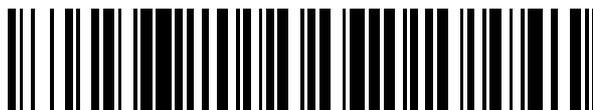


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 542 429**

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 11721549 (1)**

97 Fecha y número de publicación de la concesión europea: **06.05.2015 EP 2576818**

54 Título: **Procedimiento de secuenciación de ADN mediante hibridación**

30 Prioridad:

27.08.2010 US 377621 P
27.05.2010 EP 10305564

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
05.08.2015

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

74 Agente/Representante:

CURELL AGUILÁ, Mireia

ES 2 542 429 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

Procedimiento de secuenciación de ADN mediante hibridación.

5 La presente invención se refiere a un procedimiento rápido para determinar una secuencia de un ácido nucleico, ADN o ARN, que es útil, particularmente, para secuenciar un ácido nucleico desconocido o, alternativamente para la detección de una secuencia de ácido nucleico específica para diagnóstico.

10 Actualmente la determinación de una secuencia de ácido nucleico se encuentra en el centro de la biología molecular. Por ejemplo, una amplia gama de los fenómenos biológicos pueden evaluarse mediante una secuenciación del ADN con alta capacidad de procesamiento, por ejemplo, la variación genética, la expresión del ARN las interacciones ADN-proteicas y la conformación cromosómica (véanse como algunos 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 específica de ADN en una muestra fisiológica constituye, actualmente, la línea más importante del desarrollo de procedimientos diagnósticos, por ejemplo, para identificar la probabilidad de las bacterias para el desarrollo de resistencia a los antibióticos, anomalías genéticas, los riesgos de cáncer asociados con modificaciones genéticas y las infecciones víricas, por ejemplo, infecciones asociadas con el HIV o con varios virus de la hepatitis (véase, 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 Immunity.*, 77(7): 2840-2848, 2009).

25 La secuenciación de los ácidos nucleicos se lleva a cabo actualmente, de forma principal, mediante aplicaciones semiautomatizadas basadas capilarmente en la bioquímica de Sanger. El procedimiento clásico incluye una etapa amplificadora del ADN de interés, seguido por una de "secuenciación cíclica", en la que cada ronda de extensión del cebador, finaliza estocásticamente por la incorporación de dideoxinucleótidos marcados fluorescentemente (ddNTP). La secuencia se determina mediante separación electroforética de alta resolución de los productos de extensión monocatenarios señalizados en su extremo, en un gel polimérico basado capilarmente. La electroforesis simultánea en 96 o 384 capilares independientes proporciona un nivel limitado de paralelización.

35 La alta demanda de secuenciación de bajo coste ha concitado el desarrollo de tecnologías de secuenciación de alta capacidad de procesamiento que paralelizan el proceso de secuenciación que da lugar a miles de millones de secuencias a la vez (Shendure & Ji, *Nat Biotechnol.*, 26(10):1135-45, 2008). Las tecnologías de secuenciación de alto procesamiento pretenden disminuir el costo de la secuenciación del ADN más allá de lo que es posible con los procedimientos estándares de terminador colorante. Actualmente, esta alta capacidad de procesamiento, se alcanza con sacrificios sustanciales en la longitud y seguridad de las lecturas individuales cuando se comparan con la secuenciación de Sanger. Ejemplos de dichos nuevos procedimientos incluyen las tecnologías 454 y la Solexa. Estas tecnologías permiten la secuenciación de bibliotecas génicas de genomas enteros sin clonar en *E. Coli* o en cualquier célula huésped. Bibliotecas de fragmentos cortos de ADN flanqueados por adaptadores, capturados en la superficie de perlas, se amplifican mediante PCR de emulsión. La secuenciación se lleva a cabo utilizando la síntesis mediada por la ADN polimerasa. En el procedimiento 454 (conocido también como "pirosecuenciación"), la variedad se presenta con cada uno de los cuatro dNTPs, secuencialmente, y la cantidad incorporada se monitoriza mediante detección luminométrica del pirofosfato liberado. Una diferencia clara entre este procedimiento y el de Solexa, es que éste utiliza nucleótidos al final de la cadena. La señalización fluorescente en la base final puede eliminarse para dejar un extremo 3' desbloqueado, haciendo que la finalización de la cadena sea un proceso reversible. La tecnología SOLiD se basa en la unión de las sondas dibásicas señalizadas fluorescentemente a un cebador secuenciador hibridado a una secuencia adaptadora dentro del molde clonalmente amplificado de la biblioteca. La especificidad de la sonda dibásica se obtiene informándose de cada 1ª y 2ª base en cada reacción de unión. Ciclos múltiples de unión, detección y fragmentación, se llevan a cabo con el número de ciclos que determina la longitud eventual de lectura. En contraste con las 3 tecnologías previas, las cuales requieren todas una primera etapa de amplificación, la plataforma Helicos permite la secuenciación de moléculas únicas de ADN. Esta tecnología se basa en la utilización de un sistema de detección muy sensible de la incorporación de nucleótidos fluorescentes para interrogar directamente a las moléculas únicas del ADN mediante secuenciación por síntesis.

60 Dichos procedimientos se describen en, por ejemplo, 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 PCT nº WO2007111924; solicitud de patente US nº 2008/0020392; solicitud de patente PCT nº WO 2006/084132; solicitud de patente U.S. No. US 2009/0186349; solicitud de patente US nº US 2009/0181860; solicitud de patente US nº 2009/0181385; solicitud de patente nº US 2006/0275782; patente europea EP-B1-1141399; Shendure & 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 PCT WO 2010/016937 proporciona un procedimiento para determinar la secuencia de un ácido nucleico sometido a una fuerza.

Sin embargo, todos los procedimientos que se han desarrollado por el momento presentan serios inconvenientes. En particular, todos utilizan nucleótidos señalizados (por ejemplo fluorescentes) contribuyendo así a aumentar seriamente los costes totales. Además, todos estos nuevos procedimientos no funcionan (la plataforma Helicos) requiere amplificación de la secuencia diana antes de la secuenciación, lo que, por una parte, emplea tiempo, y por otra, aumenta la probabilidad de errores y es altamente propenso a la contaminación. Los procedimientos de la presente invención se definen en las reivindicaciones añadidas 1-21. El procedimiento según la presente invención, que se basa en las técnicas físicas y en tratamientos electrónicos, difiere de los enfoques habituales, que son químicos o bioquímicos. Sus ventajas son numerosas:

- 1) Permite la secuenciación de una molécula única y por tanto no requiere una etapa previa de amplificación (por ejemplo, mediante PCR).
- 2) Es mucho más barato que los procedimientos de la técnica, ya que se utilizan moléculas de ácidos nucleicos monocatenarios, que son mucho menos caros que los nucleótidos señalizados (con fluoróforos o algunos otros grupos). Además la cantidad de las moléculas ácido nucleicas monocatenarias estándar se reduce a un mínimo ya que la secuencia de una molécula de ácido nucleico bicatenario única está determinada. Además en algunas formas de realización por lo menos, las cadenas de sondeo pudieran reutilizarse, pues no se consumen en el proceso de secuenciación.
- 3) Permite determinar la localización (en pares de bases) de una molécula de ácido nucleico monocatenario emparejada a lo largo de un ácido nucleico bicatenario, midiendo la distancia entre los dos extremos de dicha molécula de ácido nucleico bicatenario.
- 4) Permite determinar, en un ensayo de renaturalización las distintas posiciones de hibridación de un oligonucleótido sobre una horquilla de un ácido nucleico bicatenario dado.
- 5) La medición puede repetirse periódicamente en una segunda escala de tiempo, llevando de este modo a la eliminación de positivos falsos (hibridaciones falsas parciales), a estadísticas mejoradas, permitiendo una reducción significativa en derivas instrumentales.
- 6) El experimento puede repetirse muchas veces sobre idéntica molécula, mejorando de esta forma las estadísticas y la solvencia de la medición, ya que el ácido nucleico monocatenario hibridado puede sufrir eyección (por ejemplo, reduciendo la fuerza o la fuerza iónica o utilizando una helicasa o una nucleasa) durante la realización de la fase de renaturalización.
- 7) Permite la secuenciación paralela de varias moléculas ácido nucleicas bicatenarias, ya que cada molécula puede manipularse independientemente de las otras.

Esta invención se refiere a un procedimiento para la determinación de una secuencia de ácido nucleico, en la que la renaturalización de un ácido nucleico bicatenario desnaturalizado que corresponde a dicha secuencia de ácido nucleico, está bloqueada.

Por "determinación de una secuencia de ácido nucleico", se hace referencia en la presente memoria, no sólo el desciframiento de la sucesión actual de bases en un ácido nucleico, sino también todas las actividades que llevan, 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 ácida nucleica o en la detección de una diferencia entre las secuencias de dos moléculas de ácidos nucleicos distintos.

La invención se basa en la observación de que las dos cadenas de un ácido nucleico bicatenario desnaturalizado, se rehibridarán bajo condiciones apropiadas. Si algunas moléculas se unen a cualquiera de las cadenas de dicho ácido nucleico bicatenario desnaturalizado durante la etapa de renaturalización, la rehibridación será sólo parcial. Los inventores encontraron ahora que, bajo ciertas condiciones, esta pausa en la rehibridación, permanente o transitoria, puede utilizarse para obtener información sobre la secuencia contenida en la molécula de ácido nucleico bicatenario desnaturalizado. Según la invención, es posible detectar un bloqueo de la rehibridación de la molécula de ácido nucleico bicatenario; los parámetros físicos (por ejemplo, la duración del bloqueo, la posición de éste sobre la molécula del ácido nucleico bicatenario) asociada con este bloqueo, permitirá entonces la determinación de la secuencia del ácido nucleico.

La presente invención se refiere, de este modo, a un procedimiento para determinar una secuencia de ácido nucleico, el cual incluye una etapa de detección de un bloqueo de la renaturalización de un ácido nucleico bicatenario desnaturalizado, que corresponde a dicha secuencia de ácido nucleico. Por "desnaturalización", se hace referencia en la presente memoria al procedimiento de separación de las cadenas de una molécula de ácido nucleico bicatenario, que tiene lugar cuando la mayoría de los enlaces de hidrógeno entre dichas cadenas están rotos. El proceso de desnaturalización da lugar a una molécula de ácido nucleico desnaturalizado, mediante el cual se hace referencia en la presente memoria a las dos cadenas complementarias separadas que provienen de la

desnaturalización de una molécula de ácido nucleico bicatenario. Por “renaturalización” se hace en la presente memoria referencia al proceso por el que dos cadenas complementarias separadas vuelven a formas mediante la hibridación, una doble hélice. Tal como se utiliza en la presente memoria “hibridación” es el procedimiento para establecer una interacción no covalente y secuencia-específica entre dos o más cadenas complementarias de ácidos nucleicos, a un híbrido único.

Existen diversas posibilidades conocidas, para el experto en la materia, para desnaturalizar el ácido nucleico. En una forma, la más preferida, las dos cadenas se separan sometiéndolas a una fuerza física. Por ejemplo, los extremos libres de dicho ácido nucleico bicatenario pueden ser arrastrados aparte, rompiendo así todos los enlaces entre las bases emparejadas, y abriendo el ácido nucleico bicatenario.

Así, en una forma de realización, el procedimiento de la invención se refiere a la determinación de una secuencia de ácido nucleico, el cual procedimiento comprende las etapas de:

- desnaturalización de una molécula de ácido nucleico bicatenario que corresponde a dicha secuencia de ácido nucleico, aplicando una fuerza física a dicha molécula, y
- detectando un bloqueo de la renaturalización del ácido nucleico bicatenario.

En este tipo de procedimiento de determinación secuencial, puede ser ventajoso, con objeto de facilitar el reemparejamiento, disponer que los extremos libres del ADN bicatenario (es decir, los extremos que no están unidos al soporte), se junten entre ellos de forma covalente o casi, antes de ser arrastrados aparte. En una forma de realización preferida, la molécula de ácido nucleico bicatenario es una horquilla. Si se quiere que el ácido nucleico bicatenario se represente diagramáticamente en el contexto de esta invención, es posible que se compare a una cremallera, la cual está abierta (o cerrada): la desnaturalización del ácido nucleico bicatenario es de “descompresión”; es la renaturalización, es la “recompresión”.

Los inventores observaron que, bajo ciertas condiciones, cuando una molécula se une a la molécula de ácido nucleico bicatenario desnaturalizado, la renaturalización de dicha molécula, se bloquea. La molécula unida puede ser de cualquier tipo de molécula con afinidad por una secuencia específica sobre dicha molécula de ácido nucleico bicatenario desnaturalizada, por ejemplo, un ácido nucleico, una proteína o una molécula pequeña. Sin embargo, es preferible utilizar un ácido nucleico monocatenario, ya que dicho ácido nucleico monocatenario puede hibridarse con una secuencia complementaria en una de las cadenas del ácido nucleico bicatenario desnaturalizado. Este ácido nucleico monocatenario puede ser algo largo, teniendo en cuenta que es lo bastante largo para bloquear el proceso de renaturalización. Preferentemente, la longitud del ácido nucleico monocatenario único estará comprendida entre 3 y 20 nucleótidos, más preferentemente, entre 7 y 15 e incluso, más preferentemente, entre 8 y 12.

El ácido nucleico monocatenario de la invención, puede ser, en particular una molécula de ADN o ARN, natural o modificada. Dicho ácido nucleico monocatenario puede estar constituido también por nucleótidos modificados, tal como el ácido nucleico cerrado (LNA), que son nucleótidos en los que la mitad ribosa está modificada con un puente extra, que une el 2' oxígeno y el 4' carbono, o el ácido nucleico peptídico (PNA), en el que la estructura está compuesta de unidades repetitivas de N-(2-aminoetil)-glicina, unidas mediante enlaces peptídicos.

Cuando una molécula de ácido nucleico monocatenario se añade a un ácido nucleico bicatenario desnaturalizado antes de la renaturalización, un bloqueo de la rehibridación indica que 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.

Así, el procedimiento de la invención se refiere también a un procedimiento para determinar una secuencia de ácido nucleico, el cual procedimiento incluye las etapas de:

- a) desnaturalización de una molécula de ácido nucleico bicatenario que corresponde a dicha secuencia de ácido nucleico, aplicando una fuerza física a dicha molécula;
- b) provisión de una molécula de ácido nucleico monocatenario;
- c) renaturalización de dicha molécula ácido nucleico bicatenaria en la presencia de dicha molécula de ácido nucleico monocatenario; y
- d) detección de un bloqueo de la renaturalización del ácido nucleico bicatenario.

La invención se aplica a cualquier tipo de ácido nucleico bicatenario. Muy a menudo, el ácido nucleico bicatenario será el ADN, pero se entiende que la invención también se aplica a dúplex ADN monocatenarios-ADN monocatenarios, perfectamente apareados o no, perfectamente apareados o alternativamente a dúplex ADN monocatenarios-ARN monocatenarios, apareados perfectamente o perfectamente apareados, o alternativamente, a

dúplex ARN monocatenarios-ARN monocatenarios, apareados perfectamente o no. Además el dúplex puede consistir en, por lo menos, un reemparejamiento parcial de dos cadenas únicas obtenidas de muestras de orígenes distintos. Finalmente, la invención se aplica también a las estructuras secundarias de un único ADN monocatenario o de un ARN único monocatenario.

5 En una configuración típica, las moléculas ácido nucleicas bicatenarias pueden anclarse específicamente a dos sustratos sólidos (por ejemplo, portaobjetos del microscopio, micropipeta, micropartícula). Uno de los extremos puede unirse 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 los dos extremos del ácido nucleico bicatenario cuando los soportes se mueven. Cuando la tensión es mayor que un valor umbral, las dos cadenas se separan y la molécula de ácido nucleico se desnaturaliza. La tensión aplicada está preferentemente por encima de, o es igual a 15 pN; más preferentemente, es superior o igual a 16 pN, más preferentemente es incluso superior o igual a 17 pN; en un aspecto mucho más preferido, es superior o igual a 18 pN. Esta fuerza puede variar con la temperatura, el tipo de nucleótido y de tampón, pero el experto en la materia adaptará fácilmente dicha fuerza con respecto a estos parámetros, con objeto de conseguir la separación de las dos cadenas. Por otra parte, cuando la tensión disminuye por debajo de un valor mínimo, las dos cadenas del ácido nucleico bicatenario desnaturalizado, pueden volverse a hibridizar. Para conseguir la rehibridización de dichas dos cadenas, se aplica preferentemente una tensión inferior a o igual a 12 pN; y más preferentemente, es menor o igual a 11 pN; incluso, más preferentemente, es inferior o igual a 10 pN. Muy preferentemente, el ácido nucleico bicatenario es una horquilla. Tal como se utiliza en la presente memoria, "horquilla" hace referencia a una doble hélice, en la que el extremo 5' de una cadena se une físicamente al extremo 3' de la otra cadena, mediante un bucle no emparejado. Dicha unión física puede ser covalente o no covalente. Preferentemente, dicha unión física es un enlace covalente. Así, una horquilla consta de un tronco bicatenario y de un bucle monocatenario impar. En una horquilla, los extremos de las dos cadenas que no están comprometidos en el bucle, son libres, y pueden, de esta forma se arrastrados a parte. Esto conlleva un desemparejamiento del ácido nucleico bicatenario, dando lugar así a una molécula de ácido nucleica bicatenaria desnaturalizada. Es posible abrir completamente una molécula horquilla de un ácido nucleico bicatenario, arrastrando cada extremo de dicha molécula ácido nucleica, con una fuerza superior a la de un valor umbral. Cuando la tensión aplicada a la molécula disminuye por debajo de un valor mínimo, la molécula de ácido nucleico se rehibridiza para (volver a) formar una horquilla. La presencia de una molécula de ácido nucleico monocatenario hibridizada a una de las cadenas ácido nucleicas conduce a una pausa en la rehibridización. Por tanto, la detección de dicha pausa indica que la molécula de ácido nucleico monocatenario incluye una secuencia que es complementaria, por lo menos, a parte del tronco bicatenario.

35 Es ventajoso a este respecto diseñar la secuencia y longitud del bucle, de forma que la horquilla se vuelva a plegar después de un corto intervalo, por ejemplo, 1s. A este respecto, se han descrito procedimientos en la técnica anterior, por ejemplo, en Woodside et al., *Proc. Natl. Acad. Sci. U.S.A.*, 103(16): 6190-6195, 2006). Cuando la fuerza disminuye desde la apertura al valor de la prueba, la extensión de la horquilla abierta varía, debido a la elasticidad del ADN monocatenario. El ligero retraso antes del plegamiento de la horquilla, permite al usuario determinar la extensión de la horquilla, con la misma fuerza que la utilizada para detectar el estado de bloqueo.

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

45 Se conocen técnicas que permiten que los extremos libres del ácido nucleico bicatenario se unan conjuntamente, y, en lo que sigue algunas, se describirán con más detalle.

50 Por determinación del bloqueo, se hace referencia en la presente memoria a la determinación de los parámetros físicos asociados con aquél. El más útil de estos parámetros es la posición del bloqueo sobre la molécula de ácido nucleico bicatenario, la cual posición corresponde a la de hibridación de la molécula de ácido nucleico monocatenario sobre la molécula de ácido nucleico bicatenario. Los inventores han descubierto, verdaderamente, que la posición sobre el ácido nucleico bicatenario en la que tiene lugar la pausa en la renaturalización, puede determinarse de modo preciso: la utilización de una horquilla permite al experto en la materia unos medios para determinar la distancia física entre los dos extremos libres de la horquilla en cualquier momento durante el proceso de desnaturalización/ renaturalización.

55 Por "extremo libre", se hace referencia en la presente memoria al extremo de una cadena que no está covalentemente unido a un extremo de la otra cadena; tal como se ha explicado anteriormente, estos extremos libres pueden unirse, cada uno de ellos a una superficie distinta. Por ejemplo, una de estas superficies puede ser móvil, mientras que la otra puede presentar un movimiento menor. El experto en la materia apreciará fácilmente de este modo que, para realizar la medición de la distancia entre los extremos libres del ácido nucleico bicatenario de la horquilla, es posible, simplemente, medir la distancia entre las dos superficies.

60 Esta distancia es máxima ($Z_{\text{high}}(F_{\text{open}})$) cuando la molécula de la horquilla está completamente desnaturalizada, ya que el ácido nucleico de la horquilla está entonces completamente extendido; es mínima ($Z_{\text{low}}(F_{\text{test}})$) cuando dicha molécula de la horquilla está completamente renaturalizada. Es ventajoso realizar todas las comparaciones de la longitud con la misma fuerza F_{test} , de forma que el ácido nucleico monocatenario posea idénticas propiedades

elásticas. Utilizando el retraso en el cierre del bucle, el experto en la materia puede medir Z_{high} (F_{test}). Así, la distancia entre los dos extremos libres cuando el proceso de renaturalización está en pausa temporal, puede medirse; tal como se esperaba, esta distancia z está comprendida entre Z_{high} y Z_{low} (midiéndose todas las z con $F = F_{\text{test}}$). Se deduce inmediatamente que la distancia z varía con la localización en la molécula de la horquilla de la secuencia para la que la secuencia del ácido nucleico monocatenario es complementaria. Si dicho ácido nucleico monocatenario se hibridiza con una secuencia la cual se localiza cerca de los extremos libres de la horquilla, el proceso de auto-rehibridización se bloquea justamente antes de que la horquilla completa vuelva a formarse; en este caso, Z_{pause} es mínima. Por otra parte, si dicho ácido nucleico monocatenario se hibridiza con una parte de la horquilla que esté cercana al bucle que no está aparcado, el proceso de renaturalización se detendrá en una situación en la que la horquilla está completamente o casi completamente desnaturalizada; en este caso, Z_{pause} es máximo (figura 1).

Es posible correlacionar de forma precisa una distancia física en una molécula de ácido nucleico bicatenario con diversas bases. Por ejemplo, una distancia de 1 nm corresponde a la distancia ocupada por dos nucleótidos (1 par de bases) en un ácido nucleico bajo una fuerza de 10 pN. La calibración exacta con respecto a la fuerza, es dada por la elasticidad del ácido nucleico monocatenario. Por tanto, midiendo simplemente la distancia entre los dos extremos libres de la molécula de ácido nucleico bicatenario, es posible determinar de forma precisa donde la renaturalización está bloqueada.

Así, en una forma de realización, la invención consiste en un procedimiento para determinar la secuencia de un ácido nucleico, en el que la molécula de ácido nucleico bicatenario que corresponde a la secuencia que va a determinarse, se desnaturaliza primero aplicando una fuerza física y volviendo entonces a hibridizarse en presencia de un ácido nucleico monocatenario, y detectándose la presencia de un bloqueo en la rehibridización. En un aspecto, la distancia entre los dos extremos de la molécula bicatenaria, se determina cuando el proceso de renaturalización está bloqueado. Preferentemente, la distancia entre los dos extremos de dicha molécula, se determina cuando la molécula está completamente desnaturalizada. Incluso más preferentemente, las dos distancias se comparan, determinándose la posición del bloqueo.

Otro parámetro útil asociado con el bloqueo en la renaturalización, es el período de tiempo durante el cual la renaturalización está bloqueada (al que se hace referencia en la presente memoria como la duración de la pausa en la renaturalización). Verdaderamente, es posible medir el período de tiempo durante el cual la rehibridización está bloqueada. Por ejemplo el experto en la materia puede determinar el período de tiempo durante el cual la distancia entre los dos extremos del ácido nucleico bicatenario es z , tal como se ha definido anteriormente, es decir, un valor intermedio comprendido entre Z_{high} y Z_{low} . La duración del bloqueo depende del grado de complementariedad entre las dos secuencias. Cuando más alta sea la complementariedad, más elevado será el número de enlaces que se establezcan entre las dos moléculas, y por tanto, más larga será la duración. También está claro que el tiempo de bloqueo dependerá de la longitud de la región de complementariedad entre las dos secuencias. Cuando más larga sea la región, mayor será el número de enlaces establecidos entre las dos moléculas y por tanto, más larga será la duración. Es por tanto, fácilmente concebible que bajo ciertas condiciones la duración de la pausa de renaturalización sea casi permanente. En particular, cuando el ácido nucleico monocatenario comprende más de 20, preferentemente más de 25, incluso más preferentemente más de 30 nucleótidos capaces de hibridarse con el ácido nucleico bicatenario desnaturalizado, el ácido nucleico monocatenario permanece hibridado a la horquilla bicatenaria (durante muchos minutos), incluso cuando la fuerza aplicada a dicho ácido nucleico bicatenario, disminuye a F_{test} , previniendo de esta forma la auto-rehibridización de dicha horquilla bicatenaria. En tal caso, puede ser ventajoso utilizar un enzima para eyectar la molécula de ácido nucleico monocatenario. La eyección de dicha molécula de ácido nucleico monocatenario hace, pues, posible, llevar a cabo ciclos de emparejamiento y desemparejamiento, y por tanto, mejorar la relación señal/ruido. Como ejemplos de enzimas apropiados, se pueden citar, por ejemplo, helicasas, incluyendo una helicasa UvrD, una helicasa UvrD de E. Coli, las helicasas Tte-UvrD, T7 Gp4, RecBCD, DnaB, MCM, Rep, RecQ, PcrA, T4 UvsW, la helicasa del antígeno grande T de SV40, la helicasa del herpes virus, la helicasa Sgs1 de la levadura, las helicasas dependientes DEAH_ATP y la proteína E1 de la helicasa del papilomavirus y sus homólogos. Preferentemente, se utiliza la helicasa T4 UvsW. La duración de la pausa puede también variar con las condiciones de la reacción. Dicha duración disminuirá a medida que la temperatura aumenta. Del mismo modo, las condiciones del tampón pueden también modular la duración de la pausa: por ejemplo, el magnesio, la betaina y el cloruro de tetrametilamonio (TMAC utilizado a concentraciones molares) aumentan el tiempo de bloqueo. Estos compuestos refuerzan los apareamientos AT más que los GC, reduciendo de esta forma la diferencia en la fuerza entre estos apareamientos. Sin embargo, cuando la temperatura y el tampón están fijados, la duración de la pausa dependerá sólo de la fuerza de arrastre que sobre el ácido nucleico bicatenario desnaturalizado y sobre su complementariedad con el ácido nucleico monocatenario.

Así pues, en un aspecto particular, el procedimiento de la invención comprende las etapas de:

- desnaturalizar dicha molécula de ácido nucleico bicatenario que corresponde a dicha secuencia de ácido nucleico, aplicando una fuerza física a dicha molécula;
- proporcionar una molécula de ácido nucleico monocatenario;

- renaturalizar la molécula de ácido nucleico bicatenario en presencia de dicha molécula de ácido nucleico monocatenario; y
- detectar un bloqueo de la renaturalización de dicha molécula de ácido nucleico bicatenario; y
- determinar la duración de la pausa.

En un aspecto preferido, la detección del bloqueo de la renaturalización de dicha molécula de ácido nucleico bicatenario, implica determinar la posición del bloqueo en la molécula de ácido nucleico bicatenario, tal como se ha descrito anteriormente.

En esta forma de realización particular, el procedimiento según esta invención puede utilizarse con propósitos diagnósticos para permitir, en particular, la secuenciación de regiones variables del ácido nucleico que corresponden a anomalías que se están buscando, siendo entonces la técnica similar a la descrita aquí a continuación para la secuenciación.

Sin embargo, es posible proporcionar una técnica simplificada, que se basa en la observación de que un desemparejamiento entre el oligonucleótido y la secuencia del ADN de lugar a una hibridación mucho más corta. En un primer aspecto, la renaturalización de una molécula de ácido nucleico bicatenario de la horquilla, es bloqueada por un ácido nucleico monocatenario, por cualquiera de los procedimientos que se describen anteriormente, determinándose la duración del bloqueo. En un aspecto preferido, este valor se compara a un valor de referencia. En un aspecto posteriormente preferido, el valor de referencia corresponde a la longitud de la pausa observada con un ácido nucleico monocatenario de referencia, determinado por cualquiera de los procedimientos anteriores. Con propósitos diagnósticos, por ejemplo, buscando mutaciones en el ADN genómico, la técnica puede implementarse de dos formas:

- 1) Las horquillas formadas con el ADN genómico, que comprende la búsqueda de mutaciones, se sondan con los oligonucleótidos en solución.
- 2) La horquilla que contiene la secuencia(s) con las mutaciones para búsqueda, se sondea mediante el ADN genómico que se encuentre en solución como fragmentos de ADN monocatenario de tamaño fijo. Se pondrá inmediatamente de manifiesto si el objetivo del ensayo es sólo encontrar una secuencia específica o una mutación posible en dicha secuencia, situándola en el bucle de la horquilla, lo que proporciona un esquema muy sencillo de detección. Si el oligonucleótido se hibridiza al bucle, previene completamente el repliegue de la horquilla, proporcionando un cambio muy intenso en la extensión, que puede, de esta forma, detectarse fácilmente, tal como se describe a continuación.

El procedimiento de la invención puede también utilizarse para la secuenciación directa de un ácido nucleico desconocido. El procedimiento de secuenciación de la invención permite varias formas de realización. En una primera forma de realización, se obtiene una secuenciación física con el procedimiento de la invención. Hibridando sucesivamente distintas sondas ácido nucleicas monocatenarias conocidas, a una horquilla ácido nucleico (que experimente ciclos de desnaturalización y renaturalización), se puede deducir la secuencia de dicha horquilla ácido nucleico a partir de la localización de las pausas (medidas con precisión de nm), durante la fase de renaturalización.

El experto en la materia, más que hibridar el ácido nucleico bicatenario para que va a secuenciarse con un conjunto de ácidos nucleicos monocatenarios que representan todas las combinaciones secuenciales posible, adoptará ventajosamente una estrategia que minimice el número de distintas sondas ácido nucleicas monocatenarias. Están disponibles varias opciones dependiendo de si las sondas monocatenarias, la molécula diana bicatenaria, o ambas, se optimizan.

En un aspecto, la invención se lleva a cabo con distintas sondas ácido nucleicas monocatenarias, en las que sólo son específicas un número limitado de bases, y las restantes no lo son. Por ejemplo, esta serie de sondas pueden consistir en moléculas ácido nucleicas monocatenarias de n bases, en las que todos los di-nucleótidos posibles (por ejemplo AA, AT, AG, ..., para un total de 16 combinaciones posibles), o todos los tri-nucleótidos posibles (por ejemplo AAA, AAT, AAG, ..., para un total de 64 combinaciones posibles) se unen con todas las combinaciones posibles de $n-2$ o $n-3$ nucleótidos, respectivamente, siendo n un entero, preferentemente menor o igual a 30, más preferentemente menor o igual a 20, incluso más preferentemente n es menor que igual a 8. Cuando sólo 2 o 3 bases son específicas (es decir, una serie de 16 ó 64 sondas distintas), la posición de di- o trinucleótidos se determina en cada hibridación. Esto permite mezclar una serie de moléculas e ácido nucleicas monocatenarias para reducir el número de intercambios tampónicos. Por ejemplo, en el caso de AANNNNNN, sólo cuatro series de sondas como implementadas son estrictamente necesarias en la plataforma Sólida de secuenciación desarrollada por Applied Biosystems. Los di- o trinucleótidos pueden localizarse en cualquier posición de las sondas n -mer. En una forma de realización preferida, los nucleótidos probados se localizan en el centro de los oligonucleótidos, ya que esta localización es más sensible al desemparejamiento, aumentándose la sensibilidad del procedimiento.

Una ventaja clara del procedimiento de la invención, es que dicho procedimiento permite la secuenciación de ambas

5 cadenas de la molécula bicatenaria al mismo tiempo. Verdaderamente, cada sonda hibridizará a la cadena que incluye una secuencia complementaria a la secuencia transportada por la sonda. La posición de la sonda hibridizada se determina entonces mediante el procedimiento desnaturalización/renaturalización que se describe anteriormente. Así, la secuencia de ambas cadenas puede determinarse en el mismo procesamiento, proporcionando un control interno. Para poder identificar la cadena que está unida por la sonda, es conveniente diseñar la sonda tal en la que los di- o trinucleótidos se localicen cerca del centro de las sondas, pero que se muestren ligeramente descentradas. Otra forma de realización preferida del procedimiento se refiere a donas en las que estos nucleótidos están ligeramente descentrados, de forma que el bloqueo se desviará, dependiendo de a qué cadena se unan los oligonucleótidos. Por ejemplo, un dinucleótido puede localizarse inmediatamente en el extremo 5' o en el 3' con respecto al centro de la sonda. También es posible utilizar una sonda en la que el nucleótido central sea el nucleótido de un trinucleótido más situado en el extremo 5' o en el 3'. Por ejemplo, posibles selecciones para un oligonucleótido 8-mer son NNXXNINN o NNXXXNIN. Finalmente, es también posible utilizar bases genéricas (Z) en vez de una combinación de todos los nucleótidos (N). Una base genérica (Z, tal como el 5-nitroindol o el 3-nitropirol) presenta interacciones homogéneas con la totalidad de las 4 bases y reduce la dilución del oligonucleótido.

20 La resolución de la secuenciación mediante detección mecánica de la hibridación, está limitada por la resolución espacial alcanzable en la medición de la distancia entre la perla y la superficie de anclaje. Esa resolución está determinada finalmente por la rigidez de la molécula de unión (la cual fija la amplitud del movimiento browniano la perla). Para una molécula de alrededor de 1.000 pares de bases bajo una tensión de 10 pN aproximadamente, la resolución espacial (con una segunda promediación) es de alrededor de 2 nm (es decir, alrededor de 2 pares de bases de desnaturalización ya que el ruido browniano disminuye como el cuadrado de la longitud del ADN (es decir, el cuadrado del número de nucleótidos), ajustándose bien la técnica a la secuenciación de las moléculas más cortas.

25 En otro aspecto, el ácido nucleico que va a secuenciarse se vuelve a diseñar para potenciar la determinación de la posición de la sonda hibridante. La patente US nº 6.723.513, por ejemplo, da a conocer una técnica de secuenciación que implica la ampliación de una o más bases para ayudar en la identificación posicional. En esta técnica, los pares de bases en el ácido nucleico diana están asociados con cuatro señalizaciones distintas (las señalizaciones amplificantes), que representan cada una de las 4 bases adenina, citosina, guanina y timina (o uracilo, si el ácido nucleico es un ARN). La presencia de cada base específica, adenina, citosina, guanina y timina es reemplazada entonces por la correspondiente señalización amplificadora. En una forma de realización preferida, cada señalización amplificante es un oligonucleótido de longitud específica, por ejemplo, n bases, y secuencia específica. El ácido nucleico bicatenario original puede así determinarse mediante desnaturalización/renaturalización según el procedimiento descrito anteriormente en la presencia sucesiva de oligonucleótidos complementarios a las señalizaciones amplificantes para adenina, citosina, guanina y timina. Estos oligonucleótidos se emparejarán con la cadena correspondiente del ácido nucleico bicatenario y bloquean su rehibridación en las correspondientes bases codificadas.

40 Así, en este aspecto, la invención proporciona un procedimiento para determinar la secuencia tal como se ha descrito anteriormente, en el que el ácido nucleico monocatenario es un oligonucleótido complementario para una de las señalizaciones amplificadoras. En un aspecto preferido, el procedimiento comprende una etapa posterior para determinar cada posición del bloqueo para dicho ácido nucleico monocatenario sobre la molécula de ácido nucleico bicatenario. En otro aspecto preferido todas las etapas de dicho procedimiento para la determinación de la secuencia, así como la etapa de determinación de cada posición del bloqueo, se repiten sucesivamente con cada uno de los oligonucleótidos complementarios a las identificaciones de amplificación.

50 Debido a que cada base es ampliada, es decir, es reemplazada por un oligonucleótido n -mer, la exactitud requerida para determinar la posición de la sonda hibridante necesita sólo ser inferior a n nm. Por ejemplo, si la señal de ampliación es un oligonucleótido 8-mer, la posición de la base puede determinarse de forma exacta cuando es posible determinar una distancia física entre dos extremos libres de la molécula, con una precisión menor que 8 bases, es decir, inferior a 8 bases, es decir, inferior por lo menos a 8 nm. Otra ventaja de este enfoque es que muchas perlas pueden secuenciarse en paralelo con sólo cuatro ensayos sucesivos.

55 En una segunda forma de realización, el procedimiento de la invención comprende una etapa enzimática. Una forma de realización preferida con este enfoque consiste en secuenciar la horquilla mediante hibridación sucesiva y ligado de las secuencias complementarias. Es posible, en esta forma de realización del procedimiento de la invención, determinar la secuencia de las largas moléculas ácido nucleicas bicatenarias; por largas moléculas ácido nucleicas bicatenarias, se hace referencia en la presente memoria a moléculas de más de 500 pares de bases, más preferentemente de más de 750 pares de bases, incluso más preferibles de más de 1.000 pares de bases. La técnica consiste en unir a un cebador ácido nucleico monocatenario, por su parte superior, un ácido nucleico monocatenario hibridado adyacente. La extensión del cebador se monitoriza entonces desnaturalizando y renaturalizando la molécula de ácido nucleico bicatenario de la horquilla, y detectando un bloqueo en la renaturalización, tal como se ha descrito antes. El procedimiento se repite entonces con una molécula de ácido nucleico monocatenario distinta. Según el procedimiento de la invención, no es necesaria una amplificación preliminar de la molécula de ácido nucleico bicatenario para ser secuenciada, pudiéndose llevar a cabo este procedimiento de la invención sobre una molécula de ácido nucleico bicatenario única.

En una forma de realización preferida, se utiliza una biblioteca de moléculas ácido nucleicas nonocatenarias (véase, por ejemplo, las patentes US nº 4.882.127 y nº 4.849.077). Dicha biblioteca está formada por moléculas ácido nucleicas monocatenarias de n bases, en la que todos los di-nucleótidos posibles (por ejemplo, AA, AT, AG,..., por un total de 16 combinaciones) se unen en su extremo 3' y mediante todas las combinaciones posibles de $n-2$ nucleótidos siendo n un número entero preferentemente menor que o igual a 20 más preferentemente n es menor que o igual a 12 incluso más preferentemente n es inferior que o igual a 8. En una forma de realización más preferida, los últimos nucleótidos m se fragmentan antes de llevar a cabo la próxima ronda de hibridación y unión, siendo m un entero comprendido entre 1 y $n-1$; preferentemente m es igual a $n-1$ (Mir et al., *Nucleic Acids Res.*, 37(1) e5, 2009). La utilización de una secuencia capaz de escisión permite la detección de la hibridación con un requerimiento menos exigente respecto a la exactitud de la posición de bloqueo (unos pocos nm), mientras se mantiene todavía un número bajo de etapas de síntesis. Una alternativa es utilizar oligonucleótidos, a los que les falte un fosfato en su extremo 5', de forma que sólo un oligonucleótido puede unirse en un momento determinado, antes del próximo procesamiento, se utiliza una quinasa para añadir el fosfato que falta, permitiendo de este modo la unión próxima. Repitiendo este procedimiento con cada uno de los 16 di-nucleótidos posibles, es posible detectar el aumento sucesivo de longitud de la cadena complementaria después de la unión de cada uno de los oligonucleótidos monocatenarios sucesivos. También es posible reunir los 16 nucleótidos en 4 grupos para reducir el número de ensayos, ya que cada secuencia dinucleótida se detecta 2 veces, siendo esto suficiente para determinar la secuencia. Una vez que se ha complementado el ácido nucleico bicatenario completo, mediante la biblioteca de moléculas ácido nucleicas mono catenarias, la cadena sintetizada se eyecta, (por ejemplo, con ayuda de una helicasa o de una exonucleasa), y el proceso se vuelve a iniciar con un cebador ácido nucleico monocatenario situado por encima o por debajo, que es desplazado superior o inferiormente mediante un nucleótido con respecto al cebador previo. Repitiendo el procedimiento $n-m$ veces, permite la determinación completa de la secuencia del ácido nucleico bicatenario: por ejemplo, para una biblioteca de oligómeros 8-mer, sólo 5 repeticiones del procedimiento (es decir, síntesis de una cadena complementaria), son necesarias para obtener la secuencia completa de la molécula bicatenaria cuando $m = 3$.

Mientras todos los procedimientos de la técnica anterior utilizan nucleótidos fluorescentes, el procedimiento de la invención sólo implica la detección mecánica de la extensión de la sonda. Por tanto, el procedimiento de la invención no sufre de ninguno de los inconvenientes asociados con los procedimientos de la técnica anterior. Por ejemplo, la unión exitosa de un oligómero 8-mer, representa un cambio en la extensión de la horquilla bicatenaria de 8 nm. Esto puede detectarse fácilmente con una resolución de 2 nm, que es la resolución espacial (con una segunda promediación), para una molécula de aproximadamente 1.000 pares de bases, bajo una tensión de alrededor de 10 pN. Ya que en cada etapa, se une un único oligonucleótido, su detección implica sólo la detección de un cambio relativo en la extensión, es decir, antes y después de una unión exitosa.

La aplicación del procedimiento de la invención se hecho posible, en particular por la existencia de dispositivos diseñados para llevar a cabo pruebas, en tiempo real, de la interacción del ácido nucleico al nivel único-molecular. Tal dispositivo se describe, por ejemplo, en las patentes US nº 7.052.650 y nº 7.244.391. El dispositivo que se describe utiliza trampas magnéticas para aplicar una fuerza escalar de picoNewtons sobre un lecho de tamaño micrónico y superparamagnético. Brevemente, dicho dispositivo incluye un microscopio óptico, núcleos magnéticos y un PC. Las moléculas ácido nucleicas bicatenarias se anclan en puntos diversos en un extremo a un elemento con escaso movimiento, por ejemplo, una superficie, y el otro extremo a una superficie móvil, en este caso, una perla magnética. Los núcleos magnéticos están provistos para actuar sobre la perla. En particular, los núcleos magnéticos pueden utilizarse para arrastrar el lecho lejos de la superficie. Sin embargo, la aplicación del procedimiento de la invención no está restringido al dispositivo mencionado. Cualquier dispositivo que permite que uno pueda extender completamente y entonces doblar una molécula de ácido nucleico bicatenario, mientras se monitoriza al mismo tiempo la extensión de dicha molécula, puede utilizarse para aplicar el procedimiento de la invención. Por ejemplo, las pinzas ópticas pueden utilizarse, pero requieren, sin embargo, una previa calibración de la fuerza. Y no son fácilmente paralelizables para mediciones de alta capacidad de procesamiento. Otros inconvenientes son la falta de un control torsional total del ácido nucleico, y el posible calentamiento local de la solución por el láser focalizado que puede alterar las condiciones de hibridación.

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

El conjunto ácido nucleico-perlas se inyecta en una cámara fluidica, cuya superficie se ha tratado de forma que se une el otro extremo marcado de la molécula (por ejemplo una superficie revestida con anti-Dig para unir el extremo del ácido nucleico marcado con Dig. Las perlas se anclan entonces a la superficie mediante una horquilla ácido nucleica (véase figura 1a). La distancia de la perla a la superficie es entonces monitorizada mediante varios medios conocidos por el experto en la materia: por ejemplo, los anillos de difracción de su imagen en una cámara, pueden utilizarse para deducir su distancia, o la intensidad de luz que dispersan (o emiten mediante fluorescencia) cuando se iluminan de forma evanescente, pueden utilizarse para medir su distancia. Alternativamente, el campo magnético que generan puede medirse (utilizando un sensor magnético tal como GMR o Sensores Hall), para deducir su

distancia a un sensor sobre la superficie de anclaje.

Se han descrito varias técnicas para arrastrar las perlas al anclaje sobre la molécula de ácido nucleico. Se puede utilizar la luz de un rayo láser enfocado para atrapar una perla transparente que esté cercana al punto focal. Mediante la traslación relativa del rayo con respecto a la superficie de anclaje, se puede aplicar una fuerza sobre la molécula de unión (un ensayo típico de pinzas ópticas). La fuerza ejercida, siendo proporcional al desplazamiento de la perla a partir de su posición de equilibrio, para ejercer una fuerza constante sobre la molécula de unión, requiere de un bucle de retroalimentación sobre el rayo de captura.

Para ejercer una fuerza constante sobre una perla se ha descrito la utilización de la resistencia de fricción hidrodinámica generada por un flujo alrededor de la perla, pero conlleva habitualmente una baja exactitud espacial (>100 nm). La forma de realización preferida utiliza una trampa magnética para arrastrar las perlas paramagnéticas ancladas a una superficie mediante una horquilla ácido nucleica, tal como se ha descrito anteriormente. En esta configuración, pequeños núcleos magnéticos situados por encima de la muestra, se utilizan para aplicar una fuerza constante sobre las perlas ancladas, cuya posición puede determinarse con una exactitud de < 1 nm, (dependiendo de la fuerza de arrastre y de la disipación debida a la resistencia de fricción hidrodinámica).

En cada caso, se aprecia que la horquilla de unión (hipotéticamente de elasticidad extrema) puede ser completamente desnaturada arrastrando mecánicamente las perlas con una fuerza superior a alrededor de 16 pN. Reduciendo la tensión sobre la molécula hasta por debajo de 11 pN, se deja que la horquilla se renaturalice espontáneamente (la transición a parte de la desnaturalización es reversible aunque histerética). Si, durante la fase de desnaturalización algunas moléculas en solución (tal como proteínas u oligonucleótidos complementarios del ADN, ARN, LNA o PNA) se unieron al ácido nucleico monocatenario estirado, estas moléculas bloquearán la renaturalización de la horquilla cuando la fuerza disminuye hasta por debajo de 11 pN. El principio del ensayo, es, por tanto, conmutar entre dos fuerzas: una F_{open} grande para abrir la horquillas, y una más pequeña F_{test} , que se utiliza para permitir la renaturalización y medir la extensión de la molécula en los bloques transicionales. La posición de bloqueo se relaciona con la secuencia mediante una relación lineal entre la extensión completa y la bloqueada. Para mayor seguridad la extensión completa se mide preferentemente con la prueba de fuerza F_{test} . Esta se obtiene diseñando el bucle de la horquilla de tal forma que necesite una fracción de un segundo para volver a doblarse, una vez la fuerza se haya reducido de F_{open} a F_{test} .

Para unir los ácidos nucleicos a superficies o soportes, se deben utilizar alguna de las técnicas conocidas en ese campo. Esencialmente, el ácido nucleico se ancla directamente al soporte, por ejemplo, la micro-perlas, lo que implica una funcionalización de esta superficie, por ejemplo, recubriéndole con estreptavidina, un grupo COOH, y similares, que sean capaces de reaccionar con el extremo funcionalizado del ácido nucleico.

Dichos procedimientos requieren en general, la funcionalización del ácido nucleico, especialmente de los extremos 3' y 5', es decir, injertando grupos químicos apropiados en ellos. Resulta además, preferido unir los otros dos extremos libres de la molécula mediante un bucle, con objeto de evitar que las cadenas se separen al final de la operación, de forma que ésta pueda repetirse si es apropiado. Con este propósito, pueden adoptarse distintos procedimientos.

Lo más sencillo es funcionalizar, utilizando oligonucleótidos sintéticos, uno de los extremos de un ácido nucleico bicatenario con dos funciones diferentes (biotina y amina, por ejemplo), que permitan el anclaje a dos superficies distintas pretratadas. Las dos cadenas en el otro extremo pueden unirse utilizando un nucleótido sintético parcialmente apareado en forma de un bucle. De esta forma, un ácido nucleico monocatenario emparejado, es decir, una horquilla, se produce a partir de un ácido nucleico bicatenario. La ventaja de este procedimiento se basa en su capacidad para funcionalizar una población heterogénea de fragmentos ácido nucleicos grandes (como son los obtenidos mediante fraccionamiento de un gen o de un cromosoma), que puede entonces ser analizado simultáneamente. En este caso, la muestra de ácido nucleico se fracciona utilizando dos (o más) enzimas de restricción, lo que permite que se obtenga una subpoblación con dos sitios de restricción diferentes en sus extremos que son similares en todos los fragmentos. Esto permite que los dos extremos sean tratados de modo distinto, por ejemplo, uniendo un extremo a un oligonucleótido en forma de un bucle que posee en su extremo el sitio de restricción apropiado). El inconveniente de este procedimiento se basa en la interferencia estérica entre los dos grupos funcionales adyacentes, que puede hacer dificultosa la unión a las superficies. Para resolver este problema, puede ser ventajoso añadir a cada extremo libre de la molécula horquilla una secuencia "espaciadora" de bases, al final de la cual se añade entonces un grupo funcional, no siendo complementarias las dos secuencias espaciadoras, permitiendo cada grupo funcional bastante espacio para unirse a su superficie dedicada. Más ventajosamente, la secuencia de cada secuencia espaciadora, se diseña para utilizar cebadores secuenciadores monocatenarios de secuencia conocida en el procedimiento de secuenciación de la invención. La adición de un bucle y/o de espaciadores a las moléculas ácido nucleicas bicatenarias, puede llevarse a cabo con alguno de los procedimientos que se utilizan habitualmente en biología molecular. Estos procedimientos son bien conocidos por el experto en la materia, y por lo tanto no es necesario detallarlos en la presente memoria.

Con respecto a las técnicas actuales de anclaje, existen muchas de éstas y proceden de las técnicas para el anclaje de macromoléculas (proteínas, ADN, y similares) a superficies pretratadas comercialmente disponibles. La mayoría

de estas técnicas se han desarrollado para pruebas inmunológicas y proteínas de unión (inmunoglobulinas) a superficies que llevan grupos ($--COOH$, $--NH_2$, $--OH$, y similares) capaces de reaccionar con los extremos carboxilo ($--COOH$) o amino ($--NH_2$) de las proteínas.

5 El anclaje covalente del ácido nucleico puede llevarse a cabo directamente, mediante el fosfato libre del extremo 5' de la molécula, que reaccione con una amina secundaria (superficie Covalink $--NH$ comercializada por Polylabo at Strasbourg), para formar un enlace covalente. También es posible funcionalizar ADN con un grupo amino, y entonces, proceder como con una proteína.

10 También existen superficies revestidas con estreptavidina (Perlas Dynal y similares), que permiten un anclaje casi covalente entre la estreptavidina y una molécula de ADN biotinilado. Finalmente, injertando un anticuerpo dirigido contra la digoxigenina sobre una superficie (mediante los procedimientos mencionados antes), puede anclarse un ácido nucleico funcionalizado con digoxigenina. Esto representa meramente una muestra de las muchas técnicas de anclaje que son posibles.

15 Entre las técnicas de unión y de anclaje, deben también mencionarse, por ejemplo, las descritas en la patente EP 152 886, utilizando un acoplamiento enzimático para la unión de ADN a un soporte sólido tal como celulosa.

20 La patente EP 146 815 describe también varios procedimientos de unión del ADN a un soporte. De forma similar, la solicitud de patente WO 92/16659 propone un procedimiento utilizando un polímero para unir ADN.

Naturalmente, el ácido nucleico puede unirse directamente al soporte, pero, cuando es necesario, especialmente con vistas a limitar la influencia de las superficies, el ácido nucleico puede unirse al extremo de un brazo inerte del péptido o de otra naturaleza, como es, por ejemplo, se describe en la patente EP 329 198.

25 Los ejemplos a continuación permitirán exponer otras características y ventajas de la invención que se presenta.

Leyendas de Figuras

30 Figura 1. Principio de detección de la hibridación de oligonucleótidos a su secuencia complementaria, en un ADN horquilla. El anclaje del ADN horquilla a las perlas en la superficie (a) lo desnaturaliza momentáneamente, aumentando la fuerza que arrastre a la perla a un valor superior a 16 pN. En esta fase, el fragmento complementario en solución se hibridiza a su diana en la horquilla del ADN abierto, evitando de esta forma la renaturalización de la horquilla (b) cuando la fuerza regresa a su valor inicial. El replegamiento de la horquilla presenta cuatro mesetas que se presentan en extensiones bien definidas, pero de duración variable. La meseta superior, a 73,71 nm, se asocia con la horquilla de 83 pares de bases, completamente abierta, en la prueba F_{test} , mientras que la meseta del fondo corresponde a la horquilla completamente plegada. Las dos mesetas intermedias a 25,47 nm y 35,17 nm, tienen lugar, porque dos oligos se situaron en la solución. A partir de este cambio en extensión (Z_{high-Z}), es posible deducir si a lo largo de la horquilla la secuencia complementaria se ha emparejado. Aquí, según sus posiciones, los bloques coinciden con la localización a 28,66 pares de bases y 39,60 pares de bases, estando muy de acuerdo con sus posiciones esperadas a 29 pares de bases y 40 pares de bases. Las posiciones de la meseta se evalúan mejor mediante los ajustes gaussianos para el histograma obtenido a partir de varios ciclos de apertura/cierre (aquí ~20 ciclos).

45 Figura 2. El tiempo de bloqueo depende intensamente de la longitud del oligonucleótido y de la fuerza de arrastre. A) Bloqueos de tiempo τ debidos a un oligonucleótido de 10 bases sobre una horquilla de 1.200 pares de bases. B) Histograma del tiempo de bloqueo que muestra una distribución de Poisson con un valor medio de 2 segundos. C) El tiempo de bloqueo varía con el tamaño del oligonucleótido y varía exponencialmente con la fuerza F_{test} utilizada durante la fase de prueba.

50 Figura 3. Evolución de la probabilidad de bloqueo y del tiempo de bloqueo con la concentración del oligo en el caso de un oligo de 9 bases. El tiempo de bloqueo es independiente de la concentración. La probabilidad de bloqueo presenta un K_m de 10 nM.

55 Figura 4. El tiempo de bloqueo de los oligonucleótidos que poseen 12 nucleótidos se grafica con respecto a la fuerza. Excepto la curva con símbolos circulares, todos estos oligonucleótidos tienen uno o dos desemparejamientos aunque en este último caso el bloqueo fue demasiado corto para ser medido. Si el desemparejamiento se localiza en la últimas de las primeras bases, el tiempo de bloqueo se reduce por un factor 5. Si el desemparejamiento se refiere a un par de bases AT en medio del oligonucleótido, el tiempo de bloqueo se reduce en más de 20 veces, mientras que alcanza 60 veces si se refiere a un par de bases GC. Un desemparejamiento doble reduce el tiempo de bloqueo tanto, que no puede medirse.

60 Figura 5a. Evolución del tiempo de bloqueo con temperatura para 10 oligonucleótidos ACAGCCAGCC. Típicamente, el tiempo de bloqueo disminuye en un factor 3 cuando la temperatura aumente en 10 grados.

65 Figura 5b. El tiempo de bloqueo de los oligonucleótidos que tienen 10 pares de bases, se grafican respecto a la

fuerza. Excepto la curva con símbolos circulares, todos los oligonucleótidos tienen uno o tres LNA (marcado con símbolos cuadrados). Un LNA que reemplaza al ADN aumenta el tiempo de bloqueo en más de 2 veces.

5 Figura 6. Histograma de la distribución de las extensiones de ADN en un experimento tal como el que se da a conocer en la figura 1c, donde oligonucleótidos en solución pueden emparejarse con el ADN desnaturalizado en varias posiciones a lo largo de la molécula. A partir de la posición de los picos del histograma (que está altamente correlacionado para tres moléculas diferentes es decir, distintas perlas unidas), puede deducirse la posición del híbrido a lo largo del ADN.

10 Figura 7. Histograma de posiciones de bloqueo que corresponde a los cuatro nucleótidos de 8 bases A_8 , C_8 , T_8 , G_8 para una horquilla de ADN que corresponde a una secuencia ampliada. Estas posiciones de bloqueo corresponden de forma precisa a sus posiciones esperadas. Tenemos aquí $G_8 = GCACGCAC$, $C_8 = TCGCTCGC$, $T_8 = GCCAGCCA$ y $A_8 = CCGACCGA$.

15 Ejemplos experimentales

- Preparación de ADN

20 Un fragmento de ADN bicatenario (ds) de secuencia desconocida y de un tamaño comprendido entre algunos céncimos y algunos miles de pares de bases, se une en uno de sus extremos a un bucle de ADN. Su otro extremo se une a un fragmento de dsADN, que permite la unión de sus dos cadenas a distintas superficies de revestimiento. Por ejemplo, el extremo 3' libre de una cadena, puede marcarse con biotina, permitiendo la unión a las perlas revestidas con estreptavidina, mientras que el extremo 5' en la cadena opuesta puede marcarse con digoxigenina, lo que permite su unión a superficies revestidas con un anticuerpo anti-Dig. Esta marcación del extremo puede realizarse

25 de varias formas conocidas por el experto en la materia, como la utilización de transferasa terminal para añadir nucleótidos modificados por dig o biotina, o la hibridación con oligonucleótidos marcados apropiadamente.

- Dispositivos para el alargamiento de la fuerza

30 Este constructo de ADN se incuba durante algunos minutos en una solución de perlas adecuadas (por ejemplo, las revestidas con estreptavidina), a las cuales se une mediante uno de sus extremos marcados (por ejemplo, biotina). Las perlas pueden ser transparentes si las pinzas ópticas se utilizan posteriormente para la manipulación, o magnéticas, si se utilizan trampas magnéticas o pinzas para la manipulación.

35 El conjunto ADN-perlas se inyecta en una cámara fluidica cuya superficie se ha tratado, de forma que se una al otro extremo marcado de la molécula (por ejemplo, una superficie revestida con anti-Dig para unirse al extremo del ADN marcado con Dig). Las perlas se anclan, así, a la superficie mediante una horquilla de -ADN, (véase figura 1a). La distancia de la perla a la superficie se monitoriza entonces mediante varios medios conocidos por el experto en la materia: por ejemplo, los anillos de difracción de su imagen en una cámara puede utilizarse para deducir su distancia, o la intensidad de la luz que dispersan (o emiten mediante fluorescencia), cuando se iluminan de forma evanescente, pueden utilizarse para medir su distancia. Alternativamente, el campo magnético que generan puede medirse (utilizando un sensor magnético tal como GMR o sensores Hall), para deducir su distancia a un sensor que está sobre la superficie de anclaje.

40

45 Se han descrito varias técnicas para arrastrar sobre la superficie de la molécula de ADN las perlas con objeto de anclarlas sobre él. La forma de realización preferida utiliza una trampa magnética para arrastrar perlas paramagnéticas que se anclan a una superficie mediante una horquilla de ADN. En esta configuración, pequeños elementos magnéticos situados por encima de la muestra, se utilizan para aplicar una fuerza constante sobre la perla de anclaje, cuya posición puede determinarse con una seguridad de < 1 nm (dependiendo de la fuerza de arrastre y de la disipación debida a la resistencia de fricción hidrodinámica. En esta serie de experimentos se utilizó el dispositivo que se describe en las patentes US nº 7.052.650 y nº 7.244.391. Además, si no se indica de otro modo, los experimentos realizados en ella, se llevaron a cabo en 25 mM Tris pH 7,5, 150 mM Kac, 10 mM $MgCl_2$, 0,2% BSA.

50

55 En cada caso, la horquilla puede ser desnaturalizada completamente de forma mecánica sobre las perlas con una fuerza superior a 16 pN aproximadamente. La reducción de la tensión sobre la molécula por debajo de alrededor de 11 pN, permite que la horquilla se renaturalice espontáneamente (la transición de la desnaturalización es reversible, aunque histerética). Si durante la fase de desnaturalización, algunas moléculas en solución (tales como proteínas u oligonucleótidos complementarios de ADN, RNA, LNA o PNA) se unieron al ADN (ss) monocatenario alargado, estas moléculas bloquearán transitoriamente la renaturalización de la horquilla, cuando la fuerza se disminuya por debajo de 11 pN. El principio del ensayo es la conmutación entre 2 fuerzas: una grande F_{open} , para abrir la horquilla y una más pequeña F_{test} , que se utiliza para permitir la renaturalización y para medir la extensión de la molécula en los bloqueos transitorios. La posición de bloqueo está relacionada con la secuencia, mediante una relación lineal entre la extensión completa y la bloqueada. Para mayor seguridad, la extensión total se mide preferentemente en la prueba de fuerza F_{test} . Esta se alcanza diseñando el bucle de la horquilla, de tal forma que necesita una fracción de un segundo para volver a plegar otra vez, con lo que la fuerza se reduce de F_{open} o F_{test} .

60

65

- La posición de la hibridación de un oligonucleótido puede medirse con una resolución de un par de bases

5 Midiendo la extensión de la molécula de ADN (la distancia de la perla a la superficie) durante una de estas pausas de renaturalización, es posible determinar la posición del bloqueo con precisión de nanómetro (1nm corresponde a la distancia ocupada por 2 nucleótidos (1 par de bases) en un ssADN bajo una fuerza de 10 pN). La configuración de desnaturalización muestra la proporción más grande de extensión con relación a los pares de bases (en dsADN la proporción es sólo de 0,34 nm por par de bases).

10 La exactitud de esta medición está limitada por dos contribuciones de ruido:

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

15 Pueden utilizarse distintas técnicas para medir la posición vertical de la perla. Una de las más sencillas se relaciona con video microscopia (patentes US nº 7.052.650 y nº 7.244.391). Los resultados en la Figura 1, donde se obtuvieron con este procedimiento, la resolución típica alcanza 1 nm durante un promedio de tiempo de 1 segundo. Otros procedimientos con mejor resolución se han demostrado, tales como la iluminación con láser con sensores PSD que alcanza 0,1 nm en resolución (Greenleaf and Block, *Science*, 313: 801, 2006) e iluminación con ondas evanescentes (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).

20 La limitación intrínseca en la resolución se produce por las fluctuaciones brownianas del arrastre de las perlas sobre una molécula ssADN. $\langle x^2 \rangle = 4k_B T \Delta f (6\pi\eta r) / k_{ssADN}^2(F)$, donde $k_{ssADN}(F)$ es la rigidez de una molécula ssADN, 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 medición. $k_{ssADN}(F = 10 \text{ pN}) = 0,05/N_b$ (N/m), donde N_b es el número de bases del ssADN. Para la horquilla de 84 pares de bases, esto lleva a 0,04 nm de ruido sobre un promedio de 1 segundo ($\Delta f = 1 \text{ Hz}$). El ruido más intenso en la figura 1 ($\sigma \sim 1 \text{ nm}$) se debe esencialmente al dispositivo de medida, no a las fluctuaciones intrínsecas. El ruido browniano intrínseco aumenta con el tamaño de la horquilla; una horquilla de 1.200 pares de bases produce un ruido de 0,6 nm cuando se promedia sobre un segundo.

La calidad de la hibridación se mide mediante el valor medio del tiempo de bloqueo.

35 La fuerza del bloqueo puede caracterizarse por dos parámetros: la probabilidad de bloquear P_{block} (= número de ciclos que presentan un bloqueo/número total de ciclos), y el tiempo medio de bloqueo T_{block} , P_{block} depende de k_{on} y de la concentración del oligonucleótido, mientras que T_{block} depende sólo de k_{off} , donde k_{on} y k_{off} son, respectivamente, la constante de reacción de la unión y la desunión. En la figura 2, se muestra la variación típica de T_{block} con la longitud oligonucleótida y la fuerza. Un único desemparejamiento de bases tiene un drástico efecto sobre T_{block} , equivalente a reducir la longitud del oligonucleótido en por lo menos un nucleótido, y disminuyendo el tiempo de bloqueo por un factor de 5.

40 En la práctica T_{block} y por tanto k_{off} es más fácil de medir, pues no depende de la concentración de oligonucleótidos (figura 3). Sin embargo también es posible medir k_{on} . El tiempo promedio del bloqueo, depende de la secuencia oligonucleótida, pero no sobre su posición a lo largo de la horquilla. Se estudió una secuencia que unía dos posiciones específicas a lo largo de la horquilla: el tiempo de bloqueo es el mismo para ambos bloqueos, mientras que tienen lugar en localizaciones muy distintas.

La única mutación tiene un drástico efecto sobre el tiempo de bloqueo

50 Tal como se muestra en la figura 4, un oligonucleótido de 12 bases que forma un emparejamiento perfecto con la hélice, presenta un tiempo de bloqueo muy diferente que idéntico oligonucleótido con un único desemparejamiento. En la figura 4 se muestra el tiempo de bloqueo respecto a la fuerza, para los distintos oligonucleótidos. Aumentando la fuerza, se incrementa el tiempo de bloqueo. Cuando la mutación se encuentra justamente en el primer o último nucleótido, su efecto sobre el tiempo de bloqueo es mínimo, reduciéndose por un factor 5. Tal como se espera, esta reducción depende de la naturaleza del desemparejamiento, un desemparejamiento en AT conduce típicamente a una reducción del tiempo de bloqueo de un factor de 20, mientras que un desemparejamiento en GC conduce a una reducción de un factor 60.

60 El tiempo de bloqueo se reduce drásticamente cuando el desemparejamiento se localiza en el centro del oligonucleótido

65 Como puede apreciarse en la figura 4, un desemparejamiento en el centro del oligonucleótido, provoca un bloqueo muy corto que sólo puede observarse cuando la fuerza es máxima. La reducción en el tiempo de bloqueo resultante de dicho emparejamiento excede un factor de 100 para idénticas condiciones de fuerza.

El tiempo de bloqueo depende de la temperatura y de las condiciones tampónicas

Tal como se aprecia en la figura 5a, el aumento de la temperatura reduce significativamente el tiempo de bloqueo. Las condiciones de los tampones pueden también modular el tiempo de bloqueo: el magnesio, la betaina y el cloruro de tetrametilamonio (TMAC utilizado a concentración molar) aumenta significativamente el tiempo de bloqueo comparándolo con el tampón usado en estos experimentos (25 mM Tris pH 7,5, 150 nM Kac, 10 mM MgCl₂, 0,2% BSA). Estos compuestos refuerzan los apareamientos AT más que los GC, reduciendo la diferencia en la fuerza entre estos pares.

El tiempo de bloqueo aumenta utilizando oligonucleótidos ARN o LNA

Los oligonucleótidos ARN y LNA forman híbridos más fuertes con ssADN que los oligonucleótidos de ADN. Para idéntica secuencia diana, el tiempo de bloqueo aumenta en un factor por un oligonucleótido de ARN, cuando se compara con un oligonucleótido de ADN.

Los nucleótidos LNA presentan un efecto más drástico: si un nucleótido único se convierte de un ADN a un LNA, el tiempo de bloqueo del oligonucleótido completo aumenta en un factor de 2. Convirtiendo tres bases de ADN a LNA aumenta el tiempo de bloqueo en un factor de 5. Cambiando todos los nucleótidos de ADN a LNA, como tal efecto drástico, el tiempo de bloqueo de 10 bases de los oligonucleótidos de LNA, excede en 1 hora. Reduciendo el tamaño del oligonucleótido a 6 bases de LNA, conduce a un tiempo razonable de bloqueo de 1 segundo.

Como con los oligonucleótidos de ADN, midiendo el tiempo medio de bloqueo con una de estas alternativas oligonucleótidas (LNA o ARN), se puede determinar su naturaleza: se debe a una hibridación perfecta con un oligonucleótido complementario o no, y si no, donde está el desemparejamiento (por ejemplo, en el centro del oligonucleótido hibridado o cerca de uno de sus extremos).

Longitud del oligonucleótido detectable

Ya que el tiempo de bloqueo depende exponencialmente de la longitud del oligonucleótido, este parámetro no puede variar mucho. Si el oligonucleótido es demasiado pequeño, (más pequeño que 8 bases a temperatura ambiente), el tiempo de bloqueo es demasiado pequeño para ser detectable. Si el oligonucleótido es demasiado largo (superior a 12 bases a temperatura ambiente), el tiempo de bloqueo se convierte en demasiado largo.

Los enzimas pueden estabilizar el híbrido

Añadiendo la polimerasa gp43 ADN sin NTP, aumenta el tiempo de bloqueo de los oligonucleótidos. Esto se espera, ya que el hibridado es un sustrato para la polimerasa. La gp43 polimerasa aumenta el tiempo de bloqueo de un oligonucleótido en un factor de 10.

Resumen de los parámetros de hibridación

La longitud del oligonucleótido constituye un parámetro crítico: a temperatura ambiente, la longitud de los oligonucleótidos con tiempos prácticos de bloqueo, varía entre 8 y 12 bases. Se pueden fácilmente llevar a cabo series de experimentos de desnaturalización/ renaturalización, sobre la misma molécula y medir el tiempo medio de bloqueo después de la renaturalización debido al apareamiento de los oligonucleótidos con el ADN en la fase de desnaturalización. Este tiempo depende del tamaño del oligonucleótido, de la fuerza aplicada durante la naturalización, de la temperatura y de la concentración iónica. Si el fragmento que se empareja muestra desemparejamientos, el tiempo de bloqueo se reducirá significativamente (por lo menos, 10 veces) y de una manera cuantificable. La técnica mecánica de desnaturalización/ renaturalización (descompresión/recompresión) permite así, sondear rápidamente la posición y estabilidad del apareamiento entre una secuencia oligonucleótida conocida y un fragmento de ADN de secuencia desconocida, (ver figura 1 y figura 2). Estas observaciones sugieren varias aplicaciones en la secuenciación del ADN y en el diagnóstico.

Diagnóstico y secuenciación mediante detección de la hibridación

Sondeando las horquillas del ADN que se anclan a la superficie de las perlas con distintos oligonucleótidos (introducidos sucesivamente en la cámara fluidica), se puede determinar la presencia de mutaciones posibles sobre una secuencia conocida (que dan lugar a desemparejamientos con la sonda oligonucleótida y pausas más cortas durante la renaturalización), o secuenciar un ADN desconocido determinando la posición de sondas conocidas a lo largo de la molécula (véase la figura 6).

En otro aspecto, el ácido nucleico que va a secuenciarse, se vuelve a diseñar utilizando señalizaciones de amplificación con objeto de potenciar la determinación de la posición de la sonda de hibridación. En el experimento del que se informa en la figura 7, cada aparición de cada base específica, Adenina, Citosina, Guanina y Timina, se reemplazó por la señalización amplificadora, en este caso, un oligonucleótido 8-mer. Tal como se muestra en la figura 7, las posiciones de bloqueo se corresponden perfectamente con las posiciones esperadas de la secuencia.

Listado de secuencias

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

5 <120> PROCEDIMIENTO DE SECUENCIACIÓN DE ADN MEDIANTE MANIPULACIÓN

<130> 358353D28234

10 <150> EP 10305564.6
<151> 2010-05-27

<150> US 61/377,621
<151> 2010-08-27

15 <160> 14

<170> Versión 3.3 Patentin

20 <210> 1
<211> 12
<212> ADN
<213> artificial

25 <220>
<223> Sonda

<400> 1
acagccagcc ga 12

30 <210> 2
<211> 11
<212> ADN
<213> artificial

35 <220>
<223> Sonda

<400> 2
acagccagcc g 11

40 <210> 3
<211> 10
<212> ADN
<213> artificial

45 <220>
<223> Sonda

<400> 3
acagccagcc 10

50 <210> 4
<211> 9
<212> ADN
<213> artificial

55 <220>
<223> Sonda

60 <400> 4
acagccagc 9

<210> 5
<211> 12
<212> ADN
<213> artificial

65

<220>
 <223> Diana
 5 <400> 5
 tcggctggct gt 12
 <210> 6
 <211> 12
 10 <212> ADN
 <213> artificial
 <220>
 <223> Sonda
 15 <400> 6
 agccgaccga ca 12
 <210> 7
 <211> 12
 20 <212> ADN
 <213> artificial
 <220>
 25 <223> Sonda
 <400> 7
 agccgaccga ct 12
 30 <210> 8
 <211> 12
 <212> ADN
 <213> artificial
 35 <220>
 <223> Sonda
 <400> 8
 agccgaccgt ca 12
 40 <210> 9
 <211> 12
 <212> ADN
 <213> artificial
 45 <220>
 <223> Sonda
 <400> 9
 50 agccgtccga ca 12
 <210> 10
 <211> 12
 <212> ADN
 55 <213> artificial
 <220>
 <223> Sonda
 60 <400> 10
 agccgtgccga ca 12
 <210> 11
 <211> 12
 65 <212> ADN
 <213> artificial

ES 2 542 429 T3

<220>
 <223> Sonda

5 <400> 11
 agccgtccca ca 12

<210> 12
 <211> 10
 10 <212> ADN
 <213> artificial

<220>
 <223> Sonda

15 <400> 12
 acagccagcc 10

<210> 13
 20 <211> 128
 <212> ADN
 <213> artificial

<220>
 25 <223> Sonda

<400> 13
gcacgcactc gctcgcgcca gccagccagc catcgcctgc gcacgcacgc acgcacccga 60
ccgaccgacc gagccagcca gcacgcacgc cagccaccga ccgatcgcctc gctcgcctgc 120
ccgaccga 128

30 <210> 14
 <211> 16
 <212> ADN
 <213> artificial

<220>
 <223> Sonda

<400> 14
 40 gcttcggaat gtacca 16

REIVINDICACIONES

- 5 1. Procedimiento para la determinación de una secuencia de ácido nucleico, comprendiendo dicho procedimiento las etapas de:
- 10 a) desnaturalizar una molécula de ácido nucleico bicatenario que corresponde a dicha secuencia de ácido nucleico aplicando una fuerza física a dicha molécula; en el que dicha molécula de ácido nucleico bicatenario es una horquilla, y en el que uno de los extremos del ácido nucleico bicatenario está fijado directa o indirectamente a un soporte, y en el que el otro extremo del ácido nucleico bicatenario está fijado a un soporte móvil;
- 15 b) proporcionar una molécula de ácido nucleico monocatenario;
- c) renaturalizar dicha molécula de ácido nucleico bicatenario en presencia de dicha molécula de ácido nucleico monocatenario; y
- 20 d) detectar un bloqueo de la renaturalización del ácido nucleico bicatenario, comprendiendo dicha determinación las etapas de:
- medir la distancia (z) entre los dos extremos de la molécula de ácido nucleico bicatenario que están fijados al soporte;
- 25 - medir la distancia (z_{high}) entre los dos extremos de la molécula de ácido nucleico bicatenario que están fijados al soporte, cuando dicha molécula de ácido nucleico bicatenario se desnaturaliza; y
- comparar z y z_{high} , y
- determinar la posición de la pausa.
- 30 2. Procedimiento según la reivindicación 1, en el que el ácido nucleico bicatenario se desnaturaliza en la etapa a) alejando los soportes.
- 35 3. Procedimiento según la reivindicación 2, en el que una fuerza física superior o igual a 15 pN, preferentemente superior o igual a 17 pN, más preferentemente superior o igual a 18 pN, se aplica a la molécula bicatenaria alejando los soportes.
- 40 4. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que el ácido nucleico bicatenario desnaturalizado se renaturaliza en la etapa (c) uniendo los soportes.
- 45 5. Procedimiento según la reivindicación 4, en el que la fuerza aplicada a la molécula bicatenaria se reduce a menos de o igual a 12 pN, preferentemente menos de o igual a 11 pN, más preferentemente menos de o igual a 10 pN, uniendo los soportes.
6. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que los extremos del ácido nucleico bicatenario que no están fijados a un soporte, se unen a otro covalente o no covalentemente.
- 50 7. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que las etapas a) a d) se repiten varias veces, de manera que se acumulen mediciones y aumente la relación señal/ruido.
8. Procedimiento según cualquiera de las anteriores reivindicaciones, que comprende una etapa adicional de medir la duración del bloqueo.
- 55 9. Procedimiento según la reivindicación 8, que comprende una etapa adicional de comparar la duración del bloqueo con un valor de referencia.
- 60 10. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que la molécula de ácido nucleico monocatenario se selecciona de entre una biblioteca de las moléculas de ácido nucleico monocatenario n -mer, consistiendo dicha biblioteca en la totalidad de combinaciones posibles de di- o trinucleótidos unidos con la totalidad de combinaciones posibles de $n-2$ o $n-3$ nucleótidos, siendo n respectivamente un entero preferentemente inferior o igual a 30.
11. Procedimiento según la reivindicación 10, en el que el di- o trinucleótido está ubicado en el centro de la molécula de ácido nucleico monocatenario.
- 65 12. Procedimiento según la reivindicación 10, en el que el di- o trinucleótido está ubicado excéntrico de la molécula

de ácido nucleico monocatenario.

5 13. Procedimiento según cualquiera de las reivindicaciones 1 a 9, en el que cada presencia en el ácido nucleico bicatenario de por lo menos una de las 4 bases adenina, citosina, guanina y timidina, es reemplazada por un marcador de aumento específico, siendo dicho marcador de aumento un oligonucleótido.

14. Procedimiento según la reivindicación 13, en el que el ácido nucleico monocatenario es un oligonucleótido complementario a uno de los marcadores de aumento.

10 15. Procedimiento según cualquiera de las reivindicaciones 13 o 14 que comprende una etapa adicional de determinar cada posición de bloqueo para dicho ácido nucleico monocatenario sobre la molécula de ácido nucleico bicatenario.

15 16. Procedimiento según cualquiera de las reivindicaciones 13 a 15, en el que las etapas a) a d), y la etapa adicional de determinar cada posición de bloqueo para el ácido nucleico monocatenario sobre la molécula de ácido nucleico bicatenario, se repiten sucesivamente con cada uno de los oligonucleótidos complementarios a los marcadores de aumento.

20 17. Procedimiento según cualquiera de las reivindicaciones 1 a 9, que comprende las etapas adicionales de:

i) ligar a un cebador de ácido nucleico monocatenario en dirección hacia 5' un ácido nucleico monocatenario hibridado adyacente, y

25 ii) monitorizar la extensión de dicho cebador:

α desnaturalizando y renaturalizando dicho ácido nucleico bicatenario en presencia de dicho cebador ligado, y

30 β detectando un bloqueo en la renaturalización.

18. Procedimiento según la reivindicación 17, en el que las etapas (i) a (ii) se repiten con un ácido nucleico monocatenario diferente.

35 19. Procedimiento según las reivindicaciones 17 y 18, en el que la molécula de ácido nucleico monocatenario se selecciona de entre una biblioteca de moléculas de ácido nucleico monocatenario n-mer, consistiendo dicha biblioteca en la totalidad de combinaciones posibles de dinucleótidos unidos en su extremo 3' a la totalidad de combinaciones posibles de los n-2 nucleótidos, siendo n un entero, preferentemente inferior o igual a 20.

40 20. Procedimiento según cualquiera de las reivindicaciones 17 a 19, en el que la cadena recién sintetizada es eyectada o desensamblada.

21. Procedimiento para la determinación de una secuencia de ácido nucleico según cualquiera de las reivindicaciones 17 a 20, en el que las etapas i) e ii) según la reivindicación 17 se repiten con un cebador que se desplaza hacia 5' o hacia 3' mediante un nucleótido con respecto al cebador según la reivindicación 17.

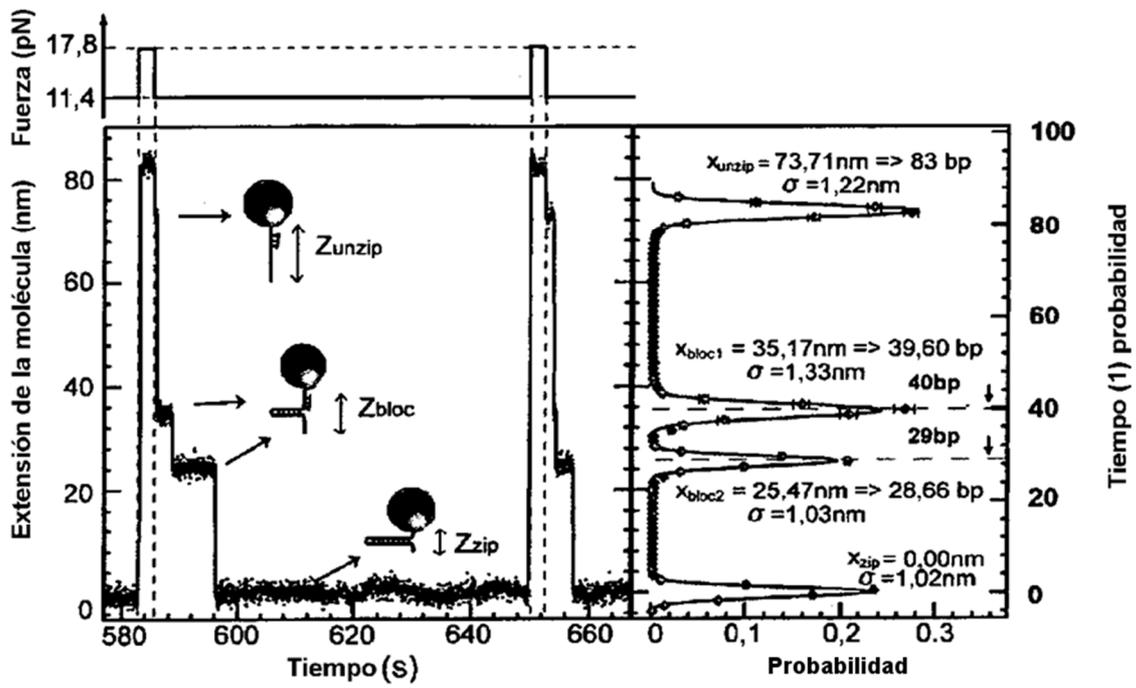


Figura 1

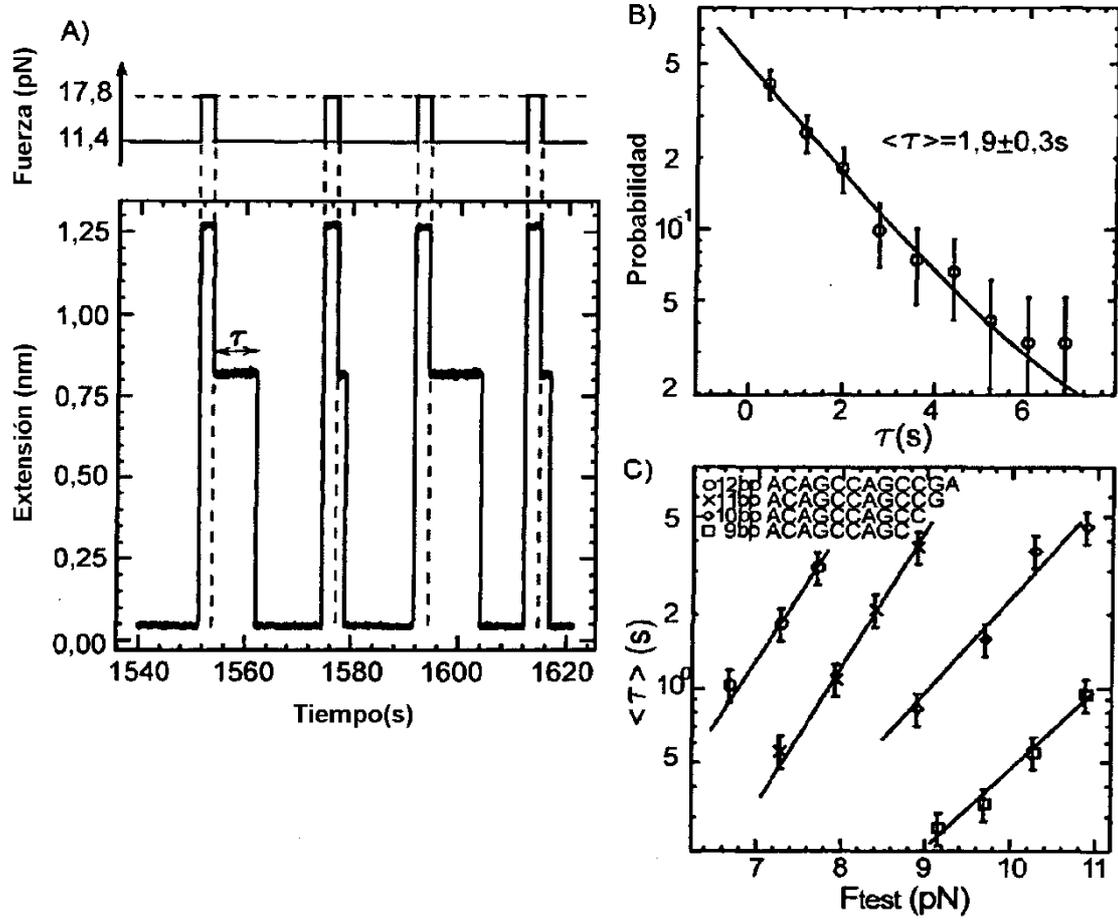


Figura 2

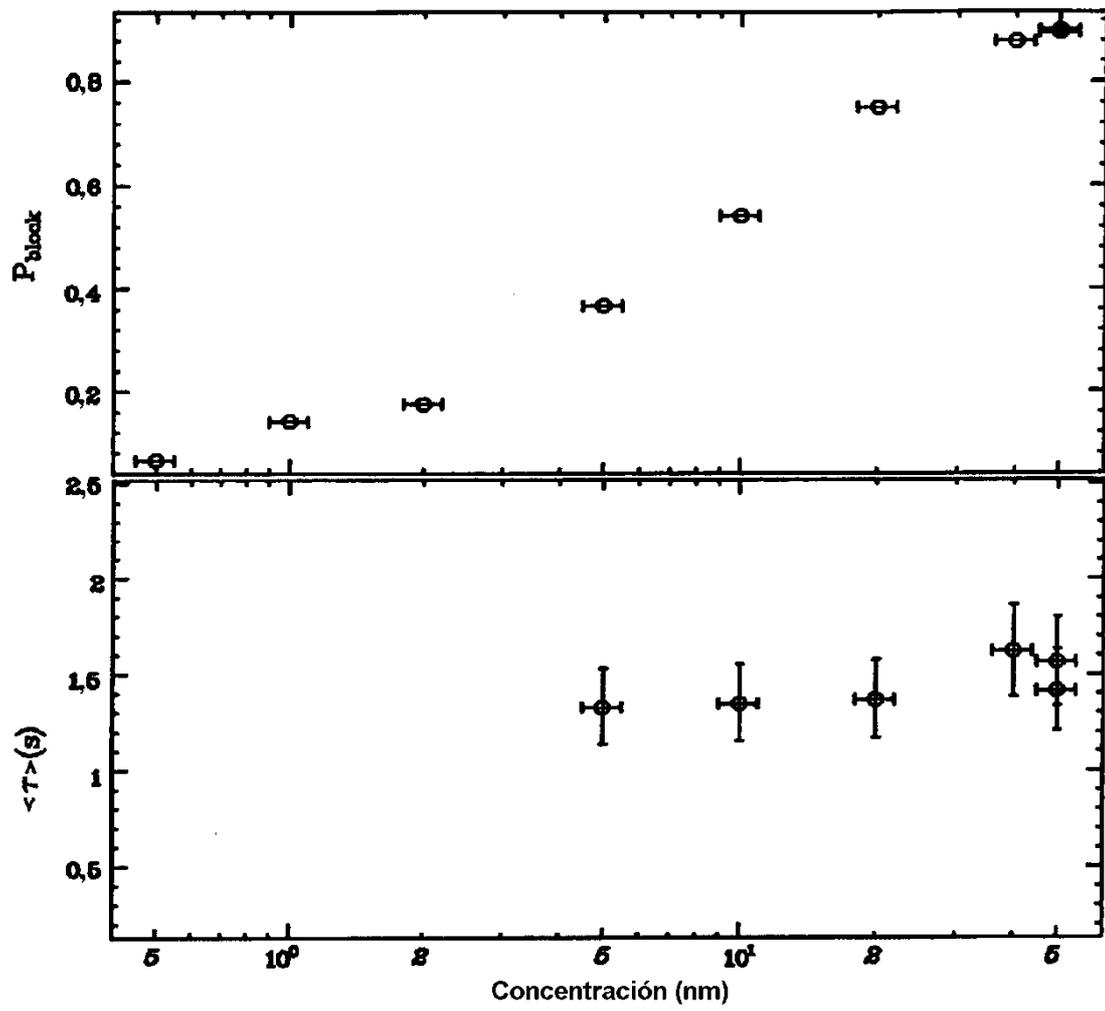


Figura 3

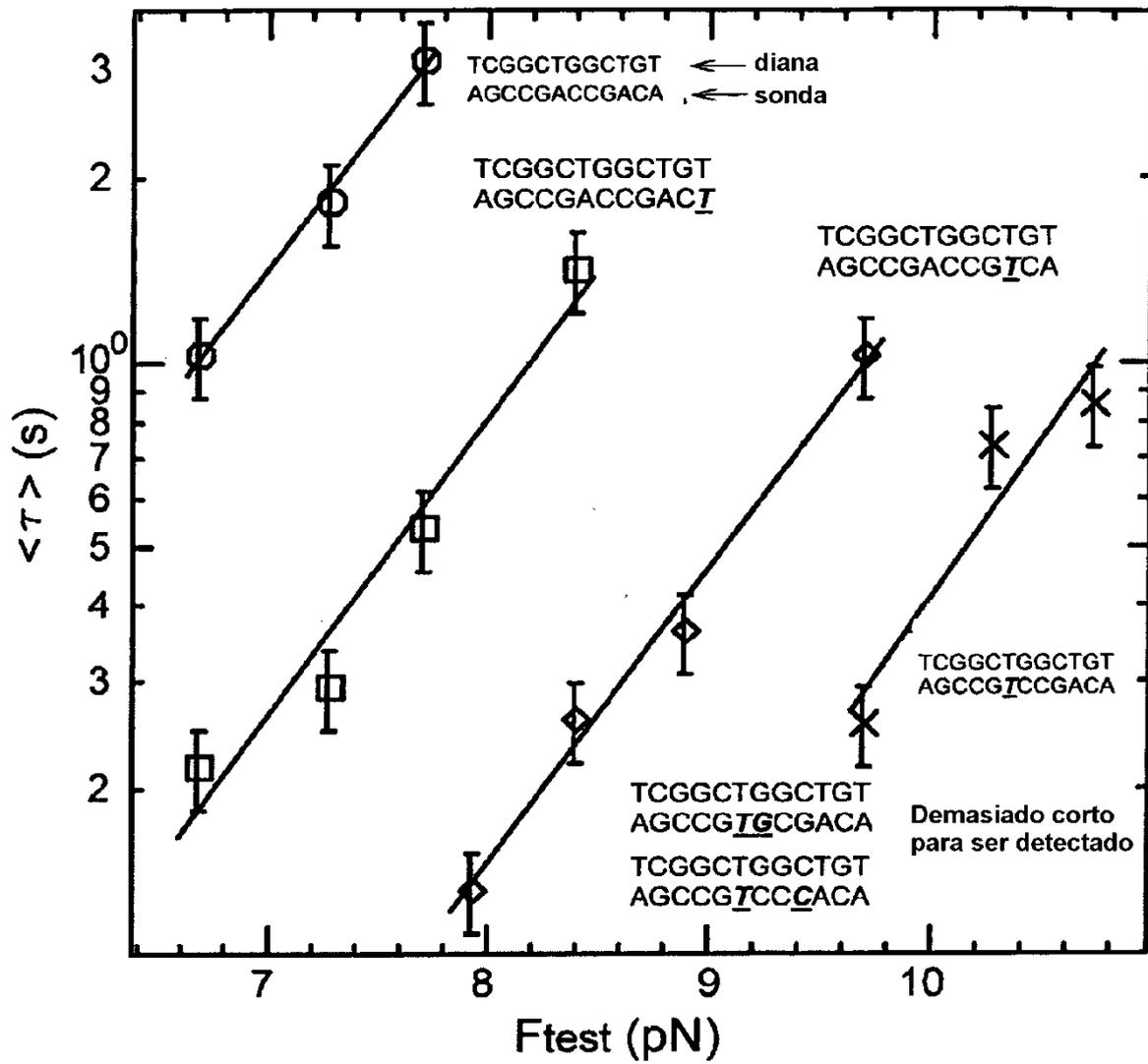


Figura 4

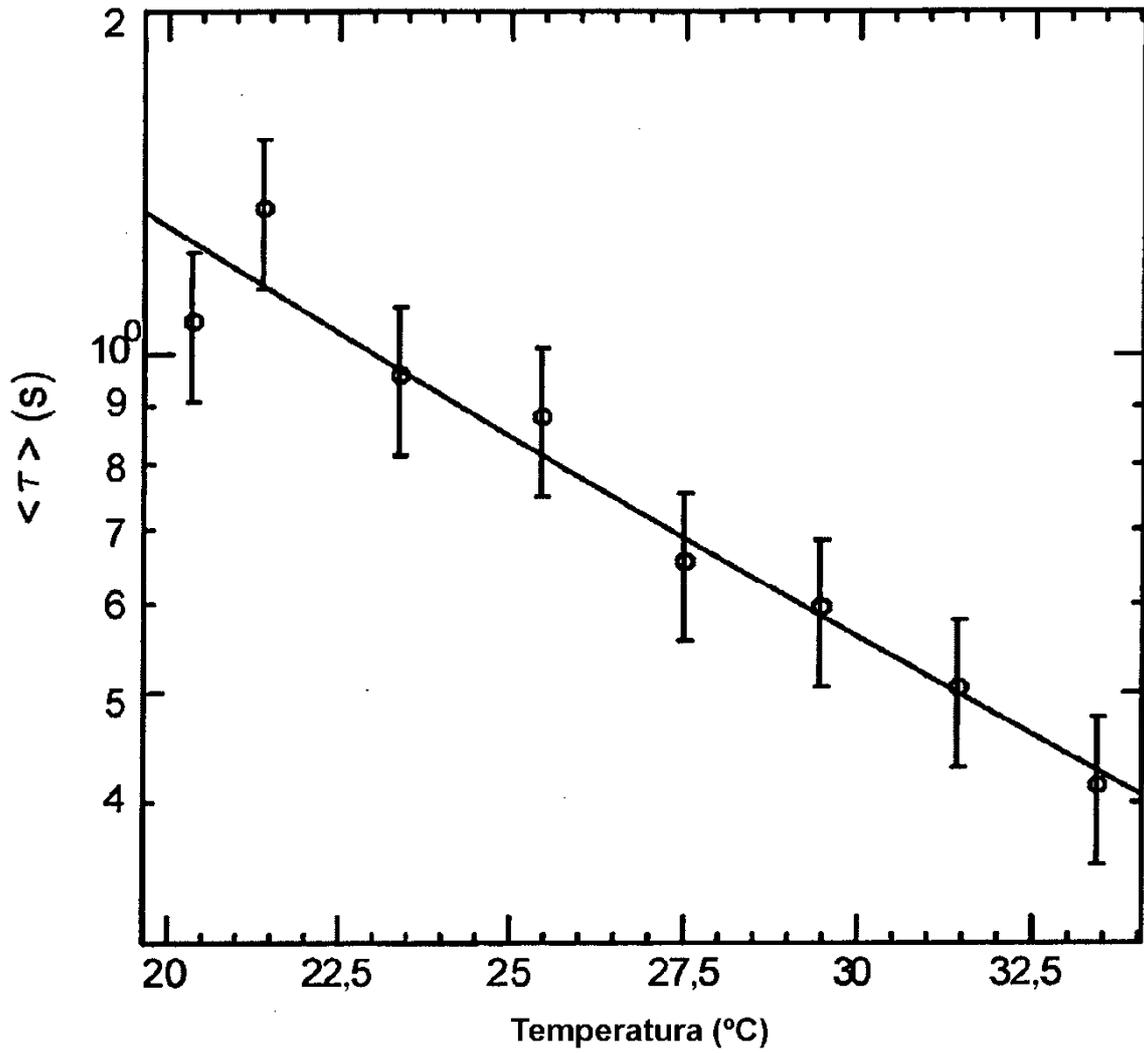


Figura 5a

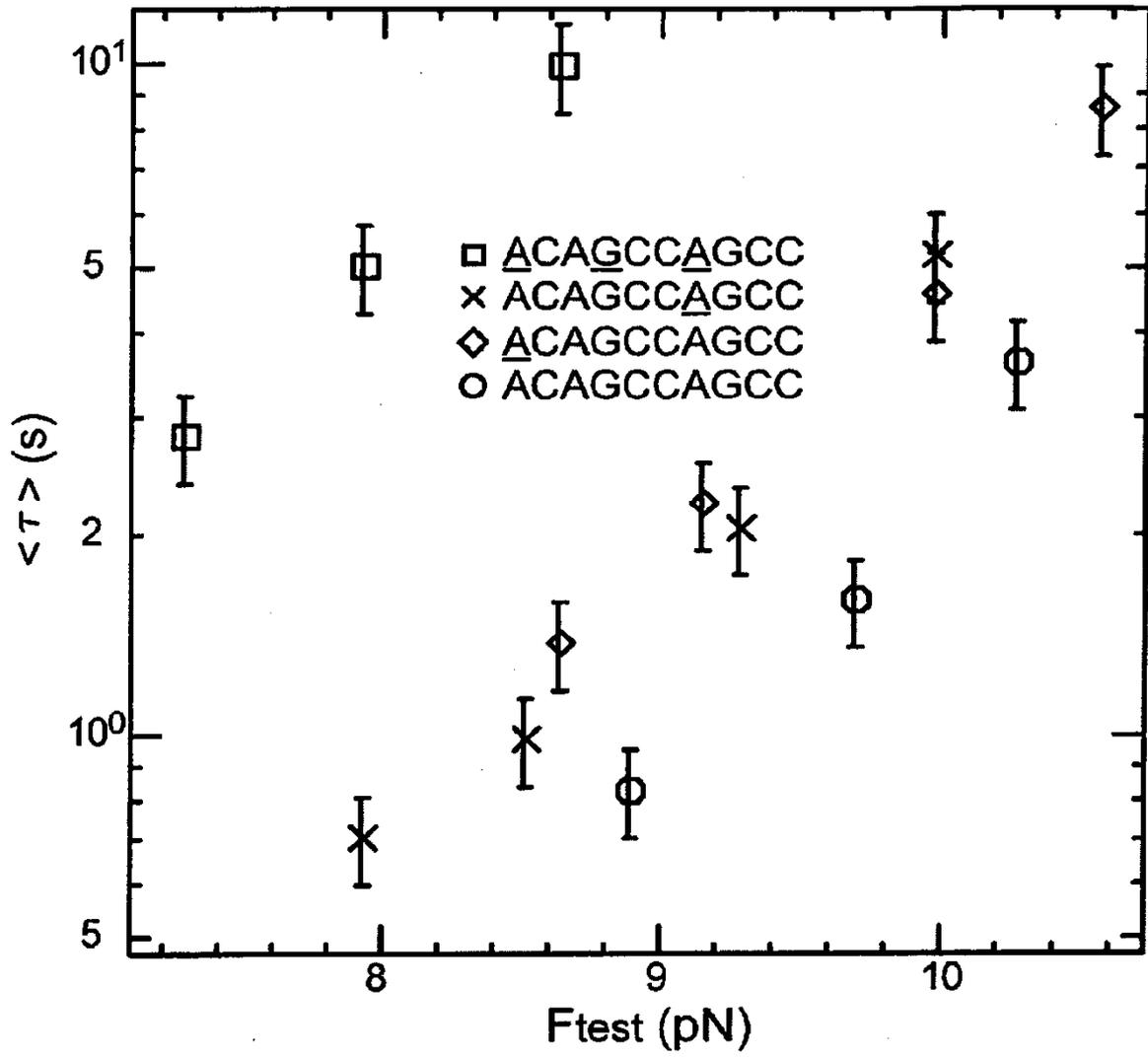


Figura 5b

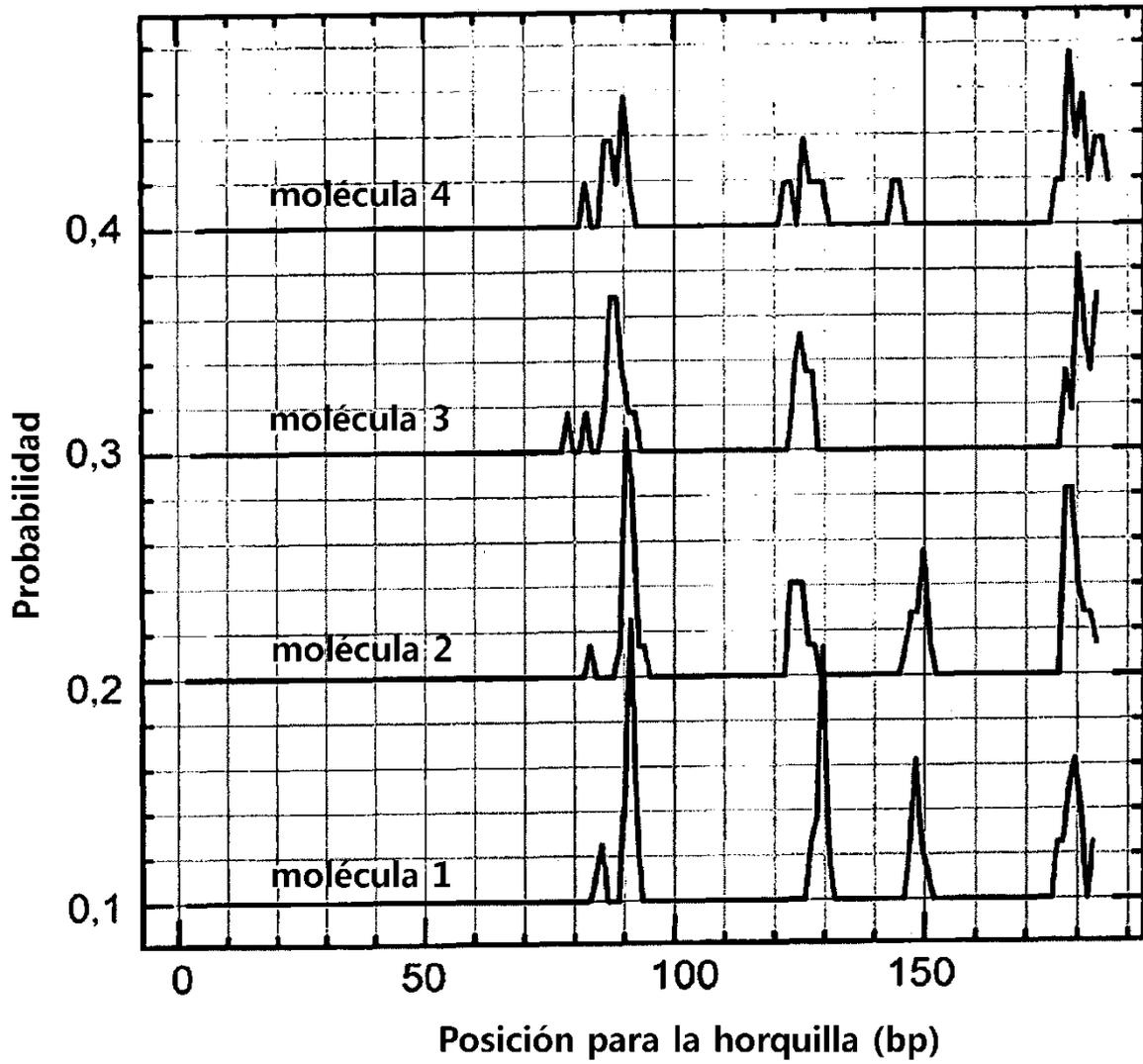


Figura 6

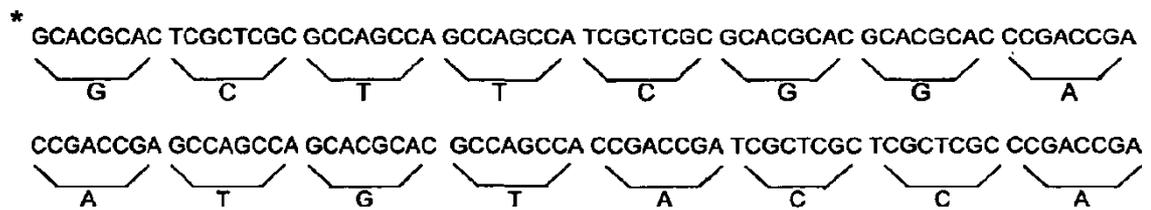
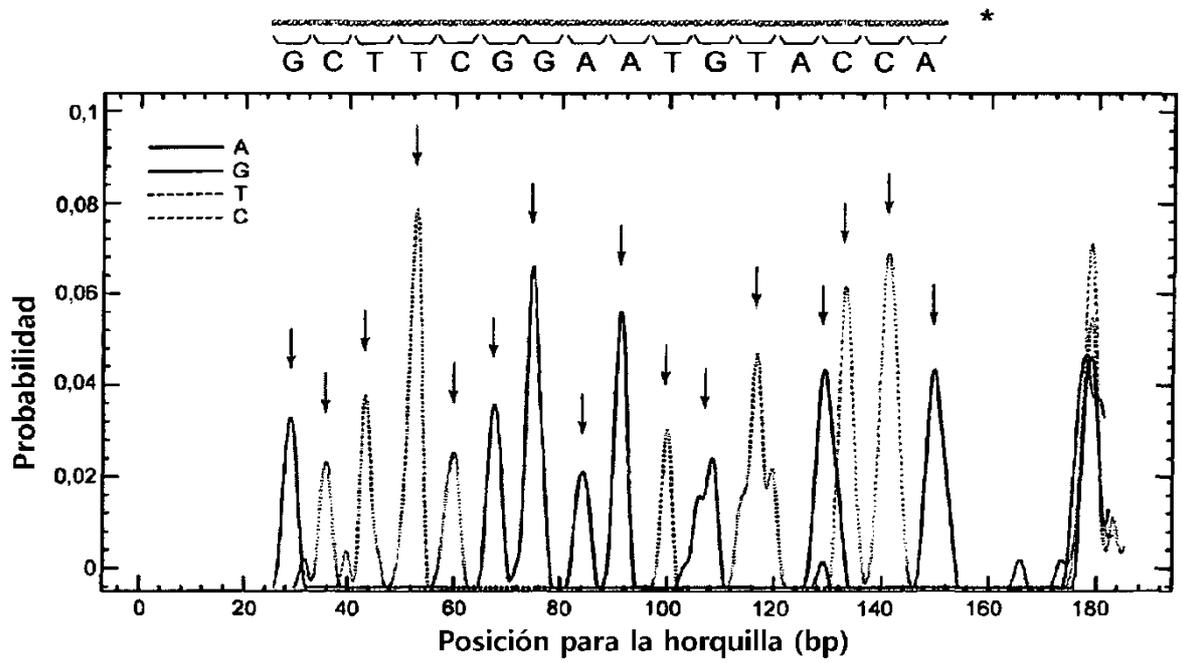


Figura 7