(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 345 137**

(51) Int. Cl.:
C12P 19/34 (2006.01)
C12Q 1/68 (2006.01)
C12M 1/34 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Número de solicitud europea: **00932575 .4**
(96) Fecha de presentación : **18.05.2000**
(97) Número de publicación de la solicitud: **1179085**
(97) Fecha de publicación de la solicitud: **13.02.2002**

(54) Título: **Procedimiento para la secuenciación de moléculas de ácido nucleico.**

(30) Prioridad: **19.05.1999 US 134827 P**

(45) Fecha de publicación de la mención BOPI: **16.09.2010**

(45) Fecha de la publicación del folleto de la patente: **16.09.2010**

(73) Titular/es:
CORNELL RESEARCH FOUNDATION, Inc.
20 Thornwood Drive, Suite 105
Ithaca, New York 14850, US

(72) Inventor/es: **Korlach, Jonas;**
Webb, Watt, W.;
Levene, Michael;
Turner, Stephen;
Craighead, Harold, G. y
Foquet, Mathieu

(74) Agente: **Curell Aguilá, Marcelino**

ES 2 345 137 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 para la secuenciación de moléculas de ácido nucleico.

Campo de la invención

La presente invención se refiere a un procedimiento para determinar la secuencia de moléculas de ácido nucleico.

Antecedentes de la invención

El objetivo que consiste en aclarar el genoma humano completo ha creado un interés en las tecnologías para la secuenciación rápida del ADN, tanto en aplicaciones a pequeña como a gran escala. Los parámetros importantes son la velocidad de secuenciación, la longitud de la secuencia que puede leerse durante una sola serie de secuenciación y la cantidad de plantilla de ácido nucleico requerida. Estos retos en investigación sugieren aspirar a secuenciar la información genética de células aisladas sin ampliación previa, y sin necesidad previa de clonar el material genético en vectores de secuenciación. Los proyectos del genoma a gran escala son actualmente demasiado costosos para ser realizados de modo realista para un gran número de organismos o pacientes. Además, a medida que aumenta el conocimiento de las bases genéticas para las enfermedades humanas, existirá una necesidad siempre creciente de secuenciación de ADN preciso y de alto rendimiento que sea asequible para aplicaciones clínicas. Los procedimientos prácticos para determinar las secuencias de pares de bases de moléculas individuales de ácidos nucleicos, preferentemente con alta velocidad y longitudes largas de lectura, proporcionarían la capacidad de medición necesaria.

Dos técnicas tradicionales para secuenciar el ADN son el método de la terminación didesoxi de Sanger (Sanger *et al*., *Proc. Natl. Acad. Sci.* U.S.A. 74: 563-5467 (1977)) y el método de degradación química de Maxam-Gilbert (Maxam y Gilbert, *Proc. Natl. Acad. Sci.* U.S.A. 74: 560-564 (1977)). Ambos procedimientos suministran cuatro muestras que contiene cada una, una familia de cadenas de ADN en las que todas las cadenas terminan en el mismo nucleótido. La electroforesis en gel de bloque ultrafino, o más recientemente la electroforesis en matriz capilar se utilizan para resolver las cadenas de diferente longitud y determinar la secuencia de nucleótidos, ya sea etiquetando de manera diferente las cadenas de cada muestra antes de la electroforesis para indicar el nucleótido terminal o introduciendo las muestras en diferentes bandas del gel o en diferentes capilares. Tanto el procedimiento de Sanger como el de Maxam-Gilbert son laboriosos y ocupan tiempo, y requieren un pretratamiento dilatado de la fuente de ADN. Se han hecho intentos para hacer uso de la espectroscopia de masas para reemplazar la etapa de electroforesis intensiva. Para el estudio de tecnologías de secuenciación existentes, véase Cheng "High-Speed-Sequence Analysis", *Progr. Biochem. Biophys*. 22: 223-227(1995).

Se han desarrollado procedimientos afines que utilizan marcadores con colorantes o fluorescentes asociadas al nucleótido terminal, donde la determinación de la secuencia se realiza también por electroforesis en gel y detectores fluorescentes automáticos. Por ejemplo, el procedimiento de ampliación de Sanger ha sido modificado recientemente para su utilización en un sistema de microsecuenciación automático que requiere solamente volúmenes de submicrolitros de reactivos y trifosfatos de didesoxiribonucleótido marcados con colorante. En la patente US nº 5.846.727 de Soper *et al.*, se lleva a cabo la detección por fluorescencia en chip con fibra óptica de un solo modo llevando la luz de excitación al canal capilar, y una segunda fibra óptica de un solo modo que recoge los fotones fluorescentes. Se estiman lecturas de las secuencias en el intervalo de 400 a 500 bases lo que no es una mejora significativa sobre la cantidad de información de secuencias obtenidas con los procedimientos de Sanger o de Maxam-Gilbert. Además, el procedimiento de Soper requiere la ampliación por RCP de la plantilla de ADN, y la purificación y la electroforesis en gel de los "escalones" de secuenciación de oligonucleótidos, antes del inicio de la reacción de separación. Estos sistemas requieren todos cantidades significativas de ADN diana. Incluso el procedimiento descrito en la patente US nº 5.302.509 de Cheeseman, que no utiliza electroforesis en gel para la determinación de secuencias, requiere por lo menos un millón de moléculas de ADN.

En una mejora reciente de una metodología de secuenciación por síntesis originalmente inventada hace diez años, la secuencias de ADN se deducen midiendo la liberación de pirofosfato en las pruebas de complejos de ADN/polimerasa con cada trifosfato de desoxirribonucleótido (dNTP) por separado y sucesivamente. Véase Ronaghi *et al.*, "A Sequencing Method Based on Real-Time Pyrophosphate", *Science* 281: 363-365 (1998) y Hyman, "A New Method of Sequencing DNA", *Anal. Biochem*. 174: 423-436(1988). Aunque utilizando nucleótidos naturales, el procedimiento necesita la sincronización de las polimerasas en las cadenas de ADN, que limitan en gran medida las longitudes de lectura de las secuencias. Se consiguieron solo aproximadamente lecturas de 40 nucleótidos y no cabe esperar que el procedimiento de detección pueda aproximarse a la sensibilidad de una sola molécula debido a la eficacia cuántica limitada de la producción de luz por la luciferasa en el procedimiento presentado por Ronaghi *et al.*, "A Sequencing Method Based on Real-Time Pyrophosphate", *Science* 281: 363-365(1998). Además, la velocidad de secuenciación global está limitada por las etapas necesarias de lavado, posteriores a las etapas químicas con objeto de identificar la presencia de pirofosfato, y por el tiempo intrínseco requerido para determinar cada par de bases que ha de secuenciarse con las cuatro bases sucesivamente. Además, se reconocieron dificultades para determinar con precisión los segmentos de homonucleótidos en las secuencias.

Los intentos previos para secuenciar *moléculas aisladas* (en general sin éxito pero originales) han utilizado exonucleasas para liberar sucesivamente bases individuales marcadas con fluorescencia como una segunda etapa una vez la ADN polimerasa ha formado una cadena complementaria completa. Véase Goodwin *et al*., "Application of Single

Molecule Detection to DNA Sequencing", *Nucleos. Nucleot*. 16: 543-550 (1997). Consiste en sintetizar una cadena de ADN marcada con cuatro análogos diferentes de dNTP fluorescentes, después de la degradación de la cadena marcada por la acción de una exonucleasa, y la detección de las bases individuales liberadas en un detector de flujo hidrodinámico. Sin embargo, tanto la polimerasa como la exonucleasa tienen que mostrar actividad en una cadena de ADN muy modificada, y la generación de una cadena de ADN sustituida con cuatro análogos dNTP fluorescentes diferentes no se ha conseguido todavía. Véase Dapprich *et al.*, "DNA Attachment to Optically Trapped Beads in Microstructures Monitored by Bead Displacement", *Bioimaging* 6: 25-32 (1998). Además, se conoce información poco precisa acerca de la relación entre el grado de marcado del ADN y la inhibición de la actividad de la exonucleasa. Véase Dörre *et al.*, "Techniques for Single Molecule Sequencing", *Bioimaging* 5: 139-152 (1997).

En una segunda estrategia que utiliza exonucleasas, se digiere ADN *natural* mientras está siendo impulsado hacia una película líquida delgada con objeto de separar espacialmente los nucleótidos escindidos. Véase Dapprich *et al.*, "DNA Attachment to Optically Trapped Beads in Microstructures Monitored by Bead Displacement", *Bioimaging* 6:25-33 (1998). Éstas a continuación se difunden a corta distancia antes de llegar a inmovilizarse en una superficie para la detección. Sin embargo, la mayoría de las exonucleasas presentan tasas de escisión dependientes de la secuencia y de la estructura, dando como resultado dificultades en el análisis de los datos e igualando conjuntos de secuencias parciales. Además, las formas de identificar las bases en la superficie de detección tienen que desarrollarse o mejorarse todavía.

Independientemente del sistema de detección, los procedimientos que utilizan exonucleasas no han sido desarrollados en los procedimientos que satisfacen la demanda actual de secuenciación rápido y de alto rendimiento. Además, la mayoría de las exonucleasas presentan velocidades de renovación relativamente lentas, y los procedimientos propuestos requieren pretratamiento dilatado, marcado y posterior inmovilización de la plantilla de ADN sobre el lecho en la corriente en circulación del fluido, todo lo cual convierte una realización más complicada en un sistema sencillo de alto rendimiento.

Se han intentado otras estrategias más directas para la secuenciación del ADN, tal como la determinación de la secuencia *espacial* de moléculas de ADN fijas y alargadas por microscopia de sonda atómica explorada. Los problemas encontrados al utilizar estos procedimientos consisten en un espaciado estrecho de las bases en la molécula de ADN (solamente 0,34 nm) y sus pequeñas diferencias fisicoquímicas para ser reconocidas por estos procedimientos. Véase Hansma *et al.*, "Reproducible Imaging and Dissection of Plasmid DNA Under Liquid with the Atomic Force Microscope", *Science* 256: 1180-1184 (1992).

En una estrategia reciente de microsecuenciación que utiliza polimerasa pero no exonucleasa, una serie de moléculas idénticas de ADN de una sola cadena (ssADN) se unen a un sustrato y se determina la secuencia repitiendo una serie de reacciones que utilizan los dNTP marcados por fluorescencia. Patente US nº 5.302.509 de Cheeseman. Sin embargo, este procedimiento requiere que cada base se añada con un marcador fluorescente y grupos de bloqueo 3'-dNTP. Una vez se añade la base y se detecta, el marcador fluorescente y el grupo de bloqueo se separan, y, a continuación, la siguiente base se añade al polímero.

El documento WO 00/36152 describe procedimientos de secuenciación que se basan en la liberación de restos de pirofosfato marcados con fluorescencia [PPi] liberados de trifosfatos de nucleósido [NTP] a medida que se crea el producto de ampliación de la polimerasa.

De este modo, los procedimientos de secuenciación actuales requieren tanto actividad de polimerasa como de exonucleasa para reducir la secuencia o se basan en la polimerasa sola con etapas adicionales de adición y separación de los dNTP bloqueados en 3'. El proyecto del genoma humano ha intensificado la demanda de secuenciación de ADN rápido, a pequeña y gran escala que permita alto rendimiento con mínimo material de partida. Continúa habiendo además necesidad de proporcionar un procedimiento para secuenciar moléculas de ácido nucleico que necesite solamente actividad de polimerasa, sin utilización de los sustituyentes de bloqueo, produciendo mayor simplicidad, miniaturización más fácil y compatibilidad para el tratamiento en paralelo de una técnica de una sola etapa.

La presente invención está dirigida a satisfacer las necesidades de superar las deficiencias en la técnica.

Sumario de la invención

La presente invención se refiere a un procedimiento de secuenciación de una molécula de ácido nucleico que tiene diversas bases nucleotídicas según el procedimiento de la reivindicación 1. Este procedimiento implica proporcionar un complejo de una enzima que polimeriza el ácido nucleico y la molécula de ácido nucleico diana orientados uno con respecto al otro en una posición adecuada para añadir un análogo nucleotídico en un sitio activo complementario con el ácido nucleico diana. Se proporcionan varios tipos de análogos nucleotídicos próximos al sitio activo, en los que cada tipo de análogo nucleotídico es complementario con un nucleótido diferente en la secuencia de ácido nucleico diana. Un análogo nucleotídico se polimeriza en un sitio activo, en el que el análogo nucleotídico que se añade es complementario del nucleótido del ácido nucleico diana, quedando preparado el análogo de ácido nucleico añadido para la adición posterior de análogos nucleotídicos. El análogo nucleotídico añadido al sitio activo se identifica como resultado de la etapa de polimerización. Se repiten las etapas de provisión de varios análogos nucleotídicos, de polimerización e identificación de modo que se determina la secuencia del ácido nucleico diana.

Se describe además un aparato adecuado para secuenciar una molécula de ácido nucleico diana. Este aparato incluye un soporte, así como una enzima que polimeriza el ácido nucleico o un cebador oligonucleotídico adecuado para unirse a una molécula de ácido nucleico diana, donde la polimerasa o el cebador oligonucleotídico está colocado sobre el soporte. Una microestructura define una zona confinada que contiene el soporte y la enzima que polimeriza el ácido nucleico o el cebador oligonucleotídico que está configurado para permitir análogos nucleotídicos marcados que no están colocados sobre el soporte para desplazarse rápidamente en la zona confinada.

Un aparato adecuado para secuenciar una molécula de ácido nucleico diana incluye un soporte sólido y una enzima que polimeriza el ácido nucleico o el cebador oligonucleotídico adecuado para hibridar una molécula de ácido nucleico diana, donde la enzima que polimeriza el ácido nucleico o el cebador oligonucleotídico está colocado sobre el soporte. Un alojamiento define una zona confinada que contiene el soporte y la enzima que polimeriza el ácido nucleico o el cebador oligonucleotídico. El alojamiento se construye para facilitar la identificación de análogos nucleotídicos marcados colocados sobre el soporte. Los guiaondas ópticos se aproximan al foco de la zona confinada que activa la radiación en la zona confinada y recogen la radiación de la zona confinada.

Se han conseguido numerosas ventajas con la presente invención. la secuenciación puede realizarse con pequeñas cantidades de ácido nucleico, con capacidad de secuenciar moléculas con la plantilla de ácido nucleico aislado que elimina la necesidad de la ampliación antes del inicio dla secuenciación. Las longitudes largas de lectura de la secuencia pueden deducirse en una serie, eliminando la necesidad de procedimientos informáticos extensos para ensamblar una secuencia completa sin separaciones de moléculas largas de la plantilla (por ejemplo, clones de cromosoma bacteriano artificial (CBA)). Para dos modos de operación de las presentes invenciones, la longitud de la lectura de las secuencias está limitada por la longitud de la plantilla que va a secuenciarse, o el procesamiento de la polimerasa, respectivamente. Utilizando los sistemas enzimáticos apropiados, por ejemplo, con proteínas accesorias para iniciar la reacción de secuenciación en los sitios específicos (por ejemplo, orígenes de la replicación) en el ácido nucleico de plantilla bicatenario, pueden eliminarse las etapas de preparación necesarias para las técnicas de secuenciación convencionales, tales como la subclonación en vectores de secuenciación.

Además, el procedimiento de secuenciación de la presente invención puede realizarse utilizando polimerasa y no exonucleasa. Esto da como resultado mayor sencillez, miniaturización más fácil y compatibilidad con el tratamiento en paralelo de una técnica de una sola etapa.

En relación con la última ventaja, algunas polimerasas presentan mayor procesividad y velocidades catalíticas que las exonucleasas, añadiéndose más de 10.000 bases antes de la disociación de la enzima para el caso de la T7 ADN polimerasa (comparada con las 3.000 bases para λ exonucleasa). En algunos casos, por ejemplo, la T7 ADN polimerasa acomplejada con T7 belicasa/primasa, los valores de procesividad son aún mayores, oscilando entre varios 100.000. Las velocidades de síntesis de ADN pueden ser muy altas, medidas *in vivo* de 1.000 bases/s. e *in vitro* de 750 bases/s. (en contraste con 12 bases/s. degradadas por la λ exonucleasa *in vitro)*. Véase Kelman *et al*., "Processivity of DNA Polymerases: Two Mechanisms, One Goal", *Structure* 6: 121-125 (1998); Carter *et al*., "The Role of Exonuclease and Beta Protein of Phage Lambda in Genetic Recombination II. Substrate Specificity and the Mode of Action of Lambda Exonuclesse", *J. Biol. Chem*. 246: 2502-2512 (1971); Taboret *et al*., "*Escherichia coli* Thioredoxin Confers Processivity on the DNA Polymerase Activity of the Gene 5 Protein of Bacteriophage T7", *J. Biol. Chem*. 262: 16212-16223 (1987); y Kovall *et al.*, "Toroidal Structure of Lambda-Exonuclease" *Science* 277: 1824-1827 (1997).

Una velocidad de incorporación de 750 bases/s. es aproximadamente 150 veces más rápida que la velocidad de secuenciación de uno de los secuenciadores de ADN ABI Prism 3700 totalmente automático de Perkin, Elmer Corp., Foster City, California, propuesto para ser utilizado en una estrategia de secuenciación por perdigonada para el genoma humano. Véase Venter *et al*., "Shotgun Sequencing of the Human Genome", *Science* 280: 1540-1542 (1998).

El pequeño tamaño del aparato que puede utilizarse para realizar el procedimiento de secuenciación de la presente invención es además muy ventajoso. La zona confinada del complejo plantilla/polimerasa puede ser suministrada por el aparato con microestructura con la posibilidad de matrices que permiten un modo de operación muy paralelo con miles de reacciones de secuenciación realizadas sucesiva o simultáneamente. Esto proporciona una herramienta rápida y ultrasensible para aplicación en investigación así como en diagnósticos médicos.

Breve descripción de los dibujos

Las figuras 1A-C representan 3 formas de realización alternativas de secuenciación según la presente invención.

Las figuras 2A-C son dibujos esquemáticos que presentan la sucesión de etapas utilizadas para secuenciar ácidos nucleicos según la presente invención.

Las figuras 3A-C representan diagramas de señales de fluorescencia frente al tiempo durante la sucesión de etapas utilizada para secuenciar el ácido nucleico según la presente invención. La figura 3C presenta la secuencia generada por estas etapas.

Las figuras 4A-D representan la estructura y dibujos esquemáticos que muestran la sucesión de etapas utilizadas para secuenciar el ácido nucleico según la presente invención en el caso en que se utilicen nucleótidos fluorescentes que llevan el marcador en la posición gamma de fosfato (representada aquí como dNTP unido en gamma).

La figura 5 representa el principio de discriminación de fluoróforos por mediciones del tiempo de debilitamiento de la fluorescencia controlada por el tiempo, que puede utilizarse para suprimir la señal de fondo según la presente invención.

La figura 6A representa un sistema para la secuenciación según la presente invención. La figura 6B es una ampliación de una parte de este sistema.

La figura 7A representa un sistema de secuenciación según la presente invención utilizando la intensificación del campo electromagnético con puntas metálicas. La figura 7B es una ampliación de una parte de este sistema.

La figura 8A presenta un sistema para la secuenciación según la presente invención utilizando aberturas del campo próximo. La figura 8B es una ampliación de una parte de este sistema.

La figura 9A presenta un sistema para la secuenciación según la presente invención utilizando nanocanales. La figura 9B es una ampliación de una parte de este sistema.

Las figuras 10A-B presentan sistemas para proporcionar reactivos a un sistema de confinamiento nanofabricado según la presente invención. En particular, la figura 10A es un dibujo esquemático que representa cómo se proporcionan reactivos y atraviesan el sistema. La figura 10B es similar pero representa este sistema en un solo chip con almohadillas para conectar el sistema a los depósitos de fluido.

Descripción detallada de la invención

La presente invención se refiere a un procedimiento de secuenciación de una molécula de ácido nucleico que presenta una pluralidad de bases nucleotídicas. Este procedimiento implica proporcionar un complejo de una enzima que polimeriza el ácido nucleico y la molécula de ácido nucleico diana orientados uno con respecto al otro en una posición adecuada para añadir un análogo nucleotídico en un sitio activo complementario al ácido nucleico diana, quedando el análogo nucleotídico añadido preparado para la adición posterior de análogos nucleotídicos. El análogo nucleotídico añadido al sitio activo se identifica como resultado de la etapa de polimerización. Se repiten las etapas de provisión de varios análogos nucleotídicos, de polimerización e identificación de modo que se determina la secuencia del ácido nucleico diana.

Un aparato adecuado para secuenciar una molécula de ácido nucleico diana incluye un soporte, así como una enzima que polimeriza el ácido nucleico o un cebador oligonucleotídico adecuado para unirse a una molécula de ácido nucleico diana, donde la polimerasa o el cebador oligonucleotídico está colocado sobre el soporte. Una microestructura define una zona confinada que contiene el soporte y la enzima que polimeriza el ácido nucleico o el cebador oligonucleotídico que está configurado para permitir análogos nucleotídicos marcados que no están colocados sobre el soporte para desplazarse rápidamente en la zona confinada.

Un aparato adicional adecuado para secuenciar una molécula de ácido nucleico diana incluye un soporte sólido y una enzima que polimeriza el ácido nucleico o el cebador oligonucleotídico adecuado para hibridar una molécula de ácido nucleico diana, donde la enzima que polimeriza el ácido nucleico o el cebador oligonucleotídico está colocado sobre el soporte. Un alojamiento define una zona confinada que contiene el soporte y la enzima que polimeriza el ácido nucleico o el cebador oligonucleotídico. El alojamiento se construye para facilitar la identificación de análogos nucleotídicos marcados colocados sobre el soporte. Los guiaondas ópticos se aproximan al foco de la zona confinada que activa la radiación en la zona confinada y recogen la radiación de la zona confinada.

La presente invención se refiere a un procedimiento de secuenciación de una molécula de ácido nucleico diana que presenta una pluralidad de bases. En su principio fundamental, el orden *temporal* de las adiciones de bases durante la reacción de polimerización se mide en una sola molécula de ácido nucleico, es decir, la actividad de una enzima, que polimeriza el ácido nucleico, denominada en adelante también polimerasa, en la molécula de ácido nucleico plantilla que va a ser secuenciada se sigue en tiempo real. La secuencia se deduce identificando qué base se está incorporando en la cadena complementaria en crecimiento del ácido nucleico diana por la actividad catalítica de la enzima que polimeriza el ácido nucleico en cada etapa en la secuencia de adiciones de bases. En la forma de realización preferida de la invención, el reconocimiento de la secuencia temporal de las adiciones de las bases, se consigue detectando la fluorescencia de análogos nucleotídicos marcados de forma apropiada ya que se incorporan en la cadena de ácido nucleico en desarrollo. La precisión del emparejamiento de bases es proporcionada por la especificidad de la enzima, con tasas de error de emparejamiento de bases falsas de 10^{-5} o inferior. Para la fidelidad enzimática, véase Johnson, "Conformational Coupling in DNA-Polymerase Fidelity", *Ann. Rev. Biochem*. 62: 685-713 (1993) y Kunkel, "DNA-Replication Fidelity", *J. Biol. Chem*. 267: 18251-18254 (1992).

La invención se aplica igualmente al secuenciación de todos los tipos de ácidos nucleicos (ADN, ARN, híbridos de ADN/ARN, etc.) utilizando un número de enzimas de polimerización (ADN polimerasas, ARN polimerasas, transcriptasas inversas, mezclas, etc.). Por consiguiente, los análogos nucleotídicos apropiados que actúan como moléculas del sustrato para la enzima que polimeriza el ácido nucleico pueden consistir en miembros de los grupos dNTP, NTP, dNTP o NTP modificados, nucleótidos peptídicos, nucleótidos peptídicos modificados o nucleótidos con esqueleto de fosfato-azúcar modificado.

Existen dos modos de operación convenientes según la presente invención. En el primer modo de operación de la invención, el ácido nucleico de la plantilla está unido a un soporte. Esto puede ser ya sea por inmovilización de (1) un cebador oligonucleotídico o (2) una molécula de ácido nucleico diana monocatenaria o (3) bicatenaria. A continuación, (1) la molécula de ácido nucleico diana está hibridada con el cebador oligonucleotídico unido, (2) un cebador oligonucleotídico está hibridado con la molécula de ácido nucleico diana inmovilizada para formar un complejo de la molécula de ácido nucleico diana cebada, o (3) una zona de reconocimiento para la polimerasa se crea en la plantilla bicatenaria (por ejemplo, por interacción con las proteínas accesorias, tal como un primate). Una enzima que polimeriza ácido nucleico en el complejo de la molécula de ácido nucleico diana cebada, se proporciona en una posición adecuada para desplazarse a lo largo de la molécula de ácido nucleico y ampliar el cebador oligonucleotídico en una zona activa. Una variedad de tipos marcados de análogos nucleotídicos, que no tienen un sustituyente de bloqueo, se suministran próximos a la zona activa, distinguiéndose cada tipo de análogo nucleotídico que es complementario de un nucleótido diferente en la secuencia de ácido nucleico diana. El cebador oligonucleotídico se amplía utilizando una enzima de polimerización del ácido nucleico para añadir un análogo nucleotídico al cebador oligonucleotídico en la zona activa, donde el análogo nucleotídico que se añade es complementario del nucleótido del ácido nucleico diana en la zona activa. Como resultado de la etapa de ampliación se identifica el análogo nucleotídico añadido al cebador oligonucleotídico. Si fuera necesario, el análogo nucleotídico marcado que se añade al cebador nucleotídico, se trata antes de que muchos análogos nucleotídicos más se incorporen en el cebador oligonucleotídico para asegurar que el análogo nucleotídico añadido al cebador oligonucleotídico no impide la detección de análogos nucleotídicos en las posteriores etapas de polimerización e identificación. Se repiten las etapas de provisión de análogos nucleotídicos marcados, que amplían el cebador oligonucleotídico, que identifican el análogo nucleotídico añadido, y que tratan el análogo nucleotídico, de modo que el cebador oligonucleotídico se amplía más y se determina la secuencia del ácido nucleico diana.

Alternativamente, el procedimiento descrito anteriormente puede llevarse a cabo acoplando en primer lugar la enzima que polimeriza el ácido nucleico a un soporte en una posición adecuada para que el complejo de la molécula de ácido nucleico diana se desplace en relación a la enzima de polimerización del ácido nucleico de modo que el complejo molecular del ácido nucleico cebado se extiende a una zona activa. En esta forma de realización, una variedad de análogos nucleotídicos marcados complementarios con el nucleótido del ácido nucleico diana en el sitio activo, se añade a medida que el complejo de ácido nucleico diana cebado se desplaza a lo largo de la enzima que polimeriza el ácido nucleico. Se repiten las etapas de provisión de análogos nucleotídicos, que amplían el cebador, que identifican el análogo nucleotídico añadido, y que tratan el análogo nucleotídico durante o después la incorporación, como se describió anteriormente, de modo que se amplía más el cebador oligonucleotídico y se determina la secuencia del ácido nucleico diana.

Las figuras 1A-C representan 3 formas de realización alternativas de secuenciación según la presente invención. En las figuras 1A un cebador de secuenciación está unido a un soporte, por ejemplo, mediante un enlace biotina-estreptavidina, con el cebador hibridado a la molécula de ácido nucleico diana y la enzima que polimeriza el ácido nucleico unida a la molécula de ácido nucleico hibridada en el sitio activo donde los análogos nucleotídicos se están añadiendo al cebador de secuenciación. En la figura 1B, la molécula de ácido nucleico diana, está unida a un soporte, con un cebador de secuenciación hibridado con la molécula de ácido nucleico plantilla y la enzima de polimerización del ácido nucleico unida a la molécula de ácido nucleico hibridada en la zona activa donde se están añadiendo los análogos nucleotídicos al cebador de secuenciación. El cebador puede añadirse antes o durante la provisión de análogos nucleotídicos. Además de estos escenarios, una molécula de ácido nucleico diana bicatenaria puede unirse a un soporte, conteniendo la molécula de ácido nucleico diana una zona de reconocimiento para unirse a la enzima de polimerización de ácido nucleico en una zona activa donde los análogos nucleotídicos se están añadiendo al cebador. Por ejemplo, dicha zona de reconocimiento puede crearse con ayuda de una proteína accesoria, tal como una ARN polimerasa o una helicasa/primasa que sintetizará un cebador corto en las zonas específicas del ácido nucleico diana y de este modo proporcionan una zona de partida para la enzima que polimeriza el ácido nucleico. Veáse Richardson, "Bacteriophage T7: Minimal Requeriment for the Replication of the Duplex DNA Molecule", *Cell* 33: 315-317 (1983). En la figura 1C, la enzima que polimeriza el ácido nucleico está unida a un soporte uniéndose en la molécula de ácido nucleico diana cebada a la zona activa donde análogos nucleotídicos se están añadiendo al cebador de secuenciación. Como en la descripción anterior, la enzima de polimerización de ácido nucleico puede asimismo estar unida a un soporte, pero siendo la molécula de ácido nucleico diana ácido nucleico bicatenario, que contiene una zona de reconocimiento para la unión de la enzima de polimerización del ácido nucleico en una zona activa donde los análogos nucleotídicos se están añadiendo a la cadena de ácido nucleico en crecimiento. Aunque las figuras 1A-C muestran solamente una reacción de secuenciación que se lleva a cabo sobre el soporte, es posible conducir una matriz de varias de dichas reacciones a zonas diferentes en un solo soporte. En esta forma de realización alternativa, cada cebador de secuenciación, ácido nucleico diana o enzima de polimerización de ácido nucleico que ha de inmovilizarse sobre el soporte sólido está moteada en esta superficie por impresión por microcontacto o estampado, por ejemplo, tal como se utiliza para la tecnología de chips de los chips de ADN, o mediante la formación de una red de puntos de enlace tratando la superficie del soporte sólido. También es posible combinar las formas de realización expuestas en la figura 1 e inmovilizar tanto la molécula de ácido nucleico diana como la enzima de polimerización del ácido nucleico próximas entre sí.

El procedimiento de secuenciación de la presente invención puede utilizarse para determinar la secuencia de la molécula de ácido nucleico, incluyendo el ADN bicatenario o monocatenario, las horquillas de ADN monocatenario, los híbridos de ADN/ARN, ARN con una zona de reconocimiento para la unión de la ARN polimerasa, u horquillas de ARN.

El cebador de secuenciación utilizado para la realización del procedimiento de la presente invención puede ser un ribonucleótido, desoxirribonucleótido, ribonucleótido modificado, desoxirribonucleótido modificado, ácido nucleico peptídico, ácido nucleico peptídico modificado, oligonucleótido con esqueleto de fosfato-azúcar modificado, y otros análogos nucleotídicos y oligonucleotídicos. Puede ser sintético o producido en la naturaleza por primasas, ARN polimerasas u otras enzimas sintetizadoras de oligonucleótidos.

La enzima que polimeriza el ácido nucleico utilizada según en la presente invención puede ser una polimerasa termoestable o una polimerasa térmicamente degradable. Los ejemplos de polimerasas termoestables adecuadas incluyen las polimerasas aisladas de *Thermus aquaticus*, *Thermus thermophilus*, *Pyrococcus woesei*, *Pyrococcus furiosus*, *Thermococcus litoralis* y *Thermotoga maritima*. Las polimerasas termodegradables útiles incluyen la ADN polimerasa de *E. coli*, el fragmento de Klenow de la ADN polimerasa de *E. coli*, T4 ADN polimerasa, T7 ADN polimerasa y otras. Los ejemplos de otras enzimas de polimerización que pueden utilizarse para determinar la secuencia de moléculas de ácido nucleico incluyen las T7, T3, SP6 ARN polimerasas de *E. coli*, y las transcriptasas inversas de VAM, M-MLV y VIH. La polimerasa puede estar unida a la secuencia de ácido nucleico diana cebada a un ácido nucleico monocatenario cebado, a un origen de replicación, a una hendidura o separación en un ácido nucleico bicatenario, a una estructura secundaria en un ácido nucleico monocaternario, a una zona de unión creada por un proteína accesoria o a un ácido nucleico monocatenario cebado.

Los materiales que son útiles en la formación del soporte incluyen, vidrio, vidrio con modificaciones en la superficie, silicio, metales, semiconductores, dieléctricos de alto índice de refracción, cristales, geles y polímeros.

En las formas de realización de las figuras 1, podría utilizarse cualquier acompañante de unión adecuada conocido por los expertos en la materia, para inmovilizar bien el cebador de secuenciación, la molécula de ácido nucleico diana o la enzima que polimeriza el ácido nucleico al soporte. También es posible el enlace no específico por adsorción. Como se representa en las figuras 1A-C, un enlace de biotina-estreptavilina es adecuado para unir el cebador de secuenciación o la molécula de ácido nucleico diana al soporte sólido. El componente de biotina de dicho enlace puede estar unido al cebador o al ácido nucleico o al soporte sólido con la estreptavidina o cualquier otra proteína que se une a la biotina, estando acoplado a la entidad opuesta.

Una estrategia para llevar a cabo esta técnica de fijación implica acoplar BIOTINA FOTOACTIVABLE™ ("PAB") (Pierce Chemical Co., Rockford, Illinois) a una superficie de la cámara utilizada para realizar el procedimiento de secuenciación de la presente invención. Este puede conseguirse por exposición a la luz a 360 mn, preferentemente a través de una pared transparente de la cámara, como se describe en Hengsakul *et al*., "Protein Patterning with a Photoactivable Derivative of Biotin", *Bioconjugate Chem*. 7: 249-54 (1996). Cuando se utiliza una nanocámara se activa la biotina en un punto limitado por difracción al microscopio óptico. Con excitación de campo próximo, la exposición puede alinearse utilizando un guiaondas para dirigir la luz al área deseada. Cuando se expone a la luz, la PAB se activa y se une por enlace covalente a la superficie interior del canal. El exceso de PAB sin unir se elimina entonces enjuagando con agua.

Alternativamente, la estreptavidina puede revestirse sobre la superficie del soporte. El cebador oligonucleotídico de ácido nucleico apropiado o la plantilla de ácido nucleico monocatenario se biotinila a continuación, creando un complejo inmovilizado de cebador de ácido nucleico-molécula diana en virtud del cebador unido a estreptavidina-biotina.

Otra estrategia para llevar a cabo el procedimiento de la presente invención consiste en utilizar ácidos nucleicos complementarios para unir el cebador de secuenciación o la molécula de ácido nucleico diana al soporte sólido. Esto puede llevarse a cabo modificando un ácido nucleico monocatenario con una secuencia principal conocida y ligando la secuencia principal conocida al cebador de secuenciación o a la molécula de ácido nucleico diana. El oligonucleótido resultante puede unirse a continuación por hibridación a un oligonucleótido acoplado al soporte y que presenta una secuencia de nucleótidos complementaria de la de secuencia principal conocida. Alternativamente, un segundo oligonucleótido puede hibridarse a un extremo de la molécula de ácido nucleico diana opuesto a este enlace al cebador oligonucleotídico. Este segundo oligonucleótido está disponible para la hibridación a una secuencia nucleica complementaria unida al soporte.

El enlace reversible o irreversible entre el soporte y el cebador oligonucleotídico y la secuencia de ácido nucleico diana puede conseguirse con los componentes de cualquier par de enlace covalente o no covalente. Otros de dichos planteamientos para inmovilizar el cebador de secuenciación o la molécula de ácido nucleico diana al soporte, incluyen un par con enlace anticuerpo-antígeno y moléculas de acoplamiento fotoactivadas.

En la forma de realización de la figura 1C, puede utilizarse cualquier técnica conocida por resultar de utilidad para inmovilizar de manera reversible o irreversible los materiales proteicos. Se ha publicado en la bibliografía que la ADN polimerasa, se inmovilizó con éxito en superficies activadas sin pérdida de actividad catalítica. Véase Yin *et al*., "Transcription. Against an Applied Force", *Science* 270:1653-1657 (1995).

Alternativamente, la proteína puede unirse a un anticuerpo que no interfiere con su actividad catalítica, como se ha publicado para la transcriptasa inversa del VIH. Véase Lennerstrand *et al*., "A Method for Combined Inmunoaffinity Purification and Assay of HIV-1 Reverse Transcriptase Activity Useful for Crude Samples", *Anal. Biochem*. 235: 141-152 (1996).

Por consiguiente, las enzimas que polimerizan el ácido nucleico pueden inmovilizarse con pérdida de función. Los anticuerpos y otras proteínas pueden diseñarse sobre superficies inorgánicas. Véase James *et al.*, "Patterned Protein layers on Solid Substrates by Thin Stamp Microcontact Printing". *Langmuir* 14:741-744 (1998) y St John *et al.*, "Diffraction-Based Cell detection Using a Microcontact Printed Antibody Grating", *Anal. Chem.* 70:1108-1111 (1998).

Alternativamente, la proteína se podría biotinilar (o marcar igualmente con otras moléculas de enlace) y a continuación unirse a una superficie de soporte revestida con estreptavidina.

En cualquiera de las formas de realización de las figuras 1A a C, el acompañante de enlace y la polimerasa o los ácidos nucleicos que inmovilizan pueden aplicarse al soporte por técnicas químicas y fotolitográficas convencionales que son bien conocidas en la técnica. Generalmente, estos procedimientos pueden llevar modificaciones convencionales en la superficie química del soporte, la incubación del soporte a diferentes temperaturas en diferentes medios y posibles etapas sucesivas de lavado e incubación de la superficie del soporte con las moléculas respectivas.

Son posibles posibilidades alternativas de colocación del complejo de polimerización, tales como por oclusión del complejo en los poros que contienen el gel demasiado pequeños para permitir el paso del complejo, pero lo suficientemente grandes para acomodar la provisión de análogos nucleotídicos. Los medios adecuados incluyen geles de agarosa, geles de poliacrilamida, materiales porosos sintéticos o nanoestructuras.

El procedimiento de secuenciación de la presente invención puede iniciarse por adición de la enzima que polimeriza el ácido nucleico a la mezcla de reacción en la forma de realización de las figuras 1A-B. Para la forma de realización de la figura 1C, el ácido nucleico cebado puede añadirse para el inicio. Pueden emplearse otros escenarios para el inicio, tales como la creación de un complejo ácido nucleico-polimerasa preformado en ausencia de iones metálicos bivalentes que forman partes integrales de las zonas activas de las polimerasas (más generalmente Mg^{2+}). La reacción de secuenciación puede iniciarse entonces añadiendo estos iones metálicos. El complejo de preiniciación de la plantilla podría encontrase también con la enzima en ausencia de nucleótidos, añadiéndose análogos nucleotídicos fluorescentes al comienzo de la reacción. Véase Huber *et al.*, "*Escherichia coli* Thioredoxin Stabilizes Complexes of Bacteriophage T7 DNA Polimerase and Primed Templates", *J. Biol. Chem.* 262:16224-16232 (1987). Alternativamente, el proceso puede iniciarse mediante la liberación de un grupo en el cebador oligonucleotídico que le protege del enlace a la enzima que polimeriza el ácido nucleico. La iluminación con rayo láser iniciaría a continuación la reacción coincidiendo con el punto de inicio de la observación.

Las figuras 2A-C son dibujos esquemáticos que muestran las sucesión de etapas utilizadas para secuenciar ácidos nucleicos según la presente invención.

En la figura 2A, están presentes análogos nucleotídicos marcados en la proximidad del complejo cebado de una enzima que polimeriza el ácido nucleico acoplada al cebador de secuenciación hibridado y a la molécula de ácido nucleico diana que están acoplados sobre el soporte sólido. Durante esta fase del proceso de secuenciación, los análogos nucleotídicos marcados se difunden o son forzados a circular a través del medio de ampliación hacia y alrededor del complejo cebado.

Según la figura 2B, una vez un análogo nucleotídico ha alcanzado la zona activa del complejo cebado, se une a él y las enzimas que polimerizan el ácido nucleico prueban si este análogo nucleotídico es complementario a la primera base abierta de la molécula de ácido nucleico diana o si representa una incompatibilidad. Esta base incompatible será rechazada con muchas probabilidades de que corresponda a la alta fidelidad de la enzima mencionada anteriormente, mientras que el análogo nucleotídico complementario se polimeriza con el cebador de secuenciación para ampliar el cebador de secuenciación.

Durante o después de añadir cada análogo nucleotídico marcado al cebador de secuenciación, se identifica el análogo nucleotídico añadido al cebador de secuenciación. Esto se consigue de la manera más eficaz proporcionando a cada análogo nucleotídico un marcador distinguible diferente. Al detectar cuál de los diferentes marcadores se añaden al cebador de secuenciación, el análogo nucleotídico correspondiente añadido al cebador de secuenciación puede identificarse y, en virtud de su naturaleza complementaria, puede determinarse la base del ácido nucleico diana que complementa el análogo nucleotídico. Una vez se consigue esto, ya no es necesario el análogo nucleotídico que se añadió al cebador de secuenciación para conservador su marcador. De hecho, la presencia continua de marcadores en bases que complementan análogos nucleotídicos en el ácido nucleico diana que ya ha sido secuenciación interferirían muy probablemente con la detección de análogos nucleotídicos sucesivamente añadidos al cebador. Por consiguiente, los marcadores añadidos al cebador de secuenciación se eliminan después de haber sido detectados, tal como se representa en la figura 2C. Esto tiene lugar preferentemente antes de que se incorporen análogos nucleotídicos adicionales en el cebador oligonucleotídico.

Repitiendo la secuencia de etapas descrita en las figuras 2A-C, se amplía el cebador de secuenciación y, como resultado, puede determinarse la secuencia completa del ácido nucleico diana. Aunque la forma de realización de inmovilización representada en las figuras 2A-C es la presentada en la figura 1A, las formas de realización de inmovilización alternativa representadas en las figuras 1B-C, podrían utilizarse asimismo en la realización de la sucesión de etapas representadas en las figuras 2A-C.

Al realizar la difusión, la incorporación y las etapas de eliminación de las figuras 2A-C, se utiliza un medio de ampliación que contiene los componentes apropiados para permitir añadir los análogos nucleotídicos al cebador de secuenciación. Los medios de ampliación adecuados incluyen, por ejemplo, una solución que contiene Tris-HCl 50 mM, pH 8,0, $MgCl_2$ 25 mM, NaCl 65 mM, DTT 3 mM (esta es la composición del medio de ampliación recomendada por el fabricante para Secuenasa, T7 ADN polimerasa mutante) y análogos nucleotídicos a una concentración apropiada que permita la identificación de la secuencia. Son posibles otros medios que son apropiados para esto y otras polimerasas, con o sin proteínas accesorias, tales como las proteínas de enlace monocatenarias. Preferentemente, la fase de ampliación se lleva a cabo a 37ºC para la mayoría de las polimerasas térmicamente degradables, aunque pueden utilizarse otras temperaturas a las cuales la polimerasa es activa.

Una vez añadido el análogo del nucleótido marcado al cebador de secuenciación, como se indicó anteriormente, el marcador específico del resto añadido debe identificarse para determinar qué tipo de análogo nucleotídico se añadió al cebador de secuenciación y, como resultado qué base complementaria de ácido nucleico diana es. Cómo se determina el marcador de la entidad añadida depende del tipo de marcador que se utilice. Para la forma de realización preferida de la invención, la exposición de las etapas de identificación se limitará a la utilización de análogos nucleotídicos portadores de restos fluorescentes. Sin embargo, otros marcadores adecuados incluyen cromóforos, enzimas, antígenos, metales pesados, sondas magnéticas, colorantes, grupos fosforescentes, materiales radioactivos, restos quimioluminiscentes, nanopartículas dispersantes o fluorescentes, restos generadores de la señal Raman, y restos de detectores electroquímicos. Dichos marcadores son conocidos en la técnica y se dan a conocer por ejemplo en Prober *et al.*, *Science* 238: 336-41 (1997); Connell *et al.*, *BioTechniques* 5(4): 342-84 (1987); Ansorge, *et al.*, *Nucleic Acids Res.* 15 (11): 4593-602 (1987); y Smith *et al.*, *Nature* 321:674 (1986).

En algunos casos, tales como para los cromóforos, fluoróforos, marcadores fosforescentes, nanopartículas o grupos de señalización Raman, es necesario someter la zona de reacción a radiación activadora para detectar el marcador. Este procedimiento se expondrá con detalle a continuación para el caso de marcadores fluorescentes. Las técnicas adecuadas para detectar el marcador fluorescente incluyen la microespectroscopia en campo lejano resuelta en el tiempo, la microespectroscopia en campo próximo, la medición de la transferencia de energía por resonancia de fluorescencia, la fotoconversión y la medición de los tiempos de permanencia de la fluorescencia. La identificación de fluoróforos, puede conseguirse por discriminación de la longitud de onda del espectro, la medición y separación de los tiempos de permanencia de la fluorescencia, la identificación de fluoróforos y/o la supresión del fondo. La identificación de fluoróforos y/o la supresión del fondo, pueden facilitarse por conexión rápida entre los modos de excitación y las fuentes de iluminación y combinaciones de los mismos.

Las figuras 3A-B, presentan gráficos de señales de fluorescencia frente al tiempo durante la sucesión de las etapas (expuestas en la figura 2) que se utilizan para realizar el procedimiento de secuenciación de la presente invención. En esencia, en este procedimiento, se distinguirá un nucleótido incorporado de los no incorporados (que se difunden al azar a través del volumen de observación o que son transferidos por convección a través de él por flujo hidrodinámico o electroforético) analizando los tiempos minúsculos de fluorescencia para cada marcador distinguible simultáneamente. Esto se consigue mediante registros de ráfagas de fotones y espectroscopia de correlación de fluorescencia resuelta en el tiempo que distingue la fluorescencia estacionaria continua del marcador incorporado (hasta la eliminación por los mecanismos expuestos a continuación) de la emisión intermitente de los fluoróforos libres. Véase Magde *et al.*, "Thermodynamic Fluctuations in a Reacting System-Measurement by Fluorescence Correlation Spectroscopy", *Phys. Rev. Lett.* 29:705-708 (1972), Kask. P. *et al.*, "Fluorescence-Intensity Distribution Analysis and its Application in Biomolocular Detection Technology", *Proc. Nat. Acad. Sci.* U.S.A 96:13756-13761 (1999) y Eggeling *et al.*, "Monitoring Conformational Dynamics of a Single Molecule by Selective Fluorescence Spectroscopy", *Proc. Nat. Acad. Sci.* U.S.A 95:1556-1561 (1998). La secuencia puede deducirse combinando los tiempos minúsculos de todos los canales de detección.

La figura 3A, presenta un gráfico de señal de fluorescencia frente al tiempo durante solo la fase de difusión de la figura 2A, suponiendo cuatro canales diferentes de detección de fluorescencia para las cuatro bases diferentes (por ejemplo, utilizando cuatro marcadores diferentes, cada uno con un espectro de emisión de fluorescencia diferente, mediante el que pueden separarse a través de filtros ópticos). Cada pico en la figura 3A, representa la ráfaga de fluorescencia resultante de la presencia de un análogo nucleotídico en el volumen de observación, siendo distinguido cada análogo nucleotídico diferente por su diferente marcador que genera picos de diferentes colores (representados en la figura 3A, por diferentes tipos de línea). La estrechez de estos picos indica que los análogos nucleotídicos tienen un tiempo de residencia breve próximo a la zona activa de secuenciación, debido a que se difunden libremente o que circulan a través del volumen de observación. Cabe esperar un pico de anchura similar para el caso de un análogo nucleotídico *incompatible* que se une temporalmente a la zona activa de la enzima que polimeriza el ácido nucleico y el rechazo posterior de la incorporación por la enzima.

La figura 3B, presenta un gráfico de señal de fluorescencia frente al tiempo durante la incorporación y posteriores y fases de eliminación de las figuras 2B-C. Como en la figura 3A, cada pico de la figura 3B representa la presencia de un análogo nucleotídico distinguiéndose cada análogo nucleotídico diferente por su marcador diferente que genera picos de colores diferentes (representados en la figura 3B por diferentes tipos de línea). La estrechez de algunos picos en la figura 3B, se refiere de nuevo a los análogos nucleotídicos que permanecen móviles en el medio de ampliación y no amplían el cebador de secuenciación. Dichos picos estrechos resultan porque estos análogos nucleotídicos tienen un tiempo de residencia breve próximo a la zona activa de secuenciación como se explicó para la figura 3A. Por otra parte, los picos más amplios corresponden a análogos nucleotídicos que tienen, en la zona activa, bases complementa-

rias en la molécula plantilla de ácido nucleico y sirven para ampliar el cebador de secuenciación. Como resultado de su inmovilización, estos análogos nucleotídicos tienen picos más amplios, debido a que permanecerán en el volumen de observación durante y después de la incorporación en la cadena de ácido nucleico en desarrollo, y de este modo continúan emitiendo fluorescencia. Su señal se termina solamente después en el tiempo como resultado de la ulterior etapa de eliminación que elimina la fluorescencia continua y que permite la identificación de episodios de incorporación ulteriores.

Desplazándose de izquierda a derecha en la figura 3B (es decir, tiempo después), la secuencia de los picos más amplios corresponde al complemento de la secuencia de la molécula plantilla de ácido nucleico. La figura 3C, presenta el resultado final y de la figura 3B, que puede conseguirse, mediante un programa informático, que detecta las ráfagas breves de fluorescencia y las descarga en el resultado final. Como resultado de dicha filtración, únicamente los picos generados por los análogos nucleotídicos inmovilizados están presentes, y se transfieren por convección a la secuencia que corresponde al complemento de la secuencia de la molécula plantilla de ácido nucleico. Esta secuencia complementaria es aquí ATACTA, por consiguiente, el orden de las bases de la molécula plantilla de ácido nucleico que se secuencia es TATGAT.

Pueden acoplarse marcadores fluorescentes a nucleótidos en varias posiciones. El acoplamiento puede hacerse con o sin un enlazador en puente al nucleótido. Convencionalmente los análogos nucleotídicos utilizados para el marcado del ácido nucleico con fluoróforos transportan el resto fluorescente acoplado a la *base* de la molécula del sustrato nucleotídico. Sin embargo, pueden acoplarse también a un resto de azúcar (por ejemplo, desoxirribosa) o alfa-fosfato. El acoplamiento al alfa-fosfato, se mostraría ventajoso, porque este tipo de enlace deja intacta la estructura interna del ácido nucleico, mientras que los fluoróforos acoplados a la base se ha observado que distorsionan la doble hélice de la molécula sintetizada y posteriormente inhiben más la actividad de la polimerasa. Véase Zhu *et al.*, "Directly Labelled DNA Probes Using Fluorescent Nucleotides with Different Lenght Linkers", *Nucleic Acids Res*. 22: 3418-3422 (1994) y Doublie *et al.*, "Crystal Structure of a Bacteriophage T7 DNA Replication Complex at 22 Angstrom Resolution", *Nature* 391: 251-258 (1998).

De este modo, los nucleótidos que contienen el grupo tiol, que se han utilizado (en forma de NTP) para los estudios de reticulación en ARN polimerasa, podrían actuar como moléculas con eje central primario para el acoplamiento de enlazadores y marcadores fluorescentes adecuados. Véase Hanna *et al.*, "Synthesis and Characterization of a New Photo-Cross-Linking CTP Analog and Its Use Photoaffinity-Labeling *Escherichia-coli* and T7-RNA Polymerases". *Nucleic Acids Res*. 21: 2073-2079 (1993).

En el caso convencional en el que el fluoróforo se acople a la base del nucleótido, está provisto por lo general de fluoróforos de un tamaño relativamente grande, tal como fluoresceína. Sin embargo, fluoróforos más pequeños, por ejemplo, pireno o colorantes de la familia de la cumarina, demostrarían ser ventajosos desde el punto de vista de ser tolerados en mayor medida por las polimerasas. De hecho, es posible sintetizar un fragmento de ADN de 7.300 pares de bases de longitud en el que un tipo de base se reemplaza completamente por el correspondiente dNTP marcado con cumarina que utiliza T7 ADN polimerasa, mientras que la enzima no puede llevar a cabo la correspondiente síntesis utilizando dNTP marcados con fluoresceína.

En todos estos casos, el fluoróforo continúa acoplado a la parte de la molécula del sustrato que se incorpora en la molécula de ácido nucleico en desarrollo durante la síntesis. Se han detectado e identificado medios adecuados para la eliminación del fluoróforo después según el esquema de secuenciación de la presente invención incluyen el fotoblanqueo del fluoróforo o la escisión fotoquímica del nucleótido y del fluoróforo, por ejemplo, la escisión de un enlace químico en el enlazador. La eliminación del marcador fluorescente de los nucleótidos ya incorporados, cuya cantidad puede ajustarse por la potencia del láser, impide la acumulación de la señal en la cadena de ácido nucleico maximizando de este modo la señal a la relación de fondo para la identificación del nucleótido. Para este esquema, el objetivo de la presente invención consiste en detectar todos los fotones de cada marcador y a continuación fotoblanquear o escindir fotoquímicamente antes o inmediatamente después de los siguientes pocos nucleótidos se incorporen para mantener los adecuados valores de señal a ruido para posteriores etapas de identificación. La fase de eliminación del proceso de la presente invención puede llevarse a cabo por cualquier procedimiento adecuado para eliminar una etiqueta sin dañar el complejo de la reacción de secuenciación.

Además de los marcadores fluorescentes que quedan en el ácido nucleico durante la síntesis, los nucleótidos que están marcados con fluorescencia o de otro modo y transportan el marcador acoplado a uno de los dos, el fosfato *beta* o *gamma* del nucleótido puede utilizarse también en el procedimiento de secuenciación de la presente invención. Anteriormente se han sintetizado compuestos análogos en forma de análogos de NTP y se ha demostrado que son excelentes sustratos para varias enzimas, incluyendo las ARN polimerasas. Véase Yarbrough *et al*., "Synthesis and Properties of Fluorescent Nucleotide Substrates for DNA-dependent RNA Polymerase", *Journal Of Biological Chemistry* 254: 12069-12073 (1979), y Chatterji *et al*., "Fluorescence Spectroscopy Analysis of Active and Regulatory Sites of RNA Polymerase", *Methods in Enzymology* 274: 456-479 (1996).

Durante la síntesis del ADN, la escisión del enlace en el nucleótido tiene lugar entre el fosfato alfa y el beta, dando lugar a que los fosfatos beta y gamma se liberen de la zona activa después de la polimerización, y el pirofosfato formado posteriormente se difunde o es transferido por convección fuera del ácido nucleico. Según la presente invención, es posible distinguir el episodio de fijación de un nucleótido y su incorporación en el ácido nucleico a partir de los episodios que implican solo la fijación (y el rechazo posterior), de un nucleótido incompatible, debido a que las

velocidades constantes de estos dos episodios son drásticamente diferentes. La etapa limitativa de la velocidad en las etapas elementales sucesivas de polimerización del ADN es un cambio de conformación de la polimerasa que puede ocurrir solamente una vez la enzima ha establecido que el nucleótido correcto (compatible) está unido a la zona activa. Por consiguiente, un episodio de un enlace incompatible de un análogo nucleotídico será mucho más corto en el tiempo que el episodio de la incorporación de la base correcta. Véase Patel *et al.*, "Pre-Steady-State Kinetic Analysis of Processive DNA Replication Including Complete Characterization of an Exonuclease-Deficient Mutant", *Biochemistry* 30: 511-525 (1991) y Wong *et al.*, "An Induced-Fit Kinetic Mechanism for DNA Replication Fidelity: Direct Measurement by Single-Turnover Kinetics". *Biochemistry* 30: 511-525 (1991). Como resultado, la fluorescencia del marcador que está acoplado al fosfato beta o gamma del análogo nucleotídico continúa siendo próxima a la de las polimerasas durante un más tiempo en el caso de que se polimerice el análogo nucleotídico, y puede distinguirse según el esquema descrito anteriormente para la figura 3. Tras la incorporación, el marcador se difundirá con el pirofosfato escindido. Este procedimiento se muestra en la figura 4. La figura 4A presenta la estructura de I-aminonaftalen-5-sulfonato (AmNS)-dUTP, ejemplo representativo de un análogo nucleotídico que lleva un marcador fluorescente acoplado al fosfato gamma, con la posición de la escisión indicada por la línea de puntos. La figura 4B-D presenta las etapas sucesivas de incorporación y la liberación del complejo pirofosfato-fluoróforo, en analogía con la figura 2. El tiempo minúsculo de fluorescencia para este esquema será el mismo que se muestra en la figura 3. De este modo, existe un esquema alternativo al expuesto anteriormente en el que fluoróforo se incorpora en primer lugar en el ácido nucleico y la señal se elimina posteriormente por fotoblanqueo o escisión fotoquímica tras la identificación del marcador.

La identificación del análogo nucleotídico específico marcado por fluorescencia que se incorpora frente al fondo de nucleotídicos no incorporados que se difunden o fluyen cerca de la enzima que polimeriza el ácido nucleico pueden aumentarse más empleando la observación de que para determinados dNTP marcados por fluorescencia (por ejemplo, cumarina-5-dGTP o AmNS-UTP la presencia de la base en forma de enlace covalente reduce de manera significativa (es decir, extingue) la fluorescencia del marcador. Véase Dhar *et al.*, "Synthesis and Characterization of Stacked and Quenched Uridine Nucleotide Fluorophores", *Journal of Biological Chemistry* 274: 14568-14572 (1999), y Draganescu *et al.*, "Fhit-Nucleotide Specificity Probed with Novel Fluorescent and Fluorogenic Substrates", *Journal of Biological Chemistry* 275: 4555-4560 (2000). La interacción entre el fluoróforo y la base extingue la fluorescencia, de modo que la molécula no es muy fluorescente en la solución por sí misma. Sin embargo, cuando dicho nucleótido fluorescente se incorpora en el ácido nucleico, el fluoróforo queda desconectado del nucleótido y no se extingue ya la fluorescencia. Para el caso de un enlace al fosfato beta o gamma del nucleótido, esto se produce de forma natural mediante la actividad catalítica de la polimerasa, en el caso de los fluoróforos unidos a la base, esto tendría que realizarse por escisión fotoquímica. La señal de fluorescencia procedente del fluoróforo escindido es mucho más brillante y puede detectarse sobre el posible fondo de la diversidad de moléculas extinguidas en la proximidad del complejo polimerasa-ácido nucleico.

Además, ya que el tiempo de permanencia de la fluorescencia de las moléculas escindidas que se difunden en la solución es mucho más breve que el tiempo de permanencia de la molécula escindida, puede conseguirse una mejora adicional de la señal al fondo utilizando iluminación pulsada y detección de fotones controlada por el tiempo. Esto se ilustra en la figura 5, que muestra las curvas de debilitamiento de la fluorescencia resueltas en el tiempo para la cumarina sola y la cumarina-dGTP, respectivamente. Debido a que la fluorescencia de la cumarina se extingue en el enlace covalente a dGTP, el tiempo de permanencia es mucho más breve que para el colorante libre solo, lo que significa que de promedio, se emiten fotones fluorescentes mucho antes después de una pulsación de excitación, por ejemplo, suministradas por un láser pulsado. Eliminando este intervalo de tiempo inmediatamente después de la pulsación de detección, que puede conseguirse, por ejemplo, con un componente de la línea de demora variable (indicado por la barra de trama cruzada con el tiempo de demora ajustable de anchura T), el umbral de respuesta del detector puede controlarse de modo que solamente se detecta la fluorescencia emitida desde el componente de debilitamiento lento, en este caso el colorante libre (o, en términos del esquema de secuenciación, el fluoróforo escindido), y de este modo el fondo de las moléculas no incoporporadas se reduce aún más. Saavedra *et al.*, "Time-Resolved Fluorimetric Detection of Terbium-Labelled Deoxyribonucleic Acid Separated by Gel Electrophoresis". *Analyst* 114: 835-838 (1989).

Los nucleótidos pueden convertirse también en fluoróforos por reacciones fotoquímicas que conllevan la formación de radicales. Esta técnica se ha utilizado con la serotonina y otras moléculas biológicamente relevantes. Véase Shear *et al.*, "Multiphoton-Excited Visible Emission by Serotonin Solutions", *Photochem. Photobiol*. 65: 931-936 (1997).

La situación fotofísica ideal sería que cada nucleótido haya generado su propia señal de fluorescencia. Desgraciadamente, el ácido nucleico y los nucleótidos individuales son fluoróforos débiles que emiten débilmente con eficacias cuánticas minúsculas y solamente en la iluminación con luz ultravioleta oscura. Sin embargo, el fluoróforo serotonina (5HT) ultravioleta natural puede fotoionizarse por absorción simultánea de 4 fotones infrarrojos, para formar un radical que reacciona con otras moléculas en estado fundamental para formar un complejo que emite fluorescencia verde brillante en la absorción de 2 fotones más. Los descubrimientos posteriores demuestran que muchas moléculas orgánicas pequeñas pueden experimentar esta conversión en multifotones.

La conocida extinción de los fluoróforos por los componentes del ácido nucleico y por los fluoróforos adyacentes así como la transferencia de la energía de resonancia puede proporcionar marcadores tolerados por la polimerasa. Furey *et al.*, "Use of Fluorescence Resonance Energy Transfer to Investigate the Conformation of DNA Substrates Bound to the Klenow Fragment", *Biochemistry* 37: 2979-2990 (1998) y Glazer *et al.*, "Energy-Transfer Fluorescent Reagents for DNA Analvses". *Curr. Op. Biotechn*. 8: 94-102 (1997).

En el montaje más eficaz de la presente invención, cada base debería distinguirse por su propio marcador de modo que la secuencia puede deducirse del resultado combinado de cuatro canales diferentes, tal como se ilustra en la figura 3C. Esto puede realizarse, por ejemplo, utilizando diferentes fluoróforos como marcadores y cuatro diferentes canales de detección, separados por fibras ópticas. Es posible distinguir los marcadores por parámetros distintos de la banda de longitud de onda de emisión tal como el tiempo de permanencia de la fluorescencia o cualquier combinación de varios parámetros para las diferentes bases. Debido a las posibles interacciones de un fluoróforo con una base, es factible emplear el mismo fluoróforo para distinguir más de una base. Como ejemplo, la cumarina-dGTP tiene un tiempo de permanencia de la fluorescencia mucho más breve que la cumarina-dCTP de modo que las dos bases podrían distinguirse por su diferencia en el tiempo de permanencia de la fluorescencia en la etapa de identificación del esquema de secuenciación, aunque llevan la misma sustancia química que el marcador fluorescente.

El procedimiento de secuenciación puede realizarse también utilizando menos de los 4 marcadores utilizados. Con 3 marcadores, la secuencia puede deducirse a partir dla secuenciación de una cadena de ácido nucleico (1) si la 4ª base puede detectarse como un retraso de tiempo en oscuridad constante entre las señales de los demás marcadores o (2) de manera inequívoca secuenciando ambas cadenas de ácido nucleico, porque en este caso se obtiene una señal de fluorescencia positiva de cada *par de bases*. Otro posible esquema que utiliza dos marcadores consiste en presentar una base marcada con un fluoróforo y las otras tres bases con otro fluoróforo. En este caso, las otras tres bases no proporcionan una secuencia sino únicamente un numero de bases que se producen entre la base específica que es identificada por el otro fluoróforo. Ciclando este fluoróforo de identificación mediante las diferentes bases en diferentes reacciones de secuenciación, la secuencia completa puede deducirse a partir de sucesivas series de secuenciación. Ampliando este esquema de utilización de dos marcadores solamente, es posible incluso obtener la secuencia completa empleando solamente dos bases marcadas por serie de secuenciación. Como apuntaron Sauer *et al.*, "Detection and Identification of Single Dye Labelled Mononucleotide Molecules Released From an Optical Fiber in Microcapillary: First Steps Towards a New Single Molecule DNA Sequencing Technique". *Phys. Chem. Chem. Phys*. 1:2471-77 (1999), la secuencia puede determinarse con 2 marcadores solos si uno realiza múltiples reacciones de secuenciación con las posibles combinaciones de los dos marcadores. Por consiguiente, en la realización del proceso de la presente invención, es deseable marcar segmentos largos de ácido nucleico por lo menos con 2 marcadores diferentes.

Cuando la secuenciación se realiza acoplando al soporte la polimerasa en lugar del ácido nucleico, es importante que la enzima sintetice segmentos largos de ácido nucleico, sin el complejo ácido nucleico/proteína que se deshace. Esto se denomina síntesis *procesiva* de ácido nucleico. Por lo menos para el sistema que utiliza T7 ADN polimerasa y dCTP reemplazadas completamente por cumarina-5-dCTP, la síntesis es completamente procesiva por lo menos en 7.300 pares de bases (una molécula de polimerasa se une a la plantilla ssADN y construye la segunda cadena completa sin desprenderse ni una vez). Con un marcador, el procedimiento de la presente invención puede realizarse observando la polimerasa en tiempo real con la resolución de pares de bases e identificando las características de la secuencia de la base, pero sin conocer las demás bases. Por consiguiente, lo más deseable sería utilizar cuatro marcadores diferentes para una velocidad y precisión mayores como se apuntó anteriormente. Sin embargo, la información de la medición de la incorporación de nucleótidos a un solo nivel molecular, tal como las velocidades de incorporación para bases individuales en un contexto de secuencia dado, puede proporcionar medios para caracterizar más la secuencia que se está sintetizando. En lo que respecta a asegurar la síntesis del proceso para el segundo modo de operación de la presente invención, pueden utilizarse proteínas accesorias para construir el complejo ácido nucleico/proteína aún más procesivo que utilizando la enzima sola que polimeriza el ácido nucleico. Por ejemplo, en condiciones óptimas, la T7 ADN polimerasa es procesiva en por lo menos 10.000 bases, mientras que en el complejo con la proteína T7 helicasa/primasa, la procesividad aumenta hasta más de 100.000 bases. Kelman *et al*., "Processivity of DNA Polymerases: Two Mechanisms, One Goal", *Structure* 6: 121-125 (1998). Una proteína de enlace monocatenaria es también una proteína accesoria adecuada. La procesividad es especialmente importante a concentraciones de análogos nucleotídicos que están por debajo del límite de saturación de una polimerasa específica, porque es sabido que los valores de procesividad de las polimerasas disminuyen al limitar las concentraciones de sustrato. Véase Patel *et al*., "Pre-Steady-State Kinetic Analysis of Processive DNA Replication Including Complete Characterization of an Exonuclease-Deficient Mutant", *Biochemistry* 30: 511-525 (1991). Otra posibilidad de asegurar la procesividad es el desarrollo o descubrimiento de una polimerasa que es totalmente procesiva en ausencia o en concentraciones de sustrato muy bajas (como es el caso, por ejemplo, para la prolongación del complejo ARN polimerasa/ADN). En el caso de que la procesividad no sea suficientemente elevada, es posible unir tanto la polimerasa como la molécula de ácido nucleico diana sobre el soporte próximo uno con otro. Esto facilitaría la reforma del complejo y la continuación de la síntesis del ADN, en caso de que el complejo de secuenciación se deshaga ocasionalmente. Pueden utilizarse también polimerasas no procesivas según la presente invención en el caso en el que el ácido nucleico diana esté unido al soporte. En la presente memoria, la misma o diferente molécula de polimerasa puede reformar el complejo y continuar la síntesis después de la disociación del complejo.

Una estrategia para llevar a cabo la presente invención se muestra en la figura 6. La figura 6A presenta un sistema de secuenciación con la solución R de reactivo colocada en la superficie 2 a la que se inmoviliza un complejo con la molécula de ácido nucleico diana cebado. Confinando la iluminación a una pequeña área próxima a la zona activa de la prolongación de la polimerasa, por ejemplo, enfocando la radiación activadora con la ayuda de lentes o de fibra óptica 6, se detectan los análogos nucleotídicos que se incorporarán en la cadena de ácido nucleico en desarrollo, debido a que están situados en la zona de iluminación. La figura 6B muestra una sección ampliada del dispositivo, con el complejo de polimerización en la zona de iluminación. Las concentraciones de sustrato se seleccionan de modo que el número de análogos nucleotídicos en el área circundante en la solución R están generalmente fuera de la zona iluminada y no se detectan.

Como se representan en la figura 6A, la fuente de iluminación 10 (por ejemplo, un láser) dirige la radiación de excitación a modo de un separador de rayo dicroico 8 a través de lentes 6 y la superficie 2 al complejo de ácido nucleico diana cebado e inmovilizado. Esto excita el marcador inmovilizado al complejo con una radiación emitida resultante que vuelve a pasar a través de la superficie 2 y de las lentes o de la fibra óptica 6. El separador de rayos dicroicos 8 permite el paso de la radiación emitida al detector (o la matriz de varios detectores) 12 que identifica el tipo de emisión. La información de la emisión detectada es dirigida a continuación al ordenador 14 donde se identifica la base nucleotídica correspondiente a la emisión y su identidad almacenada. Después de muchos ciclos de este procedimiento, el ordenador podrá generar como resultado la secuencia de la molécula de ácido nucleico diana. El correspondiente resultado de la detección corresponde de nuevo al esquema mostrado en la figura 3, como se explicó anteriormente.

Según otra forma de realización de la presente invención, la iluminación y la detección de la fluorescencia puede conseguirse construyendo el soporte para el ácido nucleico unido al final de una primera fibra óptica en modo individual que lleva la luz de excitación. Ésta y/o una segunda fibra óptica pueden utilizarse para recoger los fotones fluorescentes. Transmitiendo la radiación de longitud de onda de excitación apropiada a través de la primera fibra óptica en modo individual, el marcador fluorecerá y emitirá la frecuencia de la luz fluorescente apropiada. La luz fluorescente emitida se transmitirá en parte a la segunda fibra óptica y espectralmente separada tal como por rejillas de difracción grabadas en la fibra. El espectro de la luz retornada identifica el análogo nucleotídico específico unido. Son posibles otras técnicas para proporcionar o recoger la luz en el punto de reacción, tales como la utilización de la iluminación guiada por ondas evanescentes o por ondas, tal como la iluminación por reflexión interna total. Puede emplearse una o varias fuentes de iluminación, que suministran excitación de uno o múltiples fotones. Los detectores adecuados incluyen módulos de fotodiodo en avalancha, tubos fotomultiplicadores, cámaras CCD, chips CMOS o matrices o combinaciones de varios detectores.

Debido a que es probable que exista un límite superior a la concentración de los análogos polinucleotídicos presentes en el volumen de observación que se correlaciona con una señal permisible a la relación de fondo y capacidad para distinguir el análogo nucleotídico *específico* que está siendo incorporado en el ácido nucleico de los análogos nucleotídicos que solo se difunden alrededor de la polimerasa, es posible que el procedimiento de secuenciación de la presente invención debe realizarse a concentraciones por debajo del límite de saturación para uno o más análogos nucleotídicos.

Por ejemplo, si se utilizan ópticas convencionales limitadas a la difracción para la detección de la fluorescencia, el volumen de observación es grande de modo que tendrían que utilizarse concentraciones del sustrato en el intervalo de nanomolar para una señal de fondo aceptable. Éste es muy inferior a la k_m habitual de las polimerasas (normalmente en el intervalo de μM), a menos que se utilicen otros medios para reducir el fondo, tal como la discriminación del tiempo de permanencia como se expuso anteriormente (figura 5) o técnicas de confinamiento de volumen, como se describe a continuación, para reducir "electrónicamente" o físicamente la contribución de la fluorescencia de fondo. En un rayo láser enfocado convencionalmente, el volumen focal es aproximadamente 0,2 μm^3 (0,5 μm de diámetro, 1,5 μm en la dirección axial), correspondiente a aproximadamente 0,2 fl. Con objeto de que esté presente solamente un nucleótido fluorescente por término medio en el volumen de excitación en un momento dado, la concentración de sustrato debe reducirse a aprox. 10 nM, concentración muy inferior a los valores k_m de las ADN polimerasas (aprox. 1 a 2 μM). Veáse Polesky *et al.*, "Identification of Residues Critical for the Polymerase-Activity of the Klenow Fragment of DNA-Polymerase-I fron *Escherichia-coli*". *J. Biol. Chem.* 265: 14579-14591 (1990) y McClwe *et al.*, "The Steady State Kinetic parameters and Non-Processivity of *Escherichia coli* Deoxyribonucleic Acid Polymerase I", *J. Biol. Chem.* 250:4073-4080 (1975). De este modo, si la concentración de sustratos es muy inferior a k_m, la procesividad de la síntesis de ácido nucleico ha de asegurarse mediante una de las posibilidades mencionadas anteriormente. Alternativamente, si el volumen de observación puede reducirse, es permisible una concentración en sustrato superior, lo que aumenta naturalmente los valores de la procesividad. Por consiguiente un objetivo de la presente invención se refiere a una reducción eficaz del volumen de observación para reducir o prevenir la fluorescencia del fondo causada por nucleótidos libres marcados y aumentar la procesividad. Esto puede lograrse de numerosas maneras.

Una estrategia para reducir de ruido de fondo conlleva la intensificación del campo electromagnético cerca de los objetos con pequeños radios de curvatura.

Debido al denominado "efecto antena", la radiación electromagnética está muy intensificada en el extremo de un objeto agudo, tal como una punta metálica. Utilizando este procedimiento, el volumen que se aumenta aproximadamente corresponde a una esfera con un diámetro próximo al diámetro de la punta. Esta técnica está expuesta en Sánchez, E. J., *et al.*, "Near-Field Fluorescence Microscopy Base on Two-Photon Excitation with Metal Tips", *Phys. Rev. Lett.* 82: 4014-17 (1999).

En la realización del procedimiento de la presente invención, una enzima que polimeriza el ácido nucleico se coloca al final de una punta metálica dirigiéndose la luz de láser sobre ella, por ejemplo, con una lentes de objetivo convencional. Debido a que el volumen iluminado eficaz puede ser actualmente del orden del tamaño de la propia polimerasa prácticamente no se detectará ninguna fluorescencia a partir de los nucleótidos fluorescentes que se difunden en la solución. Además, el tiempo de residencia de las moléculas de difusión a través de dicho pequeño volumen es sumamente breve. Sin embargo, la incorporación de un nucleótido fluorescente se verá como una ráfaga de fluorescencia relativamente larga, debido a que esta molécula específica permanecerá en este pequeño volumen iluminado (hasta que sea eliminada como se expuso anteriormente).

Una estrategia para llevar a cabo esta forma de realización de la presente invención se muestra en las figuras 7A a B. La figura 7A presenta un sistema para secuenciación con intensificación del campo electromagnético con la solución de reactivo R colocada en la superficie 2 en la que se inmoviliza un complejo de molécula de ácido nucleico diana. Como se representa en la figura 7B, una punta metálica que lleva una polimerasa se coloca en la solución de reactivo R, creando una pequeña zona de iluminación alrededor de la polimerasa inmovilizada durante la iluminación por la lente 6. Confinando la iluminación en esta pequeña área, próxima a la zona activa de la prolongación de la polimerasa, se detectan análogos nucleotídicos que se incorporarán a la cadena de ácido nucleico en desarrollo, debido a que están colocados en la zona de iluminación. Por otra parte, los análogos nucleotídicos en el área circundante en la solución R están generalmente fuera de esta zona y no se detectan.

Como se representa en la figura 7A, alguna iluminación 10 (por ejemplo, un láser) dirige una o radiación de excitación de multifotones con un componente de polimerización distinto de cero paralelo a la punta a modo de un separador de rayos dicroicos 8 mediante la lente 6 y la superficie 2 al complejo del ácido nucleico diana cebado e inmovilizado. Esto excita el marcador inmovilizado al complejo con la radiación emitida resultante que vuelve a pasar a través de la superficie 2 y la lente 6. El separador 8 de rayos dicroicos permite el paso de la radiación emitida al detector 12 que identifica el tipo de emisión. La información de la emisión detectada se dirige a continuación al ordenador 14 donde la base nucleotídica correspondiente a la emisión se identifica y se almacena su identidad. Después de múltiples ciclos de este procedimiento, el ordenador podrá generar como resultado la secuencia de una molécula de ácido nucleico diana. El correspondiente resultado de detección corresponde de nuevo al esquema representado en la figura 3, como se explicó anteriormente. La principal diferencia con el caso expuesto anteriormente es que los picos cortos producidos por análogos nucleotídicos que se difunden al azar a través del volumen focal son ahora sumamente cortos, al ser tan pequeño el volumen de observación. Por consiguiente, esta estrategia de reducción del volumen de observación también produce aumento de la resolución en el tiempo con respecto a los nucleótidos incorporados frente a los no incorporados. Esto es cierto para todas las demás posibilidades de confinamiento de volumen expuestas con mayor detalle a continuación.

En la realización de este procedimiento, las puntas pueden estar formarse por varios materiales, por ejemplo, metales tales como platino, plata u oro. La fabricación de la punta puede llevarse a cabo, por ejemplo, por grabado electroquímico de alambres o por molienda con rayo iónico. Véase Sánchez, E. J., *et al.*, "Near-Field Fluorescence Microscopy Base on Two-Photon Excitation with Metal Tips", *Phys. Rev. Lett*. 82: 4014-17 (1999).

La enzima que polimeriza el ácido nucleico puede acoplarse al extremo de la punta sumergiendo la punta en una solución de moléculas de la enzima que polimeriza el ácido nucleico, aplicando un campo eléctrico a la punta con cargas que atraen la enzima que polimeriza el ácido nucleico u otras técnicas de acoplamiento (por ejemplo, con enlazadores, anticuerpos, etc.). Un modo alternativo de utilizar la intensificación del campo electromagnético para este esquema de secuenciación consiste en colocar una punta desnuda en íntima proximidad a un complejo enzimático que polimeriza el ácido nucleico inmovilizado/ácido nucleico, en lugar de tener el complejo físicamente acoplado al final de la punta. Una población de complejos podría inmovilizarse, por ejemplo en un portaobjetos de vidrio, y la punta se explora sobre la superficie hasta que se encuentra un complejo útil para la secuenciación. Se han desarrollado técnicas adecuadas para llevar a cabo este nanoposicionamiento en el campo de la microscopia de sonda de exploración.

Otra estrategia para reducir el ruido de fondo mientras se realiza el procedimiento de secuenciación de la presente invención implica la utilización de iluminación en campo próximo, como se representa en las figuras 8A-B. Aquí, como se representa en la figura 8B, el complejo de ácido nucleico diana cebado se inmoviliza en la superficie 2 aplicándose la capa opaca 16 sobre la superficie 2. Sin embargo, pequeños orificios 18 están grabados en la capa opaca 16. Cuando se iluminan desde abajo, la luz no puede penetrar completamente a través de los orificios en la solución de reactivo R porque el diámetro de los orificios 18 es más pequeño que la mitad de la longitud de onda de la luz. Sin embargo, existe alguna fuga que crea una pequeña área de luz recta por encima de la superficie 2 en el orificio 18, creando un denominado volumen de excitación en campo próximo. Como se representa en figura 8B, el complejo de ácido nucleico diana cebado está colocado en el orificio 18 donde se ilumina desde abajo. Confinando la iluminación en esta pequeña área en campo próximo, se detectan los análogos nucleotídicos incorporados, colocados en la zona de iluminación. Por otra parte, la cantidad de análogos nucleotídicos que no sirven para ampliar el cebador son pocos debido al pequeño tamaño del orificio 18 y, a la pequeña extensión detectada, se distinguen fácilmente de los análogos nucleotídicos incorporados, tal como se describió anteriormente.

El sistema para llevar a cabo esta forma de realización se presenta en la figura 8A. La fuente de iluminación 10 (por ejemplo, un láser) dirige la radiación de excitación a modo de separador de rayo dicroico 8 a través de la lente 6 y de la superficie 2 al complejo de ácido nucleico diana cebado inmovilizado. Esto excita al marcador inmovilizado en el complejo volviendo a pasar la radiación emitida resultante a través de la superficie 2 y de la lente 6. El separador de rayos dicroicos 8 permite el paso de la radiación emitida al detector 12 que identifica el tipo de emisión. La información de la emisión detectada se dirige a continuación al ordenador 14 donde se identifica la base nucleotídica correspondiente a la emisión y se almacena su identidad. Tras múltiples ciclos de este procedimiento, el ordenador podrá generar como resultado la secuencia de la molécula ácido nucleico diana.

Como alternativa adecuada, utilizando volúmenes de excitación en campo próximo, el volumen en campo próximo puede también generarse mediante la utilización de una o muchas fibras ópticas cónicas utilizadas frecuentemente en microscopia en campo próximo de exploración.

La nanofabricación es otra técnica útil en la limitación del volumen de reacción para reducir el nivel de fluorescencia de fondo. Esto implica el confinamiento del volumen de excitación a una zona con un nanocanal. Aquí, el confinamiento es posible en dos o tres dimensiones espaciales. Un hervidor de reacción con un volumen mucho más pequeño que los volúmenes focales obtenibles con ópticas de enfoque en campo lejano se fabrica en una oblea de silicio o de sílice fundida a partir de materiales ópticamente transparentes. Turner *et al.*, "Solid-State Artificial Gel DNA Electrophoresis whith an Integrated Top Layer", *Proceedings of SPIE. Micro- and Nano-Fabricated Structures and Devices for Biomedical Environmental Applications* 3258: 114-121 (1998).

La técnica aprovecha una capa de polisilicio de sacrificio para definir una calidad operativa de los canales. Stern *et al.*, "Nanochannel Fabrication for Chemical Sensors", *J. Vac. Sci. Technol.* B15: 2887-2891 (1997) y Chu *et al.*, "Silicon Nanofilter with Absolute Pore Size and High Mechanical Strength", *Proc. SFIE Int. Soc.Opt. Eng.(USA)* 2593: 9-20 (1995).

El suelo, el techo y las paredes de los canales son de nitruro de silicio, que está depositado de forma ajustada sobre una capa de sacrificio diseñada de polisilicio. La capa de sacrificio se elimina a continuación con un grabado químico húmedo de alta selectividad, dejando detrás solamente el nitruro de silicio. Esta técnica ha demostrado un control preciso de la dimensión crítica (DC) sobre un amplio intervalo de tamaños de estructuras. La altura de la capa de polisilicio puede controlarse dentro de 5 nm sobre un dispositivo completo, y las dimensiones laterales están limitadas en tamaño y control DC solamente por la técnica de litografía aplicada. La nanoestructura puede presentar una configuración punteada, acicular o resonante, para aumentar la detección del marcador.

Las figuras 9A-B muestran un sistema nanofabricado.

En la figura 9B se muestra una vista ampliada de la sección transversal del nanocanal, con reactivos R situados solamente en el área confinada 102, que está creada por las paredes de los canales 104 y 106. El complejo de la molécula de ácido nucleico diana cebado está colocado dentro del área confinada 102. Como resultado, cuando la luz de excitación atraviesa el área confinada 102, el marcador del análogo nucleotídico incorporado se excita y emite radiación que es detectada e identificada como correspondiente a una base nucleotídica particular añadida a la secuencia del cebador de extensión. Pasando los reactivos a través del área confinada 102, la cantidad de análogos nucleotídicos que no amplían el cebador son pocos en cualquier punto específico en tiempo. En pequeña medida se detectan dichas entidades móviles, se distinguen fácilmente de los restos inmovilizados tal como se describió anteriormente.

La figura 9A presenta un sistema para llevar a cabo el sistema de nanocanal.

La fuente de iluminación 10 (por ejemplo, un láser) dirige la radiación de excitación a modo de separador de rayo dicroico 8 a través de la lente 6 y del nanocanal 106 al complejo de ácido nucleico diana cebado e inmovilizado. Esto excita al marcador inmovilizado en el complejo con la radiación emitida resultante que vuelve a pasar a través de la lente 6. El separador de rayos dicroicos 8 permite el paso de la radiación emitida al detector 12 que identifica el tipo de emisión. La información de la emisión detectada se dirige a continuación al ordenador 14 donde la base nucleotídica correspondiente a la emisión se identifica y se almacena su identidad. Después de múltiples ciclos de este procedimiento, el ordenador podrá generar como resultado la secuencia de la molécula de ácido nucleico diana.

Las figuras 10A-B muestran sistemas para proporcionar reactivos al sistema de confinamiento nanofabricado. En la figura 10A, el reactivo que incluye dATP, dCTP, dGTP, dUTP, la fuente de ácido nucleico y el tampón se mantienen en depósitos separados y se conectan a través de conductos separados al colector 200 donde los reactivos se mezclan antes de enfocar el nanocanal 202. Los componentes de este sistema corriente arriba y corriente abajo del nanocanal 202 pueden combinarse, como una microestructura. En el procedimiento de pasar rápidamente a través del nanocanal 202, los reactivos se desplazan rápidamente a través de la zona de reacción 204 donde el procedimiento de secuenciación de la presente invención se lleva a cabo desde el nanocanal 202, los reactivos R restantes pasan a través de la salida 206. El sistema de la figura 10B, es en general similar al de la figura 10A, pero el sistema anterior está en un solo chip con almohadillas para conectar el sistema a depósitos de fluido. En particular, el depósito para cada uno de los reactivos está acoplado al chip 208 con almohadillas a la entrada 210 a-f, mientras que la salida para los reactivos descargados está conectada a la almohadilla 212.

Los canales nanofabricados de 75 nm de anchura y 60 nm de altura se han fabricado con excelente transparencia óptica y se han utilizado, para el control de flujo de ADN. Véase Turner *et al.*, "Solid-State Artificial Gel DNA Electrophoresis with an Integrated Top Layer", Proceedings of SPIE. Micro-and Nano-Fabricated Structures and Devices for *Biomedical Environmental Applications* 3258: 114-121 (1998), que está incorporado por la presente como referencia. Colocando el complejo de síntesis de ácido nucleico en un canal de anchura z = 25 nm, minimizando la dimensión x del rayo de láser enfocado a aprox. 300 nm y fijando la dimensión y por la anchura del canal en 100 nm, el volumen eficaz de conservación puede reducirse a 7,5 x 10^4 μm^3, correspondiente a 0,75 altolitros. Aquí, la concentración para que solamente una molécula de sustrato esté presente en el volumen de excitación asciende a 2 μM, una concentración de sustrato dentro del intervalo de la polimerización del ácido nucleico rápida y eficaz. Además, dado que existen cuatro análogos nucleotídicos diferentes, cada uno a distinguir, la concentración del sustrato eficaz para la polimerasa es cuatro veces mayor. Si se necesita un pequeño volumen eficaz de observación, la dimensión y en la dirección del flujo puede reducirse a aproximadamente 100 nm por iluminación dentro del modelo de interferencia de dos objetivos a aproximadamente ángulos axiales de 90º como en la microscopia theta. Véase Stelzer *et al.*, " A New Tool for the Observation of Embryos Other Large Specimens: Confocal Theta Fluorescence Microscopy", *J. Microscopy* 179: 1-10 (1995).

Para excitar los marcadores, la energía de activación se enfoca en la proximidad a la zona activa de la ampliación de polimerasa (es decir, donde está localizada la polimerasa). En la medida en que esta zona activa se desplaza durante la ampliación (por ejemplo, como resultado por el movimiento por la polimerasa), el foco de la energía de activación se desplaza también.

Una consideración necesaria, es la elección entre la excitación monofotónica y multifotónica de fluorescencia. La excitación multifotónica proporciona algunas ventajas poderosas, pero es más compleja y más costosa de realizar. La fluorescencia de excitación multifotónica, que utiliza la absorción simultánea de dos o más fotones, las pulsaciones infrarrojas del brillo en femtosegundo generadas por láseres bloqueados en modo de estado sólido ultrarrápido proporciona la estrategia más prometedora. Véase Denk *et al.*, "2-Photon Laser Scanning Fluorescence Microscopy", *Science* 248: 73-76 (1990).

La sensibilidad a la fluorescencia de una sola molécula, se obtiene de manera rutinaria y se puede resolver temporalmente a nivel de microsegundos con tiempos de permanencia de la fluorescencia mensurables con precisión razonable para moléculas individuales. Véase Mertz *et al.*, "Single-Molecule Detection by Two-Photon Excited Fluorescence", *Optics. Lett.* 20: 2532-2534 (1995) y Eggeling *et al.*, "Monitoring Conformational Dynamics of a Single Molecule by Selective Fluorescence Spectroscopy". *Proc. Natl. Acad. Sci.* USA 95: 1556-1561 (1998).

La señal fluorescente ideal para la secuenciación de una sola molécula consiste en ráfagas resueltas en el tiempo de fluorescencia distinguible ya que cada nucleótido está unido. Por lo tanto, en la situación ideal, una serie resuelta en el tiempo de ráfagas fluorescentes resueltas por color podrían obtenerse si los nucleótidos se unieran a intervalos distinguibles como los descritos en la figura 3. La resolución completa de la secuencia en el tiempo de episodios ofrece por consiguiente la mejor reducción del fondo y la posibilidad fehaciente para el reconocimiento de nucleótidos. Como con las polimerasas disponibles actualmente, los nucleótidos marcados se añaden más probablemente no más rápido que a intervalos de 1 milisegundo, sería posible que todos los fotones de fluorescencia detectados de cada nucleótido marcado puedan acumularse y eliminarse antes de que se una el siguiente nucleótido fluorescente. Esta secuencia ideal de ráfaga-hueco-ráfaga está realizada aunque realmente cada etapa de la cinética molecular de polimerización conlleva el procedimiento estocástico de Poisson. Para un solo procedimiento de Poisson, la demora en el tiempo más probable entre los episodios es cero aunque la demora media sería mayor de cero. Sin embargo, el procedimiento de incorporación de un solo dNTP en el ADN por la ADN polimerasa es un procedimiento multietapa sucesivo de por lo menos 5 episodios diferentes. Véase Patel *et al.*, "Pre-Steady-State Kinetic Analysis of Processive DNA Replication Including Complete Characterization of an Exonuclease-Deficient Mutant", *Biochemistry* 30: 511-525 (1991). La suma de secuencias de estas etapas producirá una demora en el tiempo más probablemente mayor de cero. Por consiguiente las ráfagas de fotones no es muy probable que se solapen.

Para los fluoróforos convencionales, aproximadamente 10^5 fotones por fluoróforo se emitirán antes del fotoblanqueo. La detección de (por lo menos) el 1% de la emisión da una incertidumbre relativa de ruido del 3%. El fondo debido a los nucleótidos libres se reduce a un nivel casi insignificante por los esquemas expuestos anteriormente, por ejemplo, limitando el tamaño del volumen focal para que contenga sólo aproximadamente un nucleótido marcado libre, con tiempos de permanencia muy breves.

El nivel de detección esperado es de aproximadamente 10^3 fotones por cada nucleótido marcado, en aproximadamente 10^{-3} s. Esta es una cantidad de recuento aceptable, $\sim 10^6$ Hz, y una cantidad de excitación del fluoróforo a aproximadamente una décima de saturación del estado excitado del singlete. Esta excitación de fluorescencia crea una ráfaga detectada de $\sim 10^3$ fotones en aproximadamente 1 ms a la longitud de onda característica de cada nucleótido marcado, dejando de promedio, una separación de aproximadamente 1 ms antes de añadir el siguiente nucleótido, muy dentro de los intervalos de tiempo medios entre la adición del nucleótido a probablemente más de un milisegundo. Los posibles solapamientos de las ráfagas pueden analizarse y resolverse mediante el tratamiento analítico de mediciones continuas de datos en secuencias coherentes con el tiempo en (en el mejor de los casos) 4 canales de la mayoría de los resultados precisos de secuenciación. Con las estadísticas de los fotones disponibles en el diseño experimental y los analizadores multicanal acoplados recientemente desarrollados y el programa informático operativo, las tasas de error pueden hacerse aceptables con 4 nucleótidos marcados o con las estrategias que implican un número más pequeño de marcadores como se expuso anteriormente.

La resolución espectral de cuatro fluoróforos que identifican los nucleótidos puede conseguirse con la excitación de dos fotones por pulsaciones infrarrojas. Los 4 fluoróforos pueden excitarse simultáneamente debido a las bandas de excitación amplias características normalmente de la excitación de dos fotones. Véase Xu *et al.*, "Multiphoton Excitation Cross-Sections of Molecular Fluorophores", *Bioimaging* 4: 198-207 (1996).

Alternativamente, pueden utilizarse múltiples fuentes de excitación en combinación o por conexión rápida para iluminar el complejo de secuenciación si es necesario. La separación espectral se realiza con filtros de interferencia convencionales pero los espectros de emisión pueden solaparse, complicando el análisis de correlación temporal y quizás acompañando la correlación cruzada de los 4 canales de color para la corrección. Si la compatibilidad de los fluoróforos con la enzima que polimeriza el ácido nucleico limita la aplicación de las series de colorantes adecuadas, una combinación de técnicas puede aplicarse para distinguir los marcadores.

Otra manera potencial de distinguir la incorporación de un nucleótido en la cadena de ácido nucleico en desarrollo, consiste en medir los cambios del tiempo de permanencia de la fluorescencia. Se ha observado que el tiempo de

permanencia de la fluorescencia de una sonda de pireno con oligonucleótido varía en función de la secuencia durante el acoplamiento del ADN. Véase Dapprich J., "Fluoreszenzdetection Molekularer Evolution (Fluorescence Detection of Molecular Evolution)". *Dissertation*, Georg-August-Univ., Goettingen, Alemania (1994).

Las interacciones fotofísicas entre el fluoróforo y la base dan como resultado tiempos de debilitamiento de la fluorescencia característicos y pueden utilizarse también para diferenciar las bases como se expuso anteriormente. La determinación del tiempo de permanencia y la discriminación en el contexto molecular individual se ha demostrado recientemente de modo que la discriminación entre las bases que se incorporan y los nucleótidos que se difunden libremente podría realizarse por mediciones del tiempo de permanencia de la fluorescencia. Véase Eggeling *et al.*, "Monitoring Conformational Dynamics of a Single Molecule by Selective Fluorescence Spectroscopy". *Proc. Natl. Acad. Sci.* USA 95: 1556-1561 (1998).

Las mediciones correlacionadas en el tiempo en cuatro canales de longitud de onda de fluorescencia pueden utilizarse con eficacia en la realización del proceso de la presente invención. El solapamiento de los espectros de emisión puede dejar que señales procedentes de un fluoróforo entren en varios canales pero la tasa de recuento relativa y la medida del tiempo de permanencia identifica al marcador correctamente. Las señales simultáneas procedentes de un nucleótido marcado incorporado y un marcador libre se pueden distinguir por la duración y la magnitud de la ráfaga, que están limitadas por el marcador libre. La ambigüedad del marcador puede reducirse más mediante la utilización de mediciones de tiempo de debilitamiento de la fluorescencia que pueden realizarse con la resolución de 0,1 ns disponibles de las demoras en el tiempo por emisión de fotones de fluorescencia después de cada pulsación de excitación del láser de femtosegundo. La emisión de fotones de fluorescencia y los propios procesos de fotoblanqueo son también procesos estocásticos pero conllevan eficiencias cuánticas suficientemente dispares que las tasas de error serían insignificantes.

En el rechazo del fondo procedente de los nucleótidos marcados que se difunden o que circulan libremente, se utiliza con ventaja el tiempo de permanencia muy breve de cualquier nucleótido libre individual en el volumen focal. El tiempo de difusión característico para un análogo nucleotídico libre a través de la dimensión abierta del volumen focal (en el peor de los casos de iluminación en campo lejano no interferométrica) será $\tau_D \sim y^2/4D\text{-}2x10^{-5}$ s., siendo y la dimensión del volumen focal y D el coeficiente difusión. Una velocidad de flujo iontoforética de 1 cm/s es suficiente para mantener sus breves ráfagas de fluorescencia en menos de 10^{-3} s y reducir el número de fotones en un orden de magnitud. Esto asegurará la discriminación frente a los nucleótidos libres e identificará la serie en el tiempo de las ráfagas que representan la secuencia de ácidos nucleicos, con la condición de que las concentraciones de análogo nucleotídico sean propiamente bajas como se expuso. Magde *et al.*, "Thermodynamic Fluctuations in a Reacting System-Measurement by Fluorescence Correlation Spectroscopy", *Phys. Rev. Lett.* 29: 705-708 (1972) y Maiti *et al.*, "Measuring Serotonin Distribution in Live Cells with Three-Photon Excitation", *Science* 275: 530-532 (1997).

La discriminación puede mejorarse utilizando técnicas de confinamiento en volumen o de detección controlado por el tiempo, como se expuso anteriormente.

La detección de la transferencia de energía por resonancia de fluorescencia (FRET) de un fluoróforo donador (por ejemplo, un donador acoplado a la polimerasa) para los aceptadores de análogos nucleotídicos adyacentes que están incorporados en la cadena de ácido nucleico en desarrollo sugiere una posibilidad más elegante de reducir el fondo de los nucleótidos incorporados. La FRET alcanza solamente distancias muy cortas que incluyen aproximadamente 20 nucleótidos y disminuye al recíproco de la sexta potencia de la distancia. La molécula de donador excitada transfiere su energía solamente a los fluoróforos aceptadores próximos, que emiten fluorescencia del aceptador espectralmente resuelta de cada nucleótido marcado a medida que se añade. Los nucleótidos ya incorporados más lejos del donador no contribuirían a la señal de fluorescencia ya que las limitaciones en la distancia y orientación de la transferencia de energía reducen el margen eficaz de observación a menos de 60 Å, eliminando de este modo con eficacia la fluorescencia de fondo de los nucleótidos no incorporados. Sin fotoblanqueo, el procedimiento requiere gran sensibilidad ya que la repetición del nucleótido abandona el margen de la FRET a la misma velocidad a la que se añaden nuevos nucleótidos, creando posiblemente ambigüedad en el reconocimiento de las secuencias. El fotoblanqueo o la escisión fotoquímica, o su combinación como se expuso anteriormente podrían resolver el problema. El fotoblanqueo de las moléculas del donador que utilizan la FRET puede evitarse si es el ácido nucleico plantilla el que se acopla y el donador que lleva la enzima que polimeriza el ácido nucleico se reemplaza periódicamente.

Una consideración final importante para el éxito de la presente invención se refiere a la estabilidad del complejo proteína/ácido nucleico en la radiación de activación, tal como los rayos de láser enfocados de manera ajustada. No es de esperar que la enzima esté afectada por la iluminación de excitación, porque se eligen longitudes de onda en las que las proteínas no se absorben, la estabilidad de la polimerasa en el rayo de láser sería lo suficientemente alta para permitir series de secuenciación precisas sobre longitudes largas de lectura. Las investigaciones previas exponiendo enzimas a la luz de láser potente han examinado el velado fotográfico y la pérdida de función. Los complejos de ARN polimerasa/ADN inmovilizados presentaban tiempos de inactivación de 82 ± 58 s. para luz de láser Nd:Y de 1047 nm de 82 a 99 mW de potencia enfocada a la proteína, correspondiente a intensidades de aproximadamente 10^8 W/cm^2. Otros estudios sobre los sistemas de actomiosina o cinesina indicaban estabilidad similar. Tanto el ADN como los enlaces de biotina-avidina han demostrado ser fotoestables en trampas ópticas. Véase Yin *et al.*, "Transcription Against an Applied Force", *Science* 270: 1653-1657 (1995), Svoboda *et al.*, "Direct Observation of Kinesin Stepping by Optical Trapping Interferometry". *Nature* 365: 721-727 (1993) y Molloy *et al.*, "Movement and Force Produced by a Single Myosin Head", *Nature* 378: 209-212 (1995).

Para la detección de la fluorescencia de los análogos nucleotídicos según la presente invención, son de esperar potencias (intensidades) de láser típicas de las mediciones de FCS, del orden de 0,1 mW (10^5 W/cm^2) para la excitación de un fotón y de 1 mW (10^6-10^7 W/cm^2) para la de dos fotones, siendo de este modo significativamente inferior que en el caso de las pinzas ópticas descritas anteriormente. La estabilidad enzimática sería por consiguiente mayor, además, con la velocidad rápida de secuenciación propuesta por este procedimiento (por ejemplo 10 pb/s) incluso 80 s. son suficientes para determinar la secuencia del ácido nucleico de 8 kb.

REIVINDICACIONES

1. Procedimiento de secuenciación de una sola molécula de ácido nucleico diana que presenta una pluralidad de bases nucleotídicas que comprende:

proporcionar un complejo de una enzima que polimeriza el ácido nucleico y la molécula de ácido nucleico diana para permitir la formación de una cadena de ácido nucleico en crecimiento complementario al ácido nucleico diana;

proporcionar una pluralidad de tipos de análogos nucleotídicos marcados, en el que cada tipo de análogo nucleotídico es complementario a un nucleótido diferente en la secuencia de ácido nucleico diana;

polimerizar la pluralidad de tipos de análogos nucleotídicos marcados incorporando la pluralidad de tipos de análogos nucleotídicos marcados en la cadena de ácido nucleico en crecimiento, en el que cada uno de los análogos incorporados puede polimerizarse con un análogo nucleotídico posterior para permitir la polimerización continua de una pluralidad de análogos nucleotídicos;

identificar la secuencia de tiempo de la incorporación de dicha pluralidad de tipos de análogos nucleotídicos marcados en la cadena de ácido nucleico en crecimiento detectando los tiempos de residencia de los análogos nucleotídicos marcados próximos a la enzima que polimeriza el ácido nucleico en tiempo real mientras que la extensión del cebador dirigida por la plantilla tiene lugar, y eliminándose el marcador tras la detección.

2. Procedimiento según la reivindicación 1, en el que se identifica una pluralidad de análogos nucleotídicos incorporados por segundo.

3. Procedimiento según la reivindicación 2, en el que se identifican hasta 100 análogos incorporados por segundo.

4. Procedimiento según la reivindicación 1, en el que la enzima que polimeriza el ácido nucleico se selecciona de entre el grupo constituido por una ADN polimerasa, una ARN polimerasa, transcriptasa inversa y mezclas de las mismas.

5. Procedimiento según la reivindicación 1, en el que la enzima que polimeriza el ácido nucleico es una polimerasa termoestable.

6. Procedimiento según la reivindicación 1, en el que la enzima que polimeriza el ácido nucleico es una polimerasa termodegradable seleccionada de entre el grupo constituido por ADN polimerasa de *E. coli*, el fragmento de Klenow de la ADN polimerasa de *E. coli*, T4 ADN polimerasa y T7 ADN polimerasa.

7. Procedimiento según la reivindicación 1, en el que la molécula de ácido nucleico diana se selecciona de entre el grupo constituido por ADN bicatenario, ADN monocatenario, horquillas de ADN monocatenario, híbridos de ADN/ARN, ARN con una zona de reconocimiento para la unión de una ARN polimerasa y horquillas de ARN.

8. Procedimiento según la reivindicación 1, en el que la enzima que polimeriza el ácido nucleico está unida al complejo con la molécula del ácido nucleico diana en un origen de replicación, en una hendidura o separación en un ácido nucleico diana bicatenario, una estructura secundaria en un ácido nucleico diana monocaternario, una zona de unión creada por un proteína accesoria seleccionada de entre una proteína de enlace monocatenaria, una primasa, y una helicasa o a un ácido nucleico monocatenario cebado.

9. Procedimiento según la reivindicación 1, en el que la enzima que polimeriza el ácido nucleico está provista de una o más proteínas accesorias seleccionadas de entre una proteína de enlace monocatenaria, una primasa, y una helicasa para modificar su actividad.

10. Procedimiento según la reivindicación 1, en el que la enzima que polimeriza el ácido nucleico es procesiva.

11. Procedimiento según la reivindicación 1, en el que la enzima que polimeriza el ácido nucleico es no procesiva.

12. Procedimiento según la reivindicación 1, en el que los análogos nucleotídicos se seleccionan de entre el grupo constituido por un ribonucleótido, un desoxirribonucleótido, un ribonucleótido modificado, un desoxirribonucleótido modificado, un nucleótido peptídico, un nucleótido peptídico modificado y un nucleótido de esqueleto central de fosfato-azúcar modificado.

13. Procedimiento según la reivindicación 1 que comprende además:

hibridar un cebador oligonucleotídico a la molécula de ácido nucleico diana antes de o durante dicha provisión de una pluralidad de análogos nucleotídicos.

14. Procedimiento según la reivindicación 13, en el que el cebador oligonucleotídico comprende nucleótidos seleccionados de entre el grupo constituido por ribonucleótidos, desoxirribonucleótidos, ribonucleótidos modificados, desoxirribonucleótidos modificados, ácidos nucleicos peptídicos, ácidos nucleicos peptídicos modificados y nucleótidos de esqueleto de fosfato-azúcar modificado.

15. Procedimiento según la reivindicación 1, en el que cada uno de los análogos nucleotídicos incorporados esta provisto de un marcador.

16. Procedimiento según la reivindicación 1, en el que el marcador se selecciona de entre el grupo constituido por cromóforos, restos fluorescentes, enzimas, antígenos, metales pesados, sondas magnéticas, colorantes, grupos fosforescentes, materiales radioactivos, restos quimioluminiscentes, nanopartículas dispersantes o fluorescentes, restos generadores de la señal Raman, y restos de detección electroquímica.

17. Procedimiento según la reivindicación 1, en el que el marcador se acopla al análogo nucleotídico en su base, resto de azúcar, alfa fosfato, beta fosfato o gamma fosfato.

18. Procedimiento según la reivindicación 1, en el que el marcador se acopla al análogo nucleotídico con un enlazador.

19. Procedimiento según la reivindicación 1, en el que el marcador se acopla al análogo nucleotídico en su terminal fosfato.

20. Procedimiento según la reivindicación 1, que comprende además:

retirar el marcador del análogo nucleotídico durante o después de dicha identificación y antes de dicha polimerización de un análogo nucleotídico posterior en una zona activa en la que el análogo nucleotídico se está añadiendo a la cadena cadena de ácido nucleico en crecimiento.

21. Procedimiento según la reivindicación 20, en el que dicha retirada se lleva a cabo blanqueando el marcador.

22. Procedimiento según la reivindicación 21, en el que dicho blanqueo se lleva a cabo por fotoblanqueo con radiación que está ajustada para provocar una retirada del marcador de referencia.

23. Procedimiento según la reivindicación 20, en el que dicha retirada se lleva a cabo por escisión del marcador del análogo nucleotídico.

24. Procedimiento según la reivindicación 23, en el que los análogos nucleotídicos marcados con beta-fosfato o con gamma-fosfato se escinden con enzimas.

25. Procedimiento según la reivindicación 1, en el que cada uno de la pluralidad de tipos de análogos nucleotídicos presentan diferentes marcadores que se distinguen entre sí durante dicha identificación.

26. Procedimiento según la reivindicación 1, en el que tres o menos de la pluralidad de tipos de análogos nucleotídicos presentan un marcador diferente.

27. Procedimiento según la reivindicación 1, en el que los diferentes tipos de análogos nucleotídicos presentan el mismo marcador pero se distinguen por las propiedades diferentes debido a la presencia de fluoróforos de base, fluoróforos extinguidos o análogos nucleotídicos fluorógenos.

28. Procedimiento según la reivindicación 1, en el que la enzima que polimeriza el ácido nucleico es portadora de un marcador y dicha identificación se lleva a cabo detectando la interacción entre el marcador y el análogo nucleotídico.

29. Procedimiento según la reivindicación 28, en el que el marcador es un donador o un aceptor de transferencia de energía por resonancia de fluorescencia.

30. Procedimiento según la reivindicación 1, en el que dicha identificación se lleva a cabo por procedimientos ópticos seleccionados de entre el grupo constituido por microespectroscopia en campo lejano, microespectroscopia en campo próximo, iluminación guiada por ondas evanescentes o por ondas, reforzamiento de nanoestructuras y las combinaciones de los mismos.

31. Procedimiento según la reivindicación 1, en el que dicha identificación se lleva a cabo utilizando la excitación monofotónica o multifotónica, la transferencia de energía por resonancia de fluorescencia o la fotoconversión.

32. Procedimiento según la reivindicación 1, en el que dicha identificación se consigue mediante discriminación de las longitudes de onda del espectro, medición y separación de los tiempos de permanencia de la fluorescencia, identificación de fluoróforos y/o supresión del fondo.

33. Procedimiento según la reivindicación 30, en el que la identificación del fluoróforo y/o la supresión del fondo utilizan la conexión rápida entre los modos de excitación y las fuentes de iluminación y las combinaciones de los mismos.

34. Procedimiento según la reivindicación 1, en el que dicha provisión de un complejo comprende:

disponer (1) un primer cebador oligonucleotídico o (2) la molécula de ácido nucleico diana en un soporte;

hibridar (1) la molécula de ácido nucleico diana con el primer cebador oligonucleotídico o (2) el primer cebador oligonucleotídico con la molécula de ácido nucleico diana dispuesta, para formar un complejo con la molécula de ácido nucleico diana cebada; y

proporcionar la enzima que polimeriza el ácido nucleico en el complejo con la molécula de ácido nucleico diana cebada.

35. Procedimiento según la reivindicación 34, en el que un segundo cebador oligonucleotídico se hibrida con la molécula de ácido nucleico diana en la zona opuesta a la hibridada por el primer cebador oligonucleotídico.

36. Procedimiento según la reivindicación 34, en el que el soporte y el cebador oligonucleotídico o la molécula de ácido nucleico diana se unen de manera reversible o irreversible con los componentes correspondientes de un par de enlace covalente o no covalente seleccionado de entre el grupo constituido por un par de enlace antígeno-anticuerpo, un par de enlace estreptavidina-biotina, moléculas de acoplamiento fotoactivadas, y un par de ácidos nucleicos complementarios.

37. Procedimiento según la reivindicación 34, en el que el cebador oligonucleotídico está dispuesto sobre el soporte y la molécula de ácido nucleico diana está hibridada con el cebador oligonucleotídico dispuesto.

38. Procedimiento según la reivindicación 34, en el que la molécula de ácido nucleico diana está dispuesta sobre el soporte y el cebador oligonucleotídico está hibridado con la molécula de ácido nucleico diana dispuesta.

39. Procedimiento según la reivindicación 1, en el que dicha provisión de un complejo comprende:

disponer, sobre un soporte, una molécula de ácido nucleico bicatenario que comprende el ácido nucleico diana y que presenta una zona de reconocimiento próxima a una zona activa en la que se está añadiendo el análogo nucleotídico a la cadena de ácido nucleico en crecimiento, y

proporcionar la enzima que polimeriza ácido nucleico sobre la molécula de ácido nucleico diana.

40. Procedimiento según la reivindicación 1, en el que dicha provisión de un complejo comprende:

disponer una enzima que polimeriza ácido nucleico sobre un soporte en una posición adecuada para que el complejo con ácido nucleico diana se desplace en relación a la enzima que polimeriza ácido nucleico.

41. Procedimiento según la reivindicación 40, en el que el soporte y la enzima que polimeriza el ácido nucleico se unen de manera reversible o irreversible con los componentes correspondientes de un par de enlace covalente o no covalente seleccionado de entre el grupo constituido por un par de enlace antígeno-anticuerpo, un par de enlace estreptavidina-biotina, moléculas de acoplamiento fotoactivadas y un par de ácidos nucleicos complementarios.

42. Procedimiento según la reivindicación 1, en el que la enzima que polimeriza el ácido nucleico o el ácido nucleico diana se dispone sobre un soporte.

43. Procedimiento según la reivindicación 1, en el que la enzima que polimeriza el ácido nucleico o el ácido nucleico diana se dispone en un gel con poros.

44. Procedimiento según la reivindicación 1, en el que el ácido nucleico diana y la enzima que polimeriza el ácido nucleico se disponen sobre un soporte sólido.

45. Procedimiento según la reivindicación 1, en el que dicha identificación se lleva a cabo reduciendo el ruido de fondo resultante de los análogos nucleotídicos libres.

46. Procedimiento según la reivindicación 45, en el que dicha identificación comprende:

dirigir la radiación de activación a una zona que corresponde sustancialmente a una zona activa en la que se está añadiendo análogo nucleotídico a la cadena de ácido nucleico en crecimiento; y

detectar el análogo nucleotídico polimerizado en la zona activa.

47. Procedimiento según la reivindicación 45, en el que dicha identificación distingue análogos nucleotídicos polimerizados en la zona activa en la que se está añadiendo análogo nucleotídico a la cadena de ácido nucleico en crecimiento procedente de los análogos nucleotídicos libres.

48. Procedimiento según la reivindicación 45, en el que dicha identificación se lleva a cabo en una zona confinada correspondiente a la zona activa en la que se está añadiendo análogo nucleotídico a la cadena de ácido nucleico en crecimiento.

49. Procedimiento según la reivindicación 48, en el que dicha identificación se lleva a cabo en una nanoestructura.

50. Procedimiento según la reivindicación 49, en el que dicha nanoestructura es una nanoestructura punteada, acicular o resonante que aumenta dicha detección.

51. Procedimiento según la reivindicación 48, en el que los análogos nucleotídicos que no se polimerizan en la zona activa se desplazan rápidamente por una microestructura hasta y desde la zona confinada correspondiente a la zona activa.

52. Procedimiento según la reivindicación 51, en el que la microestructura comprende:

una pluralidad de canales para dirigir diferentes análogos nucleotídicos a la zona confinada y

un canal de descarga para permitir que los materiales se retiren de la zona confinada, y la nanoestructura comprende:

un alojamiento que define la zona confinada y se construye para facilitar dicha identificación.

53. Procedimiento según la reivindicación 45, en el que dicha identificación se lleva a cabo por intensificación del campo electromagnético con radiación electromagnética intensificándose en la proximidad a un objeto con un radio de curvatura pequeño adyacente a la zona activa.

54. Procedimiento según la reivindicación 45, en el que dicha identificación se lleva a cabo por iluminación en campo próximo de las cavidades en las que está dispuesta la molécula de ácido nucleico diana cebada.

55. Procedimiento según la reivindicación 45, en el que dicha identificación se lleva a cabo con fibras ópticas próximas al complejo.

56. Procedimiento según la reivindicación 45, en el que dicha identificación y dicha reducción del fondo se llevan a cabo por retardo controlado por tiempo de la detección de fotones.

57. Procedimiento según la reivindicación 1, en el que dicho procedimiento se lleva a cabo secuenciando diferentes moléculas de ácido nucleico diana en una variedad de ubicaciones diferentes en una matriz.

58. Procedimiento según la reivindicación 1, en el que dicho procedimiento se lleva a cabo secuenciando simultánea o secuencialmente el mismo ácido nucleico diana y combinando la salida de dicha secuenciación.

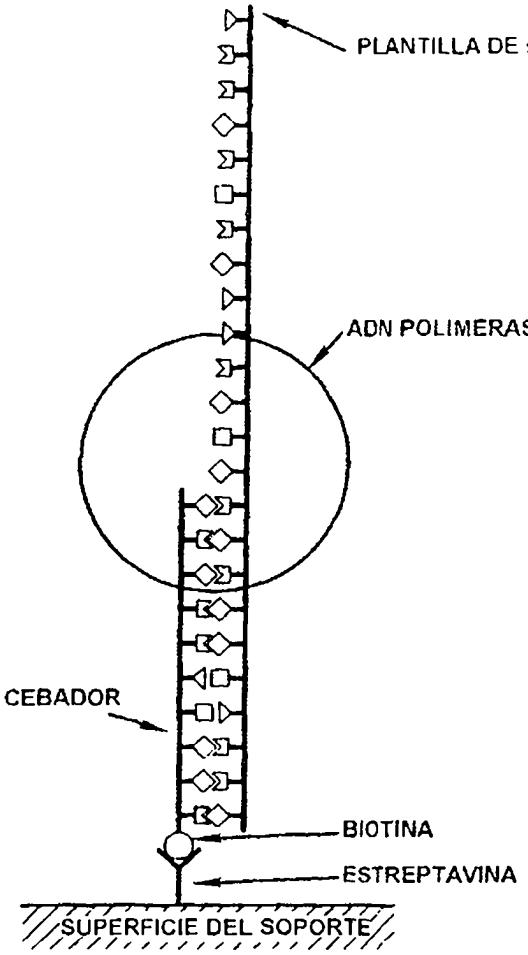

FIG. 1A

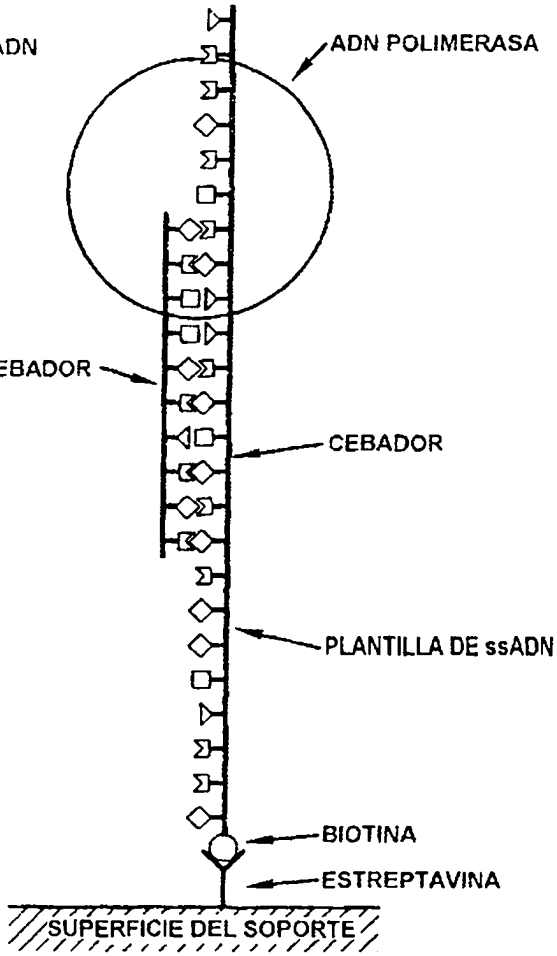

FIG. 1B

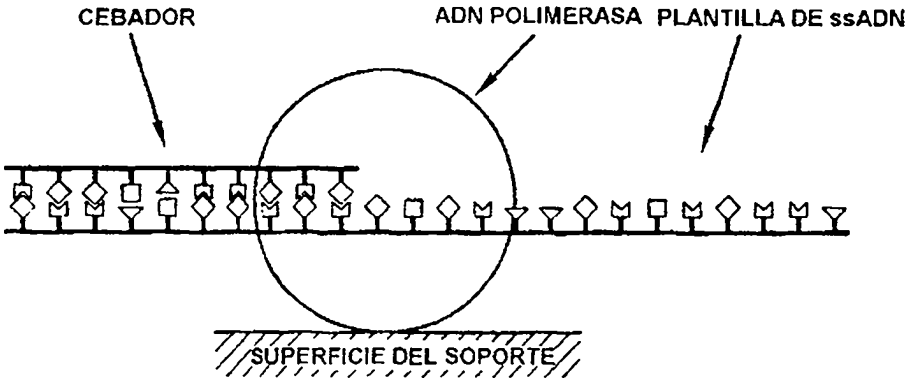

FIG. 1C

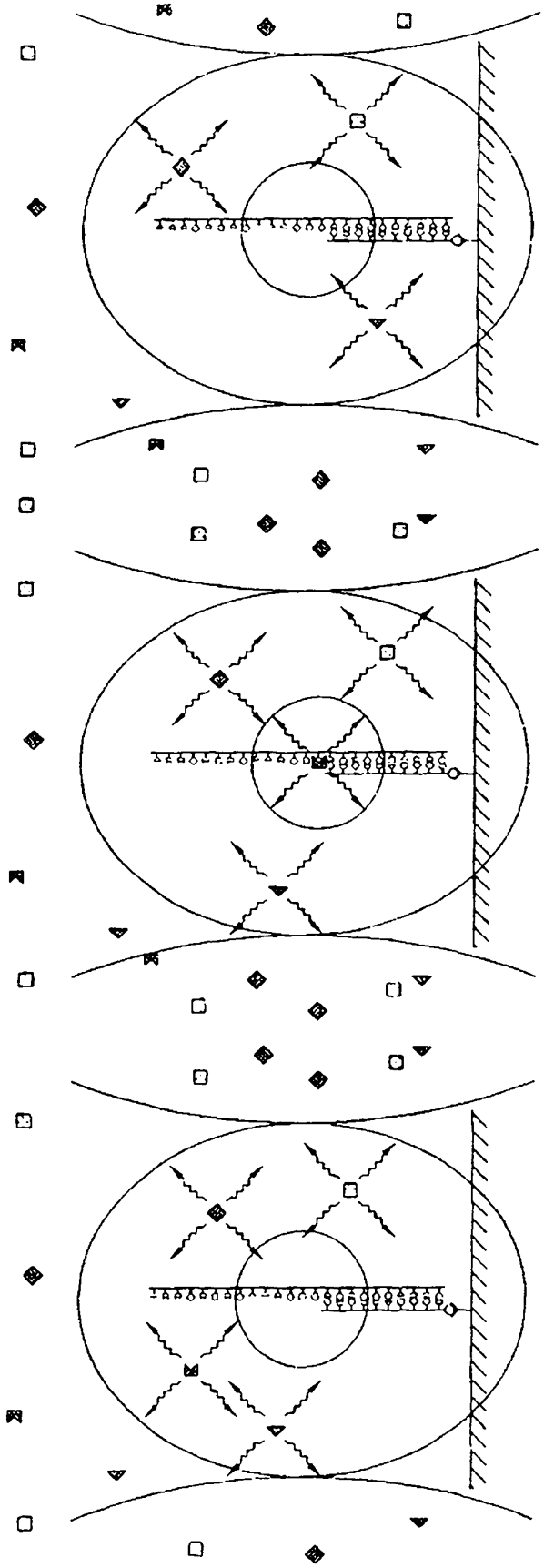

FIG. 2A

FIG. 2B

FIG. 2C

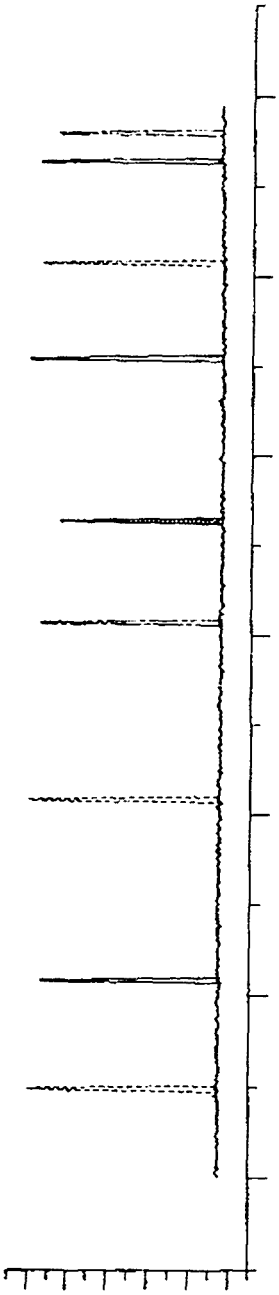

FIG. 3A

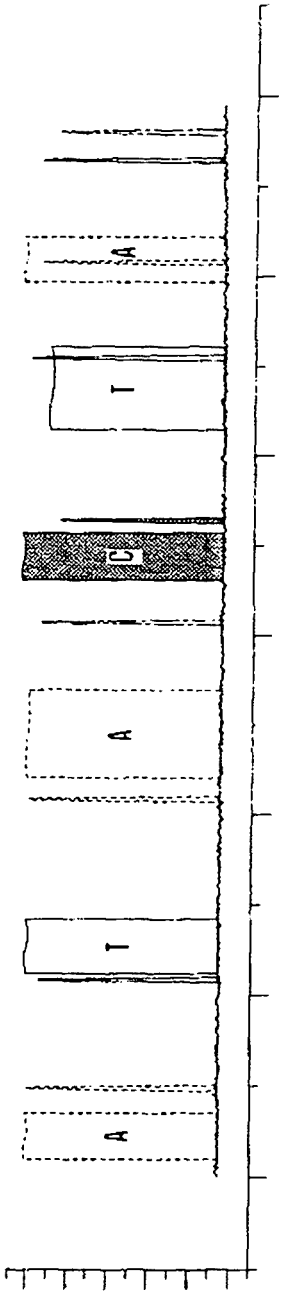

FIG. 3B

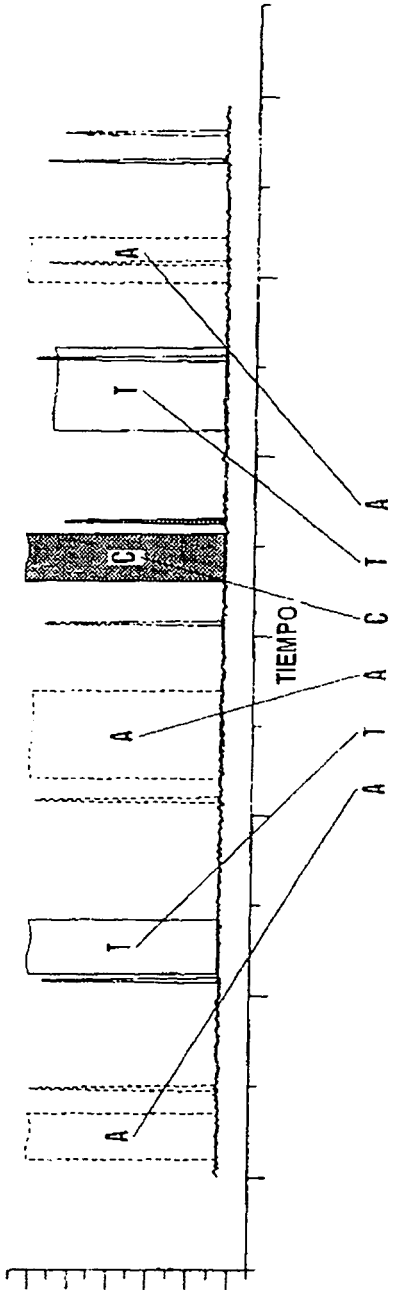

FIG. 3C

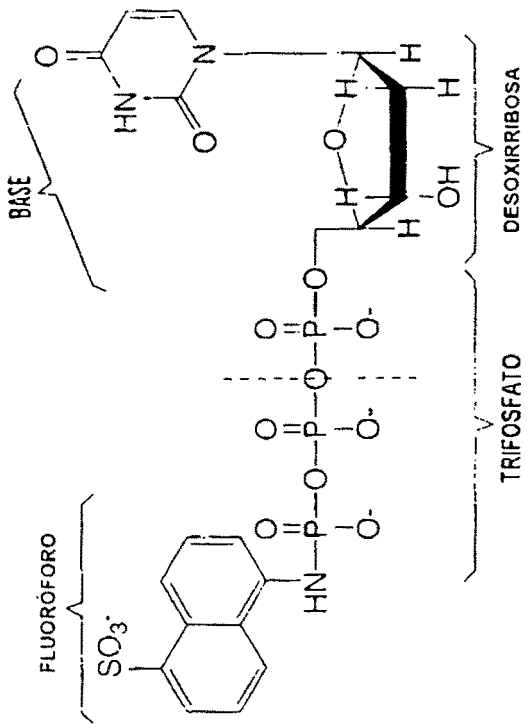

FIG. 4A

FIG. 4B

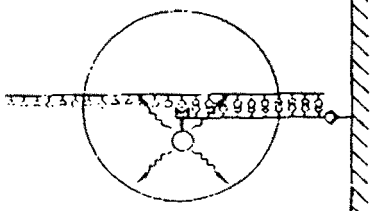

FIG. 4C

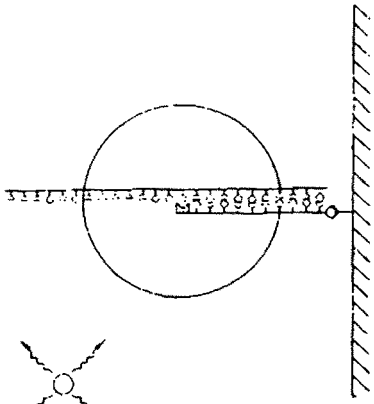

FIG. 4D

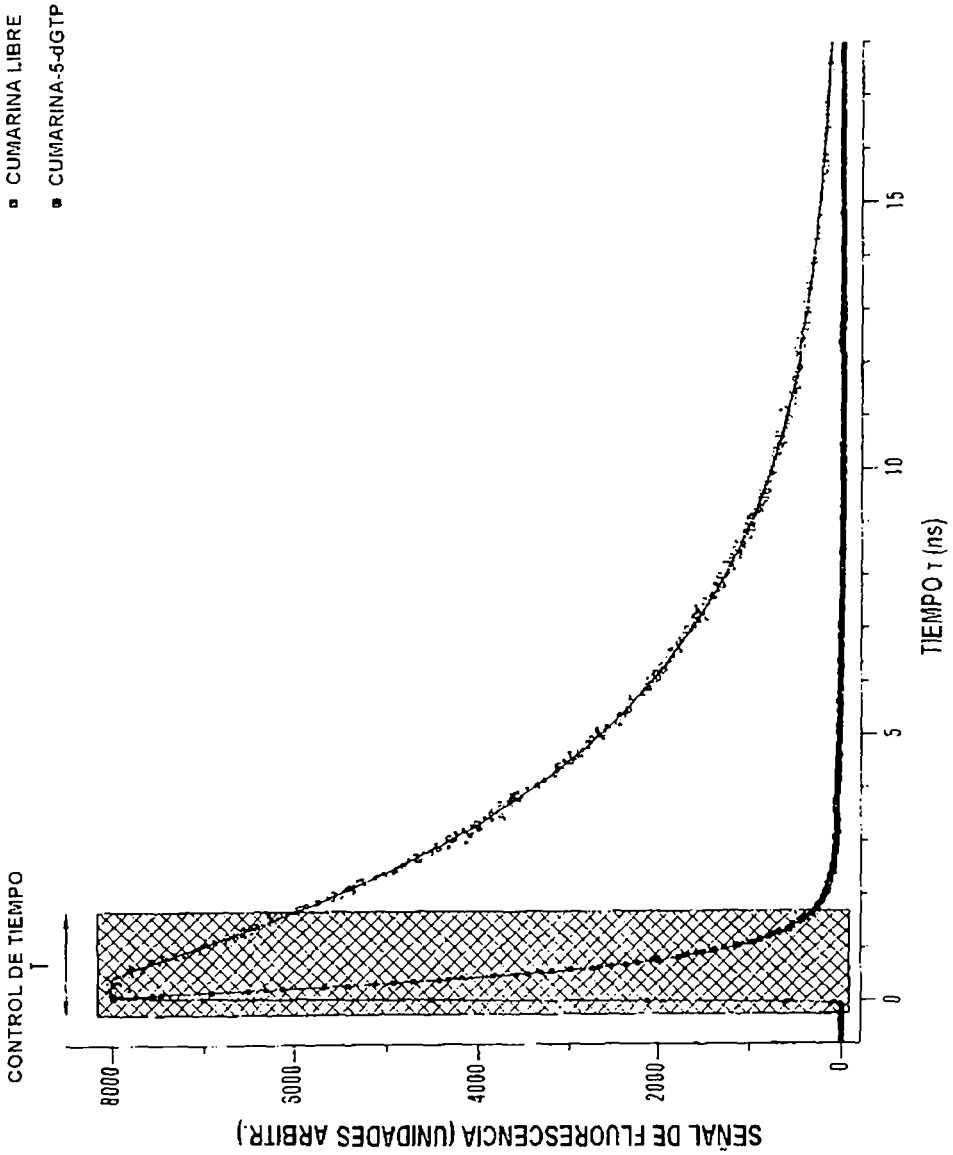
CUMARINA LIBRE
CUMARINA-5-dGTP
CONTROL DE TIEMPO
T
SEÑAL DE FLUORESCENCIA (UNIDADES ARBITR.)
8000
6000
4000
2000
0
0
5
10
15
TIEMPO τ (ns)

FIG. 5

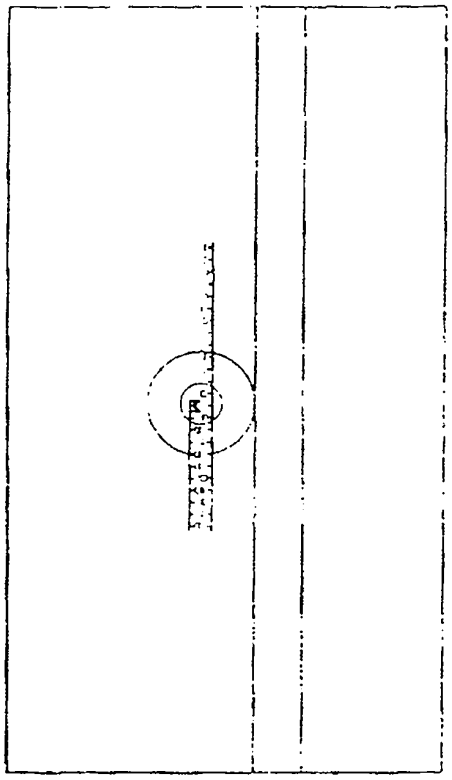

FIG. 6B

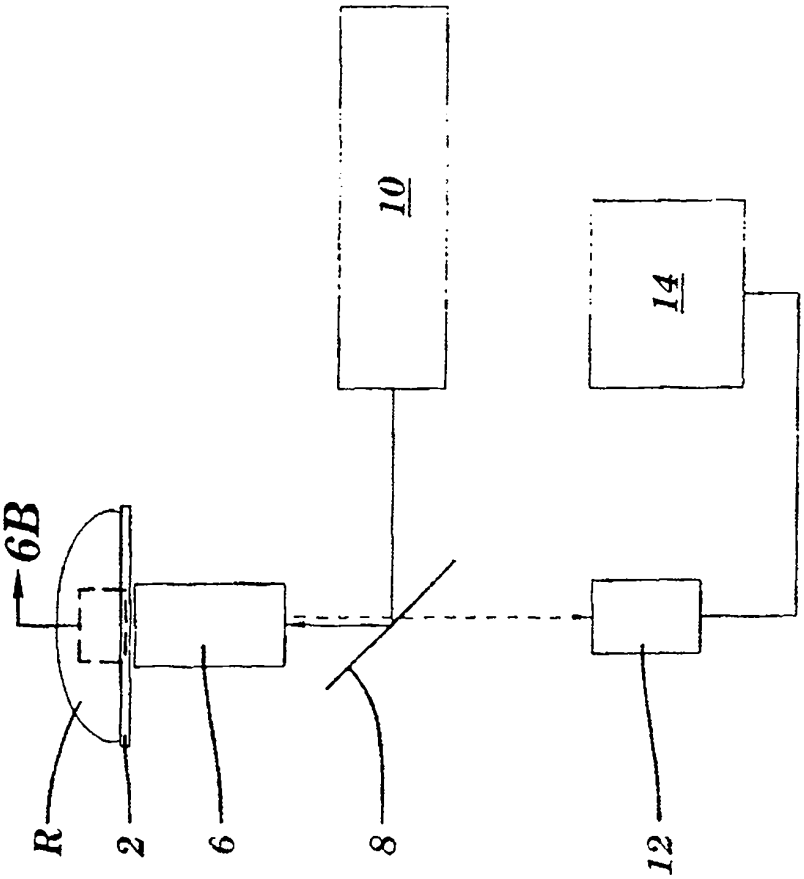

FIG. 6A

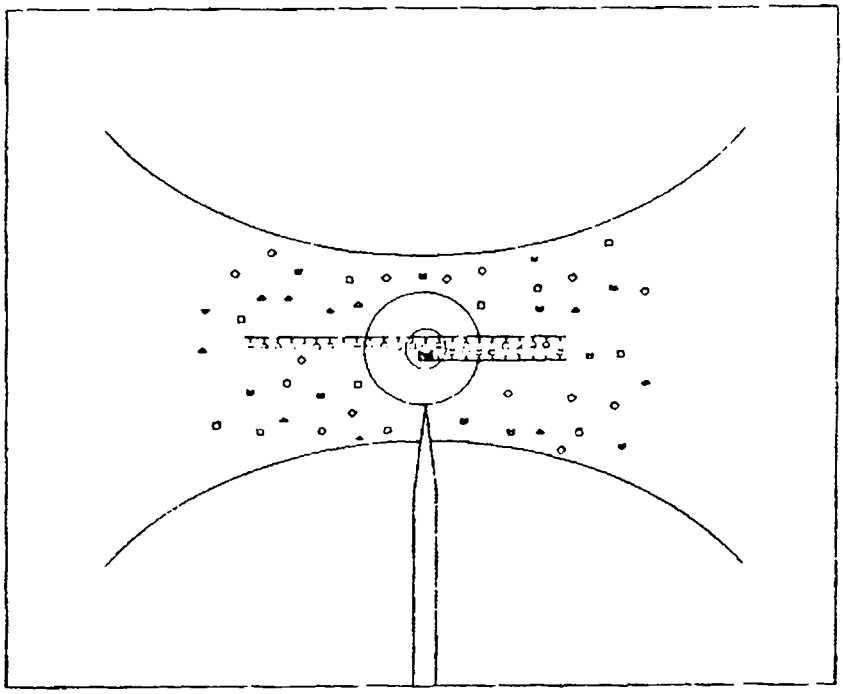

FIG. 7B

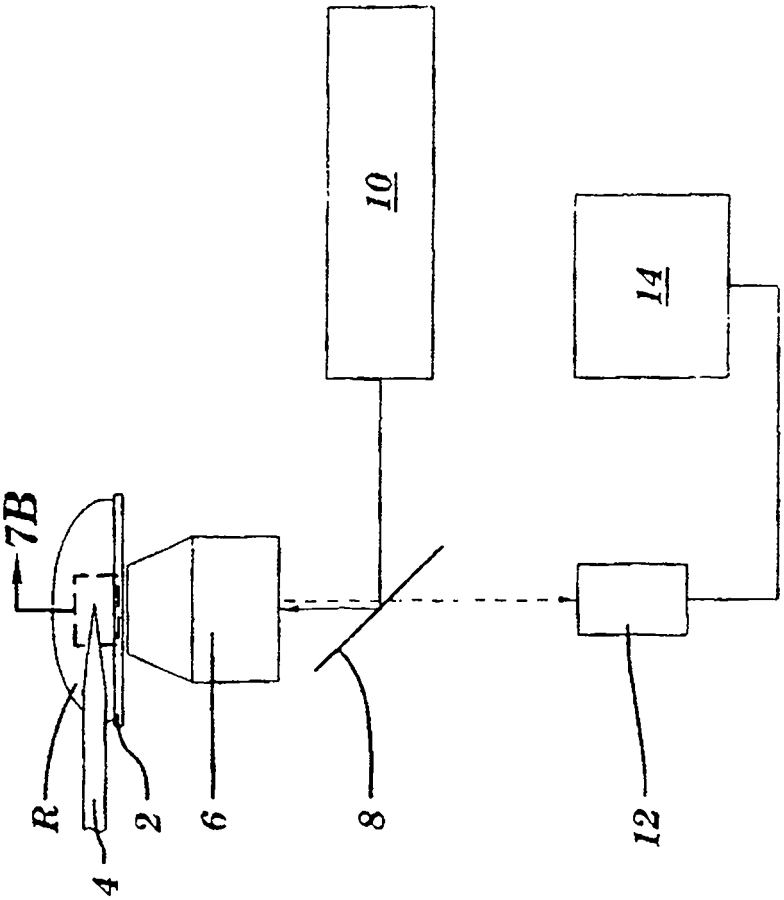

FIG. 7A

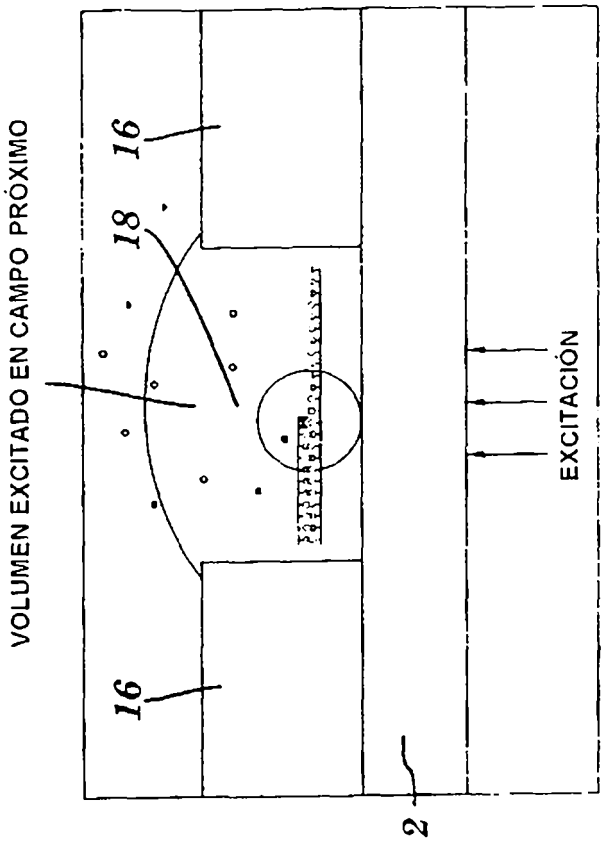
VOLUMEN EXCITADO EN CAMPO PRÓXIMO
16
18
16
EXCITACIÓN
2

FIG. 8B

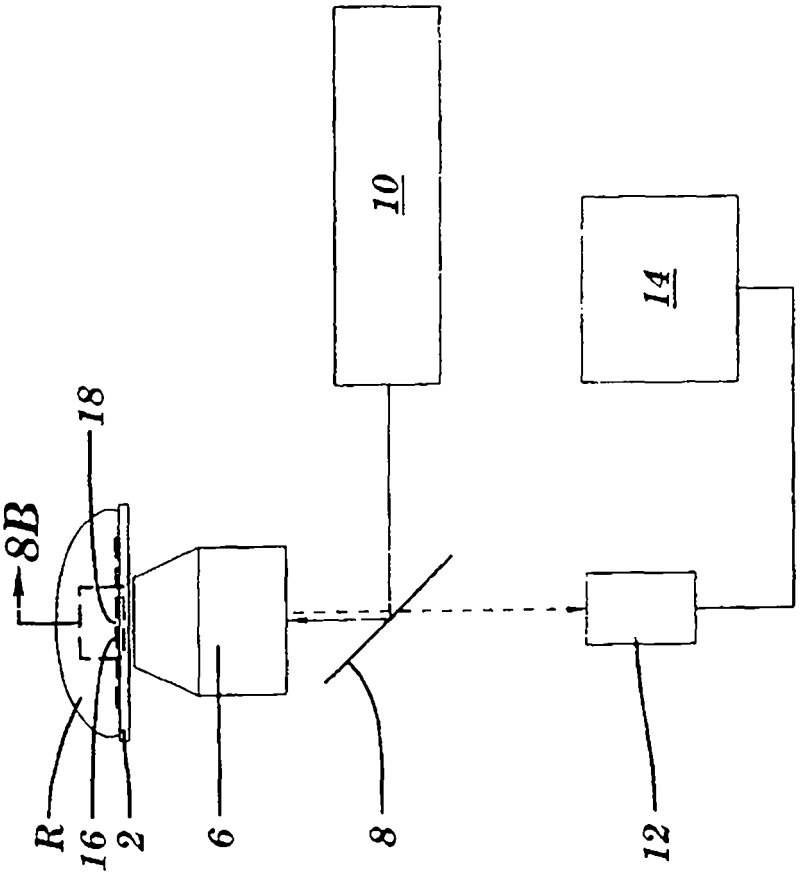
8B
18
R
16
2
6
8
10
12
14

FIG. 8A

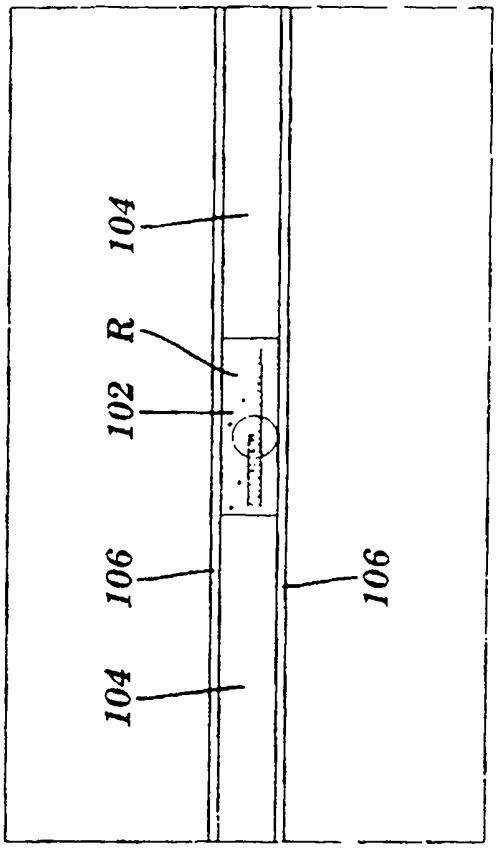

FIG. 9B

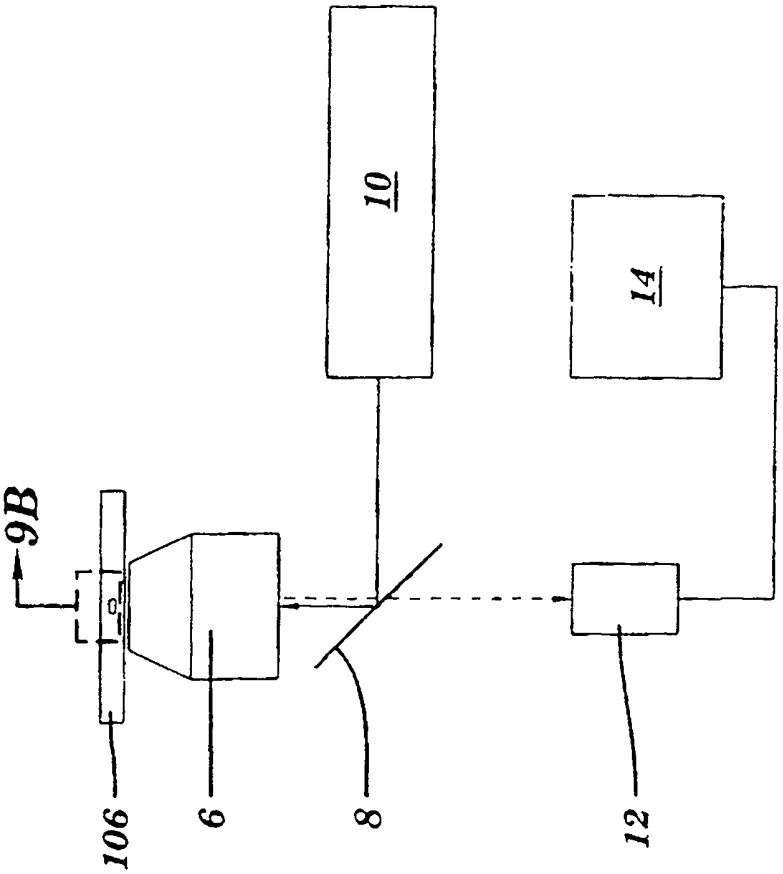

FIG. 9A

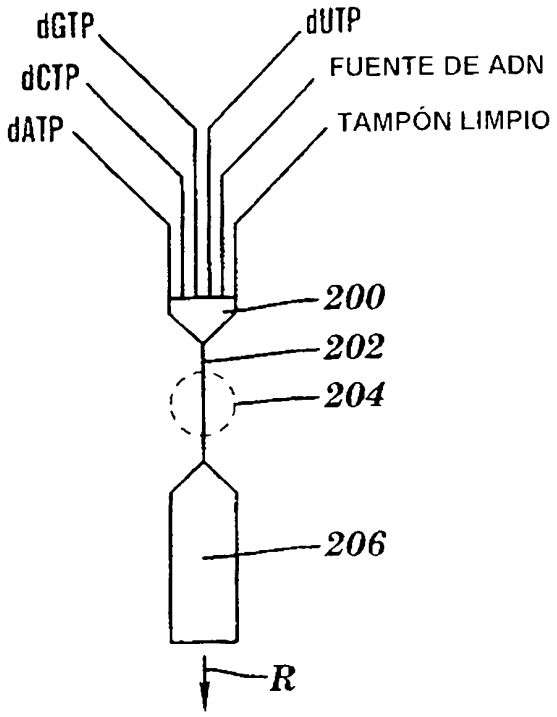

FIG. 10A

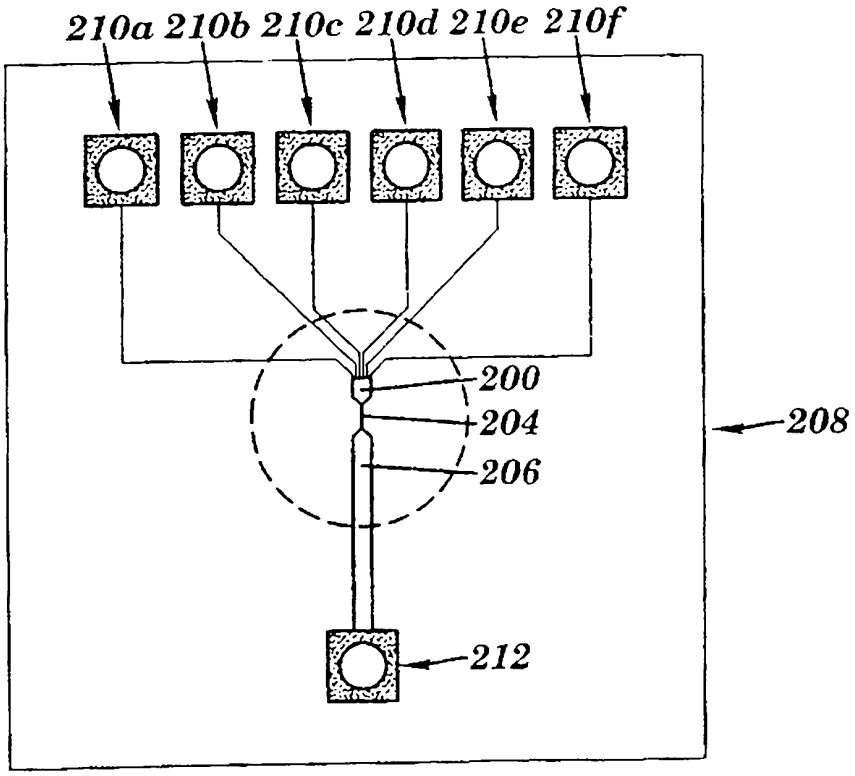

FIG. 10B