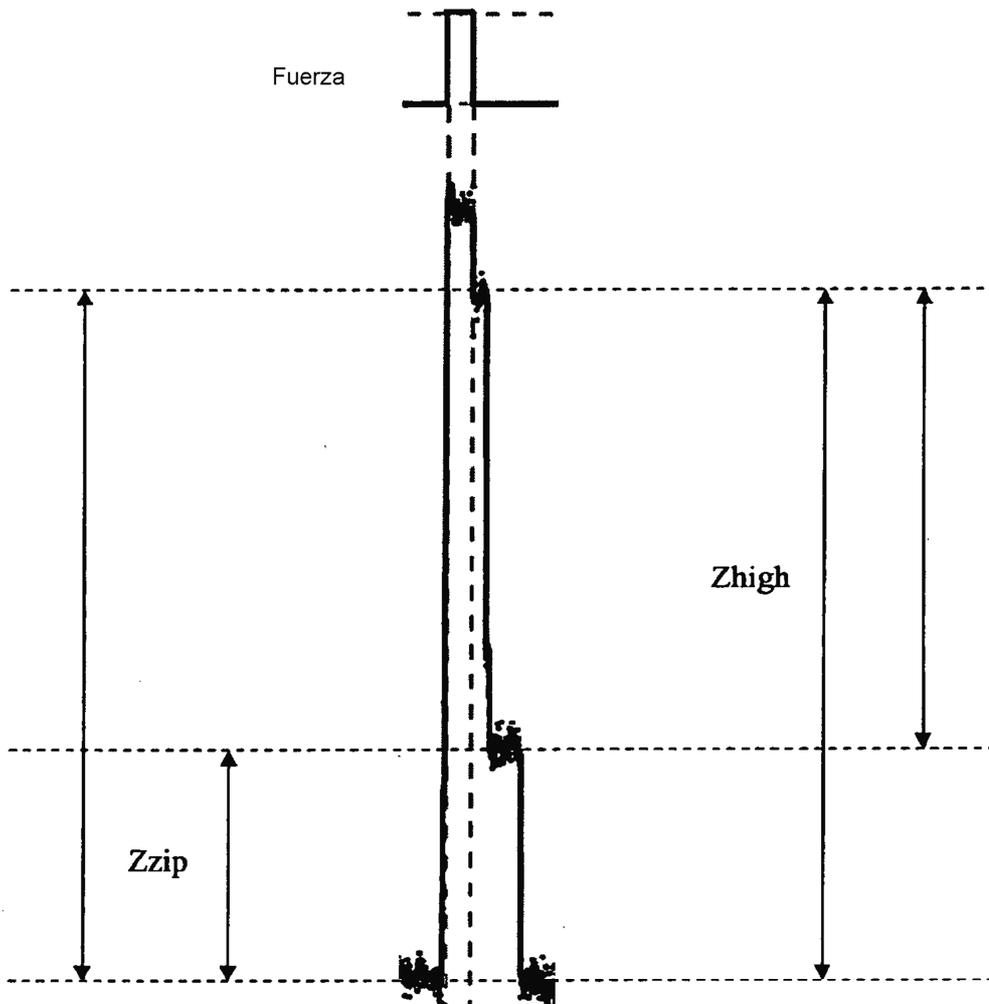


Figura 3



*
 GGATTCGCGGGTCTCTTCCTGATTGTCATCGGCTGGCTGTTCCGTTAGTTTCGAAGACTTTCGCC
 TGATCGTCCACT

Figura 4A



Úsese Zzip como referencia

Úsese Zhigh (retraso del bucle) como referencia

Figura 4B

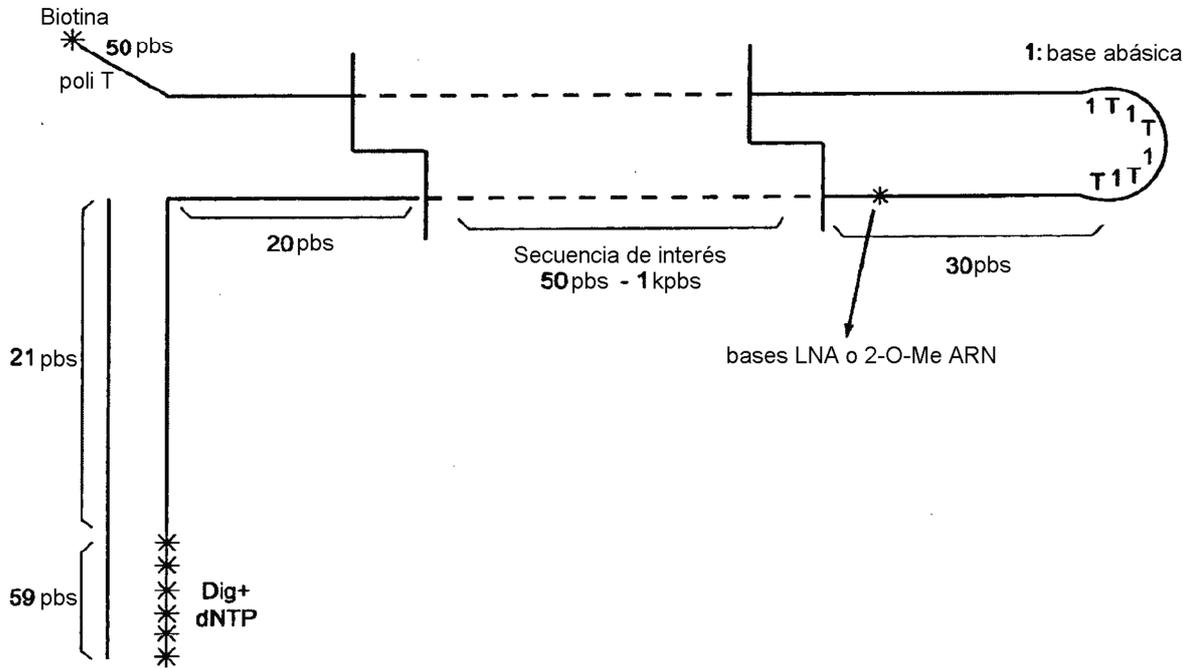
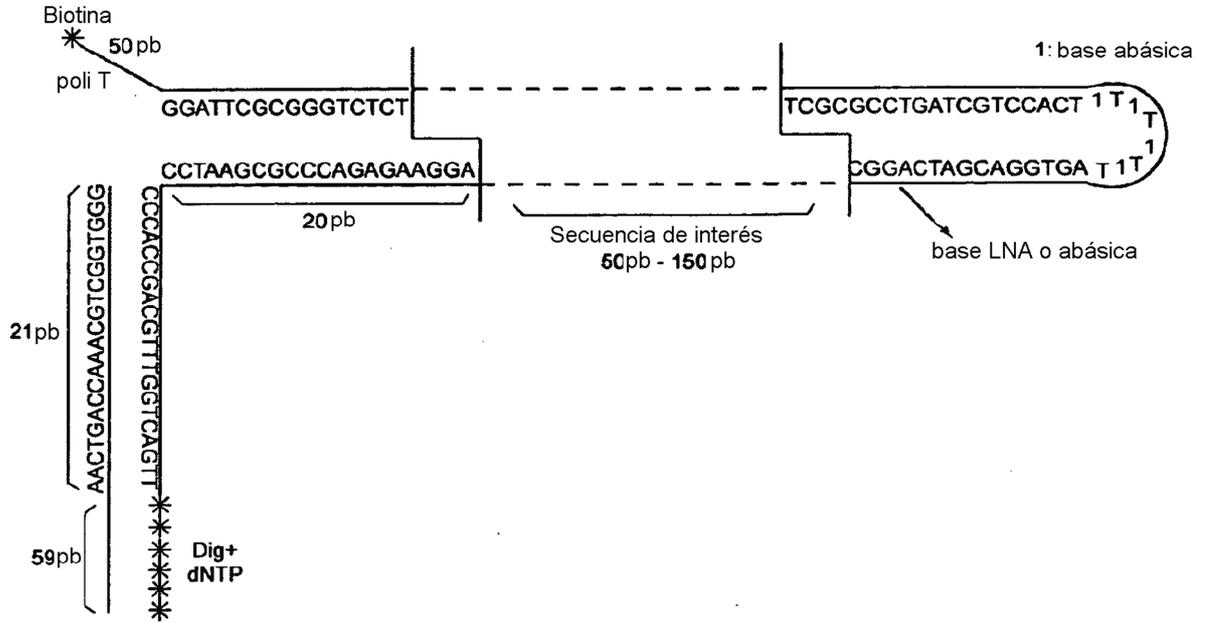


Figura 5