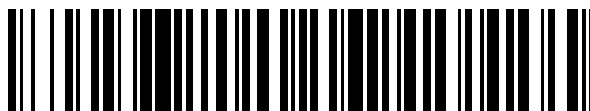


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 559 313**

51 Int. Cl.:

**C12Q 1/68** (2006.01)

**C07H 21/02** (2006.01)

**C07H 21/04** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **19.06.2008 E 08771483 (8)**

97 Fecha y número de publicación de la concesión europea: **14.10.2015 EP 2171088**

54 Título: **Secuenciación de ácidos nucleicos de alto rendimiento por expansión**

30 Prioridad:

**19.06.2007 US 945031 P**

**23.10.2007 US 981916 P**

**25.10.2007 US 305 P**

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

**11.02.2016**

73 Titular/es:

**STRATOS GENOMICS INC. (100.0%)**

**2211 Elliott Avenue, Suite 210**

**Seattle, WA 98121, US**

72 Inventor/es:

**KOKORIS, MARK STAMATIOS. y**

**MCRUER, ROBERT N.**

74 Agente/Representante:

**DE ELZABURU MÁRQUEZ, Alberto**

**ES 2 559 313 T3**

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

**DESCRIPCIÓN**

Secuenciación de ácidos nucleicos de alto rendimiento por expansión

**Antecedentes****Campo técnico**

- 5 La presente invención está generalmente relacionada con la secuenciación de ácidos nucleicos, así como con métodos relacionados con la misma.

**Descripción de la técnica relacionada**

10 Las secuencias de ácidos nucleicos codifican la información necesaria para que los seres vivos funcionen y se reproduzcan, y son esencialmente un diseño de vida. La determinación de tales secuencias es, por tanto, una herramienta útil en la investigación pura en cómo y dónde viven los organismos, así como en las ciencias aplicadas tales como el desarrollo de fármacos. En medicina, las herramientas de secuenciación pueden usarse para diagnosticar y desarrollar tratamientos para diversas patologías, que incluyen cáncer, enfermedad cardíaca, trastornos autoinmunitarios, esclerosis múltiple u obesidad. En la industria, la secuenciación puede usarse para diseñar procesos enzimáticos u organismos sintéticos mejorados. En biología, tales herramientas pueden usarse para estudiar la salud de ecosistemas, por ejemplo, y así tener una amplia gama de utilidad.

15 Una secuencia de ADN única del individuo proporciona información valiosa referente a su susceptibilidad a ciertas enfermedades. La secuencia proporcionará a los pacientes la oportunidad de cribar para la detección temprana y para recibir tratamiento preventivo. Además, dado el proyecto individual de un paciente, los profesionales clínicos serán capaces de administrar terapia personalizada para maximizar la eficacia de los fármacos y para minimizar el riesgo de una respuesta adversa a fármacos. Similarmente, la determinación del proyecto de los organismos patógenos puede conducir a nuevos tratamientos para enfermedades infecciosas y supervisión de patógenos más robusta. La secuenciación de ADN del genoma completo proporcionará la base para la medicina moderna.

20 La secuenciación de ADN es el proceso de determinación del orden de los constituyentes químicos de un polímero de ADN dado. Estos constituyentes químicos, que se llaman nucleótidos, existen en el ADN en cuatro formas comunes: desoxiadenosina (A), desoxiguanosina (G), desoxicitidina (C) y desoxitimidina (T). La secuenciación de un genoma humano diploide requiere determinar el orden secuencial de aproximadamente 6 billones de nucleótidos.

25 Actualmente, la mayor parte de la secuenciación de ADN se realiza usando el método de terminación de cadenas desarrollado por Frederick Sanger. Esta técnica, llamada secuenciación de Sanger, usa la terminación específica de secuencia de la síntesis de ADN y sustratos indicadores de nucleótidos fluorescentemente modificados para derivar la información de secuencias. Este método secuenciación una hebra de ácido nucleico diana, o longitud de lectura, de hasta 1000 bases de longitud usando una reacción en cadena de la polimerasa modificada. En esta reacción modificada la secuenciación se interrumpe aleatoriamente en tipos de bases seleccionados (A, C, G o T) y las longitudes de las secuencias interrumpidas se determinan por electroforesis en gel capilar. La longitud determina entonces qué tipo de base está localizado en esa longitud. Se producen muchas longitudes de lectura solapantes y sus secuencias se solapan usando procesamiento de datos para determinar el ajuste más fiable de los datos. Este proceso de producción de longitudes de lectura de secuencia es muy laborioso y caro y ahora está siendo sustituido por métodos nuevos que tienen mayor eficiencia.

30 Se usó el método de Sanger para proporcionar la mayoría de los datos de secuencia en el Proyecto Genoma Humano que generó la primera secuencia completa del genoma humano. Este proyecto tuvo lugar durante 10 años y casi 3 billones de dólares para completarse. Dadas estas significativas limitaciones de rendimiento y coste, es evidente que las tecnologías de secuenciación de ADN necesitarán mejorar drásticamente con el fin de lograr los objetivos establecidos propuestos por la comunidad científica. Para este fin, varias tecnologías de segunda generación, que superaron con creces las limitaciones básicas de rendimiento y coste de la secuenciación de Sanger, están consiguiendo una cuota cada vez mayor del mercado de la secuenciación. Todavía estos métodos de "secuenciación por síntesis" se quedan cortos en lograr los objetivos de rendimiento, coste y calidad requeridos por los mercados tales como la secuencia del genoma completo para la medicina personalizada.

35 Por ejemplo, 454 Life Sciences está produciendo instrumentos (por ejemplo, el secuenciador del genoma) que puede procesar 100 millones de bases en 7,5 horas con una longitud de lectura promedio de 200 nucleótidos. Su enfoque usa una variación de la reacción en cadena de la polimerasa ("PCR") para producir una colonia homogénea de ácido nucleico diana, cientos de bases en longitud, sobre la superficie de una perla. Este proceso se llama PCR en emulsión. Cientos de miles de tales perlas están entonces dispuestas sobre una "placa de picotítulo". La placa se prepara entonces para una secuenciación adicional por la cual cada tipo de base de ácido nucleico se lava secuencialmente sobre la placa. Las perlas con diana que incorporan la base producen un subproducto de pirofosfato que puede usarse para catalizar una reacción que produce luz que entonces se detecta con una cámara.

50 Illumina Inc. tiene un proceso similar que usa nucleótidos de terminación reversible y marcas fluorescentes para realizar la secuenciación de ácidos nucleicos. La longitud de lectura promedio para el analizador 1G de Illumina es

inferior a 40 nucleótidos. En lugar de usar PCR en emulsión para amplificar dianas de secuencia, Illumina tiene un enfoque para amplificar colonias por PCR sobre una superficie de matriz. Tanto los enfoques 454 como de Illumina usan una complicada amplificación por polimerasa para aumentar la intensidad de la señal, realizan mediciones de bases durante el ciclo de extensión de la secuencia limitante de la tasa y tienen longitudes de lectura limitadas debido a los errores de incorporación que degradan la señal de medición con respecto al ruido proporcionalmente a la longitud de lectura.

Applied Biosystems usa ligación de terminación reversible en vez de secuenciación por síntesis para leer el ADN. Al igual que el secuenciador del genoma 454, la tecnología usa PCR en emulsión basada en perlas para amplificar la muestra. Como la mayoría de las perlas no llevan productos de PCR, los investigadores usan a continuación una etapa de enriquecimiento para seleccionar perlas recubiertas con ADN. Las perlas recubiertas con biotina se extienden y se inmovilizan sobre una matriz de portaobjetos de vidrio cubierta con estreptavidina. A continuación, las perlas inmovilizadas se ejecutan a través de un proceso de hibridación de sondas 8-meras (cada una marcada con cuatro colorantes fluorescentes diferentes), ligación y escisión (entre las bases 5<sup>a</sup> y 6<sup>a</sup> para crear un sitio para la siguiente ronda de ligación). Cada sonda interroga a dos bases, en las posiciones 4 y 5, usando un sistema que codifica 2 bases, que se registra por una cámara. Similar al enfoque de Illumina, la longitud de lectura promedio para la plataforma Solid de Applied Biosystems es inferior a 40 nucleótidos.

Están siendo desarrollados otros enfoques para evitar el tiempo y gasto de la etapa de amplificación de la polimerasa midiendo moléculas individuales de ADN directamente. Visigen Biotechnologies, Inc. está midiendo bases fluorescentemente marcadas a medida que se secuencian incorporando un segundo fluoróforo en una ADN polimerasa manipulada por ingeniería y usando transferencia de energía de resonancia de Forster (FRET) para la identificación de nucleótidos. Esta técnica se enfrenta a los retos de separar las señales de bases que están separadas por menos de un nanómetro y por una acción de incorporación de polimerasa que tendrá variación estadística muy grande.

Un proceso que es desarrollado por LingVitae secuencia ADNc insertado en vectores plasmídicos inmovilizados. El proceso usa una enzima de restricción de clase IIS para escindir el ácido nucleico diana y ligar un oligómero en la diana. Normalmente, uno o dos nucleótidos en los nucleótidos protuberante de 5' o 3' terminales generados por la enzima de restricción determinan cuál de una biblioteca de oligómeros en la mezcla de ligación se añadirá al extremo cortado pegajoso de la diana. Cada oligómero contiene secuencias "señal" que identifican únicamente al (a los) nucleótido(s) que sustituye. Entonces se repite el proceso de escisión y ligación. La nueva molécula se secuencia entonces usando marcas específicas para los diversos oligómeros. El producto de este proceso se llama un "polímero de diseño" y siempre consiste en un ácido nucleico más largo que uno al que sustituye (por ejemplo, una secuencia diana de dinucleótido está sustituida con una secuencia de polinucleótidos "ampliada" de nada menos que 100 pares de bases). Una ventaja de este proceso es que la hebra del producto de dúplex puede amplificarse si se desea. Una desventaja es que el proceso es necesariamente cíclico y la continuidad del molde se perdería si se hicieran múltiples cortes de restricción simultáneos.

La patente de EE.UU. N°. 7.060.440 a Kless describe un proceso de secuenciación que implica incorporar oligómeros por polimerización con una polimerasa. Se usa una modificación del método de Sanger, con oligómeros terminados en el extremo como sustratos, para construir escaleras de secuenciación por electroforesis en gel o cromatografía capilar. Aunque el acoplamiento de oligómeros por ligación de extremos es muy conocido, el uso de una polimerasa para acoplar oligómeros en un proceso dirigido por molde se utilizó como una nueva ventaja.

Se espera que las técnicas de polimerización crezcan en potencia a medida que las polimerasas modificadas (y ligasas) estén disponibles mediante ingeniería genética y bioprospección, y ya se conocen métodos de eliminación de actividad de exonucleasa por modificación de polimerasas. Por ejemplo, la solicitud de patente de EE.UU. publicada 2007/0048748 a Williams describe el uso de polimerasas mutantes para incorporar nucleótidos marcados con colorante y otros nucleótidos modificados. Los sustratos para estas polimerasas también incluyen nucleótidos marcados con  $\gamma$ -fosfato. Se encontraron tanto elevada velocidad de incorporación como reducción en la tasa de error con polimerasas quiméricas y mutantes.

Además, se ha hecho un gran esfuerzo por tanto equipos académicos como industriales para secuenciar ADN nativo usando métodos no sintéticos. Por ejemplo, Agilent Technologies, Inc. junto con colaboradores universitarios, están desarrollando un método de detección de una única molécula que bobina el ADN a través de un nanoporo para hacer mediciones a medida que pasa a su través. Al igual que con Visigen y LingVitae, este método debe vencer el problema de obtener eficazmente y con exactitud distintas señales de nucleobases individuales separadas por dimensiones sub-nanométricas, además del problema de desarrollar tamaños de poro reproducibles de tamaño similar. Como tal, la secuenciación directa de ADN por detección de sus partes constituyentes tiene ahora que lograrse en un proceso de alto rendimiento debido al pequeño tamaño de los nucleótidos en la cadena (aproximadamente 4 Angstroms de centro a centro) y las limitaciones de la resolución de señal con respecto a ruido y de señales correspondientes a este respecto. La detección directa se complica adicionalmente por la estructura secundaria inherente del ADN, que no se alarga fácilmente en un polímero perfectamente lineal.

Aunque se han hecho avances significativos en el campo de la secuenciación de ADN, continúa siendo una necesidad en la materia métodos nuevos y mejorados. La presente invención cumple estas necesidades y

proporciona otras ventajas relacionadas.

5 El documento US 2002/028458 A1 desvela un método de secuenciación de toda o parte de una molécula de ácido nucleico diana que determina la secuencia de una porción de dicha molécula de ácido nucleico, además de información referente a la posición de dicha porción, y en particular proporciona un nuevo método de secuenciación que implica el aumento de una o más bases de dichas bases para ayudar en la identificación.

10 El documento WO 00/79257 A desvela un método de evaluación de una molécula de polímero que incluye residuos de monómero linealmente conectados. Dicho método incluye proporcionar una molécula de polímero en un líquido, poner en contacto el líquido con una membrana en estado sólido aislante que tiene un detector capaz de detectar características de la molécula de polímero, y hacer que la molécula de polímero atraviese una región limitada de la membrana en estado sólido de manera que monómeros de la molécula de polímero atraviesen el límite en orden secuencial, por lo que la molécula de polímero interacciona linealmente con el detector y se obtienen datos adecuados para determinar las características de la molécula de polímero.

15 El documento WO 2006/076650 A2 desvela composiciones, métodos y kits para amplificar y detectar selectivamente secuencias diana. En algunas realizaciones se desvelan una sonda circularizable y/o un par de sondas para amplificar selectivamente secuencias diana.

### Breve resumen

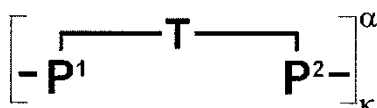
20 En términos generales, se desvelan métodos y dispositivos correspondientes y productos que vencen los retos de resolución espacial por técnicas de secuenciación de ácidos nucleicos de alto rendimiento existentes. Esto se logra codificando la información del ácido nucleico en un polímero sustituto de longitud extendida que es más fácil de detectar. El polímero sustituto (denominado en el presente documento un "Xpandómero") se forma por síntesis dirigida por molde que preserva la información genética original del ácido nucleico diana, mientras que también aumenta la separación lineal de los elementos individuales de los datos de secuencia.

25 En una realización, se desvela un método para la secuenciación de un ácido nucleico diana, que comprende: a) proporcionar una hebra hija producida por una síntesis dirigida por molde, comprendiendo la hebra hija una pluralidad de subunidades acopladas en una secuencia correspondiente a una secuencia de nucleótidos contigua de toda o una porción del ácido nucleico diana, en el que las subunidades individuales comprenden un anclaje, al menos una sonda o residuo de nucleobase, y al menos un enlace selectivamente escindible; b) escindir el al menos un enlace selectivamente escindible para dar un Xpandómero de una longitud mayor que la pluralidad de las subunidades de la hebra hija, comprendiendo el Xpandómero los anclajes y elementos indicadores para analizar la información genética en una secuencia correspondiente a la secuencia de nucleótidos contigua de toda o una porción del ácido nucleico diana; y c) detectar los elementos indicadores del Xpandómero.

30 En más realizaciones específicas, los elementos indicadores para analizar la información genética pueden asociarse a los anclajes del Xpandómero, con la hebra hija antes de la escisión del al menos un enlace selectivamente escindible, o con el Xpandómero después de la escisión del al menos un enlace selectivamente escindible. El Xpandómero puede comprender además toda o una porción de la al menos una sonda o residuo de nucleobase, y los elementos indicadores para analizar la información genética pueden asociarse a la al menos una sonda o residuo de nucleobase o pueden ser ellos mismos la sonda o residuos de nucleobase. Además, el enlace selectivamente escindible puede ser un enlace covalente, un enlace intra-anclaje, un enlace entre o dentro de las sondas o residuos de nucleobase de la hebra hija, y/o un enlace entre las sondas o residuos de nucleobase de la hebra hija y un molde diana.

En otras realizaciones, los Xpandómeros tienen las siguientes estructuras (I) a (X):

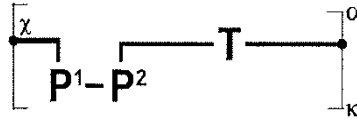
Estructura (I):



en la que

45 T representa el anclaje;  
 P<sup>1</sup> representa un primer resto de sonda;  
 P<sup>2</sup> representa un segundo resto de sonda;  
 κ representa la subunidad κ<sup>n</sup> en una cadena de m subunidades, en la que m es un número entero superior a tres; y  
 50 α representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos contigua de una porción del ácido nucleico diana.

Estructura (II):



en la que

T representa el anclaje;

5 P<sup>1</sup> representa un primer resto de sonda;

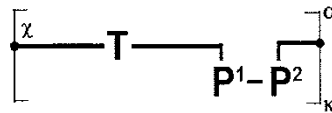
P<sup>2</sup> representa un segundo resto de sonda;

κ representa la subunidad k<sup>n</sup> en una cadena de *m* subunidades, en la que *m* es un número entero superior a tres;

10 α representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y

χ representa un enlace con el anclaje de una subunidad adyacente.

Estructura (III):



15 en la que

T representa el anclaje;

P<sup>1</sup> representa un primer resto de sonda;

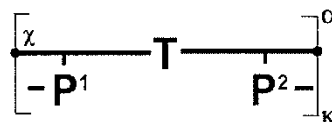
P<sup>2</sup> representa un segundo resto de sonda;

20 κ representa la subunidad k<sup>n</sup> en una cadena de *m* subunidades, en la que *m* es un número entero superior a tres;

α representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y

χ representa un enlace con el anclaje de una subunidad adyacente.

25 Estructura (IV):



en la que

T representa el anclaje;

P<sup>1</sup> representa un primer resto de sonda;

30 P<sup>2</sup> representa un segundo resto de sonda;

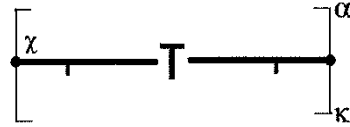
κ representa la subunidad k<sup>n</sup> en una cadena de *m* subunidades, en la que *m* es un número entero superior a tres;

α representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y

35

$\chi$  representa un enlace con el anclaje de una subunidad adyacente.

Estructura (V):



en la que

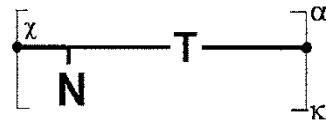
5 T representa el anclaje;

$\chi$  representa la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $m$  es un número entero superior a tres;

10  $\alpha$  representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y

$\chi$  representa un enlace con el anclaje de una subunidad adyacente.

Estructura (VI):



en la que

15 T representa el anclaje;

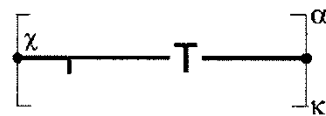
N representa un residuo de nucleobase;

$\kappa$  representa la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $m$  es un número entero superior a diez;

20  $\alpha$  representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y

$\chi$  representa un enlace con el anclaje de una subunidad adyacente.

Estructura (VII):



25 en la que

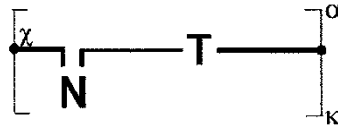
T representa el anclaje;

$\kappa$  representa la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $m$  es un número entero superior a diez;

30  $\alpha$  representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y

$\chi$  representa un enlace con el anclaje de una subunidad adyacente.

Estructura (VIII):



en la que

T representa el anclaje;

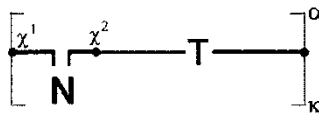
5 N representa un residuo de nucleobase;

κ representa la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $m$  es un número entero superior a diez;

10 α, representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y

χ representa un enlace con el anclaje de una subunidad adyacente.

Estructura (IX):



en la que

15 T representa el anclaje;

N representa un residuo de nucleobase;

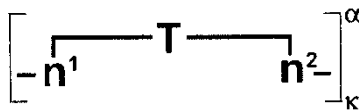
κ representa la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $m$  es un número entero superior a diez;

20 α representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos contigua de una porción del ácido nucleico diana;

χ representa un enlace con el anclaje de una subunidad adyacente; y

χ² representa un enlace inter-anclaje.

Estructura (X):



25 en la que

T representa el anclaje;

$n^1$  y  $n^2$  representa una primera porción y una segunda porción, respectivamente, de un residuo de nucleobase;

30 κ representa la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $m$  es un número entero superior a diez; y

α representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos contigua de una porción del ácido nucleico diana.

35 Además, se desvelan construcciones de sustrato de oligómero para su uso en una síntesis dirigida por molde para la secuenciación de un ácido nucleico diana.

Las construcciones de sustrato de oligómero comprenden un primer resto de sonda unido a un segundo resto de sonda, teniendo cada uno del primer y segundo restos de sonda un grupo terminal adecuado para la síntesis dirigida

por molde, y un anclaje que tiene un primer extremo y un segundo extremo con al menos el primer extremo del anclaje unido a al menos uno del primer y segundo restos de sonda, en las que la construcción de sustrato de oligómero cuando se usa en la síntesis dirigida por molde es capaz de formar una hebra hija que comprende un Xpandómero limitado y que tiene una pluralidad de subunidades acopladas en una secuencia correspondiente a la secuencia de nucleótidos contigua de toda o una porción del ácido nucleico diana, en las que las subunidades individuales comprenden un anclaje, el primer y segundo restos de sonda y al menos un enlace selectivamente escindible.

Además, se desvelan construcciones de sustrato de oligómero para su uso en una síntesis dirigida por molde para la secuenciación de un ácido nucleico diana. Las construcciones de sustrato de oligómero comprenden un residuo de nucleobase con grupos terminales adecuados para la síntesis dirigida por molde, y un anclaje que tiene un primer extremo y un segundo extremo con al menos el primer extremo del anclaje unido al residuo de nucleobase, en las que la construcción de sustrato de monómero cuando se usa en la síntesis dirigida por molde es capaz de formar una hebra hija que comprende un Xpandómero limitado y que tiene una pluralidad de subunidades acopladas en una secuencia correspondiente a la secuencia de nucleótidos contigua de toda o una porción del ácido nucleico diana, en la que las subunidades individuales comprenden un anclaje, el residuo de nucleobase y al menos un enlace selectivamente escindible.

Además, se desvelan dúplex de molde-hebra hija que comprenden una hebra hija duplexada con una hebra molde, además de a métodos de formación de los mismos a partir de la hebra molde y el oligómero o las construcciones de sustrato de oligómero.

Estos y otros aspectos de la invención serán evidentes tras la referencia a los dibujos adjuntos y tras la descripción detallada. Para este fin, se exponen diversas referencias en el presente documento que describen en más detalle ciertos procedimientos, compuestos y/o composiciones.

### Breve descripción de los dibujos

En las figuras, números de referencia idénticos identifican elementos similares. Los tamaños y posiciones relativas de los elementos en las figuras no están necesariamente dibujados a escala y algunos de estos elementos están arbitrariamente ampliados y posicionados para mejorar la legibilidad de la figura. Además, las formas particulares de los elementos como están dibujados no pretenden expresar ninguna información referente a la forma actual de los elementos particulares, y han sido únicamente seleccionadas para facilitar el reconocimiento en las figuras.

Las Figuras 1A y 1B ilustran la separación limitada entre nucleobases que debe resolverse con el fin de determinar la secuencia de nucleótidos en una diana de ácido nucleico.

Las Figuras 2A a 2D ilustran esquemáticamente varias estructuras representativas de sustratos útiles en la invención.

Las Figuras 3A, 3B y 3C son esquemas que ilustran etapas simplificadas para sintetizar un Xpandómero de un ácido nucleico diana.

La Figura 4 es un simple modelo que ilustra un dispositivo tipo nanoporo de FRET para la secuenciación de un Xpandómero.

La Figura 5 es un gráfico con canales para emisiones de fluorescencia roja, verde y azul, e ilustra cómo pueden descodificarse señales análogas en información digital que se corresponde con la información genética de secuencias codificadas en un Xpandómero. La tabla adjunta (Figura 6) muestra cómo los datos se descodifican. Empleando tres fluoróforos de multi-estado, la secuencia de bases puede leerse con alta resolución en forma digital a partir de una única molécula en tiempo real a medida que el Xpandómero se bobina a través del nanoporo.

La Figura 6 es una tabla de consulta de la que se derivan los datos de la Figura 5.

Las Figuras 7A-E son geles de productos de ligación.

La Figura 8 es una visión general de Xpandómeros oligoméricos.

La Figura 9 es una visión general de Xpandómeros monoméricos.

Las Figuras 10A a 10E representan Xpandómeros de clase I, productos intermedios y precursores en una forma simbólica y gráfica. Estos precursores se llaman Xsondas si son monofosfatos y Xmeros si son trifosfatos.

La Figura 11 es un esquema condensado de un método de síntesis de un Xpandómero por ligación en disolución usando un cebador en horquilla terminado en los extremos y construcciones de sustrato de clase I.

La Figura 12 es un esquema condensado de un método de síntesis de un Xpandómero por ligación en disolución usando un cebador en horquilla de extremo doble y construcciones de sustrato de clase I.

La Figura 13 es un esquema condensado de un método de síntesis de un Xpandómero por ligación sobre un molde inmovilizado, sin cebadores, usando construcciones de sustrato de clase I.

La Figura 14 es un esquema condensado de un método de síntesis de un Xpandómero por ligación escalonada cíclica usando construcciones de sustrato de clase I reversiblemente terminadas sobre moldes hibridados para cebadores inmovilizados.

La Figura 15 es un esquema condensado de un método de síntesis de un Xpandómero por ensamblaje promiscuo y acoplamiento químico usando construcciones de sustrato de clase I sin cebadores.

La Figura 16 es un esquema condensado de un método de síntesis de un Xpandómero por polimerización en



disolución usando un cebador en horquilla y construcciones de sustrato de trifosfato de clase I.  
 La Figura 17 es un esquema condensado de un método de síntesis de un Xpandómero sobre un molde inmovilizado usando construcciones de sustrato de trifosfato de clase I y una polimerasa.  
 Las Figuras 18A a 18E representan un Xpandómero de clase II, producto intermedio de Xpandómero y construcción de sustrato en un lenguaje simbólico y gráfico.  
 Las Figuras 19A a 19E representan un Xpandómero de clase III, producto intermedio de Xpandómero y construcción de sustrato en un lenguaje simbólico y gráfico.  
 La Figura 20 es un esquema condensado de un método de síntesis de un Xpandómero sobre un molde inmovilizado usando construcciones de sustrato de clase II que combinan hibridación y acoplamiento químico sin cebador.  
 La Figura 21 es un esquema condensado de un método de síntesis de un Xpandómero sobre un molde inmovilizado usando un cebador, construcciones de sustrato de clase II y una ligasa.  
 Las Figuras 22A a 22E representan un Xpandómero de clase IV, producto intermedio de Xpandómero y construcción de sustrato en una forma simbólica y gráfica.  
 Las Figuras 23A a 23E representan un Xpandómero de clase V, producto intermedio de Xpandómero y construcción de sustrato en una forma simbólica y gráfica.  
 La Figura 24 es un esquema condensado de un método de síntesis de un Xpandómero por polimerización en disolución usando un cebador de adaptámero y construcciones de sustrato de trifosfato de clase V.  
 La Figura 25 ilustra estructuras de desoxiadenosina (A), desoxicitosina (C), desoxiguanosina (G) y desoxitimidina (T).  
 Las Figuras 26A y 26B ilustran nucleótidos derivatizados con grupos funcionales.  
 Las Figuras 27A y 27B ilustran miembros de sonda que incorporan nucleobases derivatizadas.  
 Las Figuras 28A a 28D ilustran en más detalle sustratos de clase I-IV de la invención, que muestran aquí ejemplos de sitios de escisión de enlaces selectivamente escindibles en el esqueleto de sonda y que indican enlaces del extremo de bucle que conectan los sitios de escisión.  
 Las Figuras 29A a 29D ilustran en más detalle sustratos de clase I-IV de la invención, que muestran aquí ejemplos de sitios de escisión de enlaces selectivamente escindibles en el esqueleto de sonda y que indican enlaces del extremo de bucle que conectan los sitios de escisión.  
 La Figura 30 ilustra un método de ensamblaje de una construcción "sonda-bucle", tal como una Xsonda o un Xmero.  
 La Figura 31 ilustra un método de ensamblaje de una construcción de sustrato de clase I, en la que el bucle contiene construcciones de indicador.  
 Las Figuras 32A a 32C ilustran el uso de PEG como anclaje polimérico.  
 Las Figuras 33A a 33D ilustran poli-lisina como anclaje polimérico y construcciones dendrímeras derivadas de andamiajes de poli-lisina.  
 Las Figuras 34A a 34C ilustran métodos seleccionados de cierre del bucle de anclaje integrados con el ensamblaje de construcción de indicador de segmentos.  
 Las Figuras 35A y 35B ilustran un método de síntesis de un segmento de indicador individual por polimerización aleatorizada de bloques precursores.  
 Las Figuras 36A a 36I ilustran construcciones de indicador.  
 La Figura 37 es una tabla que muestra la composición y métodos químicos para ensamblar construcciones de indicador y sus códigos de indicador correspondientes.  
 Las Figuras 38A a 38F son adaptadores adecuados para su uso en la funcionalización de extremos de ácidos nucleicos diana.  
 La Figura 39 es un casete de adaptador para introducir un grupo funcional ANH terminal sobre un molde de ADNbc.  
 La Figura 40 es un esquema para inmovilizar y preparar un molde para la síntesis de un Xpandómero.  
 Las Figuras 41A a 41E ilustran métodos de estiramiento físico seleccionados.  
 Las Figuras 42A a 42C ilustran métodos de electroestiramiento seleccionados.  
 Las Figuras 43A a 43D ilustran métodos, reactivos y adaptadores para el estiramiento en matrices de gel.  
 Las Figuras 44A a 44C describe la construcción y uso de "marcas de arrastre".  
 La Figura 45 describe un método promiscuo basado en hibridación/ligación para sintetizar un Xpandómero.  
 Las Figuras 46A y B describen nucleobases usadas para llenar huecos.  
 Las Figuras 47A y B describen simulaciones de la aparición de huecos.  
 Las Figuras 48A y B describen simulaciones de la aparición de huecos.  
 La Figura 49 ilustra cómo los huecos se llenan con 2meros y 3meros.  
 Las Figuras 50A y B describen simulaciones del llenado de huecos usando combinaciones de 2meros y 3meros.  
 La Figura 51 ilustra el uso de adyuvantes 2meros y 3meros para romper la estructura secundaria.  
 La Figura 52 describe bases útiles como adyuvantes.  
 La Figura 53 describe sustituciones de nucleótidos usadas para reducir la estructura secundaria.  
 La Figura 54 describe un modelo de detección de nanoporos con transporte de perlas magnéticas.  
 La Figura 55 ilustra un método de detección de nanoporos convencional.  
 La Figura 56 ilustra un método de detección de nanoporos de electrodos transversos.  
 La Figura 57 ilustra un método de detección microscópica.  
 La Figura 58 ilustra la detección por microscopía electrónica.

La Figura 59 ilustra la detección usando microscopía de fuerza atómica.

Las Figuras 60A a 60E representan un Xpandómero de clase VI, producto intermedio de Xpandómero y construcción de sustrato en un lenguaje simbólico y gráfico. Estos precursores se llaman RT-NTP.

5 La Figura 61 es un esquema condensado de un método de síntesis de un Xpandómero sobre un molde inmovilizado usando construcciones de sustrato de trifosfato de clase VI reversiblemente terminadas y una polimerasa.

Las Figuras 62A a 62E representan un Xpandómero de clase VII, producto intermedio de Xpandómero y construcción de sustrato en un lenguaje simbólico y gráfico. Estos precursores se llaman RT-NTP.

10 Las Figuras 63A a 63E representan un Xpandómero de clase VIII, producto intermedio de Xpandómero y construcción de sustrato en un lenguaje simbólico y gráfico. Estos precursores se llaman RT-NTP.

Las Figuras 64A a 64E representan un Xpandómero de clase IX, producto intermedio de Xpandómero y construcción de sustrato en un lenguaje simbólico y gráfico. Estos precursores se llaman RT-NTP.

La Figura 65 es un esquema condensado de un método de síntesis de un Xpandómero sobre un molde inmovilizado usando construcciones de sustrato de trifosfato de clase IX y una polimerasa.

15 Las Figuras 66A a 66E representan un Xpandómero de clase X, producto intermedio de Xpandómero y construcción de sustrato en un lenguaje simbólico y gráfico. Estos precursores se llaman XNTP.

La Figura 67 es un esquema condensado de un método de síntesis de un Xpandómero por polimerización en disolución usando un cebador en horquilla y construcciones de sustrato de trifosfato de clase X.

### Descripción detallada

20 En la siguiente descripción se exponen ciertos detalles específicos con el fin de proporcionar un entendimiento riguroso de diversas realizaciones. Sin embargo, un experto en la materia entenderá que la invención puede ponerse en práctica sin estos detalles. En otros casos, estructuras muy conocidas no se han mostrado o descrito en detalle para evitar descripciones innecesariamente ilegibles de las realizaciones. A menos que el contexto requiera de otro modo, en toda la memoria descriptiva y reivindicaciones que siguen, la palabra “comprender” y variaciones de la misma, tales como “comprende” y “que comprende”, deben interpretarse en un sentido incluyente abierto, es decir, como “que incluye, pero no se limitan a”. Además, los encabezados en el presente documento se proporcionan para comodidad solo y no interpretan el alcance o significado de la invención reivindicada.

30 Referencia en toda esta memoria descriptiva a “una realización” significa que una característica particular, estructura o característica descrita a propósito de la realización se incluye en al menos una realización. Así, las apariciones de la expresión “en una realización” en diversos sitios en toda esta memoria descriptiva no son todas necesariamente con referencia a la misma realización. Además, las características particulares, estructuras o características pueden combinarse de cualquier manera adecuada en una o más realizaciones. Por tanto, como se usa en esta memoria descriptiva y las reivindicaciones adjuntas, las formas en singular “un”, “una”, “el” y “la” incluyen referentes plurales, a menos que el contenido imponga claramente de otro modo. Debe también observarse que el término “o” se emplea generalmente en su sentido que incluye “y/o”, a menos que el contenido imponga claramente de otro modo.

### Definiciones

Como se usa en el presente documento, y a menos que el contexto imponga de otro modo, los siguientes términos tienen los significados que se especifican a continuación.

40 “Nucleobase” es una base heterocíclica tal como adenina, guanina, citosina, timina, uracilo, inosina, xantina, hipoxantina, o un derivado heterocíclico, análogo o tautómero de la misma. Una nucleobase puede existir de forma natural o sintética. Ejemplos no limitantes de nucleobases son adenina, guanina, timina, citosina, uracilo, xantina, hipoxantina, 8-azapurina, purinas sustituidas en la posición 8 con metilo o bromo, 9-oxo-N6-metiladenina, 2-aminoadenina, 7-deazaxantina, 7-deazaguanina, 7-deaza-adenina, N4-etanocitosina, 2,6-diaminopurina, N6-etano-2,6-diaminopurina, 5-metilcitosina, 5-alquil (C3-C6)-citosina, 5-fluorouracilo, 5-bromouracilo, tiouracilo, pseudoisocitosina, 2-hidroxi-5-metil-4-triazolopiridina, isocitosina, isoguanina, inosina, 7,8-dimetilaloxazina, 6-dihidrotimina, 5,6-dihidrouracilo, 4-metil-indol, etenoadenina y las nucleobases que no existen de forma natural descritas en las patentes de EE.UU. N° 5.432.272 y 6.150.510 y solicitudes PCT WO 92/002258, WO 93/10820, WO 94/22892 y WO 94/24144, y Fasman (“Practical Handbook of Biochemistry and Molecular Biology”, pp. 385-394, 1989, CRC Press, Boca Raton, LO).

50 “Residuo de nucleobase” incluye nucleótidos, nucleósidos, fragmentos de los mismos, y moléculas relacionadas que tienen la propiedad de unirse a un nucleótido complementario. Desoxinucleótidos y ribonucleótidos, y sus diversos análogos, se contemplan dentro del alcance de esta definición. Los residuos de nucleobase pueden ser miembros de oligómeros y sondas. “Nucleobase” y “residuo de nucleobase” pueden usarse indistintamente en el presente documento y generalmente son sinónimos, a menos que el contexto imponga de otro modo.

55 “Polinucleótidos”, también llamados ácidos nucleicos, son series covalentemente enlazadas de nucleótidos en los que la posición 3' de la pentosa de un nucleótido se une por un grupo fosfodiéster a la posición 5' del siguiente. El ADN (ácido desoxirribonucleico) y el ARN (ácido ribonucleico) son polinucleótidos que se producen biológicamente en los que los residuos de nucleótidos están enlazados en una secuencia específica por enlaces fosfodiéster. Como se usa en el presente documento, los términos “polinucleótido” o “oligonucleótido” engloban cualquier compuesto de

polímero que tiene un esqueleto lineal de nucleótidos. Los oligonucleótidos, también llamados oligómeros, son polinucleótidos de cadena generalmente más corta.

- 5 “Complementario” generalmente se refiere al duplexado de nucleótidos específicos para formar pares de bases canónicas de Watson-Crick, como se entiende por aquellos expertos en la materia. Sin embargo, complementario como se denomina en el presente documento también incluye apareamiento de bases de análogos de nucleótidos, que incluyen, pero no se limitan a, 2'-desoxiinosina y 5-nitroindol-2'-desoxiribósido, que son capaces de apareamiento de bases universal con nucleótidos A, T, G o C y ácidos nucleicos bloqueados, que potencian la estabilidad térmica de los dúplex. Un experto en la materia reconocerá que la rigurosidad de hibridación es un determinante en el grado de apareamiento o desappareamiento del dúplex formado por hibridación.
- 10 “Ácido nucleico” es un polinucleótido o un oligonucleótido. Una molécula de ácido nucleico puede ser ácido desoxirribonucleico (ADN), ácido ribonucleico (ARN), o una combinación de ambos. Los ácidos nucleicos se denominan generalmente “ácidos nucleicos diana” o “secuencia diana” si son elegidos como diana para la secuenciación. Los ácidos nucleicos pueden ser mezclas o conjuntos de moléculas elegidos como diana para la secuenciación.
- 15 “Sonda” es una hebra corta de residuos de nucleobase, que se refiere generalmente a dos o más residuos de nucleobase contiguos que son generalmente monocatenarios y complementarios a una secuencia diana de un ácido nucleico. Como se integra en “Miembros de sustrato” y “Construcciones de sustrato”, las sondas pueden tener hasta 20 residuos de nucleobase de longitud. Las sondas pueden incluir residuos de nucleobase modificados y enlaces intra-nucleobase modificados en cualquier combinación. Los esqueletos de sondas pueden ligarse juntos por
- 20 cualquiera de varios tipos de enlaces covalentes, que incluyen, pero no se limitan a, enlace éster, fosfodiéster, fosforamida, fosfonato, fosforotioato, fosfortiolato, amida y cualquier combinación de los mismos. La sonda también puede tener enlaces de los extremos 5' y 3' que incluyen, pero no se limitan a, los siguientes restos: monofosfato, trifosfato, hidroxilo, hidrógeno, éster, éter, glicol, amina, amida y tioéster.
- 25 “Hibridación selectiva” se refiere a la unión complementaria específica. Polinucleótidos, oligonucleótidos, sondas, residuos de nucleobase y fragmentos de los mismos se hibridan selectivamente con hebras de ácido nucleico diana, bajo condiciones de hibridación y de lavado que minimizan la unión no específica. Como se conoce en la técnica, pueden usarse condiciones de alta rigurosidad para lograr condiciones de hibridación selectivas que favorecen un apareamiento perfecto. Condiciones para la hibridación tales como concentración de sales, temperatura, detergentes, PEG y agentes neutralizantes de GC tales como betaína pueden variarse para aumentar la rigurosidad
- 30 de la hibridación, es decir, el requisito de apareamientos exactos de C para el apareamiento de bases con G, y A para el apareamiento de bases con T o U, junto con una hebra contigua de un ácido nucleico de dúplex.
- 35 “Síntesis dirigida por molde”, “ensamblaje dirigido por molde”, “hibridación dirigida por molde”, “unión dirigida por molde” y cualquier otro proceso dirigido por molde se refieren a un proceso por el cual los residuos de nucleobase o sondas se unen selectivamente a un ácido nucleico diana complementario, y se incorporan en una hebra hija naciente. Una hebra hija producida por una síntesis dirigida por molde es complementaria a la diana monocatenaria a partir de la que se sintetiza. Debe observarse que la secuencia correspondiente de una hebra diana puede deducirse de la secuencia de su hebra hija, si ésta es conocida. “Polimerización dirigida por molde” y “ligación dirigida por molde” son casos especiales de la síntesis dirigida por molde por las cuales la hebra hija resultante se polimeriza o se liga, respectivamente.
- 40 “Contigua” indica que una secuencia continúa sin interrupción o nucleobase ausente. La secuencia contigua de nucleótidos de la hebra molde se dice que es complementaria a la secuencia contigua de la hebra hija.
- 45 “Sustratos” o “miembros de sustrato” son oligómeros, sondas o residuos de nucleobase que tienen especificidad de unión al molde diana. Los sustratos se combinan generalmente con anclajes para formar construcciones de sustrato. Sustratos de las construcciones de sustrato que forman el esqueleto primario de la hebra hija también son sustratos o miembros de sustrato de la hebra hija.
- 50 “Construcciones de sustrato” son reactivos para la síntesis dirigida por molde de hebras hijas, y se proporcionan generalmente en forma de bibliotecas. Las construcciones de sustrato generalmente contienen un miembro de sustrato para la unión complementaria a un molde diana y tanto un miembro de anclaje como sitios de unión de anclaje a los que puede unirse un anclaje. Las construcciones de sustrato se proporcionan en una variedad de formas adaptadas a la invención. Las construcciones de sustrato incluyen tanto “construcciones de sustrato oligomérico” (también llamadas “construcciones de sustrato de sonda”) como “construcciones de sustrato monomérico” (también llamadas “construcciones de sustrato de nucleobase”).
- 55 “Motivo de subunidad” o “motivo” se refiere a una subunidad de repetición de un esqueleto de polímero, teniendo la subunidad una forma global característica de las subunidades de repetición, pero teniendo también elementos específicos de especie que codifican información genética. Los motivos de residuos de nucleobase complementarios se representan en bibliotecas de construcciones de sustrato según el número de posibles combinaciones de los elementos de nucleobase de unión en la secuencia complementaria básica en cada motivo. Si los elementos de unión de nucleobase son cuatro (por ejemplo, A, C, G y T), el número de posibles motivos de combinaciones de

cuatro elementos es  $4^x$ , en la que x es el número de residuos de nucleobase en el motivo. Sin embargo, otros motivos basados en bases de apareamiento degeneradas, en la sustitución de uracilo por timidina en residuos de ribonucleobase u otros conjuntos de residuos de nucleobase, pueden conducir a mayores bibliotecas (o bibliotecas más pequeñas) de construcciones de sustrato que llevan motivo. Los motivos también se representan por construcciones de indicador específicas de especie, tales como los indicadores que constituyen un anclaje de indicador. Generalmente hay una correlación uno a uno entre el motivo de construcción de indicador que identifica una especie de sustrato particular y la complementariedad de unión y especificidad del motivo.

“Producto intermedio de Xpandómero” es un producto intermedio (también denominado en el presente documento una “hebra hija”) ensamblado de construcciones de sustrato, y se forma por un ensamblaje dirigido por molde de construcciones de sustrato usando un molde de ácido nucleico diana. Opcionalmente, se forman otros enlaces entre construcciones de sustrato confinadas que puede incluir polimerización o ligación de los sustratos, enlaces anclaje a anclaje, o enlaces anclaje a sustrato. El producto intermedio de Xpandómero contiene dos estructuras; concretamente, el Xpandómero limitado y el esqueleto primario. El Xpandómero limitado comprende todos de los anclajes en la hebra hija, pero puede comprender todo, una porción o nada del sustrato según se requiera por el método. El esqueleto primario comprende todos los sustratos confinados. Bajo la etapa de proceso en la que el esqueleto primario se fragmenta o disocia, el Xpandómero limitado ya no está limitado y es el producto de Xpandómero que se extiende a medida que los anclajes se estiran. “Hebra hija de dúplex” se refiere a un producto intermedio de Xpandómero que se hibrida o se duplexa con el molde diana.

“Esqueleto primario” se refiere a un esqueleto contiguo o segmentado de sustratos de la hebra hija. Un esqueleto primario comúnmente encontrado es el esqueleto de ribosil 5'-3' fosfodiéster de un polinucleótido nativo. Sin embargo, el esqueleto primario de una hebra hija puede contener análogos de nucleobases y análogos de oligómeros no enlazados por enlaces fosfodiéster o enlazados por una mezcla de enlaces fosfodiéster y otros enlaces de esqueleto, que incluyen, pero no se limitan a, los siguientes enlaces: enlaces de esqueleto fosforotioato, fosforotiolato, fosfonato, fosforamidato y ácido nucleico peptídico “PNA” que incluyen fosfono-PNA, serina-PNA, hidroxiprolina-PNA, y combinaciones de los mismos. Si la hebra hija está en su forma de dúplex (es decir, hebra hija de dúplex), y los sustratos no están covalentemente unidos entre las subunidades, los sustratos son, sin embargo, contiguos y forman el esqueleto primario de la hebra hija.

“Xpandómero limitado” es un Xpandómero en una configuración antes de haberse expandido. El Xpandómero limitado comprende todos los miembros de anclaje de la hebra hija. Está limitado de expandirse por al menos un enlace o enlace por anclaje que se une al esqueleto primario. Durante el proceso de expansión, el esqueleto primario de la hebra hija se fragmenta o disocia para transformar el Xpandómero limitado en un Xpandómero.

“Esqueleto de Xpandómero limitado” se refiere al esqueleto del Xpandómero limitado. Es un esqueleto covalente sintético co-ensamblado junto con el esqueleto primario en formación de la hebra hija. En algunos casos, ambos esqueletos pueden no ser discretos, sino que ambos pueden tener el mismo sustrato o porciones del sustrato en su composición. El esqueleto de Xpandómero limitado siempre comprende los anclajes, mientras que el esqueleto primario no comprende miembros de anclaje.

“Xpandómero” o “producto de Xpandómero” es una construcción molecular sintética producida por la expansión de un Xpandómero limitado, que él mismo se sintetiza por ensamblaje dirigido por molde de construcciones de sustrato. El Xpandómero está alargado con respecto al molde diana del que se produjo. Está compuesto por una concatenación de subunidades, cada subunidad un motivo, cada motivo un miembro de una biblioteca, que comprende información de secuencias, un anclaje y opcionalmente una porción, o todo del sustrato, todos los cuales se derivan de la construcción de sustrato formativo. El Xpandómero se diseña para expandirse para ser mayor que el molde diana, reduciéndose así la densidad lineal de la información de secuencias del molde diana a lo largo de su longitud. Además, el Xpandómero proporciona opcionalmente una plataforma para aumentar el tamaño y la abundancia de indicadores, que a su vez mejora la señal con respecto al ruido para la detección. Mejor densidad de información lineal y señales más fuertes aumentan la resolución y reducen los requisitos de sensibilidad para detectar y descodificar la secuencia de la hebra molde.

“Enlace selectivamente escindible” se refiere a un enlace que puede romperse bajo condiciones controladas tales como, por ejemplo, condiciones para la escisión selectiva de un enlace fosforotiolato, un enlace fotoescindible, un enlace fosforamida, un enlace 3'-O-B-D-ribofuranosilo-2', un enlace tioéter, un enlace selenoéter, un enlace sulfóxido, un enlace disulfuro, enlace desoxirribosil-5'-3' fosfodiéster, o un enlace ribosil-5'-3' fosfodiéster, además de otros enlaces escindibles conocidos en la técnica. Un enlace selectivamente escindible puede ser un enlace intra-anclaje o entre o dentro de una sonda o un residuo de nucleobase o puede ser el enlace formado por hibridación entre una sonda y una hebra molde. Los enlaces selectivamente escindibles no se limitan a enlaces covalentes, y pueden ser enlaces no covalentes o asociaciones, tales como aquellos basados en enlaces de hidrógeno, enlaces hidrófobos, enlaces iónicos, interacciones de apilamiento de anillos de enlace pi, interacciones de Van der Waals, y similares.

“Resto” es una de las dos o más partes en las que algo puede dividirse, tal como, por ejemplo, las diversas partes de un anclaje, una molécula o una sonda.

“Anclaje” o “miembro de anclaje” se refiere a un polímero o construcción molecular que tiene una dimensión generalmente lineal y con un resto terminal en cada uno de los dos extremos opuestos. Un anclaje está unido a un sustrato con un enlace en al menos un resto terminal para formar una construcción de sustrato. Los restos terminales del anclaje pueden conectarse a enlaces escindibles con el sustrato o enlaces intra-anclaje escindibles que sirven para limitar el anclaje en una “configuración limitada”. Después de sintetizar la hebra hija, cada resto terminal tiene un enlace terminal que se acopla directa o indirectamente a otros anclajes. Los anclajes acoplados comprenden el Xpandómero limitado que comprende además la hebra hija. Los anclajes tienen una “configuración limitada” y una “configuración expandida”. La configuración limitada se encuentra en construcciones de sustrato y en la hebra hija. La configuración limitada del anclaje es el precursor para la configuración expandida, como se encuentra en productos de Xpandómero. La transición de la configuración limitada a la configuración expandida produce la escisión de enlaces selectivamente escindibles que pueden estar dentro del esqueleto primario de la hebra hija o enlaces intra-anclaje. Un anclaje en una configuración limitada también se usa cuando un anclaje se añade para formar la hebra hija después del ensamblaje del “esqueleto primario”. Los anclajes pueden opcionalmente comprender uno o más indicadores o construcciones de indicador a lo largo de su longitud que pueden codificar información de secuencias de sustratos. El anclaje proporciona un medio para expandir la longitud del Xpandómero y así reducir la densidad lineal de información de secuencias.

“Construcciones de anclaje” son anclajes o precursores de anclaje compuestos de uno o más segmentos de anclaje u otros componentes arquitectónicos para ensamblar anclajes tales como construcciones de indicador, o precursores de indicador, que incluyen polímeros, copolímeros de injerto, copolímeros de bloque, ligandos de afinidad, oligómeros, haptenos, aptámeros, dendrímeros, grupos de enlace o grupo de unión por afinidad (por ejemplo, biotina).

“Elemento de anclaje” o “segmento de anclaje” es un polímero que tiene una dimensión generalmente lineal con dos extremos terminales, en los que los extremos forman enlaces terminales para concatenar los elementos de anclaje. Los elementos de anclaje pueden ser segmentos de construcciones de anclaje. Tales polímeros pueden incluir, pero no se limitan a: polietilenglicoles, poliglicoles, polipiridinas, poliisocianuros, poliisocianatos, poli(triarilmetil)metacrilatos, polialdehídos, polipirrolinonas, poliureas, fosfodiésteres de poliglicol, poliácridatos, polimetacrilatos, poliácridamidas, polivinilésteres, poliestirenos, poliamidas, poliuretanos, policarbonatos, polibutiratos, polibutadienos, polibutirolactonas, polipirrolidinonas, polivinilfosfonatos, poliácetamidas, polisacáridos, polihialuronatos, poliamidas, poliimididas, poliésteres, polietilenos, polipropilenos, poliestirenos, policarbonatos, politereftalatos, polisilanos, poliuretanos, poliéteres, poliaminoácidos, poliglicinas, poliprolinas, polilisina N-sustituida, polipéptidos, péptidos sustituidos con N en la cadena lateral, poli-glicina N-sustituida, peptoides, péptidos sustituidos con carboxilo en la cadena lateral, homopéptidos, oligonucleótidos, oligonucleótidos de ácido ribonucleico, oligonucleótidos de ácido desoxinucleico, oligonucleótidos modificados para prevenir el apareamiento de bases de Watson-Crick, análogos de oligonucleótidos, ácido policitidílico, ácido poliadenílico, ácido poliuridílico, politimidina, polifosfato, polinucleótidos, polirribonucleótidos, fosfodiésteres de polietilenglicol, análogos de péptido-polinucleótido, análogos de treosil-polinucleótido, análogos de glicol-polinucleótido, análogos de morfolino-polinucleótido, análogos de oligómeros de nucleótidos bloqueados, análogos de polipéptidos, polímeros ramificados, polímeros en peine, polímeros en estrella, polímeros dendríticos, copolímeros aleatorios, en gradiente y de bloque, polímeros aniónicos, polímeros catiónicos, polímeros que forman tallo-bucles, segmentos rígidos y segmentos flexibles.

“Ácido nucleico peptídico” o “PNA” es un análogo de ácido nucleico que tiene residuos de nucleobase adecuados para la hibridación con un ácido nucleico, pero con un esqueleto que comprende aminoácidos o derivados o análogos de los mismos.

“Ácido nucleico fosfono-peptídico” o “pPNA” es un ácido nucleico peptídico en el que el esqueleto comprende análogos de aminoácidos, tales como N-(2-hidroxietyl)fosfonglicina o N-(2-aminoetyl)fosfonglicina, y los enlaces entre unidades de nucleobases son mediante enlaces fosfonoéster o fosfonoamida.

“Ácido nucleico de serina” o “SerNA” es un ácido nucleico peptídico en el que el esqueleto comprende residuos de serina. Tales residuos pueden enlazarse mediante enlaces amida o éster.

“Ácido nucleico de hidroxiprolina” o “HypNA” es un ácido nucleico peptídico en el que el esqueleto comprende residuos de 4-hidroxiprolina. Tales residuos pueden enlazarse mediante enlaces amida o éster.

“Elemento indicador” es un elemento de señalización, complejo molecular, compuesto, molécula o átomo que también comprende una “característica de detección de indicador” asociada. Otros elementos indicadores incluyen, pero no se limitan a, donante o aceptor resonante de FRET, colorante, punto cuántico, perla, dendrímero, fluoróforo de conversión en exceso, partícula de imán, dispersante de electrones (por ejemplo, boro), masa, perla de oro, resonancia magnética, grupo ionizable, grupo polar, grupo hidrófobo. Todavía otros son marcas fluorescentes, tales como, pero no se limitan a, bromuro de etidio, SYBR Green, Texas Red, naranja de acridina, pireno, 4-nitro-1,8-naftalimida, TOTO-1, YOYO-1, cianina 3 (Cy3), cianina 5 (Cy5), ficoeritrina, ficocianina, alofocianina, FITC, rodamina, 5(6)-carboxifluoresceína, proteínas fluorescentes, DOXYL (N-oxil-4,4-dimetiloxazolidina), PROXYL (N-oxil-2,2,5,5-tetrametilpirrolidina), TEMPO (N-oxil-2,2,6,6-tetrametilpiperidina), dinitrofenilo, acridinas, cumarinas, Cy3 y Cy5 (Biological Detection Systems, Inc.), eritrosina, ácido cumárico, umbeliferona, Texas Red-rodamina, tetrametilrodamina, Rox, 7-nitrobenzo-1-oxa-1-diazol (NBD), oxazol, tiazol, pireno, fluoresceína o lantamidas; también

radioisótopos (tales como  $^{33}\text{P}$ ,  $^3\text{H}$ ,  $^{14}\text{C}$ ,  $^{35}\text{S}$ ,  $^{125}\text{I}$ ,  $^{32}\text{P}$  o  $^{131}\text{I}$ ), etidio, europio, rutenio y samario u otros radioisótopos; o marcas de masa, tales como, por ejemplo, pirimidinas modificadas en la posición C5 o purinas modificadas en la posición N7, en las que los grupos modificadores de la masa pueden ser, por ejemplos, halógeno, éter o poliéter, alquilo, éster o poliéster, o del tipo general XR, en la que X es un grupo de enlace y R es un grupo modificador de la masa, marcas quimioluminiscentes, etiquetas de espín, enzimas (tales como peroxidases, fosfatasas alcalinas, beta-galactosidasas y oxidasas), fragmentos de anticuerpos, y ligandos de afinidad (tales como un oligómero, hapteno y aptámero). La asociación del elemento indicador con el anclaje puede ser covalente o no covalente, y directa o indirecta. Asociaciones covalentes representativas incluyen enlaces de conector y de conector cero. Se incluyen enlaces al esqueleto de anclaje o a un elemento unido al anclaje tal como un dendrímero o cadena lateral. Enlaces no covalentes representativos incluyen enlaces de hidrógeno, enlaces hidrófobos, enlaces iónicos, apilamiento de anillos de enlace pi, interacciones de Van der Waals y similares. Los ligandos, por ejemplo, están asociados por unión de afinidad específica con sitios de unión sobre el elemento indicador. La asociación directa puede tener lugar en el momento de la síntesis del anclaje, después de la síntesis del anclaje, y antes o después de la síntesis de Xpandómeros.

Un “indicador” está compuesto por uno o más elementos indicadores. Los indicadores incluyen lo que se conoce como “etiquetas” y “marcas”. La sonda o residuo de nucleobase del Xpandómero puede considerarse un indicador. Los indicadores sirven para analizar la información genética del ácido nucleico diana.

“Construcción de indicador” comprende uno o más indicadores que pueden producir una señal(es) detectable(s), en la que la(s) señal(es) detectable(s) generalmente contienen información de secuencias. Esta información de señal se llama el “código de indicador” y posteriormente se descodifica en datos de secuencia genética. Una construcción de indicador también puede comprender segmentos de anclaje u otros componentes arquitectónicos que incluyen polímeros, copolímeros de injerto, copolímeros de bloque, ligandos de afinidad, oligómeros, haptenos, aptámeros, dendrímeros, grupos de enlace o grupo de unión por afinidad (por ejemplo, biotina).

“Característica de detección de indicador” referida a la “señal” describe todos los posibles elementos, propiedades o características medibles o detectables usados para comunicar la información genética de secuencias de un indicador directa o indirectamente a un dispositivo de medición. Éstos incluyen, pero no se limitan a, fluorescencia, fluorescencia de multi-longitud de onda, extinción de fluorescencia del espectro de emisión, FRET, emisión, absorbancia, reflectancia, emisión de colorante, emisión de puntos cuánticos, imagen de perlas, imagen de complejos moleculares, susceptibilidad magnética, dispersión de electrones, masa iónica, resonancia magnética, dimensión del complejo molecular, impedancia del complejo molecular, carga molecular, dipolo inducido, impedancia, masa molecular, estado cuántico, capacidad de carga, estado de espín magnético, polaridad inducible, decaimiento nuclear, resonancia o complementariedad.

“Código de indicador” es la información genética de una señal medida de una construcción de indicador. El código de indicador se descodifica para proporcionar datos de información genética específica de secuencia.

“Xsonda” es una construcción de sustrato oligomérico expansible. Cada Xsonda tiene un miembro de sonda y un miembro de anclaje. El miembro de anclaje generalmente tiene una o más construcciones de indicador. Las Xsondas con modificaciones de 5'-monofosfato son compatibles con métodos basados en ligación enzimática para la síntesis de Xpandómeros. Las Xsondas con modificaciones de conector en 5' y 3' son compatibles con los métodos basados en ligación química para la síntesis de Xpandómeros.

“Xmero” es una construcción de sustrato oligomérico expansible. Cada Xmero tiene un miembro de sustrato oligomérico y un miembro de anclaje, teniendo el miembro de anclaje generalmente una o más construcciones de indicador. Los Xmeros son 5'-trifosfatos compatibles con métodos basados en polimerasa para sintetizar Xpandómeros.

“RT-NTP” es una construcción de sustrato de nucleótido expansible modificada con 5' trifosfato (“sustrato monomérico”) compatible con la polimerización enzimática dependiente del molde. Una RT-NTP tiene un desoxirribonucleótido trifosfato modificado (“DNTP”), ribonucleótido trifosfato (“RNTP”), o un sustrato de análogo funcionalmente equivalente, denominados conjuntamente el sustrato de nucleótido trifosfato (“NTPS”). Una RT-NTP tiene dos componentes funcionales distintos; concretamente, una nucleobase 5'-trifosfato y un anclaje o precursor de anclaje. Después de la formación de la hebra hija, el anclaje se une entre cada nucleótido en las posiciones que permiten la expansión de RT controlada. En una clase de RT-NTP (por ejemplo, la clase IX), el anclaje se une después de la polimerización de RT-NTP. En algunos casos, la RT-NTP tiene un terminador de extremos reversibles y un anclaje que se reticula selectivamente directamente con anclajes adyacentes. Cada anclaje puede estar únicamente codificado con indicadores que identifican específicamente el nucleótido al que está anclado.

“XNTP” es un sustrato de nucleótido modificado con 5' trifosfato expansible compatible con polimerización enzimática dependiente del molde. Un XNTP tiene dos componentes funcionales distintos; concretamente, una nucleobase 5'-trifosfato y un anclaje o precursor de anclaje que está unido dentro de cada nucleótido en las posiciones que permiten la expansión de RE controlada por escisión intra-nucleótido.

- “Procesivo” se refiere a un proceso de acoplamiento de sustratos que es generalmente continuo y avanza con direccionalidad. Aunque no está ligado por teoría, tanto las ligasas como las polimerasas, por ejemplo, presentan comportamiento procesivo si los sustratos se añaden a una hebra hija naciente incrementalmente sin interrupción. Las etapas de hibridación y ligación, o hibridación y polimerización, no se consideran etapas independientes si el efecto neto es el crecimiento procesivo de la hebra hija naciente. Algunos, pero no todos los procesos dependientes de cebador, son procesivos.
- “Promiscuo” se refiere a un proceso de acoplamiento de sustratos que avanza a partir de múltiples puntos sobre un molde de una vez, y no es dependiente de cebador, e indica que la expansión de cadena se produce en paralelo (simultáneamente) a partir de más de un punto de origen.
- “Extensión de una única base” se refiere a un proceso de escalonado cíclico en el que los sustratos monoméricos se añaden uno por uno. Generalmente, la reacción de acoplamiento se limita de avanzar más allá de una extensión de un único sustrato en una etapa cualquiera por el uso de grupos de bloqueo reversibles.
- “Extensión de una única sonda” se refiere a un proceso de escalonado cíclico en el que sustratos oligoméricos se añaden uno por uno. Generalmente, la reacción de acoplamiento se limita de avanzar más allá de una extensión de un único sustrato en una etapa cualquiera por el uso de grupos de bloqueo reversibles.
- “Se corresponde con” o “correspondiente” se usa aquí en referencia a una secuencia monocatenaria contigua de una sonda, oligonucleótido, análogo de oligonucleótido, o hebra hija que es complementaria, y así “se corresponde con” toda o una porción de una secuencia de ácidos nucleicos diana. La secuencia complementaria de una sonda puede decirse que se corresponde con su diana. A menos que se establezca de otro modo, tanto la secuencia complementaria de la sonda como la secuencia complementaria de la diana son secuencias individualmente contiguas.
- “Resistente a nucleasa” se refiere a es un enlace que es resistente a una enzima nucleasa en condiciones en las que un enlace fosfodiéster de ADN o ARN generalmente se escindiría. Las enzimas nucleasas incluyen, pero no se limitan a, DNasa I, exonucleasa III, nucleasa de judía mungo, RNasa I y RNasa H. Un experto en este campo puede evaluar fácilmente la resistencia a nucleasas relativa de un enlace dado.
- “Ligasa” es una enzima generalmente para unir nucleótidos 3'-OH 5'-monofosfato, oligómeros, y sus análogos. Las ligasas incluyen, pero no se limitan a, ligasas dependientes de NAD<sup>+</sup> que incluyen ARNt ligasa, ADN ligasa Taq, ADN ligasa de *Thermus filiformis*, ADN ligasa de *Escherichia coli*, ADN ligasa Tth, ADN ligasa de *Thermus scotoductus*, ligasa termoestable, ADN ligasa termoestable Ampligase, ligasa tipo VanC, ADN ligasa 9°N, ADN ligasa Tsp, y ligasas novedosas descubiertas por bioprospección. Las ligasas también incluyen, pero no se limitan a, ligasas dependientes de ATP que incluyen ARN ligasa T4, ADN ligasa T4, ADN ligasa T7, ADN ligasa Pfu, ADN ligasa I, ADN ligasa III, ADN ligasa IV, y ligasas novedosas descubiertas por bioprospección. Estas ligasas incluyen isoformas no mutantes, mutantes y variantes genéticamente manipuladas.
- “Polimerasa” es una enzima generalmente para unir nucleótidos de 3'-OH 5'-trifosfato, oligómeros, y sus análogos. Las polimerasas incluyen, pero no se limitan a, ADN polimerasas dependientes de ADN, ARN polimerasas dependientes de ADN, ADN polimerasas dependientes de ARN, ADN polimerasas dependientes de ARN, ADN polimerasa T7, ADN polimerasa T3, ADN polimerasa T4, ARN polimerasa T7, ARN polimerasa T3, ARN polimerasa SP6, ADN polimerasa I, fragmento de Klenow, ADN polimerasa de *Thermophilus aquaticus*, ADN polimerasa Tth, ADN polimerasa VentR<sup>®</sup> (New England Biolabs), ADN polimerasa Deep VentR<sup>®</sup> (New England Biolabs), fragmento grande de ADN polimerasa Bst, fragmento de Stoeffel, ADN polimerasa 9°N, ADN polimerasa 9°N, ADN polimerasa Pfu, ADN polimerasa Tfl, ADN polimerasa Tth, polimerasa RepliPhi Phi29, ADN polimerasa Tli, ADN polimerasa beta de eucariota, telomerasa, polimerasa Terminator<sup>™</sup> (New England Biolabs), ADN polimerasa KOD HiFi<sup>™</sup> (Novagen), ADN polimerasa KOD1, Q-beta replicasa, transferasa terminal, transcriptasa inversa AMV, transcriptasa inversa M-MLV, transcriptasa inversa Phi6, transcriptasa inversa del VIH-1, polimerasas novedosas descubiertas por bioprospección, y polimerasas citadas en los documentos US 2007/0048748, US 6.329.178, US 6.602.695 y US 6.395.524. Estas polimerasas incluyen isoformas no mutantes, mutantes, y variantes genéticamente manipuladas.
- “Codificar” o “analizar” son verbos con referencia a transferir de un formato a otro, y se refiere a transferir la información genética de la secuencia base del molde diana en una disposición de indicadores.
- “Extragenético” se refiere a cualquier estructura en la hebra hija que no es parte del esqueleto primario; por ejemplo, un informador extragenético no es la propia nucleobase que se encuentra en el esqueleto primario.
- “Hetero-copolímero” es un material formado combinando unidades diferentes (por ejemplo, especies de subunidad de monómero) en cadenas de un “copolímero”. Los hetero-copolímeros se forman a partir de construcciones de “subunidad” discretas. Una “subunidad” es una región de un polímero compuesta por un motivo bien definido, en la que cada motivo es una especie y lleva información genética. El término hetero-copolímero también se usa en el presente documento para describir un polímero en el que todos los bloques son bloques construidos de motivos de repetición, teniendo cada motivo elementos específicos de especie. La hebra hija y el Xpandómero son ambos hetero-copolímeros, por lo que cada motivo de subunidad codifica 1 o más bases de la secuencia del molde diana y la secuencia diana entera se define adicionalmente con la secuencia de motivos.

“Soporte sólido” es un material sólido que tiene una superficie para la unión de moléculas, compuestos, células, u otras entidades. La superficie de un soporte sólido puede ser plana o no plana. Un soporte sólido puede ser poroso o no poroso. Un soporte sólido puede ser un chip o matriz que comprende una superficie, y que puede comprender vidrio, silicio, nailon, polímeros, plásticos, cerámicas, o metales. Un soporte sólido también puede ser una membrana, tal como una membrana de nailon, nitrocelulosa, o polimérica, o una placa o disco, y comprender vidrio, cerámica, metales, o plásticos, tales como, por ejemplo, poliestireno, polipropileno, policarbonato o polialómero. Un soporte sólido también puede ser una perla, resina o partícula de cualquier forma. Tales partículas o perlas pueden comprender cualquier material adecuado, tal como vidrio o cerámica, y/o uno o más polímeros, tales como, por ejemplo, nailon, politetrafluoroetileno, TEFLON™, poliestireno, poliacrilamida, Sepaharose, agarosa, celulosa, derivados de celulosa, o dextrano, y/o pueden comprender metales, particularmente metales paramagnéticos, tales como hierro.

“Bloquear reversiblemente” o “terminador” se refiere a un grupo químico que cuando se une a un segundo grupo químico en un resto previene que el segundo grupo químico entre en reacciones químicas particulares. Se conoce un amplio intervalo de grupos protectores en la química orgánica y bioorgánica sintética que son adecuados para grupos químicos particulares y son compatibles con procesos químicos particulares, que significa que protegerán a grupos particulares durante aquellos procesos y pueden ser posteriormente eliminados o modificados (véase, por ejemplo, Metzker et al. *Nucleic Acids Res.*, 22(20): 4259, 1994).

“Conector” es una molécula o resto que une dos moléculas o restos, y proporciona separación entre las dos moléculas o restos de forma que sean capaces de funcionar en su forma prevista. Por ejemplo, un conector puede comprender una cadena de hidrocarburo de diamina que se une covalentemente mediante un grupo reactivo sobre un extremo a una molécula de análogo de oligonucleótido y mediante un grupo reactivo sobre otro extremo a un soporte sólido, tal como, por ejemplo, una superficie de perla. El acoplamiento de conectores a nucleótidos y construcciones de sustrato de interés puede llevarse a cabo mediante el uso de reactivos de acoplamiento que se conocen en la técnica (véase, por ejemplo, Efimov et al., *Nucleic Acids Res.* 27: 4416-4426, 1999). Los métodos de derivatización y acoplamiento de moléculas orgánicas son muy conocidos en las técnicas de la química orgánica e bioorgánica. Un conector también puede ser escindible o reversible.

#### Visión general

En términos generales, se describen métodos y dispositivos y productos correspondientes para replicar ácidos nucleicos diana de una única molécula. Tales métodos utilizan “Xpandómeros” que permiten la secuenciación del ácido nucleico diana con elevado rendimiento y exactitud. Un Xpandómero codifica (analiza) los datos de la secuencia de nucleótidos del ácido nucleico diana en un formato linealmente expandido, mejorándose así la resolución espacial, opcionalmente con amplificación de la intensidad de señales. Estos procesos se denominan en el presente documento “secuenciación por expansión” o “SBX”.

Como se muestra en la Figura 1A, los ácidos nucleicos de dúplex nativo tienen una densidad de datos lineal extremadamente compacta; aproximadamente una separación de centro a centro de 3,4 Å entre bases apiladas secuenciales (2) de cada hebra de la hélice doble (1), y, por tanto, es tremendamente difícil obtener directamente imágenes o secuencia con cualquier exactitud y velocidad. Cuando la forma bicatenaria se desnaturaliza para formar polinucleótidos monocatenarios (3,4), las distancias de separación base a base resultantes son similares, pero el problema se agrava por dominios de la estructura secundaria.

Como se muestra en la Figura 1B, el Xpandómero (5), aquí ilustrado como una concatenación de oligómeros cortos (6,7) mantenidos juntos por anclajes T extragenéticos (8,9), es una sustitución sintética o “suplente” para la diana de ácido nucleico que va a secuenciarse. Las bases complementarias al molde se incorporan en el Xpandómero, pero los anclajes regularmente separados sirven para aumentar la distancia entre los oligómeros cortos (aquí cada uno mostrado con cuatro nucleobases representadas por círculos). El Xpandómero se prepara por un proceso en el que un producto intermedio de dúplex sintético se forma primero replicando una hebra molde. La hebra hija es única porque tiene tanto un esqueleto lineal formado por los oligómeros como un esqueleto de Xpandómero limitado comprendido de anclajes plegados. Los anclajes se abren entonces o se “expanden” para transformar el producto en una cadena de anclajes alargados. Figurativamente, la hebra hija puede visualizarse como que tiene dos esqueletos superpuestos: uno lineal (esqueleto primario) y el otro con pliegues de “acordeón” (Xpandómero limitado). La escisión selectiva de enlaces en la hebra hija permite que los pliegues de acordeón se expandan para producir un producto de Xpandómero. Este proceso se explicará más abajo en más detalle, pero debe observarse que la elección de cuatro nucleobases por oligómero y datos del anclaje como se muestra en la Figura 1B es con el fin de ilustración solo, y de ninguna forma debe interpretarse para limitar la invención.

La distancia de separación “D” entre oligómeros vecinos en el Xpandómero es ahora una variante dependiente del proceso y se determina por la longitud del anclaje T. Como se mostrará, la longitud del anclaje T se diseña en las construcciones de sustrato, los elementos estructurales a partir de los cuales se prepara el Xpandómero. La distancia de separación D puede seleccionarse para ser, por ejemplo, superior a 0,5 nm, o superior a 2 nm, o superior a 5 nm, o superior a 10 nm, o superior a 50 nm. A medida que aumenta la distancia de separación, el proceso de discriminación o “resolución” de los oligómeros individuales se vuelve progresivamente cada vez más fácil. Esto también sería cierto si, en lugar de oligómeros, nucleobases individuales de otra especie de Xpandómero



se ensartaran juntas sobre una cadena de anclajes.

Refiriéndose de nuevo a la Figura 1A, el ADN nativo se replica por un proceso de replicación semi-conservativa; cada nueva molécula de ADN es un “dúplex” de una hebra molde (3) y una hebra hija nativa (4). La información de secuencias se pasa del molde a la hebra hija nativa por un proceso de “síntesis dirigida por molde” que preserva la información genética inherente en la secuencia de los pares de bases. La hebra hija nativa recibe a su vez un molde para una hebra hija nativa de la siguiente generación, etc. Los Xpandómeros se forman por un proceso de síntesis dirigida por molde similar, que puede ser un proceso enzimático o de acoplamiento químico. Sin embargo, a diferencia del ADN nativo, una vez formados, los Xpandómeros no pueden replicarse por un proceso biológico de replicación semi-conservativa y no son adecuados para la amplificación por procesos tales como PCR. El producto de Xpandómero se diseña para limitar la estructura secundaria no deseada.

Las Figuras 2A a 2D muestran sustratos de Xpandómero de clase I (20,21,22,23) representativos. Éstos son los elementos estructurales a partir de los cuales se sintetizan los Xpandómeros. Otros sustratos de Xpandómero (diez clases se desvelan en el presente documento) se tratan en las secciones posteriores. Las construcciones de sustrato de Xpandómero mostradas aquí tienen dos componentes funcionales; concretamente, un miembro de sonda (10) y un miembro de “anclaje” (11) en una configuración de bucle. El bucle forma el anclaje “T” alargado del producto final. Únicamente por comodidad en la explicación, el miembro de sonda se representa de nuevo con cuatro residuos de nucleobase (14,15,16,17) como se muestra en la Figura 2B.

Estas construcciones de sustrato pueden estar modificadas en los extremos con grupos R, por ejemplo, como un 5'-monofosfato, 3'-OH adecuado para su uso con una ligasa (denominado en el presente documento “Xsonda”) o como un 5'-trifosfato, 3'-OH adecuado para su uso con una polimerasa (denominado en el presente documento un “Xmero”). Otros grupos R pueden ser de uso en diversos protocolos. En el primer ejemplo mostrado en la Figura 3B, los presentes inventores presentan el uso de Xsondas en la síntesis de un Xpandómero a partir de una hebra molde de un ácido nucleico diana por un proceso dependiente de ligasa.

Los cuatro residuos de nucleobase (14,15,16,17) del miembro de sonda (10) están seleccionados para ser complementarios a una secuencia contigua de cuatro nucleótidos del molde. Cada “sonda” está así diseñada para hibridarse con el molde en una secuencia complementaria de cuatro nucleótidos. Suministrando una biblioteca de muchas de tales secuencias de sonda puede formarse una réplica complementaria contigua del molde. Esta hebra hija se llama un “producto intermedio de Xpandómero”. Los productos intermedios de Xpandómero tienen formas de dúplex o monocatenarias.

El bucle de anclaje está unido al miembro de sonda (10) en el segundo y tercer residuos de nucleobase (15,16). El segundo y tercer residuos de nucleobase (15,16) también están unidos entre sí por un “enlace selectivamente escindible” (25) representado por una “V”. La escisión de este enlace escindible permite que el bucle de anclaje se expanda. Puede decirse que el anclaje linealizado “conecta” el sitio de enlace selectivamente escindible del esqueleto de polinucleótido primario de una hebra hija. La escisión de estos enlaces rompe el esqueleto primario y forma el Xpandómero más largo.

La escisión selectiva de los enlaces selectivamente escindibles (25) puede hacerse en una variedad de formas que incluyen, pero no se limitan a, escisión química de enlaces fosforotiolato, digestión por ribonucleasa de enlaces ribosil 5'-3' fosfodiéster, escisión de enlaces fotoescindibles, y similares, como se trata en mayor detalle más adelante.

La construcción de sustrato (20) mostrada en la Figura 2A tiene un único segmento de anclaje, representado aquí por una elipse (26), para la unión de elementos indicadores. Este segmento está flanqueado con segmentos de anclaje espaciadores (12,13), todos los cuales forman conjuntamente la construcción de anclaje. A este respecto pueden usarse de uno a muchos dendrímero(s), polímero(s), polímero(s) ramificado(s) o combinaciones, por ejemplo, para construir el segmento de anclaje. Para la construcción de sustrato (21) de la Figura 2B, la construcción de anclaje está compuesta por tres segmentos de anclaje para la unión de elementos indicadores (27,28,29), cada uno de los cuales está flanqueado con un segmento de anclaje espaciador. La combinación de elementos indicadores forma conjuntamente una “construcción de indicador” para producir un código de indicador digital único (para la identificación de secuencias de sonda). Estos elementos indicadores incluyen, pero no se limitan a, fluoróforos, marcas de FRET, perlas, ligandos, aptámeros, péptidos, haptenos, oligómeros, polinucleótidos, dendrímeros, estructuras de tallo-bucle, marcas de afinidad, marcas de masa y similares. El bucle de anclaje (11) de la construcción de sustrato (22) en la Figura 2C está “desnudo”. La información genética codificada en esta construcción no está codificada sobre el anclaje, pero está asociada a la sonda (10), por ejemplo, en forma de nucleótidos marcados. La construcción de sustrato (23) de la Figura 2D ilustra el principio general: como se indica por el asterisco (\*), la información de secuencias de la sonda está codificada o “se analiza” en la construcción de sustrato de una forma modificada más fácilmente detectada en un protocolo de secuenciación. Debido a que los datos de secuencia se resuelven físicamente mejor después de la escisión del enlace selectivamente escindible (25) para formar el polímero de Xpandómero linealmente alargado, el asterisco (\*) representa cualquier forma de información genética codificada para la que esto es un beneficio. El elemento o elementos bioinformáticos (\*) de la construcción de sustrato, sea cual sea su forma, pueden ser detectables directamente o pueden ser precursores a los que los elementos detectables se añaden en una etapa de marcado post-ensamblaje. En algunos casos, la

información genética está codificada en una propiedad molecular de la propia construcción de sustrato, por ejemplo, una marca de masa de multi-estado. En otros casos, la información genética está codificada por uno o más fluoróforos de los pares de donante:aceptor de FRET, o un código de barras nanomolecular, o un ligando o combinación de ligandos, o en forma de alguna otra técnica de marcado extraída de la materia. Se tratarán diversas realizaciones más abajo en más detalle.

El anclaje generalmente cumple varias funciones: (1) enlazarse secuencialmente, directa o indirectamente, a anclajes adyacentes que forman el producto intermedio de Xpandómero; (2) estirarse y expandirse para formar una cadena alargada de anclajes tras la escisión de enlaces seleccionados en el esqueleto primario o dentro del anclaje (véase la Figura 1 B); y/o (3) proporcionar una construcción molecular para incorporar elementos indicadores, también llamados “etiquetas” o “marcas”, que codifican la información de secuencias del residuo de nucleobase de su sustrato asociado. El anclaje puede diseñarse para optimizar la función codificante ajustando separaciones espaciales, abundancia, densidad informacional e intensidad de señales de sus elementos indicadores constituyentes. Un amplio intervalo de propiedades indicadoras es útil para amplificar la intensidad de señales de la información genética codificada dentro de la construcción de sustrato. La bibliografía dirigida a indicadores, códigos de barras moleculares, unión por afinidad, marcado molecular y otros elementos indicadores, es muy conocida para un experto en este campo.

Puede observarse que si cada sustrato de una construcción de sustrato contiene  $x$  nucleobases, entonces una biblioteca que representa todas las posibles combinaciones secuenciales de  $x$  nucleobases contendría  $4^x$  sondas (cuando se seleccionan las nucleobases de A, T, C o G). Si se usan otras bases puede necesitarse menos o más combinaciones. Estas bibliotecas de sustrato se diseñan de manera que cada construcción de sustrato contenga (1) una sonda (o al menos un residuo de nucleobase) complementaria a una cualquiera de las posibles secuencias diana del ácido nucleico que va a secuenciarse y (2) una construcción de indicador única que codifica la identidad de la secuencia diana a la que esa sonda (o nucleobase) particular es complementaria. Una biblioteca de sondas que contienen dos nucleobases tendría 16 miembros únicos; una biblioteca de sondas que contiene tres nucleobases tendría 64 miembros únicos, etc. Una biblioteca representativa tendría las cuatro nucleobases individuales, pero configuradas para acomodar un medio de anclaje.

La síntesis de un Xpandómero se ilustra en las Figuras 3A a 3C. El sustrato representado aquí es una Xsonda y el método puede describirse como la hibridación con ligación procesiva dependiente de cebador en disolución libre.

Pueden adaptarse muchos protocolos de biología molecular muy conocidos, tales como los protocolos para fragmentar el ADN diana y ligar adaptadores de extremo, para su uso en métodos de secuenciación y se usan aquí para preparar el ADN diana (30) para la secuenciación. Aquí, los presentes inventores ilustran, en amplios términos, lo que sería familiar para aquellos expertos en la materia, procesos para pulir los extremos de los fragmentos y ligación de extremos romos de adaptadores (31,32) diseñados para su uso con cebadores de secuenciación. Estas acciones se muestran en la etapa I de la Figura 3A. En las etapas II y III, el ácido nucleico diana se desnaturaliza y se hibrida con cebadores (33) adecuados complementarios a los adaptadores.

En la Figura 3B, la hebra molde cebada de la etapa III se pone en contacto con una biblioteca de construcciones de sustrato (36) y ligasa (L), y en la etapa IV las condiciones se ajustan para favorecer la hibridación, seguido de ligación en un 3'-OH libre de un dúplex de cebador-molde. Opcionalmente, en la etapa V, la ligasa se disocia, y en las etapas VI y VII puede reconocerse que el proceso de hibridación y ligación produce la extensión por adición acumulada de sustratos (37,38) al extremo del cebador. Aunque el cebado puede producirse a partir de adaptadores en ambos extremos de un molde monocatenario, el crecimiento de una hebra hija de Xpandómero naciente se muestra aquí que avanza a partir de un único cebador, únicamente para simplicidad. La extensión de la hebra hija se representa en las etapas VI y VII, que se repiten continuamente (incrementalmente, sin interrupción). Estas reacciones se producen en disolución libre y avanzan hasta que se haya sintetizado una cantidad suficiente de producto. En la etapa VIII se muestra la formación de un producto intermedio de Xpandómero (39) completado.

Longitudes relativamente largas de la secuencia de nucleótidos contigua pueden replicarse eficazmente de este modo para formar productos intermedios de Xpandómeros. Puede observarse que las longitudes de lectura continuas (“cántigos”) correspondientes a fragmentos de hebra molde largos pueden lograrse con esta tecnología. Será evidente para un experto en la materia que billones de estas reacciones de SBX de una única molécula pueden hacerse simultáneamente en un proceso discontinuo eficaz en un único tubo. Posteriormente, pueden secuenciarse los productos de secuenciación aleatoria de estas síntesis.

En la Figura 3C se representan las siguientes etapas del proceso de SBX. La etapa IX muestra la desnaturalización del producto intermedio de Xpandómero de dúplex, seguido de la escisión de enlaces selectivamente escindibles en el esqueleto, con los enlaces selectivamente escindibles diseñados de manera que los bucles de anclaje “se abran”, formando el producto de Xpandómero linealmente alargado (34). Tal escisión selectiva puede lograrse por cualquier número de técnicas conocidas para un experto en la materia, que incluyen, pero no se limitan a, escisión de fosfortiolato con cationes metálicos como se desvela por Mag et al. (“Synthesis and selective cleavage of an oligodeoxynucleotide containing a bridged internucleotide 5'-phosphorothioate linkage”, *Nucleic Acids Research* 19(7):1437-1441, 1991), escisión catalizada por ácido de fosforamidato como se desvela por Mag et al. (“Synthesis and selective cleavage of oligodeoxyribonucleotides containing non-chiral internucleotide phosphoramidate linkages”,

Nucleic Acids Research 17(15): 5973-5988, 1989), escisión por nucleasa selectiva de enlaces fosfodiéster como se desvela por Gut et al. ("A novel procedure for efficient genotyping of single nucleotide polymorphisms", Nucleic Acids Research 28(5): E13, 2000) y por separado por Eckstein et al. ("Inhibition of restriction endonuclease hydrolysis by phosphorothioate-containing DNA", Nucleic Acids Research, 25;17(22): 9495, 1989) y escisión selectiva del esqueleto de fosfodiéster modificado por conector fotoescindible como se desvela por Sauer et al. ("MALDI mass spectrometry analysis of single nucleotide polymorphisms by photocleavage and charge-tagging", Nucleic Acids Research 31,11 e63, 2003), Vallone et al. ("Genotyping SNPs using a UV-photocleavable oligonucleotide in MALDI-TOF MS", Methods Mol. Bio. 297:169-78, 2005) y Ordoukhanian et al. ("Design and synthesis of a versatile photocleavable DNA building block, application to phototriggered hybridization", J. Am. Chem. Soc. 117, 9570-9571, 1995).

Los refinamientos del proceso básico, tales como las etapas de lavado y ajuste de las condiciones de rigurosidad, están perfectamente dentro de la experiencia de un biólogo molecular experimentado. Variantes a este proceso incluyen, por ejemplo, inmovilización y análisis de las hebras diana, estiramiento y otras técnicas para reducir la estructura secundaria durante la síntesis del Xpandómero, marcado post-expansión, funcionalización de extremos y alternativas a la ligasa para enlazar los sustratos se tratarán en los materiales que siguen.

La síntesis de Xpandómeros se hace para facilitar la detección y secuenciación de ácidos nucleicos, y es aplicable a ácidos nucleicos de todos los tipos. El proceso es un método de "expansión" o "alargamiento" de la longitud de elementos del esqueleto (o subunidades) que codifican la información de secuencias (expandida con respecto a las pequeñas distancias nucleótido a nucleótido de ácidos nucleicos nativos) y opcionalmente también sirve para aumentar la intensidad de señales (con respecto a las señales de baja intensidad casi indistinguibles observadas para los nucleótidos nativos). Como tales, los elementos indicadores incorporados en el esqueleto sintético expandido de un Xpandómero pueden detectarse y procesarse usando una variedad de métodos de detección, que incluyen métodos de detección muy conocidos en la técnica (por ejemplo, una cámara de CCD, un microscopio de fuerza atómica, o un espectrómetro de masas regulado), además de por métodos tales como una matriz de sensor de nanoporos masivamente paralela, o una combinación de métodos. Las técnicas de detección se seleccionan basándose en la señal óptima con respecto al ruido, rendimiento, coste y factores similares.

Volviendo a la Figura 4, se muestra un simple modelo de una tecnología de detección; concretamente, un nanoporo (40) con donante de FRET (42) en una membrana (44), que se excita por luz de longitud de onda  $\lambda_1$ . Como el producto de Xpandómero (41) se alarga y es transportado a través del nanoporo (40) en la dirección de la flecha (45), se detectan explosiones en serie de emisión de longitud de onda  $\lambda_2$  de fluoróforos excitados, en la proximidad del poro. Las longitudes de onda de emisión ( $\lambda_2$ ) se separan temporalmente en función de la longitud del anclaje y la velocidad del Xpandómero que pasa a través del nanoporo. Capturando estas señales análogas y procesándolas digitalmente, la información de secuencias puede leerse directamente a partir del Xpandómero. Debe observarse que en este método de detección, el nanoporo y la membrana pueden tener muchas trayectorias a través de las cuales el Xpandómero puede translocarse. La detección de FRET requiere que haya al menos un donante de FRET excitado a lo largo de cada trayectoria. A diferencia, un nanoporo basado en contador de Coulter puede solo tener orificio de translocalización adicional a costa de la señal con respecto al ruido.

En la técnica de detección basada en nanoporos de la Figura 4, que representa cadenas de Xsondas de la estructura mostrada en la Figura 2B, las construcciones de anclaje contienen construcciones de indicador multi-elemento, como se indica por los miembros de indicador tipo caja (27,28,29) dispuestos a lo largo del anclaje. Se desvela tecnología de secuenciación de nanoporos relevante, por ejemplo, por Branton et al. en la patente de EE.UU. N.º. 6.627.067 y por Lee et al. (Lee, JW y A Meller. 2007. Rapid sequencing by direct nanoscale reading of nucleotide bases in individual DNA chains. En "New High Throughput Technologies for DNA Sequencing and Genomics, 2", Elsevier).

La Figura 5 demuestra cómo las construcciones de indicador multi-elemento, que aquí comprenden fluoróforos aceptores de FRET, aparecen a un detector posicionado en la puerta de FRET. Puede observarse en esta representación multi-canal de emisiones que señales análogas se generan a intervalos generalmente regulares de tiempo y pueden analizarse como un tipo de código digital (aquí llamado un código de indicador y, para este ejemplo, una ID de Xsonda) que revela la identidad y orden de las subunidades de Xsonda y así la secuencia genética del Xpandómero ilustrado. Pueden usarse diversas combinaciones de indicadores para crear una biblioteca de códigos de indicador que codifican secuencialmente cualquier combinación de 4 bases de A, T, G o C de la Xsonda descrita. En este ejemplo, se usan combinaciones de tres fluoróforos para producir veintidós códigos de indicador. De esta forma, se observa que la secuencia ACTG va seguida de GCCG; seguido de AAAT. Líneas de puntos verticalmente dispuestas separan los datos de fluorimetría y las subunidades correspondientes del Xpandómero (mostrado esquemáticamente). Un algoritmo interpretativo inmediatamente debajo de la representación muestra cómo las señales análogas regularmente separadas se transforman en una secuencia genética legible.

La construcción de sustrato de Xsonda ilustrada en la Figura 5, que usa una construcción de anclaje multi-elemento compuesta de tres segmentos marcados con indicador, cada uno de los cuales está flanqueado con segmentos de anclaje espaciadores, para codificar la identidad de secuencias del sustrato, se elabora adicionalmente. El primer segmento de anclaje es el código de indicador N.º 1 (lectura de izquierda a derecha), y se lee como una señal alta en el canal rojo. El segundo segmento de anclaje es el código de indicador N.º 9, y se lee como una señal verde alta y

una señal roja baja. El tercer segmento de anclaje es el código de indicador N° 8, y se lee como una señal azul baja y una señal roja baja. El código de indicador N° 1 se usa como un reloj o señal de sincronización; el código de indicador N° 9 codifica el primer resto de sonda "AC"; el código de indicador número N° 8 codifica el segundo resto "TG" de la sonda. Tomados conjuntamente, el código de indicador secuencial de "1-9-8" se corresponde con una especie particular de la Xsonda (Xsonda ID 117), que a su vez se corresponde por diseño con el fragmento de secuencia "ACTG". Tres ID de Xsonda codifican la secuencia contigua entera mostrada en la representación, "ACTGGCCGAAAT". Las emisiones de fluoróforo, la tabla para descodificar códigos de indicador y los fragmentos de secuencia, y las representaciones físicas correspondientes de las construcciones de indicador, están separadas por las líneas de puntos de las figuras según subunidades estructurales del Xpandómero de manera que pueda fácilmente observarse cómo la información de secuencias se descodifica y digitaliza.

La Figura 6 es una tabla de marcas de fluoróforo a partir de las cuales se preparó el ejemplo de la Figura 5. Esto ilustra más generalmente el uso de combinaciones de códigos de indicador multi-estado para analizar la información en forma de señales detectables. Se usan fluoróforos que tienen veintidós posibles estados de emisión para formar las construcciones de indicador de este ejemplo. Tres marcas de fluoróforo por oligómero son más que adecuadas para codificar todas las posibles combinaciones 4meras de A, T, C y G. Aumentando la longitud del anclaje, mejora la resolución entre las emisiones de marcas de fluoróforo, beneficiándose la exactitud de la etapa de detección, un principio que es generalmente aplicable.

Indicadores útiles con construcciones de anclaje de este tipo son de muchos tipos, no simplemente fluoróforos, y pueden medirse usando un ancho intervalo correspondiente de tecnologías de alto rendimiento y detección precisa, tecnologías que podrían no ser de otro modo útiles para secuenciar ácidos nucleicos nativos debido a resolución limitada. Los métodos de detección del estado de la técnica masivamente paralelos, tales como matrices de sensores de nanoporos, se facilitan por las características más medibles de Xpandómeros. Las ineficiencias en los procesos de detección de la secuenciación pueden reducirse pre-purificando lotes de Xpandómeros para eliminar productos de reacción incompletos o cortos. Se proporcionan métodos para la modificación de extremos de Xpandómeros sintetizados que puede utilizarse para tanto la purificación como bien como medio para facilitar la presentación de Xpandómeros al detector. Además, el proceso de lectura no está limitado por la limitación al encapuchado, desencapuchado, extensión de nucleótidos, marcado, u otros métodos de procesamiento simultáneos.

La Figura 7A describe un molde de dúplex parcial diseñado con un nucleótido protuberante en 5' de veinte bases para demostrar la ligación procesiva de sustratos y la ligación dirigida por molde iniciada por cebador en disolución libre. La Figura 7B es una fotografía de un gel que demuestra la ligación de los sustratos usando el formato de cebador-molde descrito en la Figura 7A. Para este ejemplo, los sustratos oligoméricos de dinucleótido de la secuencia 5' fosfato CA 3' se hibridan con el molde en presencia de un cebador y ADN ligasa T4. El nucleótido protuberante del extremo sin duplexar (si lo hay) se digiere entonces con nucleasa y los productos de ligación se separan sobre un gel de 20 % de acrilamida. La ligación produce polímeros producto que contienen subunidades demostrablemente ligadas. Como se indica por el patrón de bandeó, las reacciones positivas para ligasa que corren en los carriles 1, 3, 5, 7 y 9, que contienen moldes progresivamente más largos (4, 8, 12, 16 y 20 bases, respectivamente), demuestran claramente la ligación secuencial de sustratos 2meros (elevadas longitudes de dúplex protegidos con exonucleasa). Los carriles 2, 4, 6, 8 y 10 son controles negativos que no contienen ligasa y muestran digestión con exonucleasa completa de productos sin ligar.

La Figura 7C es un segundo gel que muestra ligación dirigida por molde de sustratos. Se ensayaron cuatro moldes de control positivo progresivamente más largos, de nuevo duplexados con un cebador de extensión (4, 8, 12 y 16 bases de molde, respectivamente). De nuevo, los sustratos oligoméricos de dinucleótido de la secuencia 5' fosfato CA 3' se hibridan con el molde en presencia de un cebador y ADN ligasa T4. El nucleótido protuberante del extremo sin duplexar (si lo hay) se digiere entonces con nucleasa y los productos de ligación se separan sobre un gel de 20 % de acrilamida. Se observa que los sustratos oligoméricos (de nuevo 2meros) se ligan al molde en los carriles 1, 2, 3 y 4, pero no en los carriles 5 y 6, en los que las hebras molde contienen un desapareamiento con el dinucleótido 5' (fosfato) CA 3' (molde del carril 5 - 5' CGCG 3'; molde del carril 6 - 5' GGGG 3').

Los resultados del gel mostrados en la Figura 7D demuestran múltiples ligaciones dirigidas por molde de una sonda de tetranucleótido modificada con bis(amino). Los modificadores de amino alifáticas fueron del enlace y composición descritos en la Figura 26. Para este ejemplo, un sustrato oligomérico de tetranucleótido de la secuencia 5' (fosfato) C (amino)A (amino)C A 3' se hibridó con una variedad de moldes complementarios progresivamente más largos (duplexados con un cebador de extensión) en presencia de un cebador y ADN ligasa T4. El nucleótido protuberante del extremo sin duplexar (si lo hay) fue entonces digerido por nucleasa y los productos de ligación se separaron sobre un gel de acrilamida al 20 %. La ligación produce polímeros producto que contienen subunidades demostrablemente ligadas. Los carriles 1 y 2 representan controles de tamaño 16mero y 20mero. Los carriles 3, 4, 5, 6, 7, 8 y 9 muestran productos de ligación para moldes complementarios progresivamente más largos (4, 6, 8, 12, 16, 18 y 20 bases de molde, respectivamente). Se observan múltiples ligaciones de tetrámero para reacciones de moldes más largas (carriles 6-9). El carril 10 muestra inhibición de ligasa esencialmente completa debido al desapareamiento de molde-sonda (molde - 5' CGCG 3').

Los resultados del gel mostrados en la Figura 7E demuestran múltiples ligaciones dirigidas por molde de una sonda de hexanucleótido modificada con bis(amino). Los modificadores de amino alifáticas fueron del enlace y composición descritos en la Figura 26. Para este ejemplo, un sustrato oligomérico de hexanucleótido de la secuencia 5' (fosfato) C A (amino)C (amino)A C A 3' se hibridó con una variedad de moldes complementarios progresivamente más largos (duplexados con un cebador de extensión) en presencia de un cebador y ADN ligasa T4. El nucleótido protuberante del extremo si duplexar (si lo hay) fue entonces digerido por nucleasa y los productos de ligación se separaron sobre un gel de acrilamida al 20 %. La ligación produce polímeros producto que contienen subunidades demostrablemente ligadas. Los carriles 1 y 2 representan controles de tamaño 16mero y 20mero. Los carriles 3, 4, 5, 6, 7, 8 y 9 muestran productos de ligación para moldes complementarios progresivamente más largos (4, 6, 8, 12, 16, 18 y 20 bases de molde, respectivamente). Se observan múltiples ligaciones de tetrámero para reacciones de moldes más largas (carriles 5-9). El carril 10 muestra inhibición de ligasa casi completa debido al desapareamiento de molde-sonda (molde - 5' CGCGCG 3').

Los sustratos incluyen tanto miembros de sonda (es decir, oligómeros como miembros de unión específicos de molde para ensamblar el producto intermedio de Xpandómero) como monómeros (es decir, miembros de nucleobase individuales como elementos de unión específicos de molde). Los presentes inventores llaman a los primeros sustratos "tipo sonda" y a los segundos sustratos "tipo monómero". Como se ilustra en la Figura 8, los Xpandómeros tipo sonda tienen cinco subgéneros básicos, mientras que la Figura 9 ilustra cinco subgéneros básicos de Xpandómeros tipo monómero. Las tablas de las Figuras 8 y 9 incluyen tres columnas: la primera que describe construcciones de sustrato, la segunda productos intermedios de Xpandómeros y la tercera los productos de Xpandómero característicos del subgénero (por fila). Las tablas se proporcionan aquí como una visión general, con métodos que hacen y usan lo mismo que se ha desvelado en mayor detalle en el presente documento más adelante. En las Figuras 8 y 9, "P" se refiere a una miembro de sonda, "T" a un miembro de anclaje (o anclaje de bucle o precursor del brazo de anclaje), "N" a un monómero (una nucleobase individual o residuo de nucleobase) y "R" a un grupo terminal.

Más específicamente, en la tabla de la Figura 8 se usa la siguiente nomenclatura:

P es un miembro de sustrato de sonda y está compuesto por  $P^1$ - $P^2$ , en la que  $P^1$  es un primer resto de sonda y  $P^2$  es un segundo resto de sonda;  
 T es un anclaje;  
 Los corchetes indican una subunidad de la hebra hija, en la que cada subunidad es un motivo de subunidad que tiene un miembro de sonda específico de especie, adicionalmente en la que dichos miembros de sonda de dichos motivos de subunidad son secuencialmente complementarios a la secuencia de nucleótidos contigua correspondiente de la hebra molde, indicada aquí  $P^1$  -  $P^2$ , y forman un esqueleto primario del producto intermedio de Xpandómero, y en la que los miembros de anclaje, opcionalmente en combinación con los restos de sonda, forman un esqueleto de Xpandómero limitado. La escisión de uno o más enlaces selectivamente escindibles dentro del producto intermedio de Xpandómero permite la expansión de las subunidades para producir un producto de Xpandómero, cuyas subunidades también se indican con corchetes;  
 $\alpha$  indica una especie de motivo de subunidad seleccionado de una biblioteca de motivos de subunidad;  
 s es un primer grupo conector unido a un primer extremo o resto de un miembro de sonda o anclaje; bajo condiciones controladas, s es capaz de reaccionar selectivamente con, directamente o mediante reticuladores, el grupo conector  $\delta$  de un extremo confinado de una subunidad adyacente para formar enlaces covalentes o equivalentemente duraderos;  
 $\delta$  es un segundo grupo conector unido a un primer extremo o resto de un miembro de sonda o anclaje; bajo condiciones controladas,  $\delta$  es capaz de reaccionar selectivamente con, directamente o mediante reticuladores, el grupo conector  $\varepsilon$  de un extremo confinado de una subunidad adyacente para formar enlaces covalentes o equivalentemente duraderos;  
 $\chi$  representa un enlace con una subunidad adyacente y es el enlace producto de la reacción de los grupos conectores  $\delta$  y  $\varepsilon$ ;  
 $\sim$  indica un enlace selectivamente escindible, que puede ser igual o diferente cuando están presentes múltiples enlaces selectivamente escindibles;  
 $R^1$  incluye, pero no se limita a, hidroxilo, hidrógeno, trifosfato, monofosfato, éster, éter, glicol, amina, amida y tioéster;  
 $R^2$  incluye, pero no se limita a, hidroxilo, hidrógeno, trifosfato, monofosfato, éster, éter, glicol, amina, amida y tioéster; y  
 $k$  indica la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $k = 1, 2, \dots$  a  $m$ , en la que  $m > 3$ , y generalmente  $m > 20$ , y preferentemente  $m > 50$ , y más preferencialmente  $m > 1000$ .

Más específicamente, y en el contexto de la tabla de la Figura 9, se usa la siguiente nomenclatura:

N es un residuo de nucleobase;  
 T es un anclaje;  
 Los corchetes indican una subunidad de la hebra hija, en la que cada subunidad es un motivo de subunidad que tiene un residuo de nucleobase específico de especie, adicionalmente en la que dichos residuos de nucleobase de dichos motivos de subunidad son secuencialmente complementarios a la secuencia de

nucleótidos contigua correspondiente de la hebra molde, indicada aquí N, y forman un esqueleto primario del producto intermedio de Xpandómero, y en la que los miembros de anclaje, opcionalmente en combinación con los residuos de nucleobase, forman un esqueleto de Xpandómero limitado. La escisión de uno o más enlaces selectivamente escindibles dentro del producto intermedio de Xpandómero permite la expansión de las subunidades para producir un producto de Xpandómero, cuyas subunidades también se indican con corchetes;

5  $n^1$  es una primera porción de un residuo de nucleobase;  
 $n^2$  es una segunda porción de un residuo de nucleobase;  
 $\varepsilon$  es un primer grupo conector unido a un primer extremo o resto de un miembro de sonda o anclaje; bajo condiciones controladas,  $\varepsilon$  es capaz de reaccionar selectivamente con, directamente o mediante reticuladores,

10 el grupo conector  $\delta$  de un extremo confinado de una subunidad adyacente para formar enlaces covalentes o equivalentemente duraderos;  
 $\delta$  es un segundo grupo conector unido a un primer extremo o resto de un miembro de sonda o anclaje; bajo condiciones controladas,  $\delta$  es capaz de reaccionar selectivamente con, directamente o mediante reticuladores,

15 el grupo conector  $\varepsilon$  de un extremo confinado de una subunidad adyacente para formar enlaces covalentes o equivalentemente duraderos;  
 $\chi$  representa un enlace con una subunidad adyacente y es el enlace producto de la reacción de grupos de enlace  $\delta$  y  $\varepsilon$ ;  
 $\chi^1$  es el enlace producto de la reacción de grupos de enlace  $\delta^1$  y  $\varepsilon^1$ ;  
 $\chi^2$  es el enlace producto de la reacción de grupos de enlace  $\delta^2$  y  $\varepsilon^2$ ;

20  $\sim$  indica un enlace selectivamente escindible, que puede ser igual o diferente cuando están presentes múltiples enlaces selectivamente escindibles;  
 $R^1$  incluye, pero no se limita a, hidroxilo, hidrógeno, trifosfato, monofosfato, éster, éter, glicol, amina, amida y tioéster;  
 $R^2$  incluye, pero no se limita a, hidroxilo, hidrógeno, trifosfato, monofosfato, éster, éter, glicol, amina, amida y

25 tioéster; y  
 $k$  indica la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $k = 1, 2, \dots$  a  $m$ , en la que  $m > 10$ , y generalmente  $m > 50$ , y normalmente  $m > 500$  o  $> 5.000$ .

#### Construcciones oligoméricas

30 Los precursores y construcciones de Xpandómero pueden dividirse en dos categorías basándose en el sustrato (oligomérico o monomérico) usado para el ensamblaje dirigido por molde. La estructura, precursores y métodos de síntesis del Xpandómero para aquellos basados en los sustratos de oligómero se tratan a continuación.

Las construcciones de sustrato son precursores reactivos para el Xpandómero y generalmente tienen un miembro de anclaje y un sustrato. El sustrato tratado aquí es un sustrato de oligómero o sonda, generalmente constituido de una pluralidad de residuos de nucleobase. Generando bibliotecas de tipo combinatorias de dos a veinte residuos de

35 nucleobase por sonda, generalmente 2 a 10 y normalmente 2, 3, 4, 5 o 6 residuos de nucleobase por sonda, se generan bibliotecas de sonda útiles como reactivos en la síntesis de precursores de Xpandómeros (construcciones de sustrato).

La sonda se describe generalmente como que tiene dos restos de sonda,  $P^1$  y  $P^2$ . Estos restos de sonda están generalmente representados en las figuras como dinucleótidos, pero en general  $P^1$  y  $P^2$  tienen cada uno al menos un residuo de nucleobase. En el ejemplo de una sonda con dos residuos de nucleobase, los restos de sonda  $P^1$  y  $P^2$  serían residuos de nucleobase individuales. El número de residuos de nucleobase para cada uno se elige, apropiadamente, para el método de síntesis del Xpandómero y puede no ser igual en  $P^1$  y  $P^2$ .

40

Para las construcciones de sustrato en las que se usan los grupos conectores  $\varepsilon$  y  $\delta$  para crear enlaces inter-subunidad, un amplio intervalo de químicas comercialmente disponibles adecuadas (Pierce, Thermo Fisher Scientific, EE.UU.) puede adaptarse para este fin. Químicas de conectores comunes incluyen, por ejemplo, ésteres de NHS con aminas, maleimidias con sulfhidrilos, imidoésteres con aminas, EDC con carboxilos para reacciones con aminas, disulfuros de piridilo con sulfhidrilos, y similares. Otras realizaciones implican el uso de grupos funcionales como hidrazida (HZ) y 4-formilbenzoato (4FB) que pueden entonces hacerse reaccionar adicionalmente para formar

45 enlaces. Más específicamente, están ampliamente disponibles (Pierce) un amplio intervalo de reticuladores (hetero- y homo-bifuncionales) que incluyen, pero no se limitan a, sulfo-SMCC (4-[N-maleimidometil]ciclohexano-1-carboxilato de sulfosuccinimidilo), SIA (yodoacetato de N-succinimidilo), sulfo-EMCS (éster de [N-e-maleimidocaproiloxi]sulfosuccinimida), sulfo-GMBS (éster de N-[g-maleimidobutirilo]sulfosuccinimida), AMAS (éster de N-(a-maleimidoacetoxi)succinimida), BMPS (N-EMCA ácido (N-e-maleimidocaproico)-éster de [ $\beta$ -maleimidopropiloxi]succinimida), EDC (clorhidrato de 1-etil-3-[3-dimetilaminopropil]carbodiimida), SANPAH (N-succinimidil-6-[4'-azido-2'-nitrofenilamino]hexanoato), SADP (N-succinimidil(4-azidofenil)-1,3'-ditiopropionato), PMPI (N-[p-Maleimidofenil]isocianato, BMPH (hidrazida del ácido N-[ $\beta$ -maleimidopropiónico], sal del ácido trifluoroacético), EMCH (hidrazida del ácido [N-e-maleimidocaproico], sal del ácido trifluoroacético), SANH (4-hidrazinonicotinato-acetonahidrazona de succinimidilo), SHTH (clorhidrato de 4-hidrazidotereftalato de succinimidilo) y C6-SFB (4-formilbenzoato de C6-succinimidilo). Por tanto, el método desvelado por Letsinger et al. ("Phosphorothioate oligonucleotides having modified internucleoside linkages", patente de EE.UU. Nº 6.242.589) puede adaptarse para

60 formar enlaces fosforotiolato.

Además, están ampliamente disponibles químicas de protección/desprotección bien establecidas para restos de conector comunes (Benoiton, "Chemistry of Peptide Synthesis", CRC Press, 2005). La protección de amino incluye, pero no se limita a, carbamato de 9-fluorenilmetilo (Fmoc-NRR'), carbamato de t-butilo (Boc-NRR'), carbamato de bencilo (Z-NRR', Cbz-NRR'), acetamida-trifluoroacetamida, ftalimida, bencilamina (Bn-NRR'), trifenilmetilamina (Tr-NRR') y bencilidenamina-p-toluenosulfonamida (Ts-NRR'). La protección de carboxilo incluye, pero no se limita a, éster metílico, éster t-butilico, éster bencilico, éster s-t-butilico y 2-alquil-1,3-oxazolona. Carbonilo incluyen, pero no se limitan a, 1,3-dioxano de dimetilacetal, y N,N-dimetilhidrazona de 1,3-ditiano. Protección de hidroxilo incluye, pero no se limita a, éter metoximetílico (MOM-OR), éter tetrahidropiránico (THP-OR), éter t-butilico, éter alílico, éter bencilico (Bn-OR), éter t-butildimetilsilílico (TBDMS-OR), éter t-butildifenilsilílico (TBDPS-OR), éster del ácido acético, éster del ácido piválico y éster del ácido benzoico.

Aunque el anclaje se representa frecuentemente como una construcción de indicador con tres grupos indicadores, diversas configuraciones de indicador pueden presentarse sobre el anclaje, y pueden comprender indicadores individuales que identifican constituyentes de sonda, indicadores individuales que identifican especies de sonda, códigos de barras moleculares que identifican especies de sondas, o el anclaje puede ser polímero desnudo (que no tiene indicadores). En el caso del polímero desnudo, los indicadores pueden ser la propia sonda, o pueden estar sobre un segundo anclaje unido a la sonda. En algunos casos, uno o más precursores de indicador están presentes sobre el anclaje, y los indicadores se unen por afinidad o se unen covalentemente tras el ensamblaje del producto de Xpandómero.

Como se trata anteriormente, la Figura 8 proporciona una visión general de construcciones oligoméricas de la invención, distinguiéndose cinco clases: clases I, II, III, IV, y V. Estas clases se aplican a tanto Xsondas como a Xmeros. Cada clase se tratará individualmente a continuación.

#### Construcciones oligoméricas de clase I

Volviendo a la Figura 10, las construcciones oligoméricas de clase I se describen en más detalle. Las Figuras 10A a 10C emplean una notación adaptada para mostrar estas moléculas como sustratos y como productos de hetero-copolímero del proceso de SBX. Las figuras se leen de izquierda a derecha, mostrando primero la construcción de sonda-sustrato (precursor oligomérico de Xpandómero), luego la hebra hija del dúplex intermedia en el centro, y a la derecha el producto de Xpandómero preparado para la secuenciación.

Como se muestra en la Figura 10A, una construcción de sustrato de clase I tiene un miembro de sonda oligomérico ( $-P^1 \sim P^2-$ ) (100) y un miembro de anclaje, T (99). El anclaje está unido por dos enlaces terminales (108,109) a restos de sonda  $P^1$  y  $P^2$ . Estas limitaciones previenen que el anclaje se alargue o expanda y así en una configuración limitada. Bajo ensamblaje dirigido por molde, los sustratos forman un dúplex con el molde diana de forma que los sustratos estén confinados.

$R^1$  y  $R^2$  son grupos terminales configurados según convenga para el protocolo de síntesis en el que se usa la construcción de sustrato. Por ejemplo,  $R^1 = 5'$ -fosfato y  $R^2 = 3'$ -OH, encontrarían uso en un protocolo de ligación, y  $R^1 = 5'$ -trifosfato y  $R^2 = 3'$ -OH para un protocolo de polimerasa I. Opcionalmente,  $R^2$  puede configurarse con un grupo de bloqueo reversible para la adición cíclica de un único sustrato. Alternativamente,  $R^1$  y  $R^2$  pueden configurarse con grupos conectores terminales para el acoplamiento químico o con grupos no conectores para un protocolo de solo hibridación.  $R^1$  y  $R^2$  pueden ser del tipo general XR, en la que X es un grupo de enlace y R es un grupo funcional.

La tilde ( $\sim$ ) en la Figura 10A y 10B indica un enlace selectivamente escindible que separa dos restos del miembro de sonda. Los enlaces terminales del anclaje están unidos a los dos restos del miembro de sonda que se separan por el enlace selectivamente escindible. El anclaje enlaza el primer resto de sonda con el segundo resto de sonda, formando un bucle que conecta el enlace selectivamente escindible. Cuando el miembro de sonda está intacto (sin escindir), el miembro de sonda puede unirse con alta fidelidad a la secuencia molde y el anclaje forma un bucle en la "configuración limitada". Cuando este enlace se escinde, el bucle de anclaje puede abrirse y el anclaje está en la "configuración expandida".

Las construcciones de sustrato son reactivos usados para el ensamblaje dependiente de molde de una hebra hija, que es una composición intermedia para producir Xpandómeros. La Figura 10B muestra la hebra hija de dúplex, un hetero-copolímero con subunidades de repetición (mostradas en corchetes). Se muestran el esqueleto primario de la hebra hija ( $-P^1 \sim P^2-$ ) y la hebra molde diana ( $-P^1 - P^2-$ ) como un dúplex (95). Cada subunidad de la hebra hija es un motivo de repetición compuesto de un miembro de sonda y un miembro de anclaje, T (99), el miembro de anclaje en configuración limitada. Los motivos tienen variabilidad específica de especie, indicada aquí por el superíndice " $\alpha$ ". Cada subunidad particular en la hebra hija se selecciona de una biblioteca de motivos por un proceso dirigido por molde y su sonda se une a una secuencia correspondiente de nucleótidos complementarios sobre la hebra molde. De esta forma, la secuencia de residuos de nucleobase de las sondas forma una copia complementaria contigua de la hebra molde diana.

La hebra hija está compuesta por un precursor de Xpandómero llamado el "Xpandómero limitado" que está adicionalmente compuesto de anclajes en la "configuración limitada". Cuando los anclajes (99) se convierten en su

“configuración expandida”, el Xpandómero limitado se convierte en el producto de Xpandómero.

Puede observarse que la hebra hija tiene dos esqueletos, un “esqueleto primario” y el esqueleto del “Xpandómero limitado”. El esqueleto primario está compuesto por sustratos de sonda contiguamente confinados. El “esqueleto de Xpandómero limitado” evita el enlace selectivamente escindible entre los restos de sonda  $P^1$  y  $P^2$  y se forma por restos de esqueleto enlazados, siendo cada resto de esqueleto un enlace lineal de  $P^1$  al anclaje a  $P^2$ , y en el que  $P^2$  puede enlazarse adicionalmente con el  $P^1$  del siguiente resto de esqueleto. Puede observarse que el esqueleto de Xpandómero limitado conecta o forma un bucle sobre los enlaces selectivamente escindibles del esqueleto primario, y seguirá covalentemente intacto cuando estos enlaces selectivamente escindibles se escindan y el esqueleto primario se fragmente.

La Figura 10C es una representación del producto de Xpandómero de clase I después de la disociación de la hebra molde y después de la escisión de los enlaces selectivamente escindibles del esqueleto primario. Métodos para la disociación de la hebra molde incluyen desnaturalización por calor, o digestión selectiva con una nucleasa, o degradación química. La hebra del producto de Xpandómero contiene una pluralidad de subunidades  $\kappa$ , en las que  $\kappa$  indica la subunidad  $\kappa^1$  en una cadena de  $m$  subunidades que constituyen la hebra hija, en la que  $\kappa = 1, 2, 3$  a  $m$ , en la que  $m > 3$ , y generalmente  $m > 20$ , y preferentemente  $m > 50$ , y más preferencialmente  $m > 1000$ . Cada subunidad está formada de un anclaje (99) y restos de sonda  $P^1$  y  $P^2$ . El miembro de anclaje T, ahora en “configuración expandida”, se observa estirado a su longitud entre los restos de sonda escindidos  $P^1$  y  $P^2$ , que siguen ligados covalentemente a las subunidades adyacentes. Cada subunidad, un motivo de subunidad  $\alpha$ , contiene información genética específica de especie establecida por el ensamblaje dirigido por molde del producto intermedio de Xpandómero (hebra hija).

La Figura 10D muestra la construcción de sustrato de la Figura 10A como modelo molecular, en la que el miembro de sonda (100) se representa arbitrariamente con cuatro residuos de nucleobase (101,102,103,104), dos de los cuales (102,103) se unen al anclaje (99) por enlaces terminales (108,109). Entre los dos enlaces terminales del anclaje está un enlace selectivamente escindible, mostrado como “V” (110) en el miembro de sonda (100). Este enlace une restos de sonda  $P^1$  y  $P^2$  referidos en la Figura 10A. El bucle de anclaje mostrado aquí tiene tres indicadores (105,106,107), que también pueden ser específicos de especies de motivos.

La Figura 10E muestra el Xpandómero producto después de la escisión de los enlaces selectivamente escindibles en el sustrato. La escisión produce la expansión del Xpandómero limitado y se indica por “E” (flechas oscuras). Los residuos (110a,110b) del enlace selectivamente escindible marcan el acontecimiento de escisión. La subunidad está indicada por líneas de puntos que agrupan verticalmente la subunidad de repetición, como se representa por los corchetes en la Figura 10C adjunta.

En el producto de Xpandómero (Figura 10E) el esqueleto primario está ahora fragmentado y no está covalentemente intacto debido a que los miembros de sonda se han escindido, separando cada  $P^1$  (92) y  $P^2$  (94). Durante el proceso de escisión, el Xpandómero limitado se libera para convertirse en el producto de Xpandómero. El Xpandómero incluye cada subunidad concatenada en la secuencia. Ligados dentro de cada subunidad están el resto de sonda  $P^1$ , el anclaje y resto de sonda  $P^2$ . Los miembros de anclaje (99) del Xpandómero que estuvieron anteriormente en la configuración limitada están ahora en la configuración expandida, sirviendo así para estirar linealmente la información de secuencias de la diana molde. La expansión de los anclajes reduce la densidad lineal de la información de secuencias a lo largo del Xpandómero y proporciona una plataforma para aumentar el tamaño y la abundancia de indicadores, que a su vez mejora la señal con respecto al ruido para la detección y descodificación de la secuencia molde.

La Figura 11 representa un esquema condensado de un método de preparación de una realización de un Xpandómero de clase I; el método ilustra la preparación y uso de sustratos y productos mostrados en las Figuras 10D y 10E. El método puede realizarse en disolución libre y se describe usando una ligasa (L) para acoplar covalentemente Xsondas confinadas. Los métodos de alivio de la estructura secundaria en el molde se tratan en una sección posterior. Condiciones adaptadas para la hibridación y ligación son muy conocidas en la técnica y tales condiciones pueden ser fácilmente optimizadas por un experto en este campo.

Las ligasas incluyen, pero no se limitan a, ligasas dependientes de  $NAD^+$  que incluyen ARNt ligasa, ADN ligasa Taq, ADN ligasa de *Thermus filiformis*, ADN ligasa de *Escherichia coli*, ADN ligasa Tth, ADN ligasa de *Thermus scotoductus*, ligasa termoestable, ADN ligasa termoestable Ampiligasa, ligasa tipo VanC, ADN ligasa 9°N, ADN ligasa Tsp, y novedosas ligasas descubiertas por bioprospección. Las ligasas también incluyen, pero no se limitan a, ligasas dependientes de ATP que incluyen ARN ligasa T4, ADN ligasa T4, ADN ligasa T7, ADN ligasa Pfu, ADN ligasa I, ADN ligasa III, ADN ligasa IV, y novedosas ligasas descubiertas por bioprospección. Estas ligasas incluyen isoformas no mutantes, mutantes, y variantes genéticamente manipuladas.

Con referencia a la Figura 11, y en la preparación para la síntesis, se proporciona un ácido nucleico diana (110) y los extremos se pulen en la preparación para la ligación de extremos romos de adaptadores. La etapa I muestra la ligación de cebadores en horquilla (120) con el ácido nucleico diana. El extremo 5' libre de los cebadores se bloquea con un grupo de bloqueo móvil (119). Los cebadores cebarán ambas hebras del ácido nucleico diana. Los adaptadores se añaden generalmente en exceso. Los grupos de bloqueo sobre los extremos calientes de los



cebadores se eliminan en la etapa II, y las dos hebras del molde se separan por desnaturalización. En la etapa III, el molde monocatenario cebado (111) se pone en contacto con una biblioteca de construcciones de sustrato (como se representa por la construcción (112) para fin de ilustración) y con ligasa, L, en condiciones permisibles para la hibridación de sustrato de sonda complementaria (113) y ligación en el extremo reactivo del cebador, como se muestra en la etapa IV. Generalmente, la hibridación y ligación se realizan a una temperatura superior a la temperatura de fusión del sustrato para reducir reacciones secundarias no específicas. Cada construcción de sustrato en este ejemplo contiene un anclaje que presenta tres indicadores. Cada sustrato de sonda tiene un enlace selectivamente escindible (indicado con una "V") entre los dos sitios de unión del anclaje. En la etapa V, una segunda construcción de sustrato (114) se añade por hibridación y ligación dirigida por molde, etc. En la etapa VI se demuestra la formación de un producto intermedio de Xpandómero completamente extendido (117). Este producto intermedio puede desnaturalizarse de la hebra molde y escindir selectivamente en los sitios de escisión mostrados, formando así un Xpandómero producto adecuado para la secuenciación. En algunas realizaciones no se necesita la desnaturalización y la hebra molde puede digerirse en su lugar.

La Figura 12 es un esquema condensado de un segundo método, aquí para la preparación de otra realización de un Xpandómero de clase I. En la preparación para la síntesis, se proporciona un ácido nucleico diana (120) y los extremos se pulen en preparación para la ligación de extremos romos de adaptadores (121,122). La etapa I muestra la ligación de precursores de cebador en horquilla doblemente bloqueados al ácido nucleico diana. Un extremo de los cebadores en horquilla del dúplex se bloquea con grupos de bloqueo móviles (125a,125b,125c,125d) previstos para prevenir la ligación y concatenación de las hebras molde o adaptadores. Los adaptadores se añaden generalmente en exceso. Los grupos de bloqueo se eliminan en la etapa II, y las dos hebras del molde se separan por desnaturalización. En la etapa III, los cebadores en horquilla se auto-hibridan, formando sitios de cebado (126,127) para la posterior ligación de construcciones de sustrato, que puede avanzar bidireccionalmente, es decir, tanto en una dirección 3' a 5' como una 5' a 3'. En la etapa IV, los moldes cebados se ponen en contacto con una biblioteca de construcciones de sustrato (128) en condiciones permisibles para la hibridación de sustratos de sonda complementaria y ligación. La ligación avanza incrementalmente (es decir, extendiendo los extremos en crecimiento con procesividad aparente) por un proceso de hibridación de sustratos de sonda complementaria y ligación en los extremos de las hebras hija nacientes. Cada construcción de sustrato en este ejemplo contiene un bucle de anclaje que presenta grupos indicadores. En la etapa V se representa la formación de un producto intermedio de Xpandómero completado (129). Opcionalmente, la hebra molde puede eliminarse por digestión con nucleasa, liberando el Xpandómero. El producto intermedio puede escindir selectivamente en los sitios de escisión mostrados, formando así un Xpandómero producto adecuado para la secuenciación. Los Xpandómeros producto se forman en disolución libre.

En la Figura 13 se muestra un método que se basa en hebras molde inmovilizadas. Aquí, las hebras molde están ancladas a una perla (u otro soporte de fase sólida) por un adaptador (131). El molde se muestra en contacto con construcciones de sustrato (132), y en la etapa I, las condiciones están adaptadas de manera que se produzca la hibridación. Puede observarse que se forman "islas" de construcciones de sustrato confinadas hibridadas. En la etapa II, la adición de ligasa, L, produce la ligación de las construcciones de sustrato confinadas, formando así múltiples secuencias contiguas de productos intermedios ligados separados por huecos. En la etapa III, las condiciones se ajustan para favorecer la disociación de material hibridado de bajo peso molecular o desapareado, y en la etapa IV, las reacciones de las etapas II a III se repiten una o más veces para favorecer la formación de productos de extensión más largos. Este proceso sin cebador se denomina en el presente documento "ligación promiscua". La ligación puede extenderse bidireccionalmente e intersecciones cortadas y empalmadas pueden sellarse con ligasa, llenando así los huecos. En la etapa V, después de la optimización de las longitudes de producto deseadas, los dúplex inmovilizados se lavan para eliminar sustrato y ligasa sin reaccionar. Entonces, en la etapa VI, las hebras hija (aquí mostradas como un producto intermedio de Xpandómero monocatenario) (138,139) se disocian del molde. La escisión selectiva de enlaces selectivamente escindibles del producto intermedio produce la formación del producto de Xpandómero (no mostrado). En esta realización, el molde inmovilizado puede volver a utilizarse. Una vez se secuencian los productos de Xpandómero, los cóntigos pueden ensamblarse por algoritmos muy conocidos para solapar y alinear los datos para construir una secuencia consenso.

Con referencia a la Figura 14, se muestra un método de uso de cebadores inmovilizados. Los moldes adaptados en los extremos (o secuencias molde aleatorias, dependiendo de la naturaleza de los cebadores inmovilizados) (142) se hibridan con los cebadores inmovilizados (140) en la etapa I. En la etapa II, los moldes inmovilizados (143) se ponen en contacto con una biblioteca de construcciones de sustrato, cuyos miembros se muestran como (144), y las condiciones se ajustan para la hibridación dirigida por molde. En este ejemplo, los extremos 3' OH de la construcción de sustrato de los miembros de sonda (grupo R) se ha sustituido (146) para bloquear reversiblemente adicionalmente la extensión. En la etapa III, los extremos confinados de la construcción de sustrato adyacente y cebador, o extremo libre de la hebra hija en crecimiento naciente, se ligan y el extremo 3' OH de la hebra hija naciente se activa eliminando el grupo R de bloqueo (146). Como se indica en las etapas IV y V, este proceso de adición cíclica escalonada puede repetirse múltiples veces. Normalmente, se usa una etapa de lavado para eliminar sustratos sin reaccionar entre cada etapa de extensión.

El proceso es así análogo a lo que se llama "extensión cíclica de una única base", pero se llamaría más apropiadamente aquí "extensión cíclica de una única sonda". Aunque se muestra la ligasa, L, el proceso puede realizarse con una ligasa, polimerasa, o por cualquier protocolo de acoplamiento químico adecuado para unir

oligómeros en una síntesis dirigida por molde. El acoplamiento químico puede producirse espontáneamente en los extremos confinados de las sondas hibridadas, o un agente de condensación puede añadirse al principio de la etapa III y cada etapa V resultante del ciclo. El grupo R de bloqueo terminal está configurado de manera que la polimerización iniciada libre no pueda producirse sobre el molde o en disolución. La etapa VI muestra la formación de un producto intermedio de Xpandómero completo (149); no puede añadirse más sustrato. Este producto intermedio puede disociarse del molde, el producto monocatenario se escinde entonces para abrir el esqueleto como se describe previamente.

Este método puede adaptarse para la secuenciación selectiva de dianas particulares en una mezcla de ácidos nucleicos, y para analizar métodos de secuenciación sobre matrices de secuenciación, por ejemplo, por selección no aleatoria de los cebadores inmovilizados. Alternativamente, pueden usarse cebadores universales o aleatorios como se muestra.

La Figura 15 describe un método de hibridación promiscua sobre un molde inmovilizado (150) (etapa I), en el que las construcciones de sustrato de la biblioteca (152) se modifican con un grupo funcional químico que es selectivamente reactivo (156), representado como un triángulo abierto, con una sonda confinada. Un detalle del grupo funcional químico de las construcciones de sustrato se muestra en la porción expandida mostrada por el círculo en trama (Figura 15a). A una cierta densidad de hibridación, el acoplamiento se inicia como se muestra en la etapa II, produciendo productos intermedios de Xpandómero de alto peso molecular enlazados por el producto reticulado (157), representado como un triángulo relleno, de la reacción de acoplamiento. Un detalle de las sondas reticuladas se muestra en la porción expandida mostrada en el círculo en trama (Figura 15b) en el producto de la etapa II. Este proceso puede ir acompañado por etapas para la disociación selectiva y eliminación de productos de bajo peso molecular y cualquier posible producto desapareado. Las químicas de acoplamiento para este método de acoplamiento químico promiscuo son conocidas para algunos expertos en la materia e incluyen, por ejemplo, las técnicas desveladas en la patente de EE.UU. N.º. 6.951.720 a Burgin et al.

En otra realización, se desvelan métodos basados en polimerasa para ensamblar Xpandómeros producto. Generalmente, los trifosfatos de sustrato (Xmeros) son el sustrato apropiado para reacciones que implican a una polimerasa. La selección de una polimerasa adecuada es parte de un proceso de optimización del protocolo experimental. Como se muestra en la Figura 16 para ilustración, y aunque no se pretende que sea limitante, se pone en contacto una mezcla de reacción que contiene un molde (160) y un cebador (161) con una biblioteca de construcciones de sustrato (162) y una polimerasa (P), en condiciones optimizadas para la polimerización dirigida por molde. En la etapa I, la polimerasa empieza a añadir procesivamente Xmeros de dinucleótido (anclajes con dos indicadores) a la hebra molde. Este proceso continúa en las etapas II y III. Cada subunidad de sonda añadida es una especie particular seleccionada por unión específica al siguiente oligómero adyacente del molde de manera que forme una copia complementaria continua del molde. Aunque no está ligado por teoría, se cree que la polimerasa ayuda en asegurar que las especies de sondas entrantes añadidas con la cadena naciente sean específicamente complementarias al siguiente segmento contiguo disponible del molde. Loeb y Patel describen ADN polimerasas mutantes con elevada actividad y fidelidad mejorada (patente de EE.UU. N.º. 6.329.178). Williams, por ejemplo, en la solicitud de patente de EE.UU. 2007/0048748 ha mostrado que las polimerasas pueden modificarse para elevada velocidad de incorporación y reducción en la tasa de error, enlazando claramente la tasa de error no con la exactitud de hibridación, sino con la procesividad de polimerasa. La etapa III produce un producto intermedio de Xpandómero completado (168). El producto intermedio de Xpandómero monocatenario se trata entonces por una manera de proceder que puede implicar la desnaturalización de la hebra molde (no mostrada). El esqueleto primario de la hebra hija se escinde selectivamente para expandir los anclajes, formando así un producto de Xpandómero adecuado para su uso en un protocolo de secuenciación, como se ha explicado previamente.

Como se muestra en la Figura 17, la síntesis dirigida por molde conducida por polimerasa de un Xpandómero puede lograrse por técnicas alternativas. Aquí, un cebador inmovilizado (170) con el que se hibrida una hebra molde procesada (171) en la etapa I. En la etapa II, la polimerasa, P, se acopla procesivamente específicamente con construcciones de sustrato complementarias (175) de una biblioteca de tales construcciones (representadas por 174) en la mezcla de reacción. Las condiciones y disoluciones de reactivo se ajustan para favorecer la actividad de polimerasa procesiva. Como se muestra aquí, la hibridación en la etapa II y la polimerización en la etapa III son actividades separadas, pero las actividades de la polimerasa no necesitan aislarse de esa forma. En la etapa IV, la adición procesiva incremental de construcciones de sustrato complementarias continúa cíclicamente (continuamente sin interrupción), produciendo el producto intermedio de Xpandómero completamente cargado (177) como se representa resultante de la etapa IV. El producto intermedio de Xpandómero puede disociarse y expandirse en preparación para su uso en un protocolo de secuenciación como se describe previamente. Obsérvese que este método también se presta a métodos de secuenciación analizados por selección de cebadores inmovilizados adecuados. Además, los métodos para estirar el molde para liberar la estructura secundaria están fácilmente adaptados a este método, y se tratan en secciones posteriores.

Polimerasas incluyen, pero no se limitan a, ADN polimerasas dependientes de ADN, ARN polimerasas dependientes de ADN, ADN polimerasas dependientes de ARN, ARN polimerasas dependientes de ARN, ADN polimerasa T7, ADN polimerasa T3, ADN polimerasa T4, ARN polimerasa T7, ARN polimerasa T3, ARN polimerasa SP6, ADN polimerasa I, fragmento de Klenow, ADN polimerasa de *Thermophilus aquaticus*, ADN polimerasa Tth, ADN polimerasa VentR<sup>®</sup> (New England Biolabs), ADN polimerasa Deep VentR<sup>®</sup> (New England Biolabs), fragmento grande

de ADN polimerasa Bst, fragmento de Stoeffel, ADN polimerasa 9°N, ADN polimerasa 9°N, ADN Polimerasa Pfu, ADN polimerasa Tfl, ADN polimerasa Tth, polimerasa RepliPHI Phi29, ADN polimerasa Tli, ADN polimerasa beta de eucariota, telomerasa, polimerasa Therminator™ (New England Biolabs ADN polimerasa), KOD HiFi™ (Novagen), ADN polimerasa KOD1, Q-beta replicasa, transferasa terminal, transcriptasa inversa AMV, transcriptasa inversa M-MLV, transcriptasa inversa Phi6, transcriptasa inversa del VIH-1, polimerasas novedosas descubiertas por bioprospección, y polimerasas citadas en los documentos US 2007/0048748, US 6.329.178, US 6.602.695 y US 6.395.524. Estas polimerasas incluyen isoformas no mutantes, mutantes, y variantes genéticamente manipuladas.

#### Construcciones oligoméricas de clase II y III

10 Con referencia a las Figuras 18A a 18E, describen construcciones oligoméricas de clase II en más detalle, que (junto con las construcciones oligoméricas de clase III isoméricas) pueden ser tanto Xsondas como Xmeros.

Las Figuras 18A a 18C se leen de izquierda a derecha, mostrando primero la construcción de sonda-sustrato (precursor oligomérico de Xpandómero), luego la hebra hija del dúplex intermedia en el centro, y a la derecha el producto de Xpandómero preparado para la secuenciación.

15 Como se muestra en la Figura 18A, una construcción de sustrato de clase II tiene un miembro de sonda oligomérico (-P<sup>1</sup>-P<sup>2</sup>-) (180) y un miembro de anclaje, T (181). El anclaje está unido por un único enlace terminal (184) de un primer resto terminal al resto de sonda P<sup>2</sup>. En el extremo distal del anclaje (186), un segundo resto terminal tiene un grupo conector δ y está posicionado próximo a R<sup>2</sup>. El segundo resto terminal también tiene una reticulación intra-anclaje escindible (187) para limitarlo a esta localización. La reticulación escindible (187) se indica por una línea de puntos, que puede indicar, por ejemplo, un enlace disulfuro. Estas limitaciones previenen que el anclaje se alargue o expanda y así está en una configuración limitada. Un segundo grupo conector ε está posicionado cerca del extremo distal (189) del miembro de sonda cerca de R<sup>1</sup>. Bajo ensamblaje dirigido por molde, los sustratos forman un dúplex con el molde diana de forma que los sustratos estén confinados. Bajo condiciones controladas, los grupos conectores δ y ε de los sustratos confinados se enlazan para formar un enlace χ entre las construcciones de sustrato adyacentes (mostradas en las Figuras 18B y 18C). Estos grupos de enlace están posicionados sobre la construcción de sustrato para limitar estas reacciones de enlace con construcciones de sustrato confinadas adyacentes. La construcción de sustrato no se enlaza preferencialmente consigo misma. Químicas de enlace y de protección/desprotección adecuadas para δ, ε y χ se detallan en la descripción de la construcción oligomérica general.

30 R<sup>1</sup> y R<sup>2</sup> son grupos terminales configurados según convenga para el protocolo de síntesis en el que se usa la construcción de sustrato. Por ejemplo, R<sup>1</sup> = 5'-fosfato y R<sup>2</sup> = 3'-OH, encontrarían uso en un protocolo de ligación, y R<sup>1</sup> = 5'-trifosfato y R<sup>2</sup> = 3'-OH para un protocolo de polimerasa I. Opcionalmente, R<sup>2</sup> puede configurarse con un grupo de bloqueo reversible para la adición cíclica de un único sustrato. Alternativamente, R<sup>1</sup> y R<sup>2</sup> pueden configurarse con los grupos conectores terminales para el acoplamiento químico o con grupos no conectores para un protocolo de solo hibridación. R<sup>1</sup> y R<sup>2</sup> pueden ser del tipo general XR, en la que X es un grupo de enlace y R es un grupo funcional.

40 Las construcciones de sustrato son reactivos usados para el ensamblaje dependiente de molde de una hebra hija, una composición intermedia para producir Xpandómeros. La Figura 18B muestra la hebra hija de dúplex, un heterocopolímero con subunidades de repetición (mostradas en corchetes). Se muestran esqueleto primario de la hebra hija (~P<sup>1</sup>-P<sup>2</sup>~) y hebra molde diana (-p<sup>1</sup>-p<sup>2</sup>-) como un dúplex (185). Cada subunidad de la hebra hija es un motivo de repetición que comprende un miembro de sonda y un miembro de anclaje. Los motivos tienen variabilidad específica de especie, indicada aquí por el superíndice α. Cada subunidad particular en la hebra hija se selecciona de una biblioteca de motivos por un proceso dirigido por molde y su sonda se une a una secuencia correspondiente de nucleótidos complementarios sobre la hebra molde. De esta forma, la secuencia de residuos de nucleobase de las sondas forma una copia complementaria contigua de la hebra molde diana.

45 Cada tilde (~) indica un enlace selectivamente escindible. El enlace interno entre los restos P<sup>1</sup> y P<sup>2</sup> de un miembro de sonda no son enlaces selectivamente escindibles, pero los enlaces inter-sonda (entre subunidades) son necesariamente selectivamente escindibles según se requiera para expandir los anclajes y el Xpandómero. En una realización, no se forma enlace directo entre las sondas de subunidades separadas, eliminándose así la necesidad de escisión selectiva posterior.

50 La hebra hija está compuesta por el precursor Xpandómero llamado el "Xpandómero limitado" que está adicionalmente compuesto de anclajes en la "configuración limitada". Cuando los anclajes se convierten en su "configuración expandida", el Xpandómero limitado se convierte en el producto de Xpandómero. Los anclajes están limitados por los enlaces χ formados por la conexión con los miembros de sonda de subunidades adyacentes y, opcionalmente, los enlaces intra-anclaje si todavía están presentes. El enlace χ une el miembro de anclaje de una primera subunidad con el extremo confinado de una segunda subunidad adyacente y se forma enlazando los grupos conectores colocados, δ de la primera subunidad y ε de la segunda subunidad.

Puede observarse que la hebra hija tiene dos esqueletos, un "esqueleto primario" y el esqueleto del "Xpandómero limitado". El esqueleto primario está compuesto por los sustratos de sonda contiguamente confinados. El "esqueleto

de Xpandómero limitado" evita el enlace selectivamente escindible entre los sustratos de subunidad y se forma por restos de esqueleto enlazados por enlace  $\chi$ , siendo cada resto de esqueleto un enlace lineal de anclaje, a P<sup>2</sup>, a P<sup>1</sup>, enlazando cada enlace  $\chi$  P<sup>1</sup> al anclaje del siguiente resto de esqueleto. Puede observarse que el esqueleto de Xpandómero limitado conecta o forma un bucle sobre los enlaces selectivamente escindibles del esqueleto primario, y seguirá covalentemente intacto cuando estos enlaces selectivamente escindibles se escindan y el esqueleto primario se fragmente.

En la Figura 18B, los grupos conectores  $\delta$  y  $\varepsilon$  se han reticulado y ahora forman un enlace  $\chi$  intra-subunidad. Después de formar el enlace  $\chi$ , el enlace intra-anclaje puede romperse, aunque se muestra aquí intacto (línea de puntos en el sustrato). Generalmente, la formación del enlace  $\chi$  depende de la proximidad del grupo conector  $\delta$  sobre la primera subunidad y la posición del grupo conector  $\varepsilon$  de una segunda subunidad confinada, de manera que estén colocados y se pongan en contacto durante o después del ensamblaje dirigido por molde de construcciones de sustrato.

En otras realizaciones, la reticulación depende solo de la hibridación con el molde para juntar los dos grupos conectores. En otras realizaciones más, el enlace  $\chi$  va precedido por acoplamiento enzimático de los miembros de sonda P a lo largo del esqueleto primario, con formación de enlaces fosfodiéster entre sondas adyacentes. En la estructura mostrada aquí, el esqueleto primario de la hebra hija se ha formado, y los enlaces inter-sustratos se representan por una tilde (~) para indicar que son selectivamente escindibles. Después de disociar o degradar la hebra molde diana, escindiendo los enlaces selectivamente escindibles (que incluyen los enlaces intra-anclaje), el Xpandómero limitado se libera y se convierte en el producto de Xpandómero.

La Figura 18C es una representación del producto de Xpandómero de clase II después de la disociación de la hebra molde y después de la escisión de los enlaces selectivamente escindibles (incluyendo aquellos en el esqueleto primario y, si no se han escindido ya, los enlaces intra-anclaje). Métodos para la disociación de la hebra molde incluyen desnaturalización por calor, o digestión selectiva con una nucleasa, o degradación química. La hebra del producto de Xpandómero contiene una pluralidad de subunidades  $\kappa$ , en la que  $\kappa$  indica la subunidad  $\kappa^n$  en una cadena de  $m$  subunidades que constituyen la hebra hija, en la que  $\kappa = 1, 2, 3$  a  $m$ , en la que  $m > 3$ , y generalmente  $m > 20$ , y preferentemente  $m > 50$ , y más preferencialmente  $m > 1000$ . Cada subunidad está formada de un anclaje, y restos de sonda P<sup>1</sup> y P<sup>2</sup>. El anclaje, T (181), se observa en su configuración expandida y se estira a su longitud entre P<sup>2</sup> y P<sup>1</sup> de subunidades adyacentes. Cada subunidad, un motivo de subunidad  $\alpha$ , contiene información genética específica de especie establecida por el ensamblaje dirigido por molde del producto intermedio de Xpandómero (hebra hija).

La Figura 18D muestra la construcción de sustrato de la Figura 18A como modelo molecular, en la que el miembro de sonda (180), representado con cuatro residuos de nucleobase (81,82,83,84), está unido al anclaje (181) por un enlace del primer resto terminal del anclaje (184). Un enlace intra-anclaje (85) de un segundo resto terminal está en el extremo distal del anclaje. Un grupo conector ( $\delta$ ) (86) también está dispuesto en el segundo resto terminal y el segundo grupo conector ( $\varepsilon$ ) (87) correspondiente está anclado con el extremo de la sonda opuesto al grupo conector ( $\delta$ ). El bucle de anclaje mostrado aquí tiene tres indicadores (78,79,80), que también pueden ser específicos de especies de motivos.

La Figura 18E muestra la construcción de sustrato después de la incorporación en el Xpandómero producto. Las subunidades se escinden y se expanden y se enlazan por enlaces  $\chi$  (88), formados enlazando los grupos conectores  $\delta$  y  $\varepsilon$  referidos en la Figura 18A. Una subunidad se indica por líneas de puntos que agrupan verticalmente la subunidad de repetición, como se representa por los corchetes en la Figura 18C adjunta. "E" de nuevo indica expansión.

En el producto de Xpandómero (Figura 18E) el esqueleto primario se ha fragmentado y no es covalentemente contiguo debido a que se ha escindido cualquier enlace directo entre las sondas de subunidades adyacentes. Mediante el proceso de escisión, el Xpandómero limitado se libera para convertirse en el producto de Xpandómero. Los miembros de anclaje que estuvieron anteriormente en la configuración limitada están ahora en la configuración expandida, sirviendo así para estirar linealmente la información de secuencias de la diana molde. La expansión de los anclajes reduce la densidad lineal de la información de secuencias a lo largo del Xpandómero y proporciona una plataforma para aumentar el tamaño y la abundancia de indicadores, que a su vez mejora la señal con respecto al ruido para la detección y descodificación de la secuencia molde.

Aunque el anclaje se representa como una construcción de indicador con tres grupos indicadores, diversas construcciones de indicador pueden estar presentes sobre el anclaje, y pueden comprender indicadores individuales que identifican constituyentes de sonda, indicadores individuales que identifican especies de sondas, códigos de barras moleculares que identifican las especies de sondas, o el anclaje puede ser un polímero desnudo. En algunos casos, uno o más precursores de indicador están presentes sobre el anclaje, y los indicadores se unen por afinidad o se unen covalentemente tras el ensamblaje del producto de Xpandómero.

Las construcciones oligoméricas de clase III, ilustradas en las Figuras 19A a 19E, sin isómeros con las construcciones de clase II tratadas anteriormente. No se incluye descripción adicional debido a que la descripción de

la clase II es adecuada para entender esta clase.

Esta clase puede servir para enfatizar que todas las clases pueden reflejarse en la aplicación de imagen especular (es decir, intercambio de los grupos  $R^1$  y  $R^2$ ). Además, esto sirve para ilustrar que las clases descritas no pretenden ser completas, sino que reflejan algunas de las muchas disposiciones posibles que engloba la presente invención.

5 La Figura 20 representa un esquema condensado de un método de preparación de una primera realización de un Xpandómero de clase II; el método ilustra la preparación y uso de sustratos y productos mostrado en las Figuras 18D y 18E. El método se realiza con químicas en fase sólida. Los métodos de alivio de la estructura secundaria en el molde se tratan en una sección posterior. Condiciones adecuadas adaptadas para la hibridación y acoplamiento químico son muy conocidas en la técnica y las condiciones pueden ser fácilmente optimizadas por un experto en este campo.

15 La etapa I de la Figura 20 muestra una mezcla de reacción que contiene un molde inmovilizado (200) y una biblioteca de reactivos de sustrato (201). Se observa que las construcciones de sustrato se unen específicamente al molde en una hibridación dirigida por molde. Las condiciones se ajustan para optimizar la complementariedad y fidelidad de la unión. Como se muestra en la inserción de la figura (véase la Figura 20a), cada construcción de sustrato confinada pone en proximidad el grupo funcional  $\delta$  (202) sobre el aspecto distal del anclaje, mostrado aquí unido con el tallo del anclaje por una reticulación intra-anclaje (203), representada por los triángulos adyacentes, y el grupo funcional  $\varepsilon$  (204) del miembro de sonda confinado.

20 En la etapa II, se produce una reacción de reticulación entre extremos proximalmente confinados hibridados de los miembros de sonda que implican a los dos grupos funcionales  $\delta$  y  $\varepsilon$ , formando así un enlace  $\chi$  anclaje a sonda inter-subunidad (205), representado como un óvalo abierto, como se muestra en la inserción de la figura (véase la Figura 20b). La hibridación se produce en paralelo en diversos sitios sobre el molde, promiscuamente, y el acoplamiento químico puede producirse en un ciclo de hibridación (etapa III), fusión rigurosa y/o lavado (etapa IV), y acoplamiento químico (etapa V). El ciclo puede repetirse para aumentar el número de subunidades contiguas ensambladas para formar el producto intermedio de Xpandómero. La etapa VI ilustra un producto intermedio de Xpandómero completado con dos hebras de producto contiguas de longitud variable. Puede emplearse un método similar con los Xpandómeros de clase III.

30 La Figura 21 ilustra un método de ligación procesiva de los sustratos de clase II sobre un molde inmovilizado. La etapa I muestra un cebador (210) que se hibrida con el molde (212), el cebador adaptado con un grupo funcional químicamente reactivo  $\varepsilon$  mostrado en la inserción de la figura como (214) (véase la Figura 21a). Una mezcla de reacción que contiene sustratos de clase II (216) se añade entonces a la etapa II. Como se muestra en la inserción de la figura (Figura 21a), estas construcciones de sustrato tienen reactividad  $\delta$  (217) y  $\varepsilon$  (214) en extremos opuestos del miembro de sonda-anclaje. Se observa que una primera construcción de sustrato se une específicamente al molde en una hibridación dirigida por molde. Las condiciones se ajustan para optimizar la complementariedad y fidelidad de la unión. Entonces se usa una ligasa para enlazar covalentemente la primera sonda al cebador (etapa II).

35 En las etapas III y IV, el proceso de hibridación y ligación procesivo de construcciones de sustrato continúa con el fin de construir el producto intermedio de Xpandómero mostrado formado en la etapa IV. Tras esto, en la etapa V, la reticulación se realiza entre los grupos  $\delta$  (217) y  $\varepsilon$  (214) (véase la Figura 21b), produciendo un enlace  $\chi$  como se representa en la Figura 21c como (219). Como se muestra en las inserciones de figuras (Figura 21b y 21c), el grupo funcional  $\delta$  (217) sobre el anclaje está limitado por una reticulación intra-anclaje (211), representada por los triángulos adyacentes, hasta que se forma el enlace  $\chi$ . El producto intermedio de Xpandómero completado se disocia opcionalmente de la hebra molde y se escinde para formar un producto de Xpandómero adecuado para la secuenciación. Puede emplearse un método similar con los Xpandómeros de clase III. Este método también puede adaptarse para su uso con una polimerasa sustituyendo construcciones de sustrato de trifosfato.

#### 45 Construcciones oligoméricas de clase IV y V

Con referencia a las Figuras 22A a 22E, se describen construcciones oligoméricas de clase IV en más detalle.

Las Figuras 22A a 22C se leen de izquierda a derecha, mostrando primero la construcción de sonda-sustrato (precursores de Xsonda o Xmero de Xpandómero), luego la hebra hija del dúplex intermedia en el centro, y a la derecha el producto de Xpandómero preparado para la secuenciación.

50 La Figura 22A muestra una construcción de sustrato de clase IV que tiene miembro de sonda oligomérico (229) con restos de sonda  $P^1$  y  $P^2$  que se unen al anclaje, T (220). El anclaje T está unido a  $P^1$  y  $P^2$  por enlace apropiado con el primer y segundo restos terminales del anclaje, respectivamente. Los grupos conectores  $\varepsilon$  del primer resto terminal y  $\delta$  del segundo resto terminal están posicionados cerca de los extremos  $R^1$  y  $R^2$  de la sonda, respectivamente (en una realización alternativa, las posiciones de los grupos funcionales pueden estar invertidas).

55 Bajo condiciones controladas, los grupos funcionales  $\delta$  (222) y  $s$  (221) reaccionarán para formar un enlace  $\chi$  como se muestra en la Figura 22B. Estos grupos de enlace están posicionados en la construcción de sustrato para limitar estas reacciones de enlace con construcciones de sustrato confinadas adyacentes. La construcción de sustrato

preferencialmente no se enlaza consigo misma. Químicas de enlace y de protección/desprotección adecuadas para  $\delta$ ,  $\varepsilon$ , y  $\chi$  se detallan en la descripción general de construcciones oligoméricas.

5  $R^1$  y  $R^2$  son grupos terminales configurados según convenga para el protocolo de síntesis en el que se usa la construcción de sustrato. Por ejemplo,  $R^1 = 5'$ -fosfato y  $R^2 = 3'$ -OH, encontrarían uso en un protocolo de ligación, y  $R^1 = 5'$ -trifosfato y  $R^2 = 3'$ -OH para un protocolo de polimerasa I. Opcionalmente,  $R^2$  puede configurarse con un grupo de bloqueo reversible para la adición cíclica de un único sustrato. Alternativamente,  $R^1$  y  $R^2$  pueden configurarse con los grupos conectores terminales para el acoplamiento químico o con grupos no conectores para un protocolo de solo hibridación.  $R^1$  y  $R^2$  pueden ser del tipo general XR, en la que X es un grupo de enlace y R es un grupo funcional.

10 Las construcciones de sustrato son reactivos usados para el ensamblaje dirigido por molde de una hebra hija, una composición intermedia para producir Xpandómeros. La Figura 22B muestra la hebra hija de dúplex, un hetero-copolímero con subunidades de repetición (mostradas en corchetes). Se muestran esqueleto primario de la hebra hija ( $-P^1 \sim P^2-$ ) y hebra molde diana ( $-P^1-P^2-$ ) como un dúplex (228). Cada subunidad de la hebra hija es un motivo de repetición que comprende un miembro de sonda y un miembro de anclaje. Los motivos tienen variabilidad específica de especie, indicada aquí por el superíndice  $\alpha$ . Cada subunidad particular en la hebra hija se selecciona de una biblioteca de motivos por un proceso dirigido por molde y su sonda se une a una secuencia correspondiente de nucleótidos complementarios sobre la hebra molde. De esta forma, la secuencia de residuos de nucleobase de las sondas forma una copia complementaria contigua de la hebra molde diana.

15 La tilde ( $\sim$ ) indica un enlace selectivamente escindible. El enlace interno entre los restos  $P^1$  y  $P^2$  de un miembro de sonda es necesariamente selectivamente escindible según se requiera para expandir los anclajes y el Xpandómero. En una realización, no se forma enlace directo entre las sondas de subunidades separadas.

20 La hebra hija está compuesta por un precursor de Xpandómero llamado el "Xpandómero limitado" que está adicionalmente compuesto de anclajes en la "configuración limitada". Cuando los anclajes se convierten en su "configuración expandida", como se muestra en la Figura 22C, el Xpandómero limitado se convierte en el producto de Xpandómero. Los anclajes están limitados por los enlaces  $\chi$  formados por conexión con el miembro de anclaje de subunidades adyacentes y por los enlaces de sonda. El enlace  $\chi$  une el miembro de anclaje de una primera subunidad con el anclaje de una segunda subunidad adyacente y se forma enlazando los grupos conectores colocados,  $\delta$  de la primera subunidad y  $\varepsilon$  de la segunda subunidad.

25 Puede observarse que la hebra hija tiene dos esqueletos, un "esqueleto primario" y el esqueleto del "esqueleto de Xpandómero limitado". El esqueleto primario está compuesto por los sustratos de sonda contiguamente confinados. El "esqueleto de Xpandómero limitado" es el enlace lineal de los anclajes en cada subunidad enlazados juntos por los enlaces  $\chi$  que evitan los sustratos de sonda de subunidad. El enlace  $\chi$  resulta de una reacción del grupo funcional  $\varepsilon$  de una primera subunidad con el grupo funcional  $\delta$  de una segunda subunidad confinada. Puede observarse que el esqueleto de Xpandómero limitado conecta o forma un bucle sobre los enlaces selectivamente escindibles del esqueleto primario, y seguirá covalentemente intacto cuando estos enlaces selectivamente escindibles se escindan y el esqueleto primario se fragmente.

30 En la Figura 22B, los grupos conectores  $\delta$  y  $\varepsilon$  se han reticulado y ahora forman un enlace  $\chi$  intra-subunidades. Generalmente, la formación del enlace  $\chi$  depende de la colocación del grupo conector  $\delta$ , en la primera subunidad, y el grupo conector  $\varepsilon$ , de una segunda subunidad confinada, de manera que se pongan en contacto durante o después del ensamblaje dirigido por molde de construcciones de sustrato.

35 En otras realizaciones, la reticulación del enlace  $\chi$  depende solo de la hibridación con el molde para juntar los dos grupos conectores. En otras realizaciones más, la formación del enlace  $\chi$  va precedida del acoplamiento enzimático de los miembros de sonda P a lo largo del esqueleto primario con enlaces fosfodiéster entre sondas adyacentes. En la estructura mostrada en la Figura 22B, el esqueleto primario de la hebra hija se ha formado, y el enlace entre los restos de sonda se representa por una tilde ( $\sim$ ) para indicar que es selectivamente escindible. Después de disociar o degradar la hebra molde diana, escindiendo los enlaces selectivamente escindibles, el Xpandómero limitado se libera y se convierte en el producto de Xpandómero como se muestra en la Figura 22C.

40 A este respecto, la Figura 22C es una representación del producto de Xpandómero de clase IV después de la disociación de la hebra molde y después de la escisión de los enlaces selectivamente escindibles del esqueleto primario. Métodos para la disociación de la hebra molde incluyen desnaturalización por calor, o digestión selectiva con una nucleasa, o degradación química. La hebra del producto de Xpandómero contiene una pluralidad de subunidades  $\kappa$ , en la que  $\kappa$  indica la subunidad  $k^1$  en una cadena de  $m$  subunidades que constituyen la hebra hija, en la que  $m > 3$ , y generalmente  $m > 20$ , y preferentemente  $m > 50$ , y más preferencialmente  $m > 1000$ . Cada subunidad está formada de un anclaje (220), y restos de sonda laterales  $P^1$  y  $P^2$ . El anclaje, T, se observa en su configuración expandida y se estira a su longitud entre subunidades adyacentes. Cada subunidad, un motivo de subunidad  $\alpha$ , contiene información genética específica de especie establecida por el ensamblaje dirigido por molde del producto intermedio de Xpandómero (hebra hija).

La Figura 22D muestra la construcción de sustrato de la Figura 22A como modelo molecular, en la que el miembro de sonda, representado con cuatro residuos de nucleobase (círculos blancos), está unido al anclaje por un enlace del primer resto terminal del anclaje. El grupo conector (221), mostrado como  $\epsilon$  en la Figura 22A, también es del primer resto terminal del anclaje. Un grupo conector (222), mostrad como  $\delta$  en la Figura 22A, está dispuesto en un segundo resto terminal en el extremo distal del anclaje (220). El bucle de anclaje mostrado aquí tiene tres indicadores (800,801,802), que también pueden ser específicos de especies de motivos. Un enlace selectivamente escindible, mostrado como una "V" (225), se localiza dentro del miembro de sonda (229).

La Figura 22E muestra la construcción de sustrato después de la incorporación en el Xpandómero producto. Las subunidades se escinden, mostradas como líneas (225a,225b), y se expanden y se enlazan por enlaces  $\chi$  (223,224), formados enlazando los grupos conectores  $\delta$  y  $\epsilon$  referidos en la Figura 22A. Una subunidad se indica por líneas de puntos que agrupan verticalmente la subunidad de repetición, como se representa por los corchetes en la Figura 22C adjunta.

En el producto de Xpandómero de la Figura 22E, el esqueleto primario se ha fragmentado y no es covalentemente contiguo debido a que se ha escindido cualquier enlace directo entre las sondas de subunidades adyacentes. Mediante el proceso de escisión, el Xpandómero limitado se libera para convertirse en el producto de Xpandómero. Los miembros de anclaje que estuvieron anteriormente en la configuración limitada están ahora en la configuración expandida, sirviendo así para estirar linealmente la información de secuencias de la diana molde. La expansión de los anclajes reduce la densidad lineal de la información de secuencias a lo largo del Xpandómero y proporciona una plataforma para aumentar el tamaño y la abundancia de indicadores, que a su vez mejora la señal con respecto al ruido para la detección y descodificación de la secuencia molde.

Aunque el anclaje se representa como una construcción de indicador con tres grupos indicadores, diversas construcciones de indicador pueden estar presentes sobre el anclaje, y pueden comprender indicadores individuales que identifican constituyentes de sonda, indicadores individuales que identifican especies de sondas, códigos de barras moleculares que identifican las especies de sondas, o el anclaje puede ser un polímero desnudo. En algunos casos, uno o más precursores de indicador están presentes sobre el anclaje, y los indicadores se unen por afinidad o se unen covalentemente tras el ensamblaje del producto de Xpandómero.

Las construcciones de sustrato de clase V son similares a las construcciones de clase IV, siendo la diferencia primaria la posición de los conectores escindibles. Las Figuras 23A a 23C se leen de izquierda a derecha, mostrando primero la construcción de sonda-sustrato (precursores de Xsonda o Xmero de Xpandómero), luego la hebra hija del dúplex intermedia en el centro, y a la derecha el producto de Xpandómero preparado para la secuenciación.

La Figura 23A ilustra una construcción de sustrato de clase V que tiene primer y segundo restos terminales de anclaje T (239) unidos con dos enlaces terminales selectivamente escindibles (234, 238) (representados como dos "˘" verticales). Estos enlaces escindibles se unen entonces al primer y segundo restos de sonda,  $P^1$  y  $P^2$ , de un miembro de sonda oligomérico (235). Los grupos conectores  $\epsilon$  (230) y  $\delta$  (231) de dicho primer y segundo restos terminales están posicionados cerca de los extremos  $R^1$  y  $R^2$  de la sonda (de nuevo, las posiciones de estos grupos funcionales pueden estar invertidas). Bajo condiciones controladas, los grupos funcionales  $\delta$  y  $\epsilon$  se hacen reaccionar para formar un enlace  $\chi$ . Estos grupos de enlace están posicionados en la construcción de sustrato para limitar estas reacciones de enlace con construcciones de sustrato confinadas adyacentes. La construcción de sustrato preferencialmente no se enlaza consigo misma. Químicas de enlace y de protección/desprotección adecuadas para  $\delta$ ,  $\epsilon$ , y  $\chi$  se detallan en la descripción general de construcciones oligoméricas.

$R^1$  y  $R^2$  son grupos terminales configurados según convenga para el protocolo de síntesis en el que se usa la construcción de sustrato. Por ejemplo,  $R^1 = 5'$ -fosfato y  $R^2 = 3'$ -OH, encontrarían uso en un protocolo de ligación como se encuentra en Xsondas, y  $R^1 = 5'$ -trifosfato y  $R^2 = 3'$ -OH para un protocolo de polimerasa I como se encuentra en Xmeros. Opcionalmente,  $R^2$  puede configurarse con un grupo de bloqueo reversible para la adición cíclica de un único sustrato. Alternativamente,  $R^1$  y  $R^2$  pueden configurarse con los grupos conectores terminales para el acoplamiento químico o con grupos no conectores para un protocolo de solo hibridación.  $R^1$  y  $R^2$  pueden ser del tipo general XR, en la que X es un grupo de enlace y R es un grupo funcional.

Las construcciones de sustrato son reactivos usados para el ensamblaje dirigido por molde de una hebra hija, una composición intermedia para producir Xpandómeros. La Figura 23B muestra la hebra hija de dúplex, un heterocopolímero con subunidades de repetición (mostradas en corchetes). Se muestran esqueleto primario de la hebra hija ( $-P^1-P^2-$ ) y hebra molde diana ( $-P^1-P^2-$ ) como un dúplex (236). Cada subunidad de la hebra hija es un motivo de repetición que comprende un miembro de sonda y un miembro de anclaje. Los motivos tienen variabilidad específica de especie, indicada aquí por el superíndice  $\alpha$ . Cada subunidad particular en la hebra hija se selecciona de una biblioteca de motivos por un proceso dirigido por molde y su sonda se une a una secuencia correspondiente de nucleótidos complementarios sobre la hebra molde. De esta forma, la secuencia de residuos de nucleobase de las sondas forma una copia complementaria contigua de la hebra molde diana.

La tilde ( $\sim$ ) indica un enlace selectivamente escindible. Los enlaces conectan restos  $P^1$  y  $P^2$  de un miembro de sonda con el anclaje y son necesariamente selectivamente escindibles según se requiera para expandir los anclajes y el

Xpandómero. En una realización, no se forma enlace directo entre las sondas de subunidades separadas.

La hebra hija está compuesta por un precursor de Xpandómero llamado el "Xpandómero limitado" que está adicionalmente compuesto de anclajes en la "configuración limitada". Cuando los anclajes se convierten en su "configuración expandida", como se muestra en la Figura 23C, el Xpandómero limitado se convierte en el producto de Xpandómero. Los anclajes están limitados por los enlaces  $\chi$  formados por conexión con los miembros de anclaje de subunidades adyacentes y por los enlaces selectivamente escindibles (234,238). El enlace  $\chi$  une el miembro de anclaje de una primera subunidad con el anclaje de una segunda subunidad adyacente y se forma enlazando los grupos conectores colocados,  $\delta$  de la primera subunidad y  $\varepsilon$  de la segunda subunidad.

Puede observarse que la hebra hija tiene dos esqueletos, un "esqueleto primario" y el esqueleto del "esqueleto de Xpandómero limitado". El esqueleto primario está compuesto por los sustratos de sonda contiguamente confinados. El "esqueleto de Xpandómero limitado" es el enlace lineal de los anclajes en cada subunidad enlazados juntos por los enlaces  $\chi$  que evitan los sustratos de sonda de subunidad. El enlace  $\chi$  resulta de una reacción del grupo funcional  $\varepsilon$  de una primera subunidad con el grupo funcional  $\delta$  de una segunda subunidad confinada. Puede observarse que el esqueleto de Xpandómero limitado conecta o forma un bucle sobre los enlaces selectivamente escindibles que conectan con el esqueleto primario, y seguirá covalentemente intacto cuando estos enlaces selectivamente escindibles se escindan y el esqueleto primario se disocie o se fragmente de otro modo.

En la Figura 23B, los grupos conectores  $\delta$  y  $\varepsilon$  se han reticulado y ahora forman un enlace  $\chi$  intra-subunidades. Generalmente, la formación del enlace  $\chi$  depende de la colocación del grupo conector  $\delta$ , en la primera subunidad, y el grupo conector  $\varepsilon$ , de una segunda subunidad confinada, de manera que se pongan en contacto durante o después del ensamblaje dirigido por molde de construcciones de sustrato.

En algunos protocolos, la reacción de reticulación depende solo de la hibridación con el molde para juntar los dos grupos reactivos. En otros protocolos, el enlace va precedido del acoplamiento enzimático de los miembros de sonda, con formación de enlaces fosfodiéster entre sondas adyacentes. En la estructura mostrada en la Figura 23B, el esqueleto primario de la hebra hija se ha formado. El anclaje, ahora unido a subunidades adyacentes por enlaces  $\chi$ , y comprende el esqueleto de Xpandómero limitado. Tras la escisión de los enlaces selectivamente escindibles ( $\sim$ ), el Xpandómero limitado se separa del esqueleto primario para convertirse en el producto de Xpandómero, y su anclajes ahora sin limitar se expanden linealmente a su longitud completa como se muestra en la Figura 23C.

A este respecto, la Figura 23C es una representación del producto de Xpandómero de clase V después de la escisión de los enlaces selectivamente escindibles que disocia el esqueleto primario. La hebra del producto de Xpandómero contiene una pluralidad de subunidades  $\kappa$ , en la que  $\kappa$  indica la subunidad  $\kappa^n$  en una cadena de  $m$  subunidades que constituyen la hebra hija, en la que  $m > 3$ , y generalmente  $m > 20$ , y preferentemente  $m > 50$ , y más preferencialmente  $m > 1000$ . Cada subunidad está formada de un anclaje, T (239), como se observa en su configuración expandida y se estira a su longitud entre subunidades adyacentes. Cada subunidad, un motivo de subunidad  $\alpha$ , contiene información genética específica de especie establecida por el ensamblaje dirigido por molde del producto intermedio de Xpandómero (hebra hija).

La Figura 23D muestra la construcción de sustrato de la Figura 23A como modelo molecular, en la que el miembro de sonda (235), representado con cuatro residuos de nucleobase (círculos blancos), está unido al primer y segundo restos terminales del anclaje por dos enlaces escindibles (234,238). El grupo conector (232), mostrado como  $\varepsilon$  en la Figura 23A, es del primer resto terminal del anclaje y el grupo conector (233), mostrado como  $\delta$  en la Figura 23A, es del segundo resto terminal del anclaje. El bucle de anclaje mostrado aquí tiene tres indicadores (237a, 237b, 237c), que también pueden ser específicos de especies de motivos.

La Figura 23E muestra la construcción de sustrato después de la incorporación en el Xpandómero producto. Las subunidades se escinden (234a, 234b, 238a, 238b) y se expanden y se enlazan por enlaces  $\chi$  (249,248), formados enlazando los grupos conectores  $\delta$  y  $\varepsilon$  referidos en la Figura 23A. Una subunidad se indica por líneas de puntos que agrupan verticalmente la subunidad de repetición, como se representa por los corchetes en la Figura 23C adjunta.

En el producto de Xpandómero de la Figura 23E, el esqueleto primario (235) se ha escindido (disociado). Mediante el proceso de escisión, el Xpandómero limitado se libera para convertirse en el producto de Xpandómero. Los miembros de anclaje que estuvieron anteriormente en la configuración limitada están ahora en la configuración expandida, sirviendo así para estirar linealmente la información de secuencias de la diana molde. La expansión de los anclajes reduce la densidad lineal de la información de secuencias a lo largo del Xpandómero y proporciona una plataforma para aumentar el tamaño y la abundancia de indicadores, que a su vez mejora la señal con respecto al ruido para la detección y descodificación de la secuencia molde.

Aunque el anclaje se representa como una construcción de indicador con tres grupos indicadores, diversas construcciones de indicador pueden estar presentes sobre el anclaje, y pueden comprender indicadores individuales que identifican constituyentes de sonda, indicadores individuales que identifican especies de sondas, códigos de barras moleculares que identifican las especies de sondas, o el anclaje puede ser un polímero desnudo. En algunos casos, uno o más precursores de indicador están presentes sobre el anclaje, y los indicadores se unen por afinidad o se unen covalentemente tras el ensamblaje del producto de Xpandómero.



La preparación y uso de un Xpandómero de clase V se ilustra en la Figura 24. En el establecimiento para la síntesis en la etapa I, un molde monocatenario (245) se pone en contacto y se hibrida con cebador de secuenciación (246). El primer ensamblaje (247) se pone entonces en contacto con una biblioteca de construcciones de sustrato de clase V y una polimerasa (etapa II). En la etapa III, los sustratos se han añadido procesivamente en una polimerización dirigida por molde. En la etapa IV, la polimerización del esqueleto primario de la hebra hija está completa y los grupos funcionales reactivos de los brazos laterales del anclaje confinado se reticulan, formando los enlaces  $\chi$  anclaje a anclaje. Finalmente, en la etapa V, los enlaces escindibles en los tallos de los bucles de anclaje se cortan, liberando el esqueleto anclaje a anclaje sintético de la hebra hija oligomérica y molde. Este Xpandómero (249) está así enteramente construido de enlaces de anclaje y se muestra que se expande espontáneamente a medida que se aleja del resto del producto intermedio sintético. Aquí, la información genética correspondiente a la secuencia de polinucleótidos diana está codificada en las subunidades contiguas de los anclajes.

#### Preparación y uso de Xmeros

Las realizaciones de clase I incluyen Xsondas y Xmeros. Las Xsondas son monofosfatos, aunque los Xmeros son trifosfatos. Los "Xmeros" son construcciones de sustrato de trifosfato de oligonucleótidos expansibles que pueden polimerizarse en una síntesis dirigida por molde dependiente de enzima de un Xpandómero. Al igual que las Xsondas, las construcciones de sustrato de Xmero tienen una forma de "sonda-bucle" característica como se ilustra en las Figuras 10A y 10C, en las que R<sup>1</sup> es 5'-trifosfato y R<sup>2</sup> es 3'-OH. Obsérvese que las construcciones de sustrato son trifosfatos de oligonucleobase o trifosfatos análogos a oligómeros, pero los miembros de sonda (es decir, el oligómero) se han modificado con una construcción de anclaje y un enlace selectivamente escindible entre los enlaces terminales del anclaje como se muestra en la Figura 10D, cuya función se ilustra adicionalmente en la Figura 10E.

Las ADN y ARN polimerasas pueden incorporar oligonucleótidos de trifosfato de dinucleótido, trinucleótido y tetranucleótido con un nivel de eficiencia y fidelidad en un proceso procesivo dependiente de cebador como se desvela en la patente de EE.UU. N°. 7.060.440 a Kless. Los trifosfatos de oligonucleótido modificados en el anclaje de longitud n (n = 2, 3, 4, o más) pueden usarse como sustratos para la incorporación basada por polimerasa en Xpandómeros. Enzimas adecuadas para su uso en los métodos mostrados en la Figura 16 y 17 incluyen, por ejemplo, ADN polimerasas dependientes de ADN, ARN polimerasas dependientes de ADN, ADN polimerasas dependientes de ARN, ARN polimerasas dependientes de ARN, ADN polimerasa T7, ADN polimerasa T3, ADN polimerasa T4, ARN polimerasa T7, ARN polimerasa T3, ARN polimerasa SP6, ADN polimerasa I, fragmento de Klenow, ADN polimerasa de *Thermophilus aquaticus*, ADN polimerasa Tth, ADN polimerasa VentR<sup>®</sup> (New England Biolabs), ADN polimerasa Deep VentR<sup>®</sup> (New England Biolabs), fragmento grande de ADN polimerasa Bst, fragmento de Stoeffel, ADN polimerasa 9°N, ADN polimerasa 9°N, ADN Polimerasa Pfu, ADN Polimerasa Tfi, ADN polimerasa Tth, polimerasa RepliPHI Phi29, ADN polimerasa Tii, ADN polimerasa beta de eucariota, telomerasa, polimerasa Terminator<sup>™</sup> (New England Biolabs), ADN polimerasa KOD HiFi<sup>™</sup> (Novagen), ADN polimerasa KOD1, Q-beta replicasa, transferasa terminal, transcriptasa inversa AMV, transcriptasa inversa M-MLV, transcriptasa inversa de Phi6, transcriptasa inversa del VIH-1, polimerasas novedosas descubiertas por bioprospección, y las polimerasas citadas en los documentos US 2007/0048748, US 6329178, US 6602695 y US 6395524. Estas polimerasas incluyen isoformas no mutantes, mutantes, y variantes genéticamente manipuladas.

La polimerización de Xmeros es un método de síntesis de Xpandómeros y se ilustra en la Figura 16, por ejemplo, en la que el sustrato 2mero se proporciona como un trifosfato. Debido a que los Xmeros se polimerizan procesivamente, las etapas de extensión, reticulación, activación de extremo y de lavado de alta rigurosidad normalmente asociadas a secuenciación cíclica por métodos de síntesis se eliminan opcionalmente con este enfoque. Así, la reacción puede realizarse en disolución. La síntesis de Xpandómeros con Xmeros también puede realizarse con moldes inmovilizados, como se ilustra en la Figura 17, en la que un trifosfato Xmero 4mero se polimeriza procesivamente en una síntesis dirigida por molde dependiente de un cebador.

Puede emplearse una variedad de métodos para la síntesis robusta de Xmeros de 5'-trifosfato. Como se describe por Burgess y Cook ("Synthesis of Nucleoside Triphosphates", Chem. Rev. 100(6):2047-2060, 2000), estos métodos incluyen (pero no se limitan a) reacciones que usan fosforamiditos de nucleósido, síntesis mediante el ataque nucleófilo de pirofosfato sobre monofosfatos de nucleósido activados, síntesis mediante el ataque nucleófilo de fosfato sobre pirofosfato de nucleósido activado, síntesis mediante ataque nucleófilo de difosfato sobre fosfato sintona activado, síntesis que implica fosfitos o fosforamiditos activados derivados de nucleósidos, síntesis que implica el desplazamiento directo de grupos salientes 5'-O- por nucleófilos de trifosfato, y métodos biocatalíticos. Un método representativo para producir sustratos de dinucleótido compatibles con polimerasa usa N-metilimidazol para activar el grupo 5'-monofosfato; la posterior reacción con pirofosfato (sal de tributilamonio) produce el trifosfato (Abramova et al., "A facile and effective synthesis of dinucleotide 5'-triphosphates", Bioorganic and Med Chem 15, 6549-6555, 2007).

Como se trata más abajo en más detalle, la construcción de anclaje de Xmero está relacionada en diseño, composición y enlace con la del anclaje usado para las Xsondas. En muchas realizaciones, la información genética está codificada sobre el anclaje y, por tanto, cada anclaje de cada construcción de sustrato es un anclaje específico de especie. La información codificada sobre el anclaje está codificada con un código de indicador que digitaliza la información genética. Por ejemplo, la codificación binaria de cinco bits en los anclajes produciría 32 códigos de

5 secuencia únicos (2<sup>5</sup>). Esta estrategia puede usarse para codificar únicamente las 16 combinaciones de dos residuos de nucleobase por miembro de sonda de una biblioteca 2mera, independientemente de la orientación del anclaje. Similar a la codificación de Xsondas, puede considerarse una variedad de estrategias de funcionalización y de marcado para Xmeros, que incluye (pero no se limita a): dendrímeros funcionalizados, polímeros, polímeros ramificados, nanopartículas y nanocristales como parte del andamiaje del anclaje, además de químicas de indicador y señales de indicador - que van a detectarse con la tecnología de detección apropiada. Las marcas específicas de base pueden introducirse (mediante la unión al anclaje) tanto antes de como después de la polimerización del Xmero, por enlace covalente o por dirigido por afinidad.

#### Diseño y síntesis de Xsondas y Xmeros

10 Una visión general de estrategias sintéticas y de escisión se presentan a continuación, empezando con los oligómeros de sonda con enlaces selectivamente escindibles, seguido de las anclaje y construcciones de anclaje de indicador.

15 Un objetivo de un método de SBX basado en Xsondas o Xmeros es ensamblar una réplica del ácido nucleico diana tan completa y eficazmente como sea posible por una síntesis dirigida por molde, generalmente un proceso o combinación de procesos seleccionados de hibridación, ligación, polimerización o reticulación química de composiciones de precursor adecuadas, llamadas aquí "sustratos". Los sustratos de Xsondas y Xmeros se suministran como bibliotecas de reactivos (por ejemplo, como partes de kits para la secuenciación) para este fin. Las bibliotecas son generalmente de naturaleza combinatoria y contienen miembros de sonda seleccionados para unirse específicamente a cualquiera o todas de las secuencias complementarias, tales como se encontrarían en un polinucleótido diana. El número de sondas requeridas en una biblioteca para este fin es una función del tamaño de la sonda. Puede considerarse que cada sonda es un fragmento de secuencia, y debe estar presente suficiente variedad de miembros de sonda para formar una copia contigua de la secuencia contigua de fragmentos de secuencias complementarias del polinucleótido diana. Para sondas en las que cada oligómero es un dímero, existen 20 16 combinaciones de especies posibles de A, T, C y G. Para sondas en las que cada oligómero es un trímero, entonces existen 64 combinaciones de especies posibles de A, T, C y G, etc. Cuando se secuencian fragmentos genómicos al azar, es probable que todas aquellas especies sean requeridas en una biblioteca de reactivos.

25 Las Xsondas y Xmeros son construcciones oligoméricas de sustrato que se dividen en cinco clases funcionales diferentes. Las construcciones oligoméricas de sustrato tienen dos componentes funcionales distintos: una oligonucleobase modificada o miembro de "sonda", y un miembro de anclaje ("T"). La sonda está unida al miembro de anclaje por una construcción de "sonda-bucle", en la que el bucle de anclaje es un precursor del miembro de anclaje linealizado del Xpandómero de producto final. Cada anclaje T puede codificarse con indicadores (comúnmente denominados "etiquetas" o "marcas"), o combinaciones de los mismos, que identifican únicamente la secuencia de sonda con la que está anclado. De esta forma, la información de secuencias del Xpandómero ensamblado se detecta más fácilmente.

30 El oligómero es la porción de sonda de la Xsonda. La sonda es una oligonucleobase modificada que tiene una cadena de x desoxirribonucleótidos, ribonucleótidos, o más generalmente, residuos de nucleobase (en la que x puede ser 2, 3, 4, 5, 6, o más). En estas discusiones, una sonda con 2, 3, 4, 5 o 6 residuos de nucleobase de longitud puede denominarse una 2mera, 3mera, 4mera, 5mera, o 6mera, respectivamente.

35 Los reactivos de construcciones de sustrato pueden sintetizarse con un esqueleto 5'-3' fosfodiéster de oligonucleótido, cuyo oligómero tiene los nucleótidos A, T, G y C (estructuras mostradas en la tabla de la Figura 25), u otros análogos de ácidos nucleicos hibridables tales como aquellos que tienen un esqueleto de péptido, esqueleto de fosfono-péptido, esqueleto de serina, esqueleto de hidroxiprolina, esqueleto mixto de péptido-fosfono-péptido, esqueleto mixto de péptido-hidroxiprolina, esqueleto mixto de hidroxiprolina-fosfono-péptido, esqueleto mixto de serina-fosfono-péptido, esqueleto de treosa, esqueleto de glicol, esqueleto de morfolino, y similares, como se conoce en la técnica. Los oligómeros de ácido desoxirribonucleico y oligómeros de ácido ribonucleico, y oligómeros mixtos de los dos, también pueden usarse como sondas. Otras bases también pueden estar sustituidas, tales como uracilo por timidina, e inosina como base degenerada. También pueden usarse residuos fragmentados de nucleobases que tienen complementariedad.

40 Una recitación más completa de bases degeneradas e inestables conocidas en la técnica incluye, pero no se limita a, xantina, hipoxantina, o un derivado heterocíclico, análogo, o tautómero de xantina e hipoxantina, 8-azapurina, purinas sustituidas en la posición 8 con metilo o bromo, 9-oxo-N<sup>6</sup>-metiladenina, 2-aminoadenina, 7-deazaxantina, 7-deazaguanina, 7-deaza-adenina, N<sup>4</sup>-etanocitosina, 2,6-diaminopurina, N<sup>6</sup>-etano-2,6-diaminopurina, 5-metilcitosina, 5-alquilil (C3-C6)-citosina, 5-fluorouracilo, 5-bromouracilo, tiouracilo, 2-hidroxi-5-metil-4-triazolopiridina, isocitosina, pseudoisocitosina, isoguanina, 7,8-dimetilaloxazina, 6-dihidrotimina, 5,6-dihidrouracilo, 4-metil-indol, etenoadenina y las nucleobases descritas en las patentes de EE.UU. N° 5.432.272 y 6.150.510, PCT publicadas WO 92/002258, WO 93/10820, WO 94/22892 y WO 94/22144, y en Fasman, Practical Handbook of Biochemistry and Molecular Biology, pp. 385-394, CRC Press, Boca Raton, LA, 1989.

55 Como se conoce en la técnica, los oligómeros pueden diseñarse para incluir modificadores de nucleótidos. En algunas realizaciones, éstos sirven de puntos de unión para el miembro de anclaje o miembros. Derivados de purina

y pirimidina adecuados para la síntesis de oligómeros derivatizados son muy conocidos en la técnica. Dos de tales bases modificadas representativas se muestran en las Figuras 26A y 26B, en las que se representan un derivado de citosina modificado con 5-amino y un residuo de guanina modificado con 8-amino.

5 Como se ilustra en las Figuras 27A y 27B, tomando un oligómero 4mero como ejemplo (aquí ilustrado como 5'-monofosfato), cualesquiera dos de las cuatro posiciones de base sobre el oligómero pueden modificarse para crear puntos de unión por químicas conocidas. Los nucleótidos modificados en los residuos de sonda 2 y 3 (sobre lados opuestos de un enlace selectivamente escindible, representado como "V") se ilustran en la Figura 27A. Esta figura ilustra un oligómero 4mero con conectores amino unidos a citosina y guanosina del oligómero. La Figura 27B ilustra un oligómero 4mero con grupos funcionales benzaldehído con la citosina y guanosina del oligómero. Los detalles son ilustrativos de métodos muy conocidos en la técnica. Para simplicidad, la mayoría de las ilustraciones proporcionadas en el presente documento supondrán 4meros, a menos que se indique lo contrario, pero se entiende que pueden emplearse otras bibliotecas de construcciones de sustrato o combinaciones de bibliotecas en la práctica de la presente invención.

### Escisión

15 Generalmente, las construcciones de sustrato de Xsondas y Xmero tienen enlaces selectivamente escindibles que permiten la expansión controlada del anclaje. Como se ha citado previamente, tal escisión selectiva puede lograrse por cualquier número de técnicas conocidas para un experto en la materia, que incluyen, pero no se limitan a, escisión del esqueleto de fosfortiolato con cationes metálicos, escisión por ácido de modificaciones del esqueleto de fosforamidoato, escisión selectiva por nucleasa de enlaces fosfodiéster estándar usando modificaciones de fosfortioato resistentes a nucleasas para la protección de esqueletos, fotoescisión de conectores de esqueleto modificado con nitrobenzilo, y reducción de enlaces disulfuro.

20 La modificación de sondas de sustrato para incluir enlaces selectivamente escindibles se ilustran en las Figuras 28A a 28D y en las Figuras 29A a 29D. La Figura 28A muestra un ejemplo de un dímero de Xsonda con un grupo ribosil 2'-OH susceptible a escisión por ribonucleasa H en un producto intermedio de Xpandómero de dúplex de ADN/ARN. El enlace es así selectivamente escindible, siempre que el (los) otro(s) nucleótido(s) en la Xsonda sean resistentes a la escisión por RNasa (por ejemplo, nucleobases de 2'-o-metil-pentosa, 2'-desoxirribosa, nucleobases de LNA "bloqueado", y nucleobases enlazadas por glicol o péptido). Estos sitios de escisión tienen otros usos, por ejemplo, un ribonucleótido en la penúltima 5'-nucleobase y un adaptador proporciona un conector escindible entre el Xpandómero y un soporte inmovilizado.

30 La Figura 28B muestra una Xsonda con enlace fosfodiéster que acopla dos nucleótidos. En esta figura, además de las Figuras 28A, 28C y 28D, los anclajes para conectar el enlace selectivamente escindible de la sonda se indican en (282) y (284). Este enlace es selectivamente escindible con nucleasa de judía mungo, nucleasa S1, DNasa I, u otras DNasas, por ejemplo, si otros enlaces que unen los anclajes de subunidad juntos son resistentes a nucleasa. La síntesis de una biblioteca 2mera, por ejemplo, con un enlace de fosfato estándar entre los puntos de unión del anclaje y el (los) enlace(s) fosfortioato en la(s) posición (posiciones) del esqueleto de nucleótidos que va(n) a quedar intactas, proporciona el patrón de escisión deseado. La Figura 28C es un dímero de Xsondas mantenido junto por un enlace 3'-fosfortiolato, y en la Figura 28D se mantiene junto por un enlace 5'-fosfortiolato. Estos enlaces son selectivamente escindibles por ataque químico, por ejemplo, con yodoetanol como se describe por Gish et al. ("DNA and RNA sequence determination based on phosphorothioate chemistry", *Science* 240(4858): 1520-1522, 1988) o por escisión con cationes metálicos divalentes como se describe por Vyle et al. ("Sequence- and strand-specific cleavage in oligodeoxyribonucleotides and DNA containing 3'-thiothymidine". *Biochemistry* 31(11): 3012-8, 1992). Otras opciones de escisión de esqueleto incluyen, pero no se limitan a, escisión fotorédox inducida por UV (como por adaptación de grupos de fotoescisión de nitrobenzilo) como se describe por Vallone et al. ("Genotyping SNPs using a UV-photocleavable oligonucleotide in MALDI-TOF MS", *Methods Mol. Bio.* 297:169-78, 2005), escisión por ácido de enlaces fosforamidoato como se describe por Obika et al. ("Acid-Mediated Cleavage of Oligonucleotide P3'→N5' Phosphoramidates Triggered by Sequence-Specific Triplex Formation", *Nucleosides, Nucleotides and Nucleic Acids* 26(8,9): 893-896, 2007) y escisión catalizada por peryodato de modificaciones del esqueleto de 3'-O-B-D-ribofuranosil-2'-desoxi como se desvela por Nauwelaerts et al. ("Cleavage of DNA without loss of genetic information by incorporation of a disaccharide nucleoside", *Nucleic Acids Research* 31(23): 6758-6769, 2003).

55 Al igual que con las Xsondas, la escisión del esqueleto de poli-Xmero para producir un Xpandómero se lleva a cabo en una variedad de formas. Como se muestra en la Figura 29A, por ejemplo, un Xmero que contiene una base de ribonucleótido digerible con RNasa puede escindirse selectivamente en aquella posición siempre que el (los) otro(s) nucleótido(s) en el Xmero sean resistentes a la escisión por RNasa (por ejemplo, nucleobases de 2'-O-metilpentosa y 2'-desoxirribosa, nucleobases de LNA "bloqueado" y nucleobases enlazadas por glicol o péptido). Para el Xmero descrito en la Figura 29A, la base en 5' es un ribonucleótido de 2'-hidroxilo estándar citidina y la base de 3' es un desoxirribonucleótido de 2' resistente a RNasa guanina. El diseño del Xmero permite la escisión selectiva por RNasa del esqueleto de Xmero para expandir un Xpandómero. Alternativamente, como se muestra en la Figura 29B, la DNasa puede usarse para escindir todos los enlaces de esqueleto protegidos no de fosfortioato. Por consiguiente, una biblioteca 2mera, por ejemplo, con un enlace fosfato estándar entre los puntos de unión de anclaje y el (los) enlace(s) fosfortioato en la(s) posición (posiciones) del esqueleto de nucleótidos que van a permanecer intactas,

proporciona el patrón de escisión deseado. La Figura 29C es un dímero de Xmero mantenido junto por un enlace 3'-fosforotiolato, y la Figura 29D se mantiene junta por un enlace 5'-fosforotiolato. Estos enlaces son selectivamente escindibles por ataque químico, por ejemplo, con yodoetanol o por escisión con cationes metálicos divalentes como se ha citado previamente. Otras opciones de escisión del esqueleto incluyen (pero no se limitan a) escisión foto-redox inducida por UV (como por adaptación de grupos de fotoescisión de nitrobenzilo) y escisión por ácido de enlaces fosforamidato, ambos de los cuales se han citado anteriormente en la Figura 28. En las Figuras 29A a 29D, los anclajes para conectar el enlace selectivamente escindible de la sonda se indican en (292) y (294).

Volviendo ahora a la Figura 30, en una primera realización general de un esquema para la síntesis de la construcción de sustrato de clase I de "sonda-bucle", dos residuos de nucleobase (círculos) en la segunda y tercera posiciones sobre la sonda se modifican para crear puntos de unión L1 y L2 para los dos extremos L1' y L2' del anclaje. El anclaje se muestra aquí como pre-ensamblado por separado y está unido al miembro de sonda en una etapa sintética (flecha). Los enlaces disulfuro intra-anclaje (representados por los dos triángulos) pueden usarse en el ensamblaje y uso de estas construcciones de sustrato. La introducción de un agente reductor al producto de Xpandómero romperá selectivamente los puentes de disulfuro que mantienen juntos el anclaje, permitiendo así la expansión del esqueleto del Xpandómero. También son útiles enlaces fotoescindibles en los anclajes de plegamiento durante el ensamblaje con posterior liberación y desplegado tras la exposición a luz.

En otras realizaciones, el esqueleto de fosfodiéster del sustrato puede modificarse para crear puntos de unión para el anclaje como se ha desvelado por Cook et al. ("Oligonucleotides with novel, cationic backbone substituents: aminoethylphosphonates", *Nucleic Acids Research* 22(24): 5416-5424, 1994), Agrawal et al. ("Site specific functionalization of oligonucleotides for attaching two different reporter groups", *Nucleic Acids Research* 18(18): 5419-5423, 1990), De Mesmaeker et al., ("Amide backbone modifications for antisense oligonucleotides carrying potential intercalating substituents: Influence on the thermodynamic stability of the corresponding duplexes with RNA- and DNA-complements", *Bioorganic & Medicinal Chemistry Letters* 7(14): 1869-1874, 1997), Shaw et al. (Boranophosphates as mimics of natural phosphodiester in DNA", *Curr Med Chem.* 8(10):1147-55, 2001), Cook et al. (patente de EE.UU. N° 5.378.825) y Agrawal ("Functionalization of Oligonucleotides with Amino Groups and Attachment of Amino Specific Reporter Groups", *Methods in Molecular Biology* Vol. 26, 1994). Los residuos de nucleobase que constituyen el miembro de sonda pueden estar sustituidos con análogos de nucleobase para alterar la funcionalidad de Xsondas. Por ejemplo, pueden usarse ácidos nucleicos bloqueados ("LNA") para aumentar la estabilidad del dúplex de sonda. Si está previsto el acoplamiento químico de Xsondas (en lugar de ligación enzimática), los extremos de sonda 5' y 3' pueden derivatizarse adicionalmente para permitir la reticulación química.

#### Diseño, composición y síntesis de construcciones de indicador

En una realización, los anclajes se codifican con "construcciones de indicador" que únicamente identifican la secuencia de residuos de nucleobase (o "sonda" de Xsondas, Xmeros, y otros sustratos de oligómero de la Figura 8) o nucleobase (como en XNTP, RT-NTP y sustratos monoméricos de la Figura 9) con los que está anclado. Los indicadores son indicadores o combinaciones de indicadores generalmente asociadas a los anclajes que sirven para "analizar" o "codificar" la información de secuencias inherente en los sustratos e inherente en el orden en el que los sustratos se incorporan en el Xpandómero. En algunas realizaciones, el anclaje es solo un espaciador y los indicadores son, o están asociados con, el sustrato.

La Figura 31 representa un método de ensamblaje del anclaje de sustrato similar al de la Figura 30, pero el anclaje pre-ensamblado incluye grupos indicadores (mostrados como las tres porciones rectangulares del anclaje) y se llama una "construcción de indicador". Las construcciones de indicador y anclajes pueden prepararse mediante una variedad de químicas de polímeros, y su uso y síntesis se trata en más detalle aquí.

En la práctica de la presente invención, los anclajes pueden servir para una variedad de funciones, por ejemplo: (1) como anclaje para enlazar secuencialmente, directa o indirectamente, anclajes adyacentes a lo largo del esqueleto de nucleobases, (2) como espaciador para estirar o expandir de manera que se forme una cadena alargada de subunidades ancladas, denominado un Xpandómero, tras la escisión del esqueleto, y/o (3) opcionalmente comprende construcciones de indicador o precursores de indicador que codifican la información de nucleobases o de secuencias oligoméricas de la construcción de sustrato individual a la que el anclaje está asociado.

Las construcciones de indicador son manifestaciones físicas de códigos de indicador, que son de naturaleza bioinformacional y digital. Los códigos de indicador analizan o codifican la información genética asociada al fragmento de secuencia de la sonda o nucleobase con la que está unida la construcción de indicador y el anclaje. Las construcciones de indicador se diseñan para optimizar la detectabilidad del código de indicador ajustando separaciones espaciales, abundancia e intensidad de señales de los indicadores constituyentes. Las construcciones de indicador pueden incorporar un amplio intervalo de elementos de señal y estructurales que incluyen, pero no se limitan a, polímeros, dendrímeros, perlas, aptámeros, ligandos y oligómeros. Estas construcciones de indicador se hacen mediante una variedad de químicas de polímero y se tratan adicionalmente más adelante.

En una realización, las construcciones de indicador están unidas a la sonda o nucleobase por un anclaje de polímero. Los anclajes pueden construirse a partir de uno o más polímeros duraderos solubles en agua o disolventes que incluyen, pero no se limitan a, el siguiente segmento o segmentos: polietilenglicoles, poliglicoles, polipiridinas,

poliisocianuros, poliisocianatos, poli(triarilmetil)metacrilatos, polialdehídos, polipirrolinonas, poliureas, fosfodiésteres de poliglicol, poliacrilatos, polimetacrilatos, poliacrilamidas, polivinilésteres, poliestirenos, poliamidas, poliuretanos, policarbonatos, polibutiratos, polibutadienos, polibutirolactonas, polipirrolidinonas, polivinilfosfonatos, poliacetamidas, polisacáridos, polihialuronatos, poliamidas, poliimidias, poliésteres, polietilenos, polipropilenos, poliestirenos, policarbonatos, politereftalatos, polisilanos, poliuretanos, poliéteres, poliaminoácidos, poliglicinas, poliprolinas, polilisina N-sustituida, polipéptidos, péptidos sustituidos con N en la cadena lateral, glicina poli-N-sustituida, peptoides, péptidos sustituidos con carboxilo en la cadena lateral, homopéptidos, oligonucleótidos, oligonucleótidos de ácido ribonucleico, oligonucleótidos de ácido desoxinucleico, oligonucleótidos modificados para prevenir el apareamiento de bases de Watson-Crick, análogos de oligonucleótidos, ácido policitidílico, ácido poliadenílico, ácido poliuridílico, politimidina, polifosfato, polinucleótidos, polirribonucleótidos, fosfodiésteres de polietilenglicol, análogos de péptido-polinucleótido, análogos de treosil-polinucleótido, análogos de glicol-polinucleótido, análogos de morfolino-polinucleótido, análogos de oligómeros de nucleótidos bloqueados, análogos de polipéptidos, polímeros ramificados, polímeros en peine, polímeros en estrella, polímeros dendríticos, copolímeros aleatorios, en gradiente y de bloque, polímeros aniónicos, polímeros catiónicos, polímeros que forman tallo-bucles, segmentos rígidos y segmentos flexibles. Tales polímeros pueden estar circularizados en puntos de unión sobre una construcción de sustrato como se describe en, por ejemplo, en las Figuras 30 y Figura 31.

El anclaje es generalmente resistente al entrelazamiento o se pliega de manera que sea compacto. Polietilenglicol (PEG), poli(óxido de etileno) (PEO), metoxipolietilenglicol (mPEG), y una amplia variedad de derivados de PEG similarmente construidos (PEG), son polímeros ampliamente disponibles que pueden utilizarse en la práctica de la presente invención. Los PEG modificados están disponibles con una variedad de reticuladores de extremos bifuncionales y heterobifuncionales y se sintetizan en un amplio intervalo de longitudes. Los PEG son generalmente solubles en agua, metanol, benceno, diclorometano, y muchos disolventes orgánicos comunes. Los PEG son polímeros generalmente flexibles que normalmente no interactúan específicamente con productos químicos biológicos.

La Figura 32A ilustra la estructura de repetición de un polímero de PEG. La Figura 32B muestra una Xsonda o Xmero con un anclaje de PEG desnudo asegurado al esqueleto de sonda por, por ejemplo, conectores terminados en amina (no mostrados) que usan químicas de conectores estándar. La Figura 32C muestra la misma construcción de sustrato después de la escisión del esqueleto de sonda en un enlace selectivamente escindible ("V"), acomodando el polímero de PEG flexiblemente el alargamiento del Xpandómero. En algunas realizaciones, los segmentos del polímero de PEG se ensamblan a trozos sobre el anclaje para proporcionar la longitud de expansión o para minimizar cuestiones estéricas, tales como, por ejemplo, en los tallos de los anclajes próximos a los enlaces de extremos terminales que conectan los brazos de anclaje con el sustrato.

Otros polímeros que pueden emplearse como anclajes, y proporcionan "andamiaje" para indicadores, incluyen, por ejemplo, poli-glicina, poli-prolina, poli-hidroxiprolina, poli-cisteína, poli-serina, ácido poli-aspartico, ácido poli-glutámico, y similares. Pueden usarse las funcionalidades de la cadena lateral para construir andamiajes ricos en grupos funcionales para capacidad o complejidad de señales añadida.

La Figura 33A muestra la estructura de poli-lisina. En las construcciones de anclaje realizadas descritas en la Figura 33B a 33D, los segmentos de anclaje de poli-lisina crean un andamiaje para la unión del indicador. En la Figura 33B, los grupos  $\epsilon$ -amino de las cadenas laterales de lisina (indicados por flechas) proporcionan funcionalidad para la unión de pluralidades de elementos indicadores a una construcción de sustrato, amplificando el código de indicador. La Figura 33C ilustra un dendrímico en estrella unido a una construcción de sustrato con cadenas laterales de poli-lisina (flechas).

La Figura 33D ilustra la carga de los oligómeros de dendrímico que puede detectarse añadiendo oligómeros complementarios marcados en una etapa de marcado post-ensamblaje y de "amplificación de señal". Esto proporciona un método útil para preparar un anclaje universal uniendo un complejo dendrímico sin marcar con múltiples grupos indicadores oligoméricos a una sonda, y luego tratando el dendrímico unido a la sonda con una selección de una o dos sondas marcadas complementarias, se obtiene un dendrímico "pintado" específico para las especies de sustrato individuales. Pueden pintarse diferentes construcciones de sonda/anclaje con diferentes sondas marcadas complementarias.

En otra realización de este enfoque, el esqueleto del sistema de indicador comprende ocho oligonucleótidos únicos que están espacialmente codificados de un modo binario usando dos indicadores fluorescentes distinguibles, cada uno excitado mediante el mismo donante de FRET. Antes de o tras el acoplamiento de las construcciones de anclaje a su construcción de sustrato respectiva, las construcciones de anclaje se codifican por secuencia por hibridación de la mezcla apropiada de elementos indicadores fluorescentes para crear el código binario específico de sonda apropiado. Se emplean variaciones de este enfoque usando dendrímicos codificados, sin marcar, polímeros, polímeros ramificados, o perlas como esqueleto del sistema de indicador. Los oligonucleótidos pueden de nuevo usarse para la codificación binaria de la construcción de indicador. Una ventaja con este enfoque es que la intensidad de señales se amplifica significativamente y que la codificación no es dependiente tras un único acontecimiento de hibridación, ambos de los cuales disminuyen la posibilidad de medición y/o error de codificación.

60

Todavía otra realización sustituye la estrategia de codificación de oligonucleótidos anteriormente descrita con ligandos heteroespecíficos unidos por afinidad para producir, por ejemplo, una construcción de indicador codificada similarmente binaria. Usando una estrategia de codificación binaria de 9 bits, esta construcción de indicador no universal, en su forma más simple, emplea solo una única química de acoplamiento para marcar simultáneamente todos los anclajes.

Dada la flexibilidad del enfoque de SBX, se usa una amplia variedad de indicadores para producir señales medibles únicas. Cada anclaje está únicamente codificado por uno o muchos segmentos de indicador distintos. El andamiaje con el que los restos de indicador están unidos puede construirse usando una amplia variedad de características estructurales existentes que incluyen, pero no se limitan a, dendrímeros, perlas, polímeros y nanopartículas. Dependiendo del esquema de codificación, pueden usarse uno o muchos andamiajes de indicador distintamente separados para el código de indicador de cada anclaje. Está disponible cualquier número de opciones para la unión directa e indirecta de restos de indicadores con el andamiaje de indicador, que incluyen (pero no se limitan a): codificación de indicador de polímero(s) químicamente reactivo(s) integrado(s) en las construcciones de anclaje; codificación de indicador de grupos superficiales químicamente reactivos sobre dendrímero(s) integrado(s) en el esqueleto de anclaje; y codificación de indicador de grupos superficiales químicamente reactivos sobre perla(s) integrada(s) en el anclaje. En este contexto, una "perla" se considera ampliamente para indicar cualquier partícula o microesfera cristalina, polimérica, de látex, o de material compuesto. Para los tres ejemplos, la abundancia de indicador puede aumentarse significativamente uniendo, con los andamiajes de indicador, polímeros que se cargan con múltiples de indicadores. Estos polímeros pueden ser tan simples como un poli-lisina de 100 residuos o más avanzado, tal como sondas oligoméricas marcadas.

También pueden usarse construcciones de anclaje de tamaño reducido. Por ejemplo, las construcciones de anclaje pueden alargarse en una etapa post-procesamiento usando métodos dirigidos para insertar unidades de separación, reduciendo así el tamaño del anclaje de indicador.

La reducción del tamaño y masa de la construcción de sustrato también puede lograrse usando anclajes sin marcar. Eliminando los indicadores voluminosos (y el andamiaje de indicador tal como dendrímeros, que para algunas realizaciones codificantes comprende más del 90 % de la masa del anclaje), pueden potenciarse la cinética de hibridación y/o de acoplamiento. Entonces pueden emplearse marcado del anclaje post-ensamblaje. Los indicadores están unidos a una o más químicas de enlace que están distribuidas a lo largo de las construcciones de anclaje usando estrategias espaciales o combinatorias para codificar la información de secuencias base. Un esquema de codificación binaria simple puede usar solo una química de enlace reactivo para el marcado del post-ensamblaje del Xpandómero. Esquemas de marcado más complicados, que pueden requerir cientos de conectores únicos, usan una estrategia basada en oligonucleótidos para el marcado de Xpandómeros. Otra realización de Xsondas o Xmeros post-marcado es usar las secuencias de nucleótidos resultantes derivadas de P<sup>1</sup> y P<sup>2</sup> (véase la Figura 10) que siguen tras la escisión y la expansión del Xpandómero para la unión del indicador por hibridación de una biblioteca de sondas marcadas. Similarmente, otras técnicas de marcado y/o de detección pueden identificar directamente los nucleótidos más espacialmente resueltos.

Se emplean anclajes y construcciones de indicador que se unen a las construcciones de sustrato con un nivel de exactitud deseado, ya que el acoplamiento erróneo conduce a ineficiencias en la detección, y también puede conducir a la terminación de polímeros o reordenamiento del código de indicador (por ejemplo, si se usa un código de indicador asimétrico). La fidelidad del proceso de SBX puede ligarse a la pureza de síntesis de las construcciones de sustrato. Tras la purificación de la construcción de anclaje/indicador para enriquecer el producto de longitud completa, la construcción puede acoplarse directamente a una sonda de oligonucleótidos heterobifuncional (codificación de indicador direccional) u homobifuncional (codificación de indicador simétrica u orientada). Al igual que con todos los métodos de síntesis de polímeros, la purificación (tamaño, afinidad, HPLC, electroforesis, etc.) se utiliza tras completarse la síntesis de construcciones de sustrato y el ensamblaje para garantizar la alta pureza de construcciones de sonda expansibles de longitud completa.

#### Síntesis de construcciones de sustratos de clase I que presentan indicadores

La síntesis de construcciones de sustrato de clase I con indicadores o precursores de indicador presentados sobre los anclajes puede llevarse a cabo en una variedad de formas. Puede usarse un proceso escalonado para ensamblar un polímero de anclaje en horquilla que está conectado cerca de los extremos de unión del anclaje de la construcción de sustrato mediante un puente de disulfuro. Esto orienta los extremos reactivos de forma que el acoplamiento del anclaje a la sonda esté altamente favorecido. Están comercialmente disponibles fosforamiditos con modificador de amino C6 para los cuatro nucleótidos (Glen Research, EE.UU.) y se usan para unir el anclaje para formar, por ejemplo, la construcción de sustrato completada. Alternativamente, puede emplearse química de conectores en forma de nucleótidos modificados con benzaldehído. La purificación por tamaño y/o afinidad es útil para enriquecer las construcciones de sustrato correctamente ensambladas.

La heterobifuncionalización de sondas puede hacerse ventajosamente sobre una matriz de soporte sólido, como es habitual para la síntesis de oligos y de péptidos, o en disolución con métodos de purificación apropiados. Están disponibles una amplia variedad de reactivos de reticulación heterobifuncionales y homobifuncionales listos para uso para modificar, por ejemplo, restos amina, carboxilo, tiol e hidroxilo, y para producir una variedad de químicas de

enlace robustas y selectivas. Como los modificadores de C6 amino están disponibles para los cuatro desoxirribonucleótidos, y pueden ponerse a disposición para los cuatro ribonucleótidos, las estrategias de funcionalización descritas aquí usan métodos de reticulación basados en amina listos para uso junto con químicas de protección/desprotección de amina bien establecidas. Sin embargo, dado la amplia variedad de químicas de fosforamido y de reticulación conocidas en la técnica, los métodos no descritos aquí también pueden considerarse y producir productos equivalentes.

La necesidad de heterobifuncionalización de la sonda puede eliminarse si la estrategia de codificación del indicador produce codificación digitalmente simétrica o usa puntos de referencia direccionales (bits de paridad) para identificar la orientación de códigos. En este caso, son suficientes dos modificadores de sonda de amina interna, ya que cualquier orientación de acoplamiento de las construcciones de indicador en el anclaje produciría una identificación de secuencias específica de sonda única.

Pueden ensamblarse uno o muchos segmentos de polímero secuencialmente usando enlaces covalentes catalizados por productos químicos (por ejemplo, reticuladores) o enzimas (por ejemplo, hibridación y ligación de sondas de ácido nucleico) para formar un anclaje circular funcionalizado en los extremos. Dado el actual estado de la técnica para los métodos de síntesis de polímeros, el enfoque de síntesis de reticulaciones químicas constituye una realización representativa. Como es habitual para muchos métodos de síntesis de polímeros, puede usarse una matriz de soporte sólido como andamiaje para la síntesis. Los segmentos de polímeros pueden ensamblarse una cada vez basándose en la funcionalización de extremos, como segmentos de pares mixtos que tienen funcionalizaciones de extremos diferentes, o como pares unidos - dos segmentos de polímero con diferentes restos terminales homobifuncionales (por ejemplo, hidrazida y amina) emparejados mediante puentes de disulfuro.

Las químicas de marcado, que incluyen tanto los restos de conector como de elemento indicador, se desarrollan y optimizan basándose en el alto rendimiento de señales y estabilidad, baja reactividad cruzada de polímeros y entrelazamiento, y la rigidez estructural (refuerzo) que estas químicas confieren al esqueleto de Xpandómero, que puede ser importante para la preparación de muestras y detección como se trata más adelante.

En la Figura 31 tratada anteriormente, la construcción de anclaje en horquilla/de indicador completa se ensambla independientemente de la sonda de oligonucleótidos y luego se une por químicas de conector homobifuncional o heterobifuncional al miembro de sonda. En una realización alternativa, como se muestra en la Figura 34, los anclajes se circularizan escalonadamente por construcción sobre secuencias de sondas inmovilizadas. La construcción de indicador y anclaje se sintetizan con segmentos de anclaje heterobifuncionales (direccionales) u homobifuncionales (simétricos u orientados) enlazados directamente a una sonda de oligonucleótidos. La secuencia de sonda mostrada en la Figura 34A es un 4mero e incluye restos de sonda  $P^1$  y  $P^2$  (el segundo y tercer círculos) separados por un enlace selectivamente escindible ("V"). Se usan técnicas de síntesis en estado sólido en la síntesis de la construcción de indicador. Esta síntesis se integra con cierre del bucle de anclaje. En la etapa I de la Figura 34A, un primer segmento de anclaje (341) con primer grupo indicador (342) se añade usando química de grupos funcionales específicos indicados por  $L1$  y  $L1'$  (el conector  $L1$  sobre uno de los restos de sonda se bloquea, como se representa por el pequeño rectángulo). En la etapa II, un segundo segmento de anclaje (344) con segundo grupo indicador (345) se añade usando química de grupos funcionales específicos indicados por  $L2'$  y  $L2'$ . En la etapa III, un tercer segmento de anclaje (346) con grupo indicador (347) se añade usando química de grupos funcionales específicos indicados por  $L2$  y  $L1'$ . En la etapa IV, y tras la eliminación del grupo de bloqueo del sitio  $L1$  sobre el resto  $P^2$  de la sonda (de nuevo representado por el rectángulo pequeño), el bucle se cierra después del acoplamiento de  $L1'$  y  $L1$ .

En otra realización más, como se ilustra en la Figura 34B, la química de conectores heterobifuncionales puede usarse de nuevo para garantizar que el anclaje esté direccionalmente posicionado sobre la sonda (aunque esto no es necesario para todas las estrategias codificantes). Una sonda 4mera se representa de nuevo con los restos de sonda  $P^1$  y  $P^2$  (el segundo y tercer círculos) separados por un enlace selectivamente escindible ("V"). En la etapa I, dos anclajes (341,344) con segmentos de indicador (342,345) se ponen en contacto con grupos funcionales  $L1$  y  $L2$  sobre  $P^1$  y  $P^2$ ; las químicas son específicas para cada anclaje. Los anclajes en esta etapa pueden estabilizarse con enlaces intra-anclaje (representados por los triángulos contiguos). En la etapa II, un tercer anclaje (346) con segmento de indicador (347) se usa para "encapuchar" los anclajes después de eliminar grupos de bloqueo (mostrados como los rectángulos pequeños) sobre los segmentos predecesores. El segmento de encapuchado también puede estabilizarse con un enlace intra-segmento (de nuevo representado como triángulos contiguos).

Volviendo ahora a la Figura 34C, se muestra una realización basada en la adición de restos de sonda  $P^1$  y  $P^2$  a enlaces terminales separados de un anclaje preformado. En la etapa I, el anclaje preformado se hace reaccionar primero con  $P^1$  poniendo en contacto  $L1$  con  $L1'$ . En la etapa II, el anclaje se hace reaccionar entonces con  $P^2$  poniendo en contacto  $L2$  con  $L2'$ . El anclaje puede estabilizarse por un enlace intra-anclaje (representado por triángulos contiguos), que pone  $P^1$  en proximidad con  $P^2$ . Los dos restos de sonda se ligan entonces en la etapa III para formar el enlace selectivamente escindible ("V") entre  $P^1$  y  $P^2$  (el segundo y tercer círculos). La ligación de  $P^1$  y  $P^2$  puede facilitarse opcionalmente duplexando dichos restos de sonda con un molde complementario.

En la Figura 35A se desvela otra realización para la síntesis de un segmento de construcción de indicador. Usando métodos químicos en estado sólido, un conector escindible (351) se ancla primero sobre un sustrato sólido (350). Un primer segmento de anclaje con un enlace reversible (352) se hace reaccionar con el conector en la etapa I, y

entonces en la etapa II se hace reaccionar con una biblioteca combinatoria de monómeros M1 y M2, mostrada en este ejemplo como una mezcla estequiométrica 4:1 de los monómeros respectivos. La síntesis de copolímeros aleatoria se hace de este modo para producir composiciones de segmento de anclaje únicas. Si se usan monómeros de péptido o de aminoácido, esto puede hacerse con química de anhídridos mixtos, por ejemplo, produciendo anclajes de péptido de copolímeros aleatorios de longitud variable (Semkin et al., "Synthesis of peptides on a resin by the mixed anhydride method", *Chemistry of Natural Compounds* 3(3):182-183, 1968; Merrifield et al., "Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide". *J. Am. Chem. Soc.* 85(14):2149-2154, 1963). En la etapa III, un elemento conector de extremos (356) bloqueado en L1 (representado como un rectángulo pequeño) se añade entonces al segmento de péptido. Tras la escisión del segmento del soporte sólido (no mostrado), el segmento de anclaje puede incorporarse en una construcción de indicador usando una variedad de métodos, algunos de los cuales se describieron en la Figura 34.

En la Figura 35B se demuestra una alternativa a esta realización. Una construcción de indicador con fragmentos de péptido aleatoriamente incorporados se sintetiza como antes en las etapas I, II y III, pero en la etapa III el grupo conector de extremos (358) está provisto de un conector heterofuncional L2. Diferente del ejemplo previo, los monómeros se proporcionan en proporciones iguales. Después de la escisión del soporte sólido (no mostrado), la construcción de indicador está disponible para la incorporación adicional en anclajes de los ejemplos previos y es un conector heterobifuncional. Las construcciones de indicador sirven para codificar la información genética del fragmento de la secuencia de sonda, como se describe más abajo en más detalle.

Los polipéptidos son construcciones de indicador útiles y también sirven de anclajes. Pueden utilizarse copolímeros aleatorios, periódicos, alternantes y de bloque junto con homopolímero, por ejemplo, para la composición del segmento de anclaje y construcción de anclaje. Los segmentos de polipéptido pueden estar funcionalizados en los extremos usando reticuladores heterobifuncionales que contienen succinimidilo para la conversión de amina en, por ejemplo, una hidrazida o 4-formilbenzoato (4FB). Los polipéptidos pueden producirse tanto por síntesis química como por clonación y expresión en exceso en sistemas biológicos (bacterias, levadura, baculovirus/insecto, mamífero). Dependiendo de la longitud deseada del segmento, los segmentos de anclaje pueden oscilar de  $N > 2$  (segmentos cortos) a  $N > 1000$  (segmentos largos). Pueden usarse químicas de protección de grupos laterales de amina según convenga. Como una alternativa a usar reticuladores listos para uso, los polipéptidos pueden sintetizarse químicamente con restos de hidrazida, 4FB y NHS directamente unidos.

También pueden usarse segmentos de polipéptido funcionalizados en los extremos usando reticuladores heterobifuncionales que contienen maleimido para la conversión de tiol en una hidrazida o 4-formilbenzoato (4FB), en la síntesis de anclajes y construcciones de indicador. Similarmente pueden emplearse segmentos de polipéptido funcionalizados en los extremos que usan reticuladores de EDC para la conversión de carboxilo en una hidrazida (HZ) o 4-formilbenzoato.

La síntesis de las cinco clases de construcción de sustrato oligomérico que tienen anclajes modificados con construcciones de indicador se logra usando los métodos sintéticos anteriores. Similarmente, la síntesis de sustratos monoméricos que tienen anclajes modificados con construcciones de indicador también puede lograrse por los mecanismos sintéticos anteriores. Las químicas para estas variantes son generalmente aplicables a los géneros de especies de Xpandómeros mostradas en las Figuras 8 y 9. Los presentes inventores vuelven ahora a estrategias codificantes y reglas para transportar la información genética en Xpandómeros con construcciones de indicador.

#### Construcciones de indicador y estrategias de código de indicador

Un "código de indicador" es una representación digital de una señal particular o secuencia señal que está integrada en los indicadores de una construcción de indicador particular. Mientras que la "construcción de indicador" es una manifestación física de información extragenética, el código de indicador es su equivalente digital.

La codificación digital requiere códigos de indicador para seguir ciertas reglas. Por ejemplo, se requieren al menos 256 códigos de indicador para identificar todas las posibles combinaciones de una biblioteca de Xsondas 4meras. Teniendo más códigos de indicador de los que son posibles, las combinaciones de construcción de indicador son ventajosas debido a que pueden usarse estados adicionales para otros fines tales como el marcado de huecos, proporcionando información posicional o identificando errores de paridad o errores de alto orden.

Pueden considerarse varias estrategias para representar físicamente un código de indicador. El anclaje puede dividirse en uno o muchos segmentos codificables, pudiendo cada uno marcarse tanto antes como después del ensamblaje de Xpandómero. Pueden usarse niveles de señal variables (cantidad de marca), longitudes (duración de señal) y formas de segmentos de construcciones de anclaje marcados para aumentar las opciones de codificación. La codificación también puede expandirse usando marcas multiplexables. Por ejemplo, usando un enfoque de etiquetado con marca de masa, puede usarse una amplia biblioteca de marcas espectralmente distintas para codificar únicamente un segmento de indicador único; 14 marcas de masa distintas usadas en combinaciones de tres marcas agrupadas sobre un único segmento de construcciones de anclaje crearían 364 espectros de 3 masas únicos. Para el anclaje multi-segmentado, el marcado post-ensamblaje del Xpandómero de la mayoría o de todo el esqueleto de anclaje puede tener el beneficio añadido de aumentar la rigidez del Xpandómero, haciendo posiblemente más fácil manipular para la detección y mejorando la estabilidad.



La Figura 36A (y también la Figura 2A) ilustra un anclaje con un único segmento de indicador de una Xsonda o un Xmero. Este enfoque se beneficia del marcado de indicador altamente multiplexable, tal como marcas de espectrómetro de masas, para producir una amplia biblioteca de salidas espectralmente distintas. Una marca de masa escindible es una molécula o complejo molecular de indicadores escindibles que puede ser fácilmente ionizado a un número mínimo de estados de ionización para producir espectros de masas precisos. Cuando se controla cuidadosamente, un espectrómetro de masas puede detectar tan solo algunos cientos de tales indicadores de esqueleto de anclaje para determinar el estado del código (aunque se requiere información posicional para distinguir un código de indicador del siguiente). Esta característica simplifica la construcción de anclaje y posiblemente acorta los requisitos de longitud del anclaje.

En una realización, usando etiquetas de marca de masa escindibles, el Xpandómero alargado puede presentarse para la detección mediante una fuente de iones de nanoporo (ionización por electropulverización, ionización química a presión atmosférica, fotoionización) o por deposición superficial (nanopeine, nanocanal, flujo laminar, electroforético y similares) seguido de ionización por desorción con láser, haz de iones, o fuentes de haces de electrones (ionización por desorción con láser asistida por matriz "MALDI", ionización por desorción con electropulverización "DESI", ionización por desorción sobre sílice "DIOS", espectrometría de masas de iones secundarios "SIMS").

Un ejemplo de construcciones de anclaje de 9 bits con 2 estados de detección ("1" y "0"), que produce 512 identidades de código, se describe en la Figura 36B. En una realización, las construcciones de anclaje consisten en los segmentos "1" y "0" que producen dos niveles de impedancia eléctrica como se mide en un detector de nanoporos tipo Coulter. En una segunda realización, las construcciones de anclaje consisten en segmentos conductores de la electricidad "1" y segmentos no conductores "0". En una tercera realización, las construcciones de anclaje consisten en segmentos fluorescentes "1" y segmentos no fluorescentes "0". Puede considerarse una pluralidad de diferentes elementos indicadores para este tipo de codificación. Para cualquiera de estos enfoques, la química de ensamblaje de anclaje y de unión de sonda puede ser idéntica - solo necesitaría cambiarse la composición del segmento de indicador. Con este simple formato, el marcado puede hacerse tanto antes como después del ensamblaje de Xpandómero. Se desea el post-marcado, ya que el segmento de indicador sin marcar es significativamente menos masivo y, como tal, es útil a una concentración mucho mayor que una construcción de indicador completamente marcada. Dependiendo de la estrategia, los segmentos de polímero pueden ser: (1) codificados mediante químicas superficiales conjugables o reactivas (por ejemplo, poli-lisina, ácido poli-glutámico), (2) no reactivos (por ejemplo, PEG, polímeros de baja reactividad), o (3) una mezcla de tanto polímeros reactivos como no reactivos. Los grupos reactivos incluyen, pero no se limitan a, aminas primarias (-NH<sub>2</sub>), carboxilos (-COOH), tioles (-SH), hidroxilos (-OH), aldehídos (-CHO) y restos de hidrazida (-R-N-N). El marcado de segmentos reactivos, que puede incluir la desprotección de grupos reactivos, puede hacerse directamente en las construcciones de sustrato, después de la formación del producto intermedio de Xpandómero, después de la escisión del esqueleto para producir el Xpandómero, o en cualquier otro momento en el proceso de SBX según convenga para producir los mejores resultados.

Pueden ser posibles niveles adicionales, como se muestra en la Figura 36C. Por ejemplo, una estrategia codificante direccional binaria (construcción de indicador no simétrico) requiere al menos ocho segmentos codificantes de indicador y un noveno segmento, un segmento de encapuchado codificable (para el cierre del bucle de anclaje y como posible punto de referencia de la construcción de sustrato central) para producir los 256 códigos mínimamente requeridos para un sustrato de cuatro bases (512 códigos si el segmento de encapuchado está codificado).

Volviendo a la Figura 36D, se muestra que una construcción de indicador de 7 bits, cada indicador con tres estados de detección, produce 2187 identidades de código detectables. El uso de espaciadores de polímero flexibles puede usarse por motivos estéricos.

En la Figura 36E se muestra un equivalente rígido de la construcción de indicador de 7 bits previa.

La Figura 36F describe un ejemplo de construcciones de anclaje de 4 segmentos rígidas con ocho estados distintos totales o combinaciones de marca por segmento. El uso de siete de las combinaciones de marca para marcar tres de las construcciones de segmentos de anclaje producirá 343 identidades de código únicas y dejará 1 segmento disponible para la identificación de límites de subunidad, paridad, u otro fin funcional. Esta realización puede usar un enfoque de marcado mixto en el que pueden incorporarse 1 a 3 marcas diferentes sobre cada segmento para producir al menos ocho combinaciones únicas por segmento como se muestra en la Figura 37, en la que se contemplan combinaciones de diferentes tipos de indicador y química de construcción de indicador. Pueden utilizarse la marca de masa y las opciones de marcado fluorescente, entre otros, como se describe. El marcado de construcciones de anclaje de ensamblaje post-ensamblaje del Xpandómero está dirigido por la abundancia e identidad de tres restos de reticulador como se describe en la Figura 37. Las longitudes del segmento de anclaje pueden estar en el orden de 100-1000 nm para las mediciones limitadas de difracción o < 100 nm para las mediciones de campo cercano, si se desea usar estas tecnologías de detección. Pueden usarse anclajes más cortos para otros métodos de detección.

Las construcciones de anclaje de 3 segmentos rígidas con 22 combinaciones de marca distintas totales por segmento se describen en la Figura 36G. El uso de 21 de las combinaciones de marca descritas en la Figura 37 para marcar dos de las construcciones de segmentos de anclaje producirá 441 identidades de construcciones de anclaje únicas y dejará un segmento disponible para la identificación de límites de subunidad de Xsondas. Esta realización usa un enfoque de marcado mixto en el que de una a tres marcas diferentes pueden incorporarse en cada segmento para producir hasta 22 combinaciones únicas por segmento (Figura 37). Pueden utilizarse la marca de masa y las opciones de marcado fluorescente, entre otros, para esta realización. Como se describe en la Figura 37, el marcado de construcciones de anclaje del post-ensamblaje del Xpandómero está dirigido por la abundancia e identidad de tres restos químicos.

10 Diseñando segmentos de indicador/de anclaje en los que la abundancia de grupos reactivos y la dimensión espacial (distancia radial desde el esqueleto de polímero) pueda variarse, pueden lograrse niveles de codificación en los que al menos son posibles tres estados de código total: Alto "2", Medio "1" y Bajo "0" (Figura 36H). Por otra parte, una estrategia de codificación de nivel dos, de tres marcas (es decir, 21 estados por indicador) solo requeriría dos segmentos codificantes de indicador para producir 441 códigos y podría usar un segmento adicional para la orientación de código (Figura 36I).

15 La codificación de indicador puede diseñarse para reducir errores inherentes con la construcción de indicador y asociados a tecnologías de detección usando varios enfoques. En el caso de la detección de nanoporos tipo Coulter, la velocidad a la que un Xpandómero pasa a través del poro y la modulación de la corriente que produce puede depender de muchos factores que incluyen, pero no se limitan a, el estado de carga de la porción del Xpandómero dentro del poro, las concentraciones de electrolito, los estados de carga superficiales del nanoporo, el potencial aplicado, los efectos de fricción que limitan el movimiento de Xpandómeros, y las dimensiones relativas de tanto el Xpandómero como el nanoporo. Si la velocidad no es predecible, la descodificación de la modulación de la corriente no puede usar tiempo (y velocidad constante) para resolver las asignaciones de medición de indicador. Una realización codificante resuelve esta cuestión usando codificación de 3 estados. La señal del indicador es la impedancia que causa una marca a la conductividad de electrolitos a través del nanoporo. Proporcionando tres posibles niveles de impedancia para un indicador, un bit de información está codificado por la transición con la siguiente marca. Por diseño, esta transición es siempre un cambio a uno de los otros dos estados. Si los tres estados están marcados A, B y C, entonces una secuencia de indicadores nunca tiene 2 A, 2 B o 2 C juntos. De esta forma, la información está codificada en transiciones de nivel y es, por tanto, independiente de la velocidad a través del nanoporo. Un esquema codificante es asignar a todas las transiciones A a B, B a C y C a A un valor de "0" mientras que a las transiciones B a A, C a B y A a C se asigna el valor "1". Por ejemplo, la secuencia detectada ABACBCA se descodifica a 0,1,1,1,0,0.

20 Aunque el momento preciso no puede resolver de forma fiable marcas secuenciales, puede ser suficiente para diferenciar la separación de la secuencia de marca sobre construcciones de un único anclaje de las del siguiente anclaje secuencial. Longitudes de anclaje de espaciadores adicionales en cualquier extremo de la secuencia de marca de indicador (en los puntos de unión de la construcción de sustrato) pueden proporcionar un gran hueco de tiempo que delinea los códigos de construcciones de anclaje en el tiempo.

25 En los casos en los que tal momento adecuado es insuficiente, puede producirse un error del desplazamiento del marco. Los errores del desplazamiento del marco resultan cuando el detector lee una serie de marcas de múltiples códigos de construcciones de anclaje (marcos), pero no delinea correctamente el inicio del código (inicio del marco). Esto produce códigos erróneos. Una realización para resolver esto es añadir más bits en el código que los que se necesitan para identificar las secuencias de bases correspondientes (que normalmente son 1 a 4 bases de longitud). Por ejemplo, se requieren ocho bases para identificar únicamente una secuencia de 4 bases. Cada par de 2 bits describe una base única. Añadiendo un bit de paridad a cada base, el código de las construcciones de anclaje aumenta a 12 bits. Una tasa de error de alta paridad (próxima al 50 %) indicaría un error del desplazamiento del marco que produciría la ausencia de un cambio de estado. Además de una ausencia de cambio de estado, otro tipo de error que puede producirse en esta tecnología de detección de nanoporos es la de un estado incorrectamente leído. Los errores de un único indicador pueden aislarse a la base particular usando el bit de paridad y puede entonces asignarse el valor "base desconocida".

30 Los Xpandómeros marcados usando tanto segmentos de impedancia eléctrica, de conducción eléctrica como fluorescentes pueden medirse en disolución usando una variedad de formatos de nanoporo, nanopeine o nanocanal. Alternativamente, los Xpandómeros pueden depositarse superficialmente como polímeros lineales espacialmente distintos usando nanopeine, nanocanal, flujo laminar o métodos de deposición electroforética, entre otros. Al igual que con el enfoque de disolución, la detección directa del Xpandómero alargado en la superficie puede hacerse midiendo las características de señal de segmentos de anclaje marcados. Dependiendo de la composición del material de deposición (conducción, aislamiento) y el sustrato subyacente (conducción, aislamiento, fluorescente), puede considerarse una variedad de tecnologías de detección para la detección de marcas.

#### Métodos de SBX adicionales

35 Varios métodos de SBX que usan construcciones oligoméricas de sustrato de clase I-V se ilustraron en las figuras previas (Figuras 11-17, 20, 21, 24). Los presentes inventores consideran ahora métodos opcionales que

complementan estos protocolos.

#### Funcionalización de extremos

La Figura 38 ilustra la preparación y uso de adaptadores de extremos diana. La Figura 38A ilustra oligonucleótidos complementarios funcionalizados en los extremos duplexados para formar un adaptador con un extremo conjugable bifuncional ("L1" y "L2") y un extremo enzimáticamente ligable (5'-fosfato y 3'-OH). Estos adaptadores también pueden diseñarse con funcionalidades adicionales. Por ejemplo, como se muestra en la Figura 38B, pueden sintetizarse adaptadores funcionalizados en los extremos con reticuladores de esqueleto anidados ("L3") y un enlace escindible ("V"). Por simplicidad, el enlace escindible V puede tener la misma química de enlaces escindibles (o enzimología) usada para liberar o expandir el Xpandómero, aunque pueden utilizarse otros conectores escindibles si se desea diferenciar entre etapas de escisión independientes. Las Figuras 38C y 38D muestran etapas para la construcción de un adaptador multifuncional de la Figura 38B. Como se muestra, puede usarse una perla magnética con un oligonucleótido anclado en la superficie complementaria a cada hebra de adaptador (dos mezclas de perlas diferentes) para ensamblar segmentos de oligonucleótidos diferencialmente modificados. Una vez ensamblados, los segmentos pueden ligarse enzimáticamente para enlazar covalentemente los segmentos hibridados. Con este enfoque, cada segmento puede modificarse individualmente de un modo no disponible para la síntesis de oligonucleótidos estándar.

La manipulación del Xpandómero puede ser útil para la eficaz presentación y detección de muestras. Por ejemplo, pueden usarse marcas de afinidad terminales para modificar selectivamente un extremo del Xpandómero para permitir el alargamiento electroforético. La unión de un modificador voluminoso de carga neutra a cualquiera del extremo 3' o 5' (no ambos) produce un arrastre electroforético sobre el Xpandómero que hace que el extremo no modificado se alargue a medida que se desplaza con el detector. Los modificadores de extremos, que incluyen, pero no se limitan a, micropérlas, nanopartículas, nanocristales, polímeros, polímeros ramificados, proteínas y dendrímeros, pueden usarse para influir en la estructura (alargamiento), posición y tasa a la que el Xpandómero se presenta al detector confiriendo propiedades diferenciadoras únicas a sus extremos tales como carga (+/-/neutra), flotabilidad (+/-/neutra), hidrofobia y paramagnetismo, por nombrar algunos. En los ejemplos proporcionados, las modificaciones de extremos producen una carga de arrastre que permite que el Xpandómero se alargue; sin embargo, también puede emplearse la estrategia opuesta en la que la modificación de extremos se usa para tirar del Xpandómero hacia y a través del detector. Con este enfoque, el tirar del modificador de extremos facilita el alargamiento del Xpandómero. El adaptador de extremos sobre la hebra molde puede también opcionalmente contener uno o más ácidos nucleicos que se usarán para sintetizar el registro del marco y las señales de validación en el Xpandómero acabado (véase la Figura 54).

La incorporación de un modificador de la afinidad puede hacerse tanto antes de, durante, como después de la síntesis de Xpandómeros (hibridación, ligación, lavado, escisión). Por ejemplo, los cebadores marcados con afinidad terminal que son complementarios a la secuencia de adaptador pueden pre-cargarse a la diana de ADNmc bajo condiciones altamente específicas antes de la síntesis de Xpandómeros. El cebador y su marca de afinidad pueden incorporarse en el Xpandómero de longitud completa y pueden usarse para modificar selectivamente su extremo. Un enfoque posiblemente más elegante es incorporar enzimáticamente modificadores de los extremos. La transferasa terminal (TdT), por ejemplo, es una polimerasa independiente de molde que cataliza la adición de desoxinucleótidos al extremo 3' hidroxilo de moléculas mono o bicatenarias de ADN. Se ha demostrado que TdT añade nucleótidos modificados (biotina) al extremo 3' (Igloi et al., "Enzymatic addition of fluorescein- or biotin-riboUTP to oligonucleotides results in primers suitable for DNA sequencing and PCR", *BioTechniques* 15, 486-497, 1993). Un amplio intervalo de enzimas son adecuadas para este fin, que incluyen (pero no se limitan a) ARN ligasas, ADN ligasas y ADN polimerasas.

En las Figuras 38E y 38F se ilustran adaptadores en horquilla. Los adaptadores en horquilla encuentran uso en la ligación de extremos romos para dar hebras molde auto-cebadoras. Una ventaja de este enfoque es que la hebra hija sigue covalentemente acoplada a la hebra molde y se re-hibrida más rápidamente tras una fusión para eliminar material sin ligar y fragmentos de bajo peso molecular. Como se muestra en la Figura 38F, estos adaptadores pueden contener funcionalidades de conector pre-formadas incorporadas para la purificación o manipulación aguas abajo, y también pueden contener sitios de escisión para la recogida más eficaz de las hebras hijas del Xpandómero.

#### Preparación y análisis del molde diana

Para realizar la secuenciación del genoma completo de ADN continuo largo, el uso de los métodos de SBX basados en Xsonda supone que el ADN se prepara en una forma manejable para la hibridación, ligación, llenado de huecos si se requiere, la expansión y medición. El proceso de ensamblaje de Xpandómeros puede mejorarse, en algunas realizaciones, por inmovilización superficial de la diana de ADN con el fin de (1) reducir la complejidad y efectos cruzados de la hibridación, (2) mejorar el lavado, (3) permitir la manipulación de diana (alargamiento) con el fin de facilitar la hibridación mejorada, y/o (4) si se usa sensor de nanoporos, proporcionar una interfase continua con el proceso de detección. Como se describe en detalle más adelante, se espera que los métodos para la preparación de molde diana, análisis y unión superficial mejoren la calidad de los datos y el ensamblaje de secuencias.

La mayoría de los métodos de secuenciación del genoma completo requieren la fragmentación del genoma diana en trozos más manejables. El cromosoma más largo en el genoma humano (cromosoma 1) tiene ~227 Mb y el cromosoma más pequeño (cromosoma 22) tiene ~36 Mb. Para la mayoría de las realizaciones altamente procesivas y continuas descritas en el presente documento, 36 Mb de secuenciación continua es demasiado largo para ser secuenciado con alta eficiencia. Sin embargo, para aprovechar la capacidad inherente de la larga longitud de lectura de SBX, como diana se eligen longitudes de fragmentos de ADN de >1 kb. Por consiguiente, pueden emplearse varias estrategias para realizar la fragmentación del genoma y preparar un conjunto de diana de ADN compatible con el método de SBX como se ha desvelado en el presente documento.

Una realización implica fragmentar el genoma total en trozos de 1-10 Kb (promedio 5 Kb). Esto puede hacerse tanto por enzimas de restricción como por cizallamiento hidrodinámico/mecánico. Los fragmentos se hacen entonces romos en los extremos y se reparan en la preparación para la ligación de extremos romos a un adaptador de secuencia ("SA") o una construcción de adaptador de secuencia-dendrímico ("SAD"). El extremo no romo de SA o la construcción SAD se diseña para ser incapaz de ligación, de manera que se prevenga el ensamblaje de multímeros de dichos adaptadores. Cualquier diana ligada a SA o SAD que esté presente puede purificarse por afinidad lejos de los adaptadores libres. En este momento, si el dendrímico de captura no se introduce con la construcción de SA, entonces esto puede hacerse (usando un exceso eficaz de dendrímico), seguido de purificación para aislar el complejo de diana-SAD. Una vez purificado, el complejo está listo para la unión a la superficie de ensayo. Para este fin, se desea que solo un único complejo diana se una por localización de reacción. Alternativamente, el dendrímico de captura puede en su lugar asociarse a la superficie de ensayo, en cuyo caso el complejo de diana-SA purificado puede unirse directamente al dendrímico que ya está localizado y unido covalentemente a la superficie. Un método similar de ensamblaje de diana de ADN sobre superficies se ha desvelado por Hong et al. ("DNA microarrays on nanoscale-controlled surface", *Nucleic Acids Research*, 33(12): e106, 2005).

Otro enfoque, llamado el método de "análisis", implica fragmentar gruesamente el genoma en trozos de 0,5-5 Mb (usando enzimas de restricción de corte raras o por cizallamiento hidrodinámico/mecánico), seguido de captura de tales fragmentos a una micromatriz modificada (o superficie de división) compuesta de sondas de captura de oligonucleótido específicas de gen/loci. Están ampliamente disponibles micromatrices a medida y conjuntos de sondas de captura de oligonucleótidos de varias fuentes comerciales (Arraylt, Euorins-Operon, Affymetric Inc.). Este análisis adicional puede proporcionar una ventaja para el ensamblaje de secuencias del extremo trasero. La captura de estos grandes fragmentos implica tanto la desnaturalización diana parcial como completa con el fin de permitir que las sondas de captura se unieran o duplexaran con ADN diana específicos. Para reducir la hibridación no específica puede ser necesario cargar la matriz de captura bajo condiciones diluidas de manera que se prevenga la hibridación cruzada entre moldes.

Cada conjunto de sondas de captura de división, que puede presentarse sobre una gran superficie, se diseña para proporcionar resolución del genoma lineal de ~3 Mb. Para proporcionar el eficaz análisis del genoma, cada conjunto de sondas de captura individual puede estar compuesto por 3 a 5 oligonucleótidos específicos de gen/loci, linealmente separados sobre el genoma por 0,5-1,5 Mb cada uno. Las sondas de captura se seleccionan para ser completamente únicas con el fragmento diana, proporcionando así tanto la especificidad como la redundancia al método. Dada la resolución de 3 Mb de un genoma de 3 Gb, este enfoque requiere una matriz de captura compuesta de sondas específicas de aproximadamente 1000 genes/loci. Los conjuntos de sondas de captura de oligonucleótidos listos para micromatriz específicos para las dianas de gen humano están ampliamente disponibles listos para su uso (Operon Biotechnologies, Huntsville AL, EE.UU.).

Una vez se han eliminado los acontecimientos de unión no específica, cada matriz de captura puede entonces avanzar como reacciones independientes a través del resto del proceso de preparación del genoma del mismo modo que se ha tratado anteriormente para el método no de análisis. Siendo la diferencia primaria que la muestra diana puede ahora analizarse posicionalmente en una superficie de ensayo de SBX o como reacciones basadas en disolución individual, reduciéndose así la complejidad del post-ensamblaje de la secuencia de adquisición de datos.

Otra realización es tener hibridación de Xsonda-diana no anclada y producción de Xpandómeros. En este caso, el ensayo de SBX se realiza en disolución libre. Este enfoque pueden usar una o una combinación de manipulaciones físicas tales como usar marcas electroforéticas, magnéticas, de arrastre, o funcionalidades terminales de flotabilidad positiva/negativa bajo condiciones estáticas o de flujo laminar, como un medio para alargar el ADN diana antes de y durante la hibridación y ligación de sondas, por ejemplo, y, para alargar o expandir el Xpandómero escindido antes de la detección. La síntesis en disolución libre de Xpandómeros, sin inmovilización, puede hacerse usando polimerasas y ligasas (con y sin cebadores) y también puede hacerse usando métodos de ligación química. Pueden usarse tanto construcciones de sustrato de trifosfato como construcciones de sustrato de monofosfato. Se concibe la síntesis simultánea de Xpandómeros de dianas de ácido nucleico múltiples y mixtas. Generalmente, las construcciones de trifosfato de sustrato son capaces de polimerización procesiva continua en disolución y pueden adaptarse a protocolos de un único tubo para la secuenciación de una única molécula masivamente en paralelo en disolución libre, por ejemplo.

Ensamblaje superficial de dianas de ácido nucleico

Un método de preparación de ácidos nucleicos diana para la secuenciación usa diana de ADN bicatenario funcionalizado en los extremos como se muestra en la Figura 39. Para este ejemplo, cada adaptador normalmente está provisto de un grupo ANH (hidrazida reactiva) útil para el procesamiento adicional, por ejemplo, usando química de enlace de 4-hidrazinonicotinato-acetona-hidrazona de succinimidilo (SANH) y un dúplex de oligonucleótidos de adaptador modificado en los extremos con amina. Similarmente, puede usarse una SANH reactiva con amina para crear restos de hidrazida reactivos. SANH se conjuga fácilmente con aldehídos como 4-formilbenzoato (4-FB) para formar enlaces hidrazona covalentes estables. El 4-formilbenzoato de C6-succinimidilo reactivo con amina (C6 SFB) puede usarse para crear un resto benzaldehído reactivo. Los oligonucleótidos complementarios funcionalizados en los extremos se duplexan para formar un extremo conjugable bifuncional (SANH y amina) y un extremo enzimáticamente ligable (3'-OH y 5'-fosfato). La ligación del adaptador crea un diana de ADNbc funcionalizado en los extremos que puede reticularse a los grupos aldehído anclados a la superficie. La Figura 39 ilustra diana de ADNbc funcionalizado en los extremos usando SANH y dúplex de oligonucleótidos de adaptador modificados en los extremos con amina.

Para muchos de los métodos de SBX descritos, las dianas de ácido nucleico pueden anclarse covalentemente a un soporte sólido recubierto plano (acero inoxidable, silicio, vidrio, oro, polímero). Como se describe en la Figura 40, los puntos de unión diana pueden producirse por derivatización de 4FB de una monocapa de SAM. El método utiliza el adaptador ANH de la Figura 39, y en la etapa I de la Figura 40, el ANH se hace reaccionar con las cabezas 4FB de la monocapa y el molde se desnaturaliza (etapa II). La otra hebra del molde, que también está marcada en los extremos, puede capturarse en una reacción separada. En la etapa III, la amina libre en el extremo 3' del molde monocatenario se hace reaccionar entonces con una perla, por ejemplo, aquí mostrada como una perla de flotabilidad, de manera que el molde diana puede estirarse. Estos complejos de captura pueden ensamblarse tanto por auto-ensamblaje aleatorio de un mezcla estequiométricamente equilibrada de polímeros funcionalizado en los extremos (ejemplo, tiol - PEG - hidrazida para la unión de diana; tiol - PEG - metoxi para el encapuchado de la monocapa auto-ensamblada "SAM") o por estampado de puntos de unión reactivos espacialmente resueltos usando técnicas litográficas. El método litográfico estampado puede producir puntos de unión diana consistentemente separados, aunque esto es difícil de hacer para la unión de una única molécula, aunque el método de auto-ensamblaje aleatorio probablemente produciría separación de la unión diana más variable, pero tiene un alto porcentaje de uniones de moléculas individuales. Están comercialmente disponibles un amplio intervalo de reticuladores monofuncionales, bifuncionales y heterobifuncionales de una variedad de fuentes. También están disponibles polímeros monofuncionales, bifuncionales y heterobifuncionales compatibles de reticulante (polietilenglicol, poli-l-lisina) de una amplia variedad de fuentes comerciales.

La densidad de diana de ADN de 1 billón de dianas sobre una superficie de  $100 \text{ cm}^2$  requeriría un promedio por área diana de  $10 \text{ um}^2$ . La separación diana en este intervalo proporciona separación diana suficiente para prevenir la significativa reacción cruzada de ácidos nucleicos diana de 5000 bases de longitud anclados a perlas (diámetro de perla 100-1000 nm; ADNbc de 5 Kb = 1700 um). El área diana puede expandirse fácilmente si se determina que la reactividad cruzada de la diana y/o perla ( $>1$  perla/diana) es inaceptablemente alta.

Alargamiento diana usando perlas o nanopartículas

Como se muestra en la Figura 40, una perla o nanopartícula (400) anclada sobre el extremo libre de ADN diana puede utilizarse para alargar y mantener la diana en su conformación monocatenaria durante, por ejemplo, la hibridación de bibliotecas de Xsonda. La retención de la diana en una conformación alargada reduce significativamente la frecuencia y estabilidad de las estructuras secundarias intramoleculares diana que se forman a las menores temperaturas. La reducción o incluso eliminación de las influencias de la estructura secundaria promueve el eficaz ensamblaje de construcciones de sustrato de alta fidelidad.

Puede emplearse una variedad de enfoques para aplicar una fuerza de alargamiento con la diana monocatenaria. Por ejemplo, perlas paramagnéticas/nanopartículas/polímeros permiten el uso del campo magnético para suministrar fuerza direccional controlada a la diana anclada a la superficie. Además, controlando la dirección de las líneas de campo magnético, pueden usarse perlas paramagnéticas/partículas para guiar y mantener la diana cargada con Xsonda sobre la superficie del sustrato durante las etapas de lavado. Secuestrando las dianas a lo largo de la superficie, puede minimizarse la pérdida de diana debido a fuerzas de cizallamiento. Como alternativa al alargamiento del campo magnético, pueden usarse perlas/partículas de flotabilidad positiva y negativa que son tanto más como menos densas que el agua, respectivamente, para proporcionar una fuerza de alargamiento. Todas las perlas/partículas se recubren superficialmente según sea necesario para minimizar interacciones no específicas (agregación de perlas, unión a sondas) y se funcionalizan para permitir la reticulación covalente con dianas modificadas con adaptador.

La Figura 41A a 41D ilustra estrategias de alargamiento de dianas basadas en perlas representativas. El alargamiento de dianas usando perlas paramagnéticas/partículas ancladas en los extremos atraídas a un campo magnético externo (B) se ilustra en la Figura 41A. El secuestro de diana sobre la superficie del sustrato (para reducir el cizallamiento diana) usando perlas paramagnéticas/partículas ancladas en los extremos atraídas a un campo magnético externo se ilustra en la Figura 41B. Las Figuras 41C y 41D ilustran el alargamiento de diana usando

perlas/partículas de flotabilidad negativa ancladas en los extremos (mayor densidad que el agua) y perlas/partículas de flotabilidad positiva (menor densidad que el agua), respectivamente. Los métodos de disolución libre para el alargamiento de dianas usando restos anclados en el extremo, por ejemplo, usando tanto una perla de flotabilidad positiva como negativa para funcionalizar extremos opuestos de una diana, proporcionan una alternativa elegante para reducir la estructura secundaria de dianas.

El uso de estas estrategias de alargamiento en la preparación de Xpandómeros se ilustra en la Figura 41E. Aquí, en la etapa I, el molde inmovilizado (416) se estira usando una perla flotante duplexada con la diana mediante el cebador de adaptámero (417). El molde se pone en contacto entonces en la etapa II con construcciones de sustrato y éstas se enlazan entonces con la diana monocatenaria, produciendo un producto intermedio de Xpandómero bicatenario (411) con construcción de sonda-bucle característica y dos esqueletos, uno mediante el esqueleto primario del polinucleótido y el otro mediante el esqueleto de Xpandómero limitado. En la etapa III, la hebra molde se desnaturaliza, y en la etapa IV el producto intermedio de Xpandómero monocatenario se escinde en enlaces selectivamente escindibles en el esqueleto primario, produciendo un desplegamiento de los bucles y alargamiento del producto de Xpandómero con esqueleto sustituto completamente extendido (419).

#### Alargamiento de dianas de ADN usando modificaciones de extremos del polímero

Una alternativa a los métodos descritos en la Figura 40 utiliza polímeros funcionalizados largos (en lugar de perlas) ligados covalentemente en los extremos libres de dianas de ADN ancladas en la superficie para alargar la diana. El electro-estiramiento y el bobinado de modificaciones de extremos del polímero a través de sustrato poroso seguido por captura (secuestro) de polímero dentro del sustrato produce ADN diana monocatenario completamente alargado significativamente libre de estructura secundaria. La Figura 42A ilustra este método, que muestra el bobinado de las hebras de diana a través de poros en un sustrato. Los sustratos porosos incluyen, pero no se limitan a, matriz de gel, óxido de aluminio poroso y membranas porosas. La captura de polímero completamente alargado puede lograrse por funcionalización química controlada en la que la reticulación o unión de polímero a sustrato poroso puede iniciarse selectivamente después del alargamiento completo de la diana. Están disponibles una variedad de estrategias de reticulación o unión. Por ejemplo, la reticulación de polímero alargado funcionalizado con carboxilo a sustrato poroso funcionalizado con amina puede lograrse con la introducción del agente de reticulación clorhidrato de 1-etil-3-[3-dimetilaminopropil]carbodiimida (EDC).

Otra realización se ilustra en la Figura 42B, es electroestirar el polímero hacia una superficie funcionalizada que reticula o se une al polímero. En este caso, el polímero no está necesariamente bobinado a través del sustrato, sino que en su lugar se reticula o se une a grupos funcionales sobre la superficie. Para este enfoque, los sustratos incluyen, pero no se limitan a, matriz de gel, óxido de aluminio poroso, membranas porosas y superficies conductoras no porosas como oro. Tanto los sustratos porosos como no porosos pueden funcionalizarse usando una variedad de químicas de reticulación (hidrazida-aldehído; oro-tiol) o de unión (biotina-estreptavidina) listas para uso.

Como alternativa, puede emplearse un sistema de apertura eléctrica encerrada que usa regiones de campo eléctrico variables (alta, baja, tierra) para capturar segmentos de polímero cargados. El método, que se ilustra en la Figura 42C, produce un tipo de jaula de Faraday de cierre que puede usarse para mantener el polímero y ADN diana en su posición alargada. Este enfoque requiere el reparto o confinamiento eficaz de campos eléctricos, y tiene la ventaja de permitir el alargamiento de dianas ajustable (durante toda la síntesis de SBX) modificando los parámetros de apertura. Fuera de la jaula de Faraday, las influencias de campos eléctricos sobre las construcciones de sustrato son mínimas.

En un método relacionado, se usa un polímero elástico para suministrar una fuerza de estiramiento constante al ácido nucleico diana menos elástico. Este método puede usarse como suplemento a cualquiera de los métodos de alargamiento descritos aquí, que incluyen perla magnética, deformación de la flotabilidad, deformación de la densidad (gravedad), o electroestiramiento. La fuerza elástica almacenada dentro del polímero proporciona una fuerza de alargamiento amortiguada más coherente con la hebra diana.

#### Sustrato diana basado en matriz de gel con electro-enderezamiento

En la Figura 43A se muestra un método alternativo para la presentación de ADN diana monocatenario (frente a la unión a soporte sólido) en el que la diana está covalentemente unida a una matriz de gel de tamizado (430). En la Figura 43b (inserción), puede observarse que las construcciones de sustrato se asocian con los moldes anclados, que se han estirado rectos. Pueden usarse campos eléctricos para el alargamiento de dianas y la presentación de construcciones de sustrato. Este enfoque tiene ventajas con respecto a otros métodos para la matriz de diana de ADN en términos de densidad de diana y alargamiento de dianas (electro-enderezamiento). Por ejemplo, la reticulación de ADN a acrilamida mediante adaptadores de extremos modificados con acrilato (unidos a oligonucleótidos) se hace rutinariamente. Acrydite™ es un fosforamidito disponible (Matrix Technologies, Inc., Hudson, NH, EE.UU.) que se ha usado ampliamente para incorporar modificaciones del extremo 5' de metacrilo a oligonucleótidos: el doble enlace en el grupo Acrydite reacciona con dobles enlaces activados de acrilamida (Kenney et al., "Mutation typing using electrophoresis and gel-immobilized Acrydite probes", *Biotechniques* 25(3):516-21, 1998). La agarosa funcionalizada con amina también está disponible lista para uso (G Biosciences, St. Louis, MO,

EE.UU.) y puede usarse similarmente para reticularse con ADN usando químicas de reticulación reactivas con amina listas para uso (Spagna et al., "Stabilization of a  $\beta$ -glucosidase from *Aspergillus niger* by binding to an amine agarose gel", *J. of Mol. Catalysis B: Enzymatic* 11 (2,3): 63-69, 2000).

5 Al igual que con los métodos de unión de soporte sólido previamente descritos, los fragmentos de ADN se funcionalizan útilmente con adaptadores de extremos para producir, por ejemplo, una modificación de 5'-Acrydite. Para garantizar la uniformidad de dianas, las dianas de ADNbc se desnaturalizan (y se mantienen en un estado desnaturalizado) mientras que se reticulan con la matriz de tamizado. La diana puede desnaturalizarse usando una variedad de técnicas existentes que incluyen, pero no se limitan a, urea, pH alcalino y fusión térmica. Una vez se ha reticulado con una matriz de gel, el ADN diana desnaturalizado puede enderezarse aplicando un campo eléctrico apropiado (similar a la electroforesis). La difusión, electroforesis y neutralización de los desnaturalizantes dentro de la matriz de gel junto con un estrecho control de temperatura, con la adición de disoluciones compatibles con la hibridación, produce un entorno propicio para la hibridación de construcciones de sustrato. Por ejemplo, pueden presentarse continuamente Xsondas a dianas ancladas por electroforesis. Este formado permite el reciclado de sondas y proporciona un buen control con respecto a la velocidad de flujo de Xsondas.

15 La densidad de dianas de una serie de matrices de gel de 5 mm x 50 mm x 20 mm produce un volumen reactivo de  $5,0 \times 10^{12} \mu\text{m}^3$ , que es sustancialmente superior a lo que puede presentarse sobre una superficie plana. La Figura 43C muestra las estructuras químicas para acilamida, bis-acrilamida, y la modificación de extremos de oligonucleótidos con Acrydite. La Figura 43D ilustra una matriz de gel de poliacrilamida con diana de ADN covalentemente unida.

20 El procesamiento enzimático de construcciones de sustrato puede hacerse dentro de la matriz de gel. Ensayos enzimáticos que usan polimerasas (PCR) y ligasas se usan rutinariamente para la modificación de oligonucleótidos covalentemente acopladas a almohadillas de gel de acrilamida (Proudnikov et al., "Immobilization of DNA in Polyacrylamide Gel for the Manufacture of DNA and DNA-Oligonucleotide Microchips", *Analytical Biochemistry* 259(1): 34-41, 1998). La ligasa, por ejemplo, se suministra a dianas de Xsonda ensambladas dentro de la matriz de gel tanto por difusión pasiva, electroforesis, como una combinación de ambas. Se consideran un amplio intervalo de densidades de matriz y tamaños de poro, sin embargo, la difusión pasiva se beneficia de matriz de densidad relativamente baja y poros grandes. El suministro activo de ligasa por electroforesis puede lograrse por unión de modificadores de la carga (polímeros y dendrímeros) a la enzima para potenciar su migración a través de la matriz.

25 Una variación del método de matriz de gel es someter a electroforesis los polinucleótidos diana monocatenarios modificados en los extremos a través de una matriz de gel. Las modificaciones de extremo pueden usarse para crear arrastre sobre un extremo de la diana. Puede realizarse una modificación de extremo catiónica en el otro extremo de la diana de manera que la arrastre mediante el gel. También puede inducirse un gradiente o zona de temperatura en el gel de manera que se rampee o module la rigurosidad de hibridación a medida que las dianas progresan a través de la matriz de gel. Puede hacerse procesamiento enzimático dentro de la matriz de gel o después de que los complejos hibridados de diana-sustrato salgan de la matriz de gel. El proceso pueden repetirse según sea necesario para producir la longitud de Xpandómero promedio deseada.

30 La Figura 44 ilustra el uso de marcas de arrastre para la post-síntesis de manipulación de Xpandómeros. La marca de arrastre (representada como un diamante) sirve de equivalente de disolución de una técnica de estiramiento y transporte (Meagher et al., "Free-solution electrophoresis of DNA modified with drag-tags at both ends", *Electrophoresis* 27(9):1702-12, 2006). En la Figura 44A, las marcas de arrastre están unidas por química de conectores a adaptadores funcionalizados sobre los extremos 5' del molde monocatenario (L1' de la marca de arrastre se combina con L1 del molde). En la etapa mostrada, la adición procesiva de construcciones de sustrato se muestra ya en progreso. En la Figura 44B, la marca de arrastre se añade hibridando con un adaptador complementario. En la Figura 44C, el Xpandómero se trata primero de manera que se una un conector a un adaptador de 3' (mostrado como el cuadrado pequeño con L1'), y entonces el conector (L1) se usa en la etapa II para unir la marca de arrastre al mismo. La transferasa terminal, que extiende un extremo 3'-OH de ADN mono o bicatenario, polimeriza los trifosfatos de nucleótido modificados por conector tales como nucleótidos biotinilados. También pueden usarse otras enzimas para añadir bases u oligómeros modificados. En la Figura 44C, una marca de arrastre se añade al extremo 3' del producto intermedio de Xpandómero. Las marcas de arrastre pueden incluir, pero no se limitan a, nanopartículas, perlas, polímeros, polímeros ramificados y dendrímeros.

#### Huecos y llenado de huecos con hibridación y ligación de construcciones de sustrato

Pueden emplearse múltiples variantes en los métodos de ligación para tratar la gestión de huecos y el llenado de huecos. En una realización, se usa ligación escalonada cíclica, en la que sondas se ensamblan secuencialmente de un cebador anclado en la superficie duplexado con el polinucleótido diana. En otra realización, se producen hibridación y ligación "promiscua" de construcciones de sustrato simultáneamente a través de toda la secuencia de ADN diana entera, generalmente sin un cebador.

Para tanto los enfoques de ligación cíclica como promiscua, es posible que, por ejemplo, alguna porción de las construcciones de sustrato 4meras (por ejemplo, el 25 %) y alguna porción de las construcciones de sustrato 5meras (por ejemplo, el 20 %) se hibride adyacentemente de manera que se permita la ligación. De estos dúplex

adyacentes, algún porcentaje puede contener secuencias desapareadas que producirán tanto error de medición (si están ligadas) como fallo al ligarse. La mayoría de los casos del fallo de la ligación (debido a desapareamiento) es sin consecuencia, ya que las sondas sin ligar se eliminan antes de la siguiente ronda de hibridación. El ciclo de hibridación / lavado / ligación / lavado puede repetirse varios cientos de veces, si se desea (6 minutos/ciclo = 220 ciclos/día y 720 ciclos/3 días; 10 minutos /ciclo = 144 ciclos/día y 432 ciclos/3 días).

En el proceso cíclico (ilustrado con Xsondas en la Figura 14), las Xsondas se ensamblan secuencialmente desde un extremo de un cebador anclado en la superficie duplexado con el ADN diana; una Xsonda por ciclo. Si esto es el caso, la longitud de lectura diana puede calcularse usando las siguientes suposiciones: el 25 % de hibridación adyacente (de 4 meros) con el 20 % de fidelidad de dúplex de sonda perfecto produciría solo 20 ligaciones de 4meros secuenciales ( $400 \times 25\% \times 20\%$ ) después de 400 ciclos. Usando estas suposiciones, un ensayo de 400 ciclos (1,5 días a 6'/ciclo) produciría una longitud de producto promedio de 80 nucleótidos.

En el proceso promiscuo (ilustrado con Xsondas de clase I en la Figura 45), si se deja que las reacciones de hibridación y ligación de Xsondas se produzcan espontáneamente y simultáneamente a lo largo del molde diana de ADN, puede lograrse la replicación de un molde de ADN mucho más largo usando muchos menos ciclos. Tras cada ciclo de hibridación y lavado, puede usarse ligación para conectar los dúplex restantes que son tanto adyacentes como el 100 % correctos. Puede usarse otro lavado más riguroso para eliminar productos de ligación más pequeños (todos 8meros y menos; algunos 12meros), además de todos los 4meros sin ligar. Con reacciones de ligación que se producen durante toda la secuencia diana de 1.000-10.000 nucleótidos, la procesividad del ensayo aumenta espectacularmente. Como el enfoque promiscuo para la hibridación y ligación de Xsondas es independiente de la longitud del molde diana de ADN, la mayoría del molde diana puede replicarse en una fracción de los ciclos requerida para el método en serie. Pueden usarse tiempos de ciclo más largos para compensar limitaciones cinéticas si un ajuste tal aumenta la fidelidad y/o cantidad de reacciones de hibridación/ligación.

La Figura 45 ilustra una progresión secuencial del proceso promiscuo. La etapa I ilustra la hibridación de Xsondas 4meras en múltiples loci a lo largo de la diana de ADN. La etapa II ilustra la hibridación y ligación de Xsondas adyacentes; las Xsondas ligadas se estabilizan y así siguen hibridadas aunque las Xsondas sin ligar se fundan. La etapa III y la etapa IV ilustran otro ciclo de hibridación térmico seguido de ligación y fusión térmica de Xsondas sin ligar. Cada ciclo extiende preferencialmente las cadenas de Xsondas ligadas existentes. Como se ilustra en la etapa IV, después de ciclos repetidos, la diana de ADN se satura con Xsondas que dejan huecos, más cortos que una longitud de sonda, en los que no se ha producido duplexado. Para completar el Xpandómero, las Xsondas se enlazan a través de los huecos por medios enzimáticos o químicos como se ilustran en la etapa V.

#### Llenado de huecos estándar

Como se ilustra con Xsondas en la etapa V de la Figura 45, después de completarse el proceso de ciclos de hibridación y ligación, los huecos de secuencia pueden llenarse para producir un Xpandómero continuo. Los huecos a lo largo del esqueleto de ADN diana pueden llenarse usando procesos de llenado de huecos basados en ADN polimerasa/ligasa bien establecidos (Stewart et al., "A quantitative assay for assessing allelic proportions by iterative gap ligation", *Nucleic Acids Research* 26(4):961-966, 1998). Estos huecos se producen cuando cadenas de Xsondas adyacentes se encuentran las unas con las otras y tienen una longitud de huecos de 1, 2 o 3 nucleótidos entre ellas (suponiendo una Xsonda de clase I 4mera como se ilustra en la Figura 45). El llenado de huecos también puede hacerse mediante reticulación química (Burgin et al.). Después de completarse el llenado de huecos, puede procesarse el complemento de ADN diana, que está compuesto principalmente de Xsondas ligadas con cargas periódicas de 1, 2 o 3 nucleótidos (purificación, escisión, modificación de extremo, marcado de indicador) según convenga para producir un Xpandómero medible. Como los productos de ensayo de SBX se preparan y se purifican en procesamiento por lotes antes de la etapa de detección, la detección es eficaz y no está limitada por ningún procesamiento bioquímico simultáneo.

Con el fin de diferenciar los huecos de la señal de indicador específica de nucleótido, pueden usarse desoxinucleótido trifosfatos modificados (que tanto ya está marcados como son capaces de ser marcados tras el ensayo) para identificar la secuencia de huecos. Desoxinucleótido trifosfatos adecuados se ilustran en las Figuras 46A (bases modificadas con conector) y 46B (bases modificadas con biotina).

La longitud y frecuencia de huecos dependen de varias variables de síntesis, que incluyen número de ciclos, rigurosidad de hibridación, estrategia de bibliotecas (de secuenciación aleatoria frente a sub-biblioteca con ajuste estequiométrico), longitud de molde diana (100b - 1 Mb) y densidad de reacción (0,1-10 B). La frecuencia de huecos y la longitud de huecos pueden reducirse significativamente utilizando condiciones de rigurosidad máximas, siempre que las condiciones sean compatibles con el intervalo de tiempo de ensayo especificado. La longitud de diana y la densidad de reacción también son factores importantes con respecto a tanto la fidelidad del alineamiento como la probabilidad de llenado de huecos.

El proceso promiscuo de hibridación puede usar métodos de ciclado térmico para mejorar la rigurosidad de hibridación y para aumentar la frecuencia de alineamientos de construcciones de sustrato adyacentes. La hibridación, ligación y la fusión térmica continúan bajo un ciclado térmico preciso rutinario hasta que la mayoría de la secuencia diana se duplexa con sonda. Los dúplex de sonda-diana no específicos débilmente unidos pueden



eliminarse por una simple etapa de lavado, de nuevo bajo control térmico preciso. La ligación enzimática o química puede realizarse para enlazar cualquier construcción oligomérica adyacentemente hibridada a lo largo del ADN diana, generando secuencias más largas y más estabilizadas. La ligación enzimática tiene el beneficio añadido de proporcionar una comprobación cruzada de fidelidad del dúplex de adición, ya que la sonda-diana desapareada no están eficazmente ligadas. Usando un segundo lavado bajo condiciones de ciclo térmico preciso, las construcciones de sustrato sin ligar pueden fundir el esqueleto de ADN; quedarán productos intermedios más largos. Las secuencias ligadas más largas crecen en múltiples loci a lo largo del ADN diana hasta que la replicación esté prácticamente completa. El proceso de hibridación/lavado/ligación/lavado se repite de cero a muchas veces hasta que la mayoría de molde diana se ha replicado.

El ciclado de temperatura acoplado con la estabilización de la construcción de sustrato debido al apilamiento de bases puede utilizarse para enriquecer en acontecimientos de hibridación adyacentes. Las condiciones de ciclado térmico, similares en concepto a la "PCR con rampa decreciente de temperatura", pueden usarse para eliminar dúplex no adyacentes menos estables. Por ejemplo, ciclando repetitivamente entre la temperatura de hibridación y la temperatura de fusión superior ("T<sub>m</sub>") calculada para la biblioteca de oligos (como se determina por fusión de sonda individual - es decir, no apilamiento de bases estabilizado) pueden seleccionarse positivamente acontecimientos de hibridación de sondas adyacentes. Con eficiencia de enriquecimiento suficiente, el número de ciclos puede reducirse significativamente. Esto puede realizarse en condiciones de un baño único o de multi-baño.

Con un tamaño de biblioteca pequeño (256 construcciones de sustrato 4meras, por ejemplo), la biblioteca puede repartirse en sub-bibliotecas basándose en un intervalo de temperatura de fusión incluso más estrecho. Esta estrategia puede beneficiarse significativamente del método si el sesgo de T<sub>m</sub> no puede controlarse mediante la modificación de bases, adyuvantes de disolución de hibridación, y similares. Además, la estequiometría de la sonda puede ajustarse (por arriba o por abajo) para compensar cualquier sesgo residual que todavía pueda existir en la biblioteca/sub-bibliotecas.

Se realizó modelado informático para evaluar la estadística de la aparición de huecos y las longitudes de construcciones de sustrato consecutivamente conectadas que se hibridan sobre un molde de ADN diana. El modelo simula un ciclo térmico de hibridación/ligación con una biblioteca completa de 256 Xsondas 4meras. Los resultados del modelo que se presentan aquí se basan en el siguiente proceso modelo:

i) La etapa de hibridación/ligación simula Xsondas 4meras aleatorias que cumplen un molde de ADN aleatoriamente secuenciado en posiciones aleatorias y que se hibridan si se aparean. Ninguna Xsonda puede solapar ningún nucleótido. La hibridación continúa hasta que todas las localizaciones en la diana se hibridan más de 3 nucleótidos de longitud.

ii) La etapa de fusión térmica se simula eliminando todas las Xsondas que son en cadenas más cortas de M Xsondas 4meras de longitud (en la que  $m = 2, 3, 4, 5 \dots$ ). Las cadenas más largas siguen sobre el molde de ADN. Una "Cadena" se define aquí como múltiples Xsondas consecutivas sin huecos entre ellas.

iii) Repetir el ciclo definido por i) y ii) de manera que se construyan Xsondas a partir de los loci existentes de cadenas de multi-Xsonda. El ciclado se detiene cuando no se produce cambio entre 2 ciclos consecutivos.

La Figura 47 y la Figura 48 tienen cada una dos gráficos que ilustran la estadística de ejecutar el modelo 100 veces en un molde de ADN aleatorio. Para fines de procesamiento, la diana de ADN se eligió para tener 300 nucleótidos de longitud (pero la estadística aquí tratada no se cambia si la longitud del molde es 5000 nucleótidos). La Figura 47 muestra los resultados para  $M = 3$ , en la que cadenas más cortas de  $M=3$  "funden" el molde de ADN. La Figura 47A muestra la estadística de las longitudes de cadena en el último ciclo. La distribución parece ser casi exponencial oscilando de 12mero ( $M = 3$ ) hasta algunos 160meros largos con una media de ~28 nucleótidos de longitud. La Figura 47B muestra que se necesita un intervalo de 5 a 20 ciclos para completar el molde de ADN usando este método de "crecimiento" de cadena y que el 90 % de las ejecuciones se completan después de 12 ciclos. Las Figuras 48A y 48B muestran el mismo tipo de información para el caso en el que  $M=4$ , una fusión térmica más rigurosa. En este caso, la longitud promedio de cadena aumenta a ~40 nucleótidos, pero requiere ~ 18 ciclos para completar el 90 % de las ejecuciones. Las longitudes de cadena más largas reducen la probabilidad de tener un hueco como era de esperar. La probabilidad media de no llenar una posición de nucleótido para estos datos de  $M = 4$  tiene una media de 0,039 (aproximadamente 1 en 25).

#### Llenado de huecos con construcciones de sustrato 3meras y 2meras

Una extensión de la realización de hibridación 4mera básica es usar construcciones de sustrato 3meras y 2meras para llenar los huecos de secuencia de 3 bases y 2 bases, respectivamente. La Figura 49 muestra un ejemplo de llenado de huecos por adición secuencial o simultánea de Xsondas más cortas. Si la mayoría del molde diana se duplexa satisfactoriamente con Xsondas tras la hibridación 4mera, la mayoría de los huecos restantes tienen tanto 1, 2 como 3 bases de longitud. El tamaño promedio del Xpandómero puede aumentarse significativamente si puede llenarse un porcentaje de moderado a alto de los huecos de 2 y 3 bases y ligarse con Xsondas más pequeñas. Como la hibridación de estas sondas más pequeñas se hace después de duplexarse la mayor parte de la diana, la rigurosidad de temperatura puede reducirse para permitir el duplexado de 3meros y luego 2meros, respectivamente. La estabilidad del dúplex de las Xsondas más pequeñas aumenta debido a la estabilización por apilamiento de bases. Tras la ligación, los huecos restantes se llenan usando una polimerasa para insertar nucleótidos, seguido de

nuevo por ligación para enlazar todos los 5'-fosfatos y 3'-hidroxilos adyacentes restantes. Los nucleótidos modificados con conector incorporado en polimerasa permiten el marcado de huecos, tanto antes como después de la incorporación. Estas marcas pueden o no pueden identificar la base. Basándose en modelos estadísticos, resumidos en los diagramas de la Figura 50, los polímeros de Xpandómero de longitud completa tienen más del 95 % de cobertura de la secuencia diana (superior a 16 bases). Según el modelo, el uso de llenado de huecos de sondas 3meras/2meras hace bajar el porcentaje de bases sin marcar un factor de 6 (suponiendo incorporación de alta eficiencia), que, para una secuencia diana de 5000 bases, produciría 5000 bases de secuencia principalmente contigua con menos de 50 huecos de bases individuales sin discriminar a través de todo el replicón diana. Como tal, se identifica a 16X cobertura superior al 99 % de las secuencias de hueco sin identificar.

El llenado de huecos con estas Xsondas 3meras y 2meras más pequeñas, por ejemplo, extiende la longitud del Xpandómero promedio tres veces. Los diagramas que resumen los resultados de modelado estadístico como se muestra en las Figuras 50A y 50B indican que los Xpandómeros de huecos rellenos 3meros/2meros de longitud promedio, bajo las condiciones descritas, están en el intervalo de 130 bases de secuencia. Además, con 5000 gigabases de secuencia de Xpandómero sintetizada de este modo, que es el equivalente a convertir solo 2,5 ng de secuencia diana en Xpandómero, la selección de tamaño del 10 % más largo de fragmentos de Xpandómero de esa población produciría 500 gigabases de secuencia con longitudes promedio en el intervalo de 381 bases, aunque la selección de tamaño del 2 % más largo de fragmentos de Xpandómero produciría 100 gigabases de secuencia con longitudes promedio en el intervalo de 554 bases. Estas longitudes de fragmento pueden lograrse sin la necesidad de llenado de huecos de bases individuales con polimerasa. Si se hace hibridación de Xsondas 4meras sin ningún llenado de huecos, los modelos estadísticos indican que las longitudes promedio para el 10 % y 2 % más largos de los fragmentos serían 106 y 148 bases, respectivamente.

#### Adición de adyuvantes para reducir la estructura secundaria diana

El ADN diana monocatenario anclado en la superficie largo puede presentar un diana de hibridación desafiante debido a estructuras secundarias intramoleculares. La adición de una perla anclada en los extremos, como se describe en la Figura 41E, reduce la aparición de estas formaciones intramoleculares, y las desestabiliza cuando se producen. Sin embargo, puede encontrarse necesario disminuir adicionalmente la formación de estructuras secundarias, que puede bloquear eficazmente la hibridación de construcciones de sustrato, con la adición de adyuvantes.

La adición de sondas de oligonucleótidos 2meras y/o 3meras sin marcar a, por ejemplo, una mezcla de hibridación de Xsondas puede servir para este fin. Para eliminar la posibilidad de incorporación en el Xpandómero, estas sondas se sintetizan para ser no ligables (sin 5'-fosfato o 3'-hidroxilo). Debido a su pequeño tamaño, no es probable que los adyuvantes 2meros y/o 3meros formen dúplex estables a las temperaturas de hibridación de Xsondas 4meras; sin embargo, su presencia en la mezcla de reacción a concentraciones de moderadas a altas reduce la frecuencia y estabilidad de la estructura secundaria diana bloqueando débilmente y transitoriamente el acceso de secuencias de nucleótidos intramoleculares que pueden de otro modo duplexarse. La Figura 51 es una ilustración de cómo la adición de adyuvantes 2meros o 3meros inhibe la formación de estructuras secundarias (para simplicidad solo, en la figura no se mostraron Xsondas 4meras). Acoplado con la fuerza de alargamiento proporcionada por el anclaje de perlas, los adyuvantes disminuyen significativamente la frecuencia de formación de estructuras secundarias.

Estos adyuvantes de dinucleótido y/o trinucleótido pueden estar compuestos de nucleótidos estándar, nucleótidos modificados, nucleótidos universales (5-nitroindol, 3-nitropirrol y desoxiinosina), o cualquier combinación de los mismos, para crear todas las combinaciones de secuencia necesarias. La Figura 52 muestra sustituciones de bases universales comunes que se usaron para este fin.

Otro enfoque para reducir la estructura secundaria diana es replicar ADN diana para producir una diana de ADNc sintético que tiene una estabilidad de estructura secundaria reducida en comparación con ADN nativo. La Figura 53 enumera algunos análogos de nucleótidos, como se ilustra en la solicitud de EE.UU. 2005/0032053 ("Nucleic acid molecules with reduced secondary structure"), que pueden incorporarse en la diana de ADNc sintética descrita. Se ha demostrado que estos análogos, que incluyen (pero no se limitan a) N4-etildesoxicitidina, 2-aminoadenosina-5'-monofosfato, 2-tiouridina-monofosfato, inosina-monofosfato, pirrolopirimidina-monofosfato y 2-tiocitidina-monofosfato, reducen la estructura secundaria de ADNc. Este enfoque puede usarse junto con alargamiento de dianas y adyuvantes de hibridación para reducir la estructura secundaria.

#### Detección y medición

Como se ha mencionado previamente, el Xpandómero puede marcarse y medirse por cualquier número de técnicas. El potencial de salida masiva de datos del método de SBX coincide bien con las matrices de sensor basadas en nanoporos o tecnologías equivalentes. En una realización, la matriz de nanoporos puede servir de fuente de ionización sobre el extremo delantero de un espectrómetro de masas, en el que los códigos de indicador sobre el Xpandómero son marcas de espectroscopía de masas escindibles. Otras realizaciones implican el uso de un sensor de nanoporos tal como impedancia eléctrica/conductancia o FRET.

Una realización de detección usa una matriz de nanoporos como se representa en la Figura 54. En esta realización, el Xpandómero se ensambla usando métodos descritos previamente, excepto que la diana de ADN no está anclada a un soporte sólido inmovilizado, sino que está anclada a una perla magnética que tiene su sonda de anclaje bobinada a través de un sustrato de nanoporo. Además, un multitud de fluoróforos donantes de FRET, para dos longitudes de onda de excitación, están anclados con la entrada del nanoporo (mostrada como pequeños cuadrados). Los fluoróforos aceptores de FRET constituyen los indicadores incorporados sobre el Xpandómero. Después de escindir las Xsondas ligadas del Xpandómero, el Xpandómero se estira por una fuerza que puede ser electrostática, magnética, gravitacional y/o mecánica y puede facilitarse por la perla usada para extender la diana de ADN original (mostrada unida con la parte superior del Xpandómero). En la etapa de medición final, el Xpandómero se saca a través del nanoporo aplicando una fuerza magnética sobre la perla magnética mostrada debajo de la estructura de nanoporo. Los fluoróforos donantes en la entrada del nanoporo se excitan con una fuente de luz ( $\lambda_1$ ) y como los indicadores pasan proximales con los fluoróforos donantes, se excitan y emiten su fluorescencia distintiva ( $\lambda_2$ ) que descodifica la secuencia de nucleótidos asociada.

La Figura 54 también ilustra una realización de codificar la secuencia en los códigos de indicador. Sobre cada anclaje hay 4 sitios de indicador, cada uno cargado con una combinación de dos tipos de fluoróforo aceptor de FRET. Esto proporciona cuatro estados medibles sobre cada sitio usando nivel de intensidad relativa. Si la fluorescencia de fluoróforos aceptores fuera roja y verde, la codificación de estos estados con los 4 nucleótidos puede ser: A: verde>roja, C: solo roja, G: solo verde y T: roja > verde.

En otra realización, el Xpandómero se marca con indicadores de marca de masa que se miden usando un espectrómetro de masas. El Xpandómero se dosifica en un capilar estrecho que alimenta los indicadores secuencialmente en un ionizador de electropulverización. Para permitir la medición del espectrómetro de masas de los indicadores de marca de masa discretos, las marcas de masa pueden ser fotoescindidas del andamiaje de indicador justo antes de, durante, o después de la ionización de la marca de masa. Puede usarse espectrómetro de masas basado en sector magnético, cuadrupolo, o tiempo de vuelo ("TOF") para la detección de marcas de masa. El instrumento solo requiere distinguir un número limitado de marcas de masa separada. Esto puede usarse para mejorar la sensibilidad y rendimiento del instrumento. Por tanto, el empleo de un instrumento que tiene múltiples canales para realizar la ionización y detección en paralelo aumenta el rendimiento en órdenes de magnitud.

Una realización de detección del enfoque de espectrómetro de masas es un detector de masas de TOF multi-canal que puede leer >100 canales simultáneamente. Un instrumento adecuado usaría una fuente de ionización multi-canal que alimenta Xpandómero en múltiples canales a una concentración y tasa que maximiza el uso del canal, maximizando así la tasa de salida de datos de buena calidad. Una fuente de ionización tal requiere tener separación adecuada de los segmentos de indicador de marca de masa. Las marcas de masa pueden ser fotoescindidas a medida que surgen de un nanoporo y se ionizan. Los requisitos de dispersión no son altos, de manera que un corto tubo de vuelo es todo lo que se requiere. La salida de medición extremadamente alta es posible con un enfoque de detección de espectrómetro de masas multi-canal. Por ejemplo, una matriz de 100 canales de iones de nanoporo que lee a la tasa de 10.000 códigos de indicador por segundo (con 4 mediciones de nucleótido por código de indicador) lograría un rendimiento del instrumento de > 4 Mbases/segundo.

En otra realización, el nanoporo se usa de un modo similar a un contador Coulter (Figura 55). En esta implementación, la densidad de carga del Xpandómero está diseñada para ser similar a la del ADN nativo. Los indicadores están diseñados para producir 3 niveles de impedancia como se mide en, por ejemplo, el detector de nanoporos. Los Xpandómeros se presentan en una disolución libre que tiene altas concentraciones de electrolito, tales como KCl 1 M. El nanoporo tiene 2 a 15 nm de diámetro y tiene 4 a 30 nm de longitud. Para lograr buena resolución de las marcas de construcciones de anclaje, se eligen para estar próximas a la misma longitud o más largas. El diámetro de las marcas puede ser diferente para los 3 niveles. El segmento de Xpandómero cerca de los enlaces de la construcción de anclaje no tiene indicadores (por ejemplo, PEG), este segmento tendrá un nivel de impedancia particular. Uno de los tres niveles de indicador puede ser equivalente. Pueden producirse diferentes distintivos de impedancia, que implican tanto el nivel de impedancia como la respuesta temporal, variando las longitudes de segmento, densidad de carga y densidad molecular. Por ejemplo, para lograr 3 niveles de impedancia diferentes, los segmentos de construcciones de anclaje marcadas pueden codificarse químicamente para acoplarse a uno de tres tipos de polímero diferentes, cada uno con una longitud y densidad de carga diferente.

Cuando el Xpandómero pasa a través del detector de nanoporos, la corriente se modula según qué tipo de marca esté presente. La cantidad de polímero cargado que reside en el nanoporo afecta a tanto la corriente de especies de electrolitos como a la velocidad de translocalización.

La detección basada en polímeros por nanoporos se demuestra en las patentes de EE.UU. N° 6.465.193 y 7.060.507, por ejemplo, y como era de esperar se muestra que los parámetros físicos de un polímero modulan la salida eléctrica de un nanoporo.

En otra realización de un aparato de detección basado en nanoporos (Figura 56), los electrodos laterales fijados con el nanoporo se usan para medir la impedancia o conductividad de lado a lado en el nanoporo, aunque se aplica un voltaje a través de la película de soporte sólido. Esto tiene la ventaja de separar la función de translocalización de la función de impedancia. A medida que el Xpandómero es transportado a través del nanoporo, de nuevo se mide la

modulación de corriente. Se emplean técnicas microfluídicas y de micropipeteado, junto con marcas de arrastre, perlas magnéticas, técnicas de estiramiento electroforético, etc., con el fin de transportar el Xpandómero a través del nanoporo. Por ejemplo, la electroforesis en disolución libre marcada en los extremos, también llamada ELFSE, es un método de rotura del equilibrio de la carga con respecto a la fricción de ADN de arrastre libre que puede usarse para la electroforesis de Xpandómeros en disolución libre (Slater et al., "End-labeled free-solution electrophoresis of DNA", *Electrophoresis* 26: 331-350, 2005).

Los métodos de anclaje, estiramiento, marcado y medición de fragmentos grandes de ADN están bien establecidos (Schwartz et al., "A single-molecule barcoding system using nanoslits for DNA analysis", *PNAS*, 104(8):2673-2678, 2007; y Blanch et al., "Electrokinetic Stretching of Tethered DNA", *Biophysical Journal* 85: 2539-2546, 2003). Sin embargo, la resolución de nucleobases individuales para los fines de secuenciación del genoma completo de ácidos nucleicos nativos está más allá de las capacidades de estas técnicas. En la Figura 57 a 59 se representan varios métodos de detección de Xpandómero de "molécula individual". En la Figura 57 se representa un microscopio y cualquiera de una variedad de técnicas de obtención de imágenes directas que se aprovechan de la mayor resolución espacial de la estructura del Xpandómero que puede concebirse. Una molécula de Xpandómero se coloca plana sobre una superficie generalmente plana y se escanea a lo largo de su longitud. Ejemplos incluyen microscopía óptica y microscopía de fluorescencia. El uso de microscopía de alta resolución está limitado en la mayoría de los casos a >100 nm de resolución que todavía es posible con anclajes largos (100 nm por indicador, por ejemplo). Puede lograrse mayor resolución (que requiere anclajes <100 nm por indicador, por ejemplo) usando técnicas tales como microscopía de fluorescencia de reflexión interna total (TIRF: Starr, T.E. et al., *Biophys. J.* 80, 1575-84, 2001), guías de onda de modo cero (Levine M.J. et al., *Science* 299, 682-85, 2003) o microscopía óptica de campo cercano (NSOM: de Lang et al., *J. Cell Sci.* 114, 4153-60 2001) o microscopía de barrido láser confocal. En el caso de FRET localizada, las interacciones de excitación pueden localizarse a < 10 nm, conduciendo a longitudes de anclaje de ~10 nm/indicador.

La detección y análisis de grandes moléculas de ADN por microscopía electrónica está bien establecida (Montoliu et al., "Visualization of large DNA molecules by electron microscopy with polyamines: application to the analysis of yeast endogenous and artificial chromosomes", *J. Mol. Bio.* 246(4):486-92, 1995), sin embargo, la secuenciación de polinucleótidos precisa y de alto rendimiento usando estos métodos ha demostrado ser difícil. En la Figura 58 se indican conceptualmente microscopía electrónica de transmisión (TEM) y de barrido (SEM) para la detección de un Xpandómero. Aquí se usa un haz de electrones enfocado para barrer un Xpandómero, que es de nuevo generalmente plano sobre una superficie. Los aspectos de la estructura del Xpandómero sirven para descodificar la información genética sobre el esqueleto. Pueden usarse fijación de especímenes y técnicas de recubrimiento por pulverización, que permiten la obtención de imágenes de características individuales y del tamaño de átomos de moléculas, para potenciar el aumento.

También son de interés espectroscopía de conductancia túnel de electrones de apertura por nanoelectrodos, en la que un haz de electrones túnel se modula entre dos puntas de nanoelectrodo por el transporte del Xpandómero entre las puntas (Lee et al., "Nanoelectrode-Gated Detection of Individual Molecules with Potential for Rapid DNA Sequencing", *Solid State Phenomena* 121-123: 1379-1386, 2007). El Xpandómero perturba la corriente túnel por su efecto de barrido-conducción, que puede amplificarse sobre ADN nativo por uso de indicadores adecuados. Esta técnica tiene la ventaja de que se evita la fijación de especímenes y el requisito de vacío, y en teoría, pueden emplearse matrices masivamente paralelas de puertas de electrodos para leer muchos Xpandómeros en paralelo.

En la Figura 59 se ilustra conceptualmente microscopía de fuerza atómica. En una simple realización, un nanotubo montado sobre un voladizo sensible se barre a través de una superficie y las fuerzas atractivas y repulsivas entre la sonda y la superficie de la muestra se traducen en un cuadro topológico de la superficie que se barre. Esta técnica puede lograr resolución muy alta, pero tiene velocidades de barrido relativamente lentas (M. Miles, *Science* 277, 1845-1847 (1997)). La microscopía electrónica túnel de barrido (STM) es una tecnología relacionada para la obtención de imágenes de superficies; la sonda, sin embargo, no toca la superficie, sino que se mide una corriente de túnel entre la superficie y la sonda. Aquí, el Xpandómero puede depositarse plano sobre una superficie y barrerse físicamente con la punta de la sonda, parecido a una aguja fonográfica en una grabación.

#### Ensamblaje de secuencias

La secuencia de referencia del genoma humano publicada (u otra secuencia de referencia) puede usarse como herramienta de alineamiento para ayudar en el ensamblaje de cantidades masivas de datos de secuencias producidos con SBX. A pesar de la probable inclusión de pequeños huecos de secuencia, posicionalmente identificados, las capacidades de longitud de lectura largas descritas para SBX basados en Xpandómeros simplifica y mejora la fidelidad del ensamblaje de secuencias del genoma completo. Como se trata anteriormente, el proceso puede simplificarse adicionalmente dividiendo fragmentos contiguos en localizaciones dimensionalmente confinadas sobre la superficie de la reacción de ensayo. En esta realización, el método de análisis puede reducir espectacularmente el tiempo y error de ensamblaje.

Construcciones monoméricas

La Figura 9 proporciona una visión general de construcciones monoméricas de la invención. Se distinguen un total de cinco clases que incluyen cuatro clases de RT-NTP (VI, VII, VIII y IX) y una clase de XNTP (X). Cada clase se tratará individualmente a continuación.

- 5 Las construcciones monoméricas de clases VI a X se distinguen de las construcciones oligoméricas de clase I a V en que usan un único residuo de nucleobase como sustrato. En la siguiente descripción, N, se refiere a cualquier residuo de nucleobase, pero normalmente es un nucleótido trifosfato o análogo en el presente documento. Tiene puntos de unión sobre un anclaje (también descrito en el presente documento), por ejemplo, a los anillos heterocíclicos de la base, con el grupo ribosa, o con el  $\alpha$ -fosfato del residuo de nucleobase. Como se describe, el método primario de síntesis dirigida por molde usa polimerasa, pero cualquier método que pueda realizar la síntesis dirigida por molde es apropiado, que incluye métodos de ligación química y enzimática.

15 Para las construcciones de sustrato en las que se usan grupos conectores  $\sigma$  y  $\delta$  para crear enlaces inter-subunidad, puede adaptarse un amplio intervalo de químicas comercialmente disponibles adecuadas (Pierce, Thermo Fisher Scientific, EE.UU.) para este fin. Las químicas de conectores comunes incluyen, por ejemplo, ésteres de NHS con aminas, maleimidadas con sulfhidrilos, imidoésteres con aminas, EDC con carboxilos para reacciones con aminas, piridildisulfuros con sulfhidrilos, y similares. Otras realizaciones implican el uso de grupos funcionales como hidrazida (HZ) y 4-formilbenzoato (4FB), que pueden entonces hacerse reaccionar adicionalmente para formar enlaces. Más específicamente, un amplio intervalo de reticuladores (hetero- y homo-bifuncionales) están ampliamente disponibles (Pierce), que incluyen, pero no se limitan a, sulfo-SMCC (4-[N-maleimidometil]ciclohexano-1-carboxilato de sulfosuccinimidilo), SIA (yodoacetato de N-succinimidilo), sulfo-EMCS (éster de [N-e-maleimidocaproiloxi]sulfosuccinimida), sulfo-GMBS (éster de N-[g-maleimidobutiriloxi]sulfosuccinimida), AMAS (éster de N-(a-maleimidoacetoxi)succinimida), BMPS (N-EMCA ácido (N-e-maleimidocaproico)-éster de [ $\beta$ -maleimidopropiloxi]succinimida), EDC (clorhidrato de 1-etil-3-[3-dimetilaminopropil]carbodiimida), SANPAH (N-succinimidil-6-[4'-azido-2'-nitrofenilamino]hexanoato), SADP (N-succinimidil(4-azidofenil)-1,3'-ditiopropionato), PMPI (N-[p-Maleimidofenil]isocianato, BMPH (hidrazida del ácido N-[ $\beta$ -maleimidopropiónico], sal del ácido trifluoroacético), EMCH (hidrazida del ácido [N-e-maleimidocaproico], sal del ácido trifluoroacético), SANH (4-hidrazinonicotinato-acetonahidrazona de succinimidilo), SHTH (clorhidrato de 4-hidrazidotereftalato de succinimidilo) y C6-SFB (4-formilbenzoato de C6-succinimidilo). Por tanto, el método desvelado por Letsinger et al. ("Phosphorothioate oligonucleotides having modified internucleoside linkages", la patente de EE.UU. N° 6.242.589) puede adaptarse para formar enlaces fosforotiolato.

Además, están ampliamente disponibles químicas de protección/desprotección bien establecidas para restos de conector comunes (Benoiton, "Chemistry of Peptide Synthesis", CRC Press, 2005). La protección de amino incluye, pero no se limita a, carbamato de 9-fluorenilmetilo (Fmoc-NRR'), carbamato de t-butilo (Boc-NRR'), carbamato de bencilo (Z-NRR', Cbz-NRR'), acetamida-trifluoroacetamida, ftalimida, bencilamina (Bn-NRR'), trifenilmetilamina (Tr-NRR') y bencilidenamina-p-toluenosulfonamida (Ts-NRR'). La protección de carboxilo incluye, pero no se limita a, éster metílico, éster t-butílico, éster bencilico, éster s-t-butílico y 2-alkil-1,3-oxazolona. Carbonilo incluyen, pero no se limitan a, 1,3-dioxano de dimetilacetal, y N,N-dimetilhidrazona de 1,3-ditiano. Protección de hidroxilo incluye, pero no se limita a, éter metoximetílico (MOM-OR), éter tetrahidropiranílico (THP-OR), éter t-butílico, éter alílico, éter bencilico (Bn-OR), éter t-butildimetilsilílico (TBDMS-OR), éter t-butildifenilsilílico (TBDPS-OR), éster del ácido acético, éster del ácido piválico y éster del ácido benzoico.

En el presente documento, el anclaje se representa frecuentemente como una construcción de indicador con tres grupos indicadores, diversas construcciones de indicador pueden estar presentes sobre el anclaje, y pueden comprender un único indicador para identificar el sustrato, múltiples indicadores para identificar el sustrato, o el anclaje puede ser polímero desnudo (que no tiene indicadores). Obsérvese que pueden usarse indicadores para la sincronización de la detección, corrección de errores, redundancia, u otras funciones. En el caso del polímero desnudo, los indicadores pueden ser el propio sustrato, o pueden estar sobre un segundo anclaje unidos al sustrato. En algunos casos, uno o más precursores de indicador están presentes sobre el anclaje, y los indicadores se unen por afinidad o se unen covalentemente tras el ensamblaje del producto de Xpandómero.

Anteriormente se desvelan estrategias de codificación de indicadores, y se tratan adicionalmente más adelante. Por ejemplo, la codificación binaria de dos bit de cada monómero produciría cuatro secuencias de código único (11,10, 01, 00), que pueden usarse para identificar cada base de secuencia (adenina "A", citosina "C", guanina "G", timidina "T"), suponiendo que el acoplamiento del sustrato sea direccional. Si no tiene dirección, entonces un tercer bit proporciona codificación inequívoca. Alternativamente, una construcción de indicador único multiplexada de 4 estados proporciona un código de indicador único para cada base de secuencia. Puede considerarse una variedad de estrategias de funcionalización y de marcado para construcciones de anclaje, que incluyen, por ejemplo, dendrímeros funcionalizados, polímeros, polímeros ramificados, nanopartículas y nanocristales como parte de un andamiaje de indicador, además de químicas de indicador con una característica de detección que va a detectarse con la tecnología de detección apropiada que incluye, por ejemplo, fluorescencia, emisores o excitadores de FRET, densidad de carga, tamaño o longitud. Pueden incorporarse marcas específicas de base en el anclaje como sustratos marcados tanto antes como después del ensamblaje del Xpandómero. Una vez el Xpandómero se ha liberado completamente y alargado, los códigos de indicador pueden detectarse y analizarse usando una variedad

de métodos de detección.

Bibliotecas de sustratos adecuados como sustratos monoméricos incluyen (pero no se limitan a) ATP modificado, GTP, CTP, TTP y UTP.

#### Construcciones monoméricas de clase VI

5 La Figura 60 describe construcciones de sustrato monomérico de clase VI (un tipo de RT-NTP) en más detalle. Las Figuras 60A a 60C se leen de izquierda a derecha, mostrando primero la construcción de sustrato monomérico (precursor de Xpandómero que tiene un único residuo de nucleobase), luego la hebra hija del dúplex intermedia en el centro, y a la derecha el producto de Xpandómero preparado para la secuenciación.

10 Como se muestra en la Figura 60A, las construcciones de sustrato monomérico de clase VI tienen un anclaje, T (600), unidos por un enlace (601) de un primer resto terminal a un residuo de nucleobase del sustrato, N. Un grupo conector,  $\epsilon$ , está dispuesto en el primer resto terminal (603) del anclaje próximo a  $R^1$ . En el extremo distal (604) del anclaje, un segundo resto terminal con un segundo grupo conector,  $\delta$ , está posicionado próximo a  $R^2$ . El segundo resto terminal del anclaje está asegurado al primer resto terminal en proximidad con la nucleobase por una reticulación selectivamente escindible intra-anclaje (o por otra limitación). La reticulación escindible intra-anclaje (605) se indica aquí por línea de puntos, que puede indicar, por ejemplo, un enlace disulfuro o un conector fotoescindible. Esta limitación previene que el anclaje se alargue o expanda y se dice que está en su "configuración limitada". Bajo ensamblaje dirigido por molde, los sustratos forman un dúplex con el molde diana de forma que los sustratos estén confinados. Bajo condiciones controladas, los grupos conectores colocados  $\delta$  y  $\epsilon$  de los sustratos confinados se enlazan para formar un enlace entre las construcciones de sustrato adyacentes. Los grupos conectores  $\delta$  y  $\epsilon$  de una construcción de sustrato monomérico no forman un enlace intra-sustratos debido a limitaciones del posicionamiento. Químicas de enlace y de protección/desprotección adecuadas para  $\delta$ ,  $\epsilon$ , y  $\chi$  se detallan en la descripción general de construcciones monoméricas.

15  $R^1$  y  $R^2$  son grupos terminales configurados según convenga para el protocolo de síntesis en el que se usa la construcción de sustrato. Por ejemplo,  $R^1 = 5'$ -fosfato y  $R^2 = 3'$ -OH, encontrarían uso en un protocolo de ligación, y  $R^1 = 5'$ -trifosfato y  $R^2 = 3'$ -OH para un protocolo de polimerasa I. Opcionalmente,  $R^2$  puede configurarse con un grupo de bloqueo reversible para la adición cíclica de un único sustrato. Alternativamente,  $R^1$  y  $R^2$  pueden configurarse con los grupos conectores terminales para el acoplamiento químico o con grupos no conectores para un protocolo de solo hibridación.  $R^1$  y  $R^2$  pueden ser del tipo general XR, en la que X es un grupo de enlace y R es un grupo funcional.

20 Durante el ensamblaje, la construcción de sustrato monomérico se polimeriza primero sobre el extremo extensible de la hebra hija naciente por un proceso de polimerización dirigida por molde usando un molde monocatenario como guía. Generalmente, este proceso se inicia a partir de un cebador y avanza en la dirección de 5' a 3'. Generalmente, se usa una ADN polimerasa u otra polimerasa para formar la hebra hija, y las condiciones se seleccionan de manera que se obtenga una copia complementaria de la hebra molde. Posteriormente, el grupo conector  $\delta$ , que está ahora  
35 colocado con el grupo conector  $\epsilon$  del anclaje de subunidad adyacente, se reticula selectivamente para formar un enlace  $\chi$ , que es un enlace inter-anclaje inter-subunidad. Los enlaces  $\chi$  unen los anclajes en una cadena continua, formando un producto intermedio llamado la "hebra hija de dúplex", como se muestra en la Figura 60B. Después de formarse el enlace  $\chi$ , puede romperse el enlace intra-anclaje.

40 La hebra hija de dúplex (Figura 60B) es un hetero-copolímero con subunidades mostradas en corchetes. El esqueleto primario ( $\sim N \sim$ ) $\kappa$ , hebra molde ( $-N'$ -) $\kappa$  y anclaje (T) se muestran como una hebra hija duplexada, en la que  $\kappa$  indica una pluralidad de subunidades de repetición. Cada subunidad de la hebra hija es un "motivo" de repetición y los motivos tienen variabilidad específica de especie, indicada aquí por el superíndice  $\alpha$ . La hebra hija se forma a partir de especies de la construcción de sustrato monomérico seleccionadas por un proceso dirigido por molde de una biblioteca de especies de motivos, uniéndose el sustrato de monómero de cada especie de construcción de  
45 sustrato a un nucleótido complementario correspondiente sobre la hebra molde diana. De esta forma, la secuencia de residuos de nucleobase (es decir, esqueleto primario) de la hebra hija es una copia complementaria contigua de la hebra molde diana.

50 Cada tilde ( $\sim$ ) indica un enlace selectivamente escindible mostrado aquí como los enlaces inter-sustratos. Éstos son selectivamente escindibles para liberar y expandir los anclajes (y el Xpandómero) sin degradar el propio Xpandómero.

55 La hebra hija está compuesta por un precursor de Xpandómero llamado el "Xpandómero limitado" que está adicionalmente compuesto de anclajes en la "configuración limitada". Cuando los anclajes se convierten en su "configuración expandida", el Xpandómero limitado se convierte en el producto de Xpandómero. Los anclajes están limitados por los enlaces  $\chi$  inter-subunidad, la unión de sustrato y, opcionalmente, los enlaces intra-anclaje si todavía están presentes. El enlace  $\chi$  une el primer resto terminal del anclaje de una primera subunidad con el segundo resto terminal del anclaje en el extremo confinado de una segunda subunidad y se forma enlazando los grupos conectores colocados,  $\epsilon$  de la primera subunidad y  $\delta$  de la segunda subunidad.

Puede observarse que la hebra hija tiene dos esqueletos, un “esqueleto primario” y el esqueleto del “Xpandómero limitado”. El esqueleto primario está compuesto por los sustratos monoméricos contiguamente confinados y polimerizados. El “esqueleto de Xpandómero limitado” evita el enlace selectivamente escindible entre los sustratos de monómero y se forma por restos de esqueleto enlazados con enlaces  $\chi$ , siendo cada resto de esqueleto un anclaje. Puede observarse que el esqueleto de Xpandómero limitado conecta mediante los enlaces selectivamente escindibles del esqueleto primario, y seguirá covalentemente intacto cuando estos enlaces selectivamente escindibles se escindan y el esqueleto primario se fragmente.

El enlace  $\chi$  de anclaje (reticulación de los grupos conectores  $\delta$  y  $\varepsilon$ ) va generalmente precedido de acoplamiento enzimático de los sustratos de monómero para formar el esqueleto primario, con, por ejemplo, enlaces fosfodiéster entre bases adyacentes. En la estructura mostrada aquí, el esqueleto primario de la hebra hija se ha formado, y el inter-sustrato, se representan por una tilde (~) para indicar que son selectivamente escindibles. Después de disociar o degradar la hebra molde diana, escindiendo los enlaces selectivamente escindibles (que incluyen los enlaces intra-anclaje), el Xpandómero limitado se libera y se convierte en el producto de Xpandómero. Métodos para la disociación de la hebra molde incluyen desnaturalización por calor, o digestión selectiva con una nucleasa, o degradación química. Un método de escisión de selección usa digestión con nucleasa en la que, por ejemplo, enlaces fosfodiéster del esqueleto primario se digieren por una nucleasa y enlaces anclaje a anclaje son resistentes a nucleasa.

La Figura 60C es una representación del producto de Xpandómero de clase VI después de la disociación de la hebra molde y después de la escisión de los enlaces selectivamente escindibles (incluyendo aquellos en el esqueleto primario y, si no se han escindido ya, los enlaces intra-anclaje). La hebra del producto de Xpandómero contiene una pluralidad de subunidades  $\kappa$ , en la que  $\kappa$  indica la subunidad  $k^n$  en una cadena de  $m$  subunidades que constituyen la hebra hija, en la que  $\kappa = 1, 2, 3$  a  $m$ , en la que  $m > 10$ , generalmente  $m > 50$ , y normalmente  $m > 500$  o  $> 5.000$ . Cada subunidad está formada de un anclaje en su configuración expandida y se estira a su longitud entre los enlaces  $\chi$  de subunidades adyacentes. El sustrato lateral está unido al anclaje en cada subunidad. Cada subunidad, un motivo de subunidad  $\alpha$ , contiene información genética específica de especie establecida por el ensamblaje dirigido por molde del producto intermedio de Xpandómero (hebra hija).

La Figura 60D muestra la construcción de sustrato de la Figura 60A como modelo molecular, en la que el miembro de sustrato de monómero, representado con un residuo de nucleobase (606), está unido al anclaje por un enlace (607) del primer resto terminal del anclaje. También está dispuesto sobre el primer resto terminal un grupo conector (608), mostrado como  $\varepsilon$  en la Figura 60A. Un segundo grupo conector (609), mostrado como  $\delta$  en la Figura 60A, está dispuesto en el segundo resto terminal en el extremo distal del anclaje. Un enlace intra-anclaje selectivamente escindible (602), representado por los triángulos contiguos, se muestra que limita el anclaje enlazando el primer y segundo resto terminal. Los grupos conectores  $\varepsilon$  y  $\delta$  están posicionados para no interactuar y para alinearse preferentemente cerca de los lados de  $R^1$  y  $R^2$  del sustrato, respectivamente. El bucle de anclaje (600) mostrado aquí tiene tres indicadores (590,591,592), que también pueden ser específicos de especies de motivos.

La Figura 60E muestra la construcción de sustrato después de la incorporación en el Xpandómero producto. Las subunidades se escinden y se expanden y se enlazan por enlaces  $\chi$  (580,581), representados como un óvalo abierto, formados enlazando los grupos conectores  $\delta$  y  $\varepsilon$  referidos en la Figura 60A. Una subunidad se indica por líneas de puntos que agrupan verticalmente la subunidad de repetición, como se representa por los corchetes en la Figura 60C adjunta.

En el producto de Xpandómero de la Figura 60E, el esqueleto primario se ha fragmentado y no es covalentemente contiguo debido a que cualquier enlace directo entre los sustratos de subunidades adyacentes se ha escindido. Mediante el proceso de escisión, el Xpandómero limitado se libera para convertirse en el producto de Xpandómero. Los miembros de anclaje que estuvieron anteriormente en la configuración limitada están ahora en la configuración expandida, sirviendo así para estirar linealmente la información de secuencias de la diana molde. La expansión de los anclajes reduce la densidad lineal de la información de secuencias a lo largo del Xpandómero y proporciona una plataforma para aumentar el tamaño y la abundancia de indicadores, que a su vez mejora la señal con respecto al ruido para la detección y descodificación de la secuencia molde.

Aunque el anclaje se representa como una construcción de indicador con tres grupos indicadores, diversas construcciones de indicador pueden estar presentes sobre el anclaje, y pueden comprender indicadores individuales que identifican monómero o el anclaje puede ser un polímero desnudo. En algunos casos, uno o más precursores de indicador están presentes sobre el anclaje, y los indicadores se unen por afinidad o se unen covalentemente tras el ensamblaje del producto de Xpandómero.

Volviendo ahora a la Figura 61, se muestra un método de extensión de una única base (SBE) con construcciones de sustrato monomérico. Los moldes diana adaptados a los extremos (o secuencias de molde diana aleatorias, dependiendo de la naturaleza de los cebadores inmovilizados) se hibridan primero con los cebadores inmovilizados. Antes de la etapa I, los moldes inmovilizados (611) se ponen en contacto con una biblioteca de construcciones de sustrato monoméricas, un miembro de la cual se muestra para ilustración (612), y polimerasa (P). Las condiciones se ajustan para la polimerización dirigida por molde. En este ejemplo, los extremos 5' del primer sustrato monomérico ( $R^2$  de las Figuras 60A y 60D) se polimerizan para iniciar la hebra hija naciente. El sustrato se ha

sustituido en los extremos 3' ( $R^1$  de las Figuras 60A y 60D), para bloquear reversiblemente la extensión adicional. Esto se muestra en más detalle en la Figura 61a aumentada (líneas de puntos), en la que el sustrato monomérico se muestra confinando el cebador. Este sustrato es un 5'-trifosfato, y se forma un enlace fosfodiéster con el cebador por acción de la polimerasa. Debe observarse que el grupo funcional reactivo ( $\delta$  de la Figura 60A) mostrado en la Figura 61c normalmente está desactivado (mediante reacción u otros medios) en el primer sustrato añadido (como se muestra en la Figura 61a por el rectángulo).

En la etapa I de la Figura 61, el grupo de bloqueo se elimina (sin rectángulo) para permitir la adición de otro monómero (extensión). Los métodos para bloquear reversiblemente el extremo 3' incluyen el uso de Pd para catalizar la eliminación de un grupo alilo para regenerar un extremo 3'-hidroxilo viable, o el uso de nucleótidos terminados con 3'-O-(2-nitrobencilo), en los que los hidroxilos activos pueden regenerarse por exposición a una fuente de UV para fotoescindir el resto terminal como se describe por Ju et al. ("Four-color DNA sequencing by synthesis using cleavable fluorescent nucleotide reversible terminators", PNAS 26;103(52): 19635-19640, 2006, y "Four-color DNA sequencing by synthesis on a chip using photocleavable fluorescent nucleotides", PNAS 26;102(17):5926-31, 2005). El 3'-OH regenerado se muestra en más detalle en la Figura 61 b en vista aumentada. En esta vista, los extremos confinados de la construcción de sustrato adyacente y cebador se polimerizan y el extremo 3'OH de la hebra hija naciente se activa eliminando el grupo de bloqueo. En la etapa II, el molde se pone en contacto con una biblioteca de construcciones de sustrato monomérico, en la que los grupos funcionales  $\delta$  son reactivos bajo condiciones controladas, y un sustrato de monómero se polimeriza con la hebra hija naciente. Esta polimerización coloca los grupos  $\epsilon$  (608) y  $\delta$  (609) de las construcciones de sustrato confinadas como se muestra en la Figura 61 c. En la etapa III, bajo condiciones controladas, estos grupos reaccionan para formar un enlace  $\chi$  (580), como se muestra en más detalle en la Figura 61 d.

Como se indica, el ciclado a través de las etapas II, III y IV extiende la hija naciente cada vez por sustrato adicional (construcción). Normalmente, puede usarse una etapa de lavado para eliminar reactivos sin reaccionar entre etapas. El proceso es así análogo a lo que se llama en la bibliografía, "extensión de una única base" cíclica. El proceso se muestra con polimerasa, P, pero puede adaptarse para una ligasa o un protocolo químico de ligación adecuado para unir construcciones de sustrato en una síntesis dirigida por molde. La etapa V muestra el producto intermedio de la hebra hija para el Xpandómero (605). Este producto intermedio puede disociarse del molde y el cebador, por ejemplo, con una nucleasa que ataca el esqueleto primario de la hebra hija, aliviando así los anclajes limitados y liberando el producto de Xpandómero.

Al igual que con todos los métodos de SBE, el eficaz lavado entre ciclos es útil para reducir reacciones secundarias no deseables. Para facilitar adicionalmente la incorporación de bases individuales mediante regiones de molde con alta estructura secundaria, la temperatura de extensión puede variarse a través de cada ciclo de extensión y/o aditivos o adyuvantes, tales como betaína, TMACL, o PEG, también pueden añadirse para neutralizar los efectos de la estructura secundaria (como se hace en protocolos convencionales de extensión de polimerasa). Y, si fuera necesario, la estequiometría de las especies de construcción de sustrato puede variarse para compensar el sesgo de la reacción que favorece ciertas bases, tales como C o G.

Un método alternativo para producir Xpandómeros de clase VI es hacer la polimerización procesiva basada en polimerasa. ADN y ARN polimerasas, además de cualquier enzima que funcione similarmente, demostraron catalizar la polimerización precisa de RT-NTP, grupos R de bloqueo reversiblemente terminales ausentes, pueden considerarse para este enfoque.

Se conoce un amplio intervalo de químicas de reticulación en este campo, y son útiles para la formación de enlaces  $\chi$ . Éstos incluyen el uso de ésteres de NHS con aminas, maleimidias con sulfhidrilos, imidoésteres con aminas, EDC con sulfhidrilos y carboxilos para reacciones con aminas, piridildisulfuros con sulfhidrilos, etc. Otras realizaciones implican el uso de grupos funcionales como hidrazida (HZ) y 4-formilbenzoato (4FB), que pueden entonces hacerse reaccionar adicionalmente para enlazar subunidades. En una opción, dos químicas de enlace diferentes,  $\epsilon_1/\delta_1$  y  $\epsilon_2/\delta_2$  (también denominadas L1/L1' y L2/L2' en otras partes de este documento) que reaccionan para formar enlaces  $\chi_1$  y  $\chi_2$ , respectivamente, y pueden usarse para funcionalizar diferencialmente dos conjuntos de construcciones de sustrato de RT-NTP. Por ejemplo, si se realiza un ciclo de SBE con un conjunto de RT-NTP funcionalizado con  $\epsilon_1/\delta_2$ , el siguiente ciclo de SBE usaría el conjunto de  $\epsilon_2/\delta_1$ , haciendo que una reacción de  $\delta_2/\epsilon_2$  colocados formara  $\chi_1$ . La activación ordenada de los pares de reticulación es útil para minimizar los errores de reticulación y reacciones secundarias no deseadas.

#### Construcciones monoméricas de clase VII

Las moléculas de clase VII son análogos de las de clase VI descritas previamente. La diferencia primaria es que el enlace escindible es entre el anclaje y el sustrato en lugar de entre los sustratos. En la Figura 62, las construcciones de sustrato monomérico de clase VII (un tipo de RT-NTP) se desvelan en más detalle. Las Figuras 62A a 62C se leen de izquierda a derecha, mostrando primero la construcción de sustrato monomérico (precursor de Xpandómero que tiene un único residuo de nucleobase), luego la hebra hija del dúplex intermedia en el centro, y a la derecha el producto de Xpandómero preparado para la secuenciación.



Como se muestra en la Figura 62A, las construcciones de sustrato monomérico de clase VII tienen un anclaje, T (620), unido por un conector selectivamente escindible (621) a un primer resto terminal a un residuo de nucleobase del sustrato, N. Otro grupo conector,  $\epsilon$  (622), está dispuesto en el primer resto terminal del anclaje próximo a  $R^1$ . En el extremo distal del anclaje T un segundo resto terminal con un segundo grupo conector,  $\delta$  (623), está posicionado próximo a  $R^2$ . El segundo resto terminal del anclaje está asegurado al primer resto terminal en proximidad con la nucleobase por una reticulación selectivamente escindible intra-anclaje (624) (o por otra limitación). La reticulación escindible intra-anclaje se indica aquí por línea de puntos, que puede indicar, por ejemplo, un enlace disulfuro o un conector fotoescindible. Esta limitación previene que el anclaje se alargue o expanda y está en una "configuración limitada". Bajo ensamblaje dirigido por molde, los sustratos forman un dúplex con el molde diana de forma que los sustratos estén confinados. Bajo condiciones controladas, los grupos conectores colocados  $\delta$  y  $\epsilon$  de los sustratos confinados se enlazan para formar un enlace entre las construcciones de sustrato adyacentes. Los grupos conectores  $\delta$  y  $\epsilon$  de una construcción de sustrato monomérico no forman un enlace intra-sustratos debido a limitaciones del posicionamiento. Químicas de enlace y de protección/desprotección adecuadas para  $\delta$ ,  $\epsilon$  y  $\chi$  se detallan en la descripción general de construcciones monoméricas.

$R^1$  y  $R^2$  son grupos terminales configurados según convenga para el protocolo de síntesis en el que se usa la construcción de sustrato. Por ejemplo,  $R^1 = 5'$ -fosfato y  $R^2 = 3'$ -OH, encontrarían uso en un protocolo de ligación, y  $R^1 = 5'$ -trifosfato y  $R^2 = 3'$ -OH para un protocolo de polimerasa I. Opcionalmente,  $R^2$  puede configurarse con un grupo de bloqueo reversible para la adición cíclica de un único sustrato. Alternativamente,  $R^1$  y  $R^2$  pueden configurarse con los grupos conectores terminales para el acoplamiento químico o con grupos no conectores para un protocolo de solo hibridación.  $R^1$  y  $R^2$  pueden ser del tipo general XR, en la que X es un grupo de enlace y R es un grupo funcional.

Durante el ensamblaje, la construcción de sustrato monomérico se polimeriza primero sobre el extremo extensible de la hebra hija naciente por un proceso de polimerización dirigida por molde usando un molde monocatenario como guía. Generalmente, este proceso se inicia a partir de un cebador y avanza en la dirección de  $5'$  a  $3'$ . Generalmente, se usa una ADN polimerasa u otra polimerasa para formar la hebra hija, y las condiciones se seleccionan de manera que se obtenga una copia complementaria de la hebra molde. Posteriormente, se provoca que el grupo conector  $\delta$ , que está ahora colocado con el grupo conector  $\epsilon$  del anclaje de subunidad adyacente, se reticule y forme un enlace  $\chi$ , que es un enlace inter-anclaje inter-subunidad. Los enlaces  $\chi$  unen los anclajes en una cadena continua, formando un producto intermedio llamado la "hebra hija de dúplex", como se muestra en la Figura 62B. Después de formarse el enlace  $\chi$ , puede romperse el enlace intra-anclaje.

La hebra hija de dúplex (Figura 62B) es un hetero-copolímero con subunidades mostradas en corchetes. El esqueleto primario (-N)- $\kappa$ , hebra molde (-N')- $\kappa$  y anclaje (T) se muestran como una hebra hija duplexada, en la que  $\kappa$  indica una pluralidad de subunidades de repetición. Cada subunidad de la hebra hija es un "motivo" de repetición y los motivos tienen variabilidad específica de especie, indicada aquí por el superíndice  $\alpha$ . La hebra hija se forma a partir de especies de la construcción de sustrato monomérico seleccionadas por un proceso dirigido por molde de una biblioteca de especies de motivos, uniéndose el sustrato de monómero de cada especie de construcción de sustrato a un nucleótido complementario correspondiente sobre la hebra molde diana. De esta forma, la secuencia de residuos de nucleobase (es decir, esqueleto primario) de la hebra hija es una copia complementaria contigua de la hebra molde diana.

La tilde (~) indica un enlace selectivamente escindible mostrado aquí como el anclaje al conector de sustrato. Éstos son selectivamente escindibles para liberar y expandir los anclajes (y el Xpandómero) sin degradar el propio Xpandómero.

La hebra hija está compuesta por un precursor de Xpandómero llamado el "Xpandómero limitado" que está adicionalmente compuesto de anclajes en la "configuración limitada". Cuando los anclajes se convierten en su "configuración expandida", el Xpandómero limitado se convierte en el producto de Xpandómero. Los anclajes están limitados por los enlaces  $\chi$  inter-subunidad, el enlace escindible al sustrato y, opcionalmente, los enlaces intra-anclaje si todavía están presentes. El enlace  $\chi$  une el primer resto terminal del anclaje de una primera subunidad con el segundo resto terminal del anclaje en el extremo confinado de una segunda subunidad y se forma enlazando los grupos conectores colocados,  $\epsilon$  de la primera subunidad y  $\delta$  de la segunda subunidad.

Puede observarse que la hebra hija tiene dos esqueletos, un "esqueleto primario" y un "esqueleto de Xpandómero limitado". El esqueleto primario está compuesto por los sustratos monoméricos contiguamente confinados y polimerizados. El esqueleto de Xpandómero limitado evita el enlace selectivamente escindible que conecta con el sustrato y se forma por restos de esqueleto enlazados por enlace  $\chi$ , siendo cada resto de esqueleto un anclaje. Puede observarse que el esqueleto de Xpandómero limitado conecta mediante los enlaces selectivamente escindibles conectados al esqueleto primario, y seguirá covalentemente intacto cuando estos enlaces selectivamente escindibles se escindan y el esqueleto primario se disocie o fragmente.

El enlace  $\chi$  de anclaje (reticulación de los grupos conectores  $\delta$  y  $\epsilon$ ) va generalmente precedido de acoplamiento enzimático de los sustratos de monómero para formar el esqueleto primario con, por ejemplo, enlaces fosfodiéster entre bases adyacentes. En la estructura mostrada aquí, el esqueleto primario de la hebra hija se ha formado, y el

inter-sustrato, se representan por una tilde (~) para indicar que son selectivamente escindibles. Después de disociar o degradar la hebra molde diana, escindiendo los enlaces selectivamente escindibles (que incluyen los enlaces intra-anclaje), el Xpandómero limitado se libera y se convierte en el producto de Xpandómero. Métodos para la disociación de la hebra molde incluyen desnaturalización por calor, o digestión selectiva con una nucleasa, o degradación química. Un método de escisión de selección usa digestión con nucleasa en la que, por ejemplo, enlaces fosfodiéster del esqueleto primario se digieren por una nucleasa y enlaces anclaje a anclaje son resistentes a nucleasa.

La Figura 62C es una representación del producto de Xpandómero de clase VII después de la disociación de la hebra molde y después de la escisión de los enlaces selectivamente escindibles (incluyendo aquellos unidos al esqueleto primario y, si no se han escindido ya, los enlaces intra-anclaje). La hebra del producto de Xpandómero contiene una pluralidad de subunidades  $\kappa$ , en la que indica la subunidad  $\kappa^n$  en una cadena de  $m$  subunidades que constituyen la hebra hija, en la que  $\kappa = 1, 2, 3$  a  $m$ , en la que  $m > 10$ , generalmente  $m > 50$ , y normalmente  $m > 500$  o  $> 5.000$ . Cada subunidad está formada de un anclaje (620) en su configuración expandida y se estira a su longitud entre los enlaces  $\chi$  de subunidades adyacentes. El esqueleto primario se ha eliminado completamente. Cada subunidad, un motivo de subunidad  $\alpha$ , contiene información genética específica de especie establecida por el ensamblaje dirigido por molde del producto intermedio de Xpandómero (hebra hija).

La Figura 62D muestra la construcción de sustrato de la Figura 62A como modelo molecular, en la que el miembro de sustrato de monómero, representado con un residuo de nucleobase (626), enlazado por un enlace selectivamente escindible (625) al primer resto terminal del anclaje. También está dispuesto sobre el primer resto terminal un grupo conector (629), mostrado como  $\varepsilon$  en la Figura 62A. Un segundo grupo conector (628), mostrado como  $\delta$  en la Figura 62A, está dispuesto en el segundo resto terminal en el extremo distal del anclaje. Se muestra un enlace intra-anclaje selectivamente escindible (627), representado por los triángulos contiguos, que limita el anclaje enlazando el primer y segundo resto terminal. Los grupos conectores  $\varepsilon$  (629) y  $\delta$  (628) están posicionados para no interactuar y para alinearse preferentemente cerca de los lados de  $R^1$  y  $R^2$  del sustrato, respectivamente. El bucle de anclaje mostrado aquí tiene tres indicadores (900,901,902), que también pueden ser específicos de especies de motivos.

La Figura 62E muestra la construcción de sustrato después de la incorporación en el Xpandómero producto. Las subunidades se escinden y se expanden y se enlazan por enlaces  $\chi$  (910), representados como un óvalo abierto, formados enlazando los grupos conectores  $\delta$  y  $\varepsilon$  referidos en la Figura 62A. Una subunidad se indica por líneas de puntos que agrupan verticalmente la subunidad de repetición, como se representa por los corchetes en la Figura 62C adjunta.

En el producto de Xpandómero de la Figura 62E, el esqueleto primario se ha disociado o fragmentado y está separado del Xpandómero. Mediante el proceso de escisión, el Xpandómero limitado se libera para convertirse en el producto de Xpandómero. Los miembros de anclaje que estuvieron anteriormente en la configuración limitada están ahora en la configuración expandida, sirviendo así para estirar linealmente la información de secuencias de la diana molde. La expansión de los anclajes reduce la densidad lineal de la información de secuencias a lo largo del Xpandómero y proporciona una plataforma para aumentar el tamaño y la abundancia de indicadores, que a su vez mejora la señal con respecto al ruido para la detección y descodificación de la secuencia molde.

Aunque el anclaje se representa como una construcción de indicador con tres grupos indicadores, diversas construcciones de indicador pueden estar presentes sobre el anclaje, y pueden comprender indicadores individuales que identifican monómero o el anclaje puede ser un polímero desnudo. En algunos casos, uno o más precursores de indicador están presentes sobre el anclaje, y los indicadores se unen por afinidad o se unen covalentemente tras el ensamblaje del producto de Xpandómero.

#### Construcciones monoméricas de clase VIII

Las moléculas de clase VIII son análogos de la clase VI descrita previamente. La diferencia primaria es que el grupo conector,  $\varepsilon$ , está conectado directamente al sustrato en lugar de al anclaje. En la Figura 63, los presentes inventores describen construcciones de sustrato monomérico de clase VIII (un tipo de RT-NTP) en más detalle. Las Figuras 63A a 63C se leen de izquierda a derecha, mostrando primero la construcción de sustrato monomérico (precursor de Xpandómero que tiene un único residuo de nucleobase), luego la hebra hija del dúplex intermedia en el centro, y a la derecha el producto de Xpandómero preparado para la secuenciación.

Como se muestra en la Figura 63A, las construcciones de sustrato monomérico de clase VII tienen un anclaje, T (630), unido por un enlace (631) de un primer resto terminal a un residuo de nucleobase del sustrato, N. En el extremo distal del anclaje (632), un segundo resto terminal con un segundo grupo conector,  $\delta$ , está posicionado preferencialmente próximo a  $R^2$ . El segundo resto terminal del anclaje está asegurado al primer resto terminal en proximidad con la nucleobase por una reticulación selectivamente escindible intra-anclaje (o por otra limitación). La reticulación escindible intra-anclaje (633) se indica aquí por línea de puntos, que puede indicar, por ejemplo, un enlace disulfuro o un conector fotoescindible. Esta limitación previene que el anclaje se alargue o expanda y se dice que está en su "configuración limitada". Un grupo conector,  $\varepsilon$  (635), está unido al sustrato de monómero preferencialmente próximo a  $R^1$ . Bajo ensamblaje dirigido por molde, los sustratos forman un dúplex con el molde

diana de forma que los sustratos estén confinados. Bajo condiciones controladas, los grupos conectores  $\delta$  y  $\varepsilon$  de los sustratos confinados está colocados y se enlazan para formar un enlace entre las construcciones de sustrato adyacentes. Los grupos conectores  $\delta$  y  $\varepsilon$  de una construcción de sustrato monomérico no forman un enlace intra-sustratos debido a limitaciones del posicionamiento. Químicas de enlace y de protección/desprotección adecuadas para  $\delta$ ,  $\varepsilon$  y  $\chi$  se detallan en la descripción general de construcciones monoméricas.

$R^1$  y  $R^2$  son grupos terminales configurados según convenga para el protocolo de síntesis en el que se usa la construcción de sustrato. Por ejemplo,  $R^1 = 5'$ -fosfato y  $R^2 = 3'$ -OH, encontrarían uso en un protocolo de ligación, y  $R^1 = 5'$ -trifosfato y  $R^2 = 3'$ -OH para un protocolo de polimerasa I. Opcionalmente,  $R^2$  puede configurarse con un grupo de bloqueo reversible para la adición cíclica de un único sustrato. Alternativamente,  $R^1$  y  $R^2$  pueden configurarse con los grupos conectores terminales para el acoplamiento químico o con grupos no conectores para un protocolo de solo hibridación.  $R^1$  y  $R^2$  pueden ser del tipo general XR, en la que X es un grupo de enlace y R es un grupo funcional.

Durante el ensamblaje, la construcción de sustrato monomérico se polimeriza primero sobre el extremo extensible de la hebra hija naciente por un proceso de polimerización dirigida por molde usando un molde monocatenario como guía. Generalmente, este proceso se inicia a partir de un cebador y avanza en la dirección de  $5'$  a  $3'$ . Generalmente, se usa una ADN polimerasa u otra polimerasa para formar la hebra hija, y las condiciones se seleccionan de manera que se obtenga una copia complementaria de la hebra molde. Posteriormente, se provoca que el grupo conector  $\delta$ , que está ahora colocado con el grupo conector  $\varepsilon$  del anclaje de subunidad adyacente, se reticule y forme un enlace  $\chi$ , que es u enlace inter-subunidad. Los enlaces  $\chi$  proporcionan u segundo enlace entre subunidades (los enlaces inter-sustratos polimerizados so los primeros) y forman un producto intermedio llamado la "hebra hija de dúplex", como se muestra en la Figura 63B.

La hebra hija de dúplex (Figura 63B) es un hetero-copolímero con subunidades mostradas en corchetes. El esqueleto primario ( $\sim N\sim$ ) $\kappa$ , hebra molde ( $-N'$ ) $\kappa$  y anclaje (T) se muestran como una hebra hija duplexada, en la que  $\kappa$  indica una pluralidad de subunidades de repetición. Cada subunidad de la hebra hija es un "motivo" de repetición y los motivos tienen variabilidad específica de especie, indicada aquí por el superíndice  $\alpha$ . La hebra hija se forma a partir de especies de la construcción de sustrato monomérico seleccionadas por un proceso dirigido por molde de una biblioteca de especies de motivos, uniéndose el sustrato de monómero de cada especie de construcción de sustrato a un nucleótido complementario correspondiente sobre la hebra molde diana. De esta forma, la secuencia de residuos de nucleobase (es decir, esqueleto primario) de la hebra hija es una copia complementaria contigua de la hebra molde diana.

Cada tilde ( $\sim$ ) indica un enlace selectivamente escindible mostrado aquí como los enlaces inter-sustratos. Éstos son necesariamente selectivamente escindibles para liberar y expandir los anclajes (y el Xpandómero) sin degradar el propio Xpandómero.

La hebra hija está compuesta por un precursor de Xpandómero llamado el "Xpandómero limitado" que está adicionalmente compuesto de anclajes en la "configuración limitada". Cuando los anclajes se convierten en su "configuración expandida", el Xpandómero limitado se convierte en el producto de Xpandómero. Los anclajes están limitados por los enlaces  $\chi$  inter-subunidad, la unión de sustratos y, opcionalmente, los enlaces intra-anclaje si todavía están presentes. El enlace  $\chi$  une el sustrato de una primera subunidad con el anclaje del segundo resto terminal en el extremo confinado de una segunda subunidad y se forma enlazando los grupos conectores colocados,  $\varepsilon$  de la primera subunidad y  $\delta$  de la segunda subunidad.

Puede observarse que la hebra hija tiene dos esqueletos, un "esqueleto primario" y el esqueleto del "Xpandómero limitado". El esqueleto primario está compuesto por los sustratos monoméricos contiguamente confinados y polimerizados. El "esqueleto de Xpandómero limitado" evita el enlace selectivamente escindible entre los sustratos de monómero y se forma por restos de esqueleto enlazados por enlace  $\chi$ , siendo cada resto de esqueleto un anclaje enlazado a un sustrato que entonces se enlaza con el siguiente anclaje de resto de esqueleto con un enlace  $\chi$ . Puede observarse que el esqueleto de Xpandómero limitado conecta mediante los enlaces selectivamente escindibles del esqueleto primario, y seguirá covalentemente intacto cuando estos enlaces selectivamente escindibles se escindan y el esqueleto primario se fragmente.

El enlace  $\chi$  de anclaje (reticulación de los grupos conectores  $\delta$  y  $\varepsilon$ ) va generalmente precedido de acoplamiento enzimático de los sustratos de monómero para formar el esqueleto primario, con, por ejemplo, enlaces fosfodiéster entre bases adyacentes. En la estructura mostrada aquí, el esqueleto primario de la hebra hija se ha formado, y los enlaces inter-sustratos se representan por una tilde ( $\sim$ ) para indicar que son selectivamente escindibles. Después de disociar o degradar la hebra molde diana, escindiendo los enlaces selectivamente escindibles (que incluyen los enlaces intra-anclaje), el Xpandómero limitado se libera y se convierte en el producto de Xpandómero. Métodos para la disociación de la hebra molde incluyen desnaturalización por calor, o digestión selectiva con una nucleasa, o degradación química. Un método de escisión de selección usa digestión con nucleasa en la que, por ejemplo, enlaces fosfodiéster del esqueleto primario se digieren por una nucleasa y enlaces anclaje a anclaje son resistentes a nucleasa.

La Figura 63C es una representación del producto de Xpandómero de clase VIII después de la disociación de la hebra molde y después de la escisión de los enlaces selectivamente escindibles (incluyendo aquellos en el esqueleto primario y, si no se han escindido ya, los enlaces intra-anclaje). La hebra del producto de Xpandómero contiene una pluralidad de subunidades  $\kappa$ , en la que indica la subunidad  $\kappa^n$  en una cadena de  $m$  subunidades que constituyen la hebra hija, en la que  $\kappa = 1, 2, 3$  a  $m$ , en la que  $m > 10$ , generalmente  $m > 50$ , y normalmente  $m > 500$  o  $> 5.000$ . Cada subunidad está formada de un anclaje en su configuración expandida y se estira a su longitud entre los enlaces  $\chi$  de subunidades adyacentes. El sustrato lateral está unido al anclaje en cada subunidad. Cada subunidad, un motivo de subunidad  $\alpha$ , contiene información genética específica de especie establecida por el ensamblaje dirigido por molde del producto intermedio de Xpandómero (hebra hija).

La Figura 63D muestra la construcción de sustrato de la Figura 63A como modelo molecular, en la que el miembro de sustrato de monómero, representado con un residuo de nucleobase (634), está unido al anclaje por un enlace (631) del primer resto terminal del anclaje. Un segundo grupo conector (639), mostrado como  $\delta$  en la Figura 63A, está dispuesto en el segundo resto terminal en el extremo distal del anclaje. Se muestra un enlace intra-anclaje selectivamente escindible (633), representado por los triángulos contiguos, que limita el anclaje enlazando el primer y segundo resto terminal. También unidos al sustrato está un grupo conector (638), mostrado como  $\varepsilon$  en la Figura 63A. Los grupos conectores  $\varepsilon$  (638) y  $\delta$  (639) están posicionados para no interactuar y para alinearse preferentemente cerca de los lados de  $R^1$  y  $R^2$  del sustrato, respectivamente. El bucle de anclaje mostrado aquí tiene tres indicadores (900,901,902), que también pueden ser específicos de especies de motivos. La reticulación intra-anclaje escindible (633) se muestra en la Figura 63D y la Figura 63E que está posicionada más próxima al sustrato que  $\delta$  (639). En las Figuras 60D y Figura 60E, las posiciones están cambiadas. Este posicionamiento puede estar en cualquier sitio y en una realización ambas funciones de conector pueden estar sobre un único grupo multifuncional.

La Figura 63E muestra la construcción de sustrato después de la incorporación en el Xpandómero producto. Las subunidades se escinden y se expanden y se enlazan por enlaces  $\chi$  (970,971), representados como un óvalo abierto, formados enlazando los grupos conectores  $\delta$  y  $\varepsilon$  referidos en la Figura 63A. Una subunidad se indica por líneas de puntos que agrupan verticalmente la subunidad de repetición, como se representa por los corchetes en la Figura 63C adjunta.

En el producto de Xpandómero de la Figura 63E, el esqueleto primario se ha fragmentado y no es covalentemente contiguo debido a que cualquier enlace directo entre los sustratos de subunidades adyacentes se ha escindido. Mediante el proceso de escisión, el Xpandómero limitado se libera para convertirse en el producto de Xpandómero. Los miembros de anclaje que estuvieron anteriormente en la configuración limitada están ahora en la configuración expandida, sirviendo así para estirar linealmente la información de secuencias de la diana molde. La expansión de los anclajes reduce la densidad lineal de la información de secuencias a lo largo del Xpandómero y proporciona una plataforma para aumentar el tamaño y la abundancia de indicadores, que a su vez mejora la señal con respecto al ruido para la detección y descodificación de la secuencia molde.

Aunque el anclaje (630) se representa como una construcción de indicador con tres grupos indicadores, diversas construcciones de indicador pueden estar presentes sobre el anclaje, y pueden comprender indicadores individuales que identifican monómero o el anclaje puede ser un polímero desnudo. En algunos casos, uno o más precursores de indicador están presentes sobre el anclaje, y los indicadores se unen por afinidad o se unen covalentemente tras el ensamblaje del producto de Xpandómero.

#### 40 Construcciones monoméricas de clase IX

Una construcción de sustrato de clase IX se distingue de la otra RT-NTP debido a que tiene dos puntos de unión de anclaje a los que un anclaje libre está unido después de ensamblarse el esqueleto primario. La Figura 64 describe construcciones de sustrato monomérico de clase IX (a tipo de RT-NTP) en más detalle. Las Figuras 64A a 64C se leen de izquierda a derecha, mostrando primero la construcción de sustrato monomérico (precursor de Xpandómero que tiene un único residuo de nucleobase), luego la hebra hija del dúplex intermedia en el centro, y a la derecha el producto de Xpandómero preparado para la secuenciación.

Como se muestra en la Figura 64A, la construcción de sustrato monomérico de clase IX tiene un residuo de nucleobase del sustrato, N, con dos sitios de unión de anclaje, los grupos conectores  $\delta_1$  y  $\delta_2$ . También se muestra un anclaje libre, T (640), con los grupos conectores  $\varepsilon_1$  y  $\varepsilon_2$  de un primer y segundo resto terminal del anclaje. El primer y segundo restos terminales del anclaje limitan el libre anclaje por una reticulación selectivamente escindible intra-anclaje (647) y sirve para colocar los grupos conectores  $\varepsilon_1$  y  $\varepsilon_2$ . La reticulación escindible se indica aquí por línea de puntos, puede indicar, por ejemplo, un enlace disulfuro o un conector fotoescindible. Esta limitación previene que el anclaje se alargue o expanda y se dice que está en su "configuración limitada". Los grupos conectores,  $\delta$  y  $\delta_2$ , están unidos al sustrato de monómero orientado próximo a  $R^1$  y  $R^2$ , respectivamente. Bajo síntesis dirigida por molde, los sustratos forman un dúplex con el molde diana y el grupo conector de una construcción de sustrato, por ejemplo  $\delta_1$  y el grupo conector de la construcción de sustrato confinada, por ejemplo  $\delta_2$ , están colocados. Bajo condiciones controladas, estos conectores colocados se ponen en contacto con los conectores  $\varepsilon$  colocados en el extremo del anclaje libre. Se producen dos reacciones de enlaces selectivos,  $\varepsilon_1$  y  $\delta_1$  para formar  $\chi^1$  y  $\varepsilon_2$  y  $\delta_2$  para formar  $\chi^2$ ; los sustratos adyacentes están ahora conectados por el anclaje limitado. Químicas de enlace y de

protección/desprotección adecuadas para  $\delta$ ,  $\varepsilon$  y  $\chi$  se detallan en la descripción general de construcciones monoméricas.

$R^1$  y  $R^2$  son grupos terminales configurados según convenga para el protocolo de síntesis en el que se usa la construcción de sustrato. Por ejemplo,  $R^1 = 5'$ -fosfato y  $R^2 = 3'$ -OH, encontrarían uso en un protocolo de ligación, y  $R^1 = 5'$ -trifosfato y  $R^2 = 3'$ -OH para un protocolo de polimerasa I. Opcionalmente,  $R^2$  puede configurarse con un grupo de bloqueo reversible para la adición cíclica de un único sustrato. Alternativamente,  $R^1$  y  $R^2$  pueden configurarse con los grupos conectores terminales para el acoplamiento químico o con grupos no conectores para un protocolo de solo hibridación.  $R^1$  y  $R^2$  pueden ser del tipo general XR, en la que X es un grupo de enlace y R es un grupo funcional.

5 Durante el ensamblaje, la construcción de sustrato monomérico (sin anclaje) se polimeriza sobre el extremo extensible de la hebra hija naciente por un proceso de polimerización dirigida por molde usando un molde monocatenario como guía. Generalmente, este proceso se inicia a partir de un cebador y avanza en la dirección de 5' a 3'. Generalmente, se usa una ADN polimerasa u otra polimerasa para formar la hebra hija, y las condiciones se seleccionan de manera que se obtenga una copia complementaria de la hebra molde. El grupo conector  $\delta_1$  está  
10 ahora colocado con el grupo conector  $\delta_2$  del sustrato de monómero adyacente. Después de la polimerización del esqueleto primario, los anclajes libres se reticular para formar enlaces  $\chi^1$  y enlaces  $\chi^2$  entre los dos extremos de anclaje y dos de los sustratos adyacentes.

En una realización, los anclajes libres no tienen información de secuencias y se llaman “desnudos”. En este caso, puede usarse química de un único conector de manera que los grupos conectores  $\delta_1$  y  $\delta_2$  sean los mismos y los  
20 grupos conectores  $\varepsilon_1$  y  $\varepsilon_2$  sean los mismos con un tipo de enlace  $\chi$ .

En la realización en la que el anclaje libre comprende información de tipo base (como en los indicadores), hay anclajes libres específicos de especie. Generalmente, hay cuatro tipos de base que requerirían cuatro tipos de anclaje libre con la información de bases correspondiente. Pueden usarse varios métodos para enlazar las especies de anclaje libre a su base correcta. En un método, se usan cuatro químicas de enlace heteroespecífico para  
25 diferenciar adicionalmente el par de conectores  $\delta_1$  y  $\varepsilon_1$ , ahora se expresan como  $\delta_{1\alpha}$  y  $\varepsilon_{1\alpha}$  en la que  $\alpha$  es uno de cuatro tipos y forman tipos de enlace  $\chi^{1\alpha}$ .  $\delta_{1\alpha}$  y  $\delta_{1\alpha}$  de un tipo  $\alpha$  se enlazarán solo entre sí. En este método, se provoca que el par de conectores  $\delta_2$  y  $\varepsilon_2$  se enlacen solo después de formarse  $\delta_{1\alpha}$  y  $\varepsilon_{1\alpha}$ . En un segundo método, se usan diferentes grupos de protección selectivamente desprotegibles, en los que cada grupo de protección está asociado a un tipo de conector, para bloquear selectivamente los grupos conectores. En un primer ciclo, no hay  
30 protección en un tipo de base y su tipo de anclaje libre asociado es  $\varepsilon_1$  unido a  $\delta_1$  para formar un enlace  $\chi^1$ . En un segundo ciclo, un tipo de grupo de protección se elimina de un tipo de base y su tipo de anclaje libre asociado es  $\varepsilon_1$  unido a  $\delta_1$  para formar un enlace  $\chi^1$ . Este último ciclo se repite para los dos siguientes tipos de base y después de completarse, se provoca que el par de conector  $\delta_2$  y  $\varepsilon_2$  se enlacen. Entre cada etapa se incluyen etapas de lavado para reducir los errores de unión. Sin pérdida de generalidad, la restante descripción será en términos de enlaces  $\chi^1$   
35 y enlaces  $\chi^2$ . Obsérvese que después de unirse el anclaje y formarse los enlaces  $\chi$ , puede romperse el enlace intra-anclaje.

Como se muestra en la Figura 64B, los enlaces  $\chi^1$  proporcionan un enlace entre subunidades (además de los enlaces inter-sustratos polimerizados) y forman un producto intermedio llamado la “hebra hija de dúplex”. El esqueleto primario ( $\sim N \sim$ ) $\kappa$ , hebra molde ( $-N'$ ) $\kappa$  y anclaje (T) se muestran como una hebra hija duplexada, en la que  $\kappa$   
40 indica una pluralidad de subunidades de repetición. Cada subunidad de la hebra hija es un “motivo” de repetición y los motivos tienen variabilidad específica de especie, indicada aquí por el superíndice  $\alpha$ . La hebra hija se forma a partir de especies de la construcción de sustrato monomérico seleccionadas por un proceso dirigido por molde de una biblioteca de especies de motivos, uniéndose el sustrato de monómero de cada especie de construcción de sustrato a un nucleótido complementario correspondiente sobre la hebra molde diana. De esta forma, la secuencia de residuos de nucleobase (es decir, esqueleto primario) de la hebra hija es una copia complementaria contigua de la hebra molde diana.  
45

Cada tilde ( $\sim$ ) indica un enlace selectivamente escindible mostrado aquí como los enlaces inter-sustratos. Éstos son necesariamente selectivamente escindibles para liberar y expandir los anclajes (y el Xpandómero) sin degradar el propio Xpandómero.

50 La hebra hija está compuesta por un precursor de Xpandómero llamado el “Xpandómero limitado” que está adicionalmente compuesto de anclajes en la “configuración limitada”. Cuando los anclajes se convierten en su “configuración expandida”, el Xpandómero limitado se convierte en el producto de Xpandómero. Los anclajes están limitados por los enlaces  $\chi$  a sustratos adyacentes y, opcionalmente, los enlaces intra-anclaje (si todavía están presentes).

55 Puede observarse que la hebra hija tiene dos esqueletos, un “esqueleto primario” y el esqueleto del “Xpandómero limitado”. El esqueleto primario está compuesto por los sustratos monoméricos contiguamente confinados y polimerizados. El “esqueleto de Xpandómero limitado” evita el enlace selectivamente escindible entre sustratos de monómero y se forma por restos de restos de esqueleto enlazados con enlace  $\chi^1$ , siendo cada resto de esqueleto un

anclaje enlazado a un sustrato (por un enlace  $\chi^2$ ) que entonces se enlaza con el siguiente anclaje de resto de esqueleto con un enlace  $\chi^1$ . Puede observarse que el esqueleto de Xpandómero limitado conecta mediante los enlaces selectivamente escindibles del esqueleto primario, y seguirá covalentemente intacto cuando estos enlaces selectivamente escindibles se escindan y el esqueleto primario se fragmente.

5 El enlace  $\chi$  de anclaje (reticulación de los grupos conectores  $\delta$  y  $\epsilon$ ) va generalmente precedido de acoplamiento enzimático de los sustratos de monómero para formar el esqueleto primario, con, por ejemplo, enlaces fosfodiéster entre bases adyacentes. En la estructura mostrada aquí, el esqueleto primario de la hebra hija se ha formado, y los enlaces inter-sustratos se representan por una tilde (~) para indicar que son selectivamente escindibles. Después de disociar o degradar la hebra molde diana, escindiendo los enlaces selectivamente escindibles (que incluyen los enlaces intra-anclaje), el Xpandómero limitado se libera y se convierte en el producto de Xpandómero. Métodos para la disociación de la hebra molde incluyen desnaturalización por calor, o digestión selectiva con una nucleasa, o degradación química. Un método de escisión de selección usa digestión con nucleasa en la que, por ejemplo, enlaces fosfodiéster del esqueleto primario se digieren por una nucleasa y enlaces anclaje a anclaje son resistentes a nucleasa.

15 La Figura 64C es una representación del producto de Xpandómero de clase IX después de la disociación de la hebra molde y después de la escisión de los enlaces selectivamente escindibles (incluyendo aquellos en el esqueleto primario y, si no se han escindido ya, los enlaces intra-anclaje). La hebra del producto de Xpandómero contiene una pluralidad de subunidades  $\kappa$ , en la que indica la subunidad  $\kappa^n$  en una cadena de  $m$  subunidades que constituyen la hebra hija, en la que  $\kappa = 1, 2, 3$  a  $m$ , en la que  $m > 10$ , generalmente  $m > 50$ , y normalmente  $m > 500$  o  $> 5.000$ . Cada subunidad está formada de un anclaje (640) en su configuración expandida enlazada a un sustrato de monómero y con el enlace  $\chi^1$  con la siguiente subunidad adyacente. Cada subunidad, un motivo de subunidad  $\alpha$ , contiene información genética específica de especie establecida por el ensamblaje dirigido por molde del producto intermedio de Xpandómero (hebra hija).

25 La Figura 64D muestra la construcción de sustrato de la Figura 64A como modelo molecular, en la que el miembro de sustrato de monómero, representado con un residuo de nucleobase (641), tiene dos enlaces (642,643), mostrados en la Figura 64a como  $\delta_1$  y  $\delta_2$ , que serán puntos de unión para el anclaje libre. El anclaje libre se muestra con dos grupos conectores (644,645), mostrados como  $\epsilon_1$  y  $\epsilon_2$  en la Figura 64A, sobre el primer y segundo restos terminales del anclaje. Se muestra un enlace intra-anclaje selectivamente escindible (647) que limita el anclaje enlazando el primer y segundo restos terminales. Los grupos conectores están posicionados para fomentar la reticulación de los extremos de anclaje entre los sustratos de monómero y prevenir la reticulación a través de un sustrato de monómero.

30 El bucle de anclaje mostrado aquí tiene tres indicadores (900,901,902), que también pueden ser específicos de especies de motivos, pero requiere un método para enlazarse correctamente con las bases correctas en el esqueleto primario.

35 La Figura 64E muestra la construcción de sustrato después de la incorporación en el Xpandómero producto. Las subunidades se escinden y se expanden y se enlazan por enlaces  $\chi^1$  inter-anclajes (930,931), y enlaces  $\chi^2$  inter-anclajes (932,933). Cada subunidad es un anclaje enlazado a un sustrato de monómero y conectado sobre el siguiente enlace  $\chi^1$ . Una subunidad se indica por líneas de puntos que agrupan verticalmente la subunidad de repetición, como se representa por los corchetes en la Figura 64C adjunta.

40 En el producto de Xpandómero de la Figura 64E, el esqueleto primario se ha fragmentado y no es covalentemente contiguo debido a que cualquier enlace directo entre los sustratos de subunidades adyacentes se ha escindido. Mediante el proceso de escisión, el Xpandómero limitado se libera para convertirse en el producto de Xpandómero. Los miembros de anclaje que estuvieron anteriormente en la configuración limitada están ahora en la configuración expandida, sirviendo así para estirar linealmente la información de secuencias de la diana molde. La expansión de los anclajes reduce la densidad lineal de la información de secuencias a lo largo del Xpandómero y proporciona una plataforma para aumentar el tamaño y la abundancia de indicadores, que a su vez mejora la señal con respecto al ruido para la detección y descodificación de la secuencia molde.

45 Aunque el anclaje se representa como una construcción de indicador con tres grupos indicadores, diversas construcciones de indicador pueden estar presentes sobre el anclaje, y pueden comprender indicadores individuales que identifican monómero o el anclaje puede ser un polímero desnudo. En algunos casos, uno o más precursores de indicador están presentes sobre el anclaje, y los indicadores se unen por afinidad o se unen covalentemente tras el ensamblaje del producto de Xpandómero.

50 La Figura 65 demuestra un método de síntesis de Xpandómeros usando construcciones de sustrato RT-NTP de clase IX. Primero se selecciona un molde diana (650) y se hibrida con un cebador inmovilizado. En la etapa I, el cebador se extiende por síntesis dirigida por molde de una hebra hija. Este proceso continúa en la etapa II, y en la Figura 65a, se muestra una vista ampliada (flecha de puntos) de la hebra hija. Se ilustran la hebra molde, cebador y construcciones de sustrato de nucleobase de clase IX polimerizadas (sin anclajes), cada construcción de sustrato con funcionalidades químicas, representados como los símbolos de cerradura y llave, para la adición química de reactivos de anclaje. Las funcionalidades se seleccionan de manera que cada monómero tenga un sitio de unión del

anclaje específico de base y un sitio de unión universal. En la etapa III se introducen anclajes en horquilla con cuatro conectores específicos de especie. Los anclajes se enlazan al esqueleto primario en un modo de base específica según los conectores específicos de base como se muestra en la Figura 65b, una vista ampliada (flecha de puntos). Aquí, círculos blancos y negros indican la química de unión química universal y las formas de diamante y tenedor indican las químicas de unión específica de base. En la etapa IV y la Figura 65c, una vista ampliada (flecha de puntos), después de completarse todo el enlace específico de base, provoca que los conectores universales formen enlaces. Obsérvese que los conectores universales sobre ambos anclajes y el esqueleto primario están en estrecha proximidad a su homólogo de conector específico de base para evitar errores de enlace. Se muestra que la unión química de los reactivos de anclaje se ha completado, y se ha formado un Xpandómero limitado sobre el molde. El Xpandómero se libera entonces (no mostrado) por disociación del molde, y escindiendo los enlaces selectivamente escindibles (de esqueleto primario y enlaces intra-anclaje).

#### Construcciones monoméricas de clase X

Las construcciones de sustrato de clase X, también llamadas XNTP, se diferencian de RT-NTP en que el anclaje está contenido dentro de cada sustrato de monómero para formar un anclaje intra-sustratos, teniendo cada sustrato XNTP un enlace selectivamente escindible dentro del sustrato que, una vez escindido, permite la expansión del anclaje limitado. En la Figura 66, los presentes inventores describen construcciones de sustrato monomérico de clase X en más detalle. Las Figuras 66A a 66C se leen de izquierda a derecha, mostrando primero la construcción de sustrato monomérico (precursor de Xpandómero que tiene un único residuo de nucleobase), luego la hebra hija del dúplex intermedia en el centro, y a la derecha el producto de Xpandómero preparado para la secuenciación.

Como se muestra en la Figura 66A, la construcción de sustrato monomérico de clase X tiene un residuo de nucleobase del sustrato, N, que tiene dos restos (662,663) separados por un enlace selectivamente escindible (665), uniéndose cada resto a un extremo de un anclaje (660). Los extremos del anclaje pueden unirse a modificaciones del grupo conector sobre el heterociclo, el grupo ribosa, o el esqueleto de fosfato. El sustrato de monómero también tiene un sitio de escisión intra-sustrato posicionado dentro del esqueleto de fosforribosilo de forma que la escisión proporcione la expansión del anclaje limitado. Por ejemplo, para sintetizar un monómero de ATP de clase X, el conector de amino sobre 8-[(6-amino)hexil]-amino-ATP o N6-(6-amino)hexil-ATP puede usarse como un primer punto de unión del anclaje, y, un conector de esqueleto mixto, tal como la modificación no conectora (N-1-aminoalquil) fosforamidato o (2-aminoetil) fosfonato, puede usarse como un segundo punto de unión del anclaje. Además, una modificación de esqueleto conectora tal como un fosforamidato (3' O-P-N 5') o un fosfortiolato (3' OP-S 5'), por ejemplo, puede usarse para la escisión química selectiva del esqueleto primario.

$R^1$  y  $R^2$  son grupos terminales configurados según convenga para el protocolo de síntesis en el que se usa la construcción de sustrato. Por ejemplo,  $R^1 = 5'$ -trifosfato y  $R^2 = 3'$ -OH para un protocolo de polimerasa I. El  $R^1$  5' trifosfato puede incluir modificaciones de esqueleto mixtas, tales como un fosfonato de aminoetil o 3'-O-P-S-5' fosfortiolato, para permitir el enlace del anclaje y la escisión del esqueleto, respectivamente. Opcionalmente,  $R^2$  puede configurarse con un grupo de bloqueo reversible para la adición cíclica de un único sustrato. Alternativamente,  $R^1$  y  $R^2$  pueden configurarse con los grupos conectores terminales para el acoplamiento químico.  $R^1$  y  $R^2$  pueden ser del tipo general XR, en la que X es un grupo de enlace y R es un grupo funcional.

Durante el ensamblaje, la construcción de sustrato monomérico se polimeriza sobre el extremo extensible de la hebra hija naciente por un proceso de polimerización dirigida por molde usando un molde monocatenario como guía. Generalmente, este proceso se inicia a partir de un cebador y avanza en la dirección de 5' a 3'. Generalmente, se usa una ADN polimerasa u otra polimerasa para formar la hebra hija, y las condiciones se seleccionan de manera que se obtenga una copia complementaria de la hebra molde.

Como se muestra en la Figura 66B, los residuos de nucleobase se polimerizan una subunidad con la siguiente y forman un producto intermedio llamado la "hebra hija de dúplex". El esqueleto primario (-N)-k, hebra molde (-N')-k y anclaje (T) se muestran como una hebra hija duplexada, en la que k indica una pluralidad de subunidades de repetición. Cada subunidad de la hebra hija es un "motivo" de repetición y los motivos tienen variabilidad específica de especie, indicada aquí por el superíndice  $\alpha$ . La hebra hija se forma a partir de especies de la construcción de sustrato monomérico seleccionadas por un proceso dirigido por molde de una biblioteca de especies de motivos, uniéndose el sustrato de monómero de cada especie de construcción de sustrato a un nucleótido complementario correspondiente sobre la hebra molde diana. De esta forma, la secuencia de residuos de nucleobase (es decir, esqueleto primario) de la hebra hija es una copia complementaria contigua de la hebra molde diana.

La ("V") mostrada en la Figura 66B encima del residuo de nucleobase indica un enlace selectivamente escindible que divide el sustrato en el primer y segundo restos. Tras la escisión, el primer resto (669) de una subunidad seguirá enlazado con el segundo resto (668) de la subunidad adyacente y dentro de una subunidad, cada resto se conectará, uno con el otro por el anclaje. Estos son necesariamente selectivamente escindibles para liberar y expandir los anclajes (y el Xpandómero) sin degradar el propio Xpandómero.

La hebra hija tiene dos esqueletos, un "esqueleto primario" y el esqueleto del "Xpandómero limitado". El esqueleto primario está compuesto por los sustratos monoméricos contiguamente confinados y polimerizados. El "esqueleto de Xpandómero limitado" evita el enlace selectivamente escindible dentro del sustrato de monómero y se forma por los

enlaces inter-sustratos que enlazan los restos de esqueleto, siendo cada resto de esqueleto un anclaje enlazado a dos restos de los residuo de nucleobase todavía intactos. El esqueleto de Xpandómero limitado conecta mediante el enlace selectivamente escindible dentro de cada monómero, y seguirá covalentemente intacto cuando estos enlaces selectivamente escindibles se escindan y los monómeros se fragmenten en porciones  $n^1$  y  $n^2$  mostradas en la Figura 66C.

La escisión va precedida del acoplamiento enzimático de los sustratos de monómero para formar el esqueleto primario, con, por ejemplo, enlaces de esqueleto fosfodiéster o mixtos entre bases adyacentes. En la estructura mostrada aquí, el esqueleto primario de la hebra hija se ha formado. Después de disociar o degradar la hebra molde diana y escindir los enlaces selectivamente escindibles, el Xpandómero limitado se libera y se convierte en el producto de Xpandómero. Métodos para la disociación de la hebra molde incluyen, por ejemplo, desnaturalización por calor.

La Figura 66C es una representación del producto de Xpandómero de clase X después de la disociación de la hebra molde y después de la escisión de los enlaces selectivamente escindibles. La hebra del producto de Xpandómero contiene una pluralidad de subunidades  $\kappa$ , en la que  $\kappa$  indica la subunidad  $k^1$  en una cadena de  $m$  subunidades que constituyen la hebra hija, en la que  $\kappa = 1, 2, 3$  a  $m$ , en la que  $m > 10$ , generalmente  $m > 50$ , y normalmente  $m > 500$  o  $> 5.000$ . Cada subunidad está formada de un anclaje en su configuración expandida enlazada a porciones  $n^1$  y  $n^2$  de un sustrato de monómero, y cada subunidad está enlazada con la siguiente por los enlaces de polimerización de monómeros. Cada subunidad, un motivo de subunidad  $\alpha$ , contiene información genética específica de especie establecida por el ensamblaje dirigido por molde del producto intermedio de Xpandómero (hebra hija).

La Figura 66D muestra la construcción de sustrato como modelo molecular, en la que el miembro de nucleobase (664) está unido a un primer y segundo resto de la nucleobase, cada resto con un sitio de unión al anclaje (662,663). El anclaje (660) comprende grupos indicadores (900,901,902). Un enlace selectivamente escindible que separa los dos restos de la nucleobase se indica por una "V" (665).

La Figura 66E muestra el producto de Xpandómero. Las subunidades comprenden el anclaje expandido (660) unido a las porciones de nucleobase (669,668), mostrada como  $n^1$  y  $n^2$  en la Figura 66C, cada subunidad unida por enlaces inter-nucleobases. Mediante el proceso de escisión, el Xpandómero limitado se libera para convertirse en el producto de Xpandómero. Los miembros de anclaje que estuvieron anteriormente en la configuración limitada están ahora en la configuración expandida, sirviendo así para estirar linealmente la información de secuencias de la diana molde. La expansión de los anclajes reduce la densidad lineal de la información de secuencias a lo largo del Xpandómero y proporciona una plataforma para aumentar el tamaño y la abundancia de indicadores, que a su vez mejora la señal con respecto al ruido para la detección y descodificación de la secuencia molde.

Aunque el anclaje se representa como una construcción de indicador con tres grupos indicadores, diversas construcciones de indicador pueden estar presentes sobre el anclaje, y pueden comprender indicadores individuales que identifican monómero o el anclaje puede ser un polímero desnudo. En algunos casos, uno o más precursores de indicador están presentes sobre el anclaje, y los indicadores se unen por afinidad o se unen covalentemente tras el ensamblaje del producto de Xpandómero.

La Figura 67 muestra un método de ensamblaje de un Xpandómero con las construcciones de sustrato XNTP de clase X. En la primera vista se usa un cebador en horquilla para cebar un molde y el molde se pone en contacto con una polimerasa y sustratos de monómero de clase X. Se muestra en la etapa I que la polimerización extiende la hebra hija naciente procesivamente por adición dirigida por molde de sustratos de monómero. La vista ampliada (Figura 67a) ilustra esto en más detalle. Se muestra que la hebra hija complementaria a la hebra molde está compuesta de sustratos de nucleobase modificados con sitio de escisión interno ("V"). En la etapa II, el proceso de formación del producto intermedio de Xpandómero se completa, y en la etapa III, un proceso de escisión, disociación del producto intermedio y la expansión de la hebra hija se muestra en proceso. En la vista ampliada (Figura 67b) se muestra que la escisión interna de los sustratos de nucleobase alivia la limitación sobre los anclajes, que se expanden, alargando el esqueleto del Xpandómero.

#### EJEMPLO 1.

#### SÍNTESIS DE UNA XSONDA 2MERA QUIMÉRICA "CA" CON ENLACE INTERNUCLEÓTIDICO DE RIBOSIL-5'-3' SELECTIVAMENTE ESCINDIBLE

Las construcciones oligoméricas de sustrato están compuestas de miembros de sonda y miembros de anclaje y tienen una construcción general de "sonda-bucle". La síntesis del miembro de sonda se lleva a cabo usando métodos bien establecidos de síntesis de oligómeros en fase sólida. En estos métodos, la adición de nucleobases a una cadena de sonda naciente sobre una resina se lleva a cabo con, por ejemplo, química de fosforamidito (las patentes de EE.UU. Nº 4.415.732 y 4.458.066), y pueden sintetizarse económicamente miligramos o gramos de oligómero sintético usando sintetizadores automatizados fácilmente disponibles. La síntesis de oligonucleótidos en fase sólida típica implica realizar reiterativamente cuatro etapas: desprotección, acoplamiento, encapuchado y oxidación. Sin embargo, al menos un enlace en la sonda de una construcción de sustrato de Xsonda de clase I es un enlace selectivamente escindible, y al menos dos restos de sonda se modifican para la aceptación de un miembro



de anclaje. El enlace selectivamente escindible se localiza entre los restos de sonda seleccionados para la unión del anclaje (es decir, “entre” no debe limitarse a que significa “entre miembros de nucleobase adyacentes” debido a que el primer y segundo puntos de la unión del anclaje solo necesitan estar posicionados en cualquier parte sobre un primer y segundo resto de la sonda, respectivamente, estando los restos unidos por el enlace selectivamente escindible). En este ejemplo, un enlace internucleotídico de ribosil-5'-3', que es selectivamente escindible por ribonucleasa H, es el enlace selectivamente escindible y los dos puntos de unión del anclaje son el primer y segundo residuos de nucleobase de una sonda 2mera.

La síntesis de Xsondas modificadas con conector se logra usando fosforamiditos comercialmente disponibles, por ejemplo, de Glen Research (Sterling, VA, EE.UU.), BioGenex (San Ramon, CA, EE.UU.), Dalton Chemical Laboratories (Toronto, Canadá), Thermo Scientific (EE.UU.), Link Technologies (RU), y otros, o pueden sintetizarse a medida. Pueden usarse métodos de síntesis bien establecidos para preparar la sonda descrita en la que la nucleobase de 3', que para este ejemplo es un modificador de amino C6-desoxiadenosina, se une primero a un soporte universal usando 5'-dimetoxitritil-N6-benzoil-N8-[6-(trifluoroacetilamino)-hex-1-il]-8-amino-2'-desoxiadenosina-3'-[(2-cianoetil)-(N,N-diisopropil)]-fosforamidito, seguido de la adición de un modificador de amino C6-citidina usando 5'-dimetoxitritil-N-dimetilformamidina-5-[N(trifluoroacetilaminohexil)-3-acrilimido]-citidina,3'-[(2-cianoetil)-(N,N-diisopropil)]-fosforamidito, en la que la 5'-citidina es un ribonucleótido. La adición de un reactivo de fosforilación química seguido de métodos estándar de escisión, desprotección y de purificación completa la síntesis. El producto de dinucleótido es un 5'-fosfato (aminoC6-citosina) (aminoC6-desoxiadenosina) 3' con un enlace ribosil-5', 3' centralmente escindible y conectores de amino sobre cada base.

El anclaje de Xsonda para este ejemplo se construye a partir de éster de bis-N-succinimidil-[pentaetilenglicol] (Pierce, Rockford IL; Producto N° 21581). Las aminas conectoras del oligómero de pCA modificado se reticulán con bis(NHS)PEG5 según las instrucciones del fabricante. Se obtiene un producto del peso molecular esperado para la construcción circularizada de PEG-sonda.

## EJEMPLO 2.

### SÍNTESIS DE UNA XSONDA 4MERA “TATA” CON ENLACE FOSFOROTIOLATO SELECTIVAMENTE ESCINDIBLE

Pueden sintetizarse Xsondas 4meras con un enlace fosforotiolato como el enlace selectivamente escindible. Para el siguiente ejemplo se describe la síntesis de un tetranucleótido 5' fosfato (dT) (aminoC6-dA) (dT) (aminoC6-dA) 3'.

Primero se prepara una 5' mercapto-desoxitimidina como se describe por Mag et al. (“Synthesis and selective cleavage of an oligodeoxynucleotide containing a bridged internucleotide 5'-phosphorothioate linkage”, Nucl Acids Res 19:1437-41, 1991). Se hace reaccionar timidina con dos equivalentes de cloruro de p-toluenosulfonilo en piridina a temperatura ambiente, y el 5'-tosilato resultante se aísla por cristalización en etanol. El tosilato se convierte en una 5'-(S-tritil)-mercapto-5'-desoxi-timidina con cinco equivalentes de tritilolato de sodio (preparado *in situ*). El nucleótido 5'-(S-tritil)-mercapto-timidina se purifica y se hace reaccionar con 2-cianoetoxi-bis-(N,N-diisopropilamino-fosfano) en presencia de tetrazol para preparar el elemento estructural de 3'-O-fosforamidito.

Para empezar la síntesis automatizada, el modificador de amino C6-desoxiadenosina se une primero a un soporte universal usando 5'-dimetoxitritil-N6-benzoil-N8-[6-(trifluoroacetilamino)-hex-1-il]-8-amino-2'-desoxiadenosina-3'-[(2-cianoetil)-(N,N-diisopropil)]-fosforamidito, seguido de la adición de mercaptotimidina fosforamidito preparado anteriormente. Antes de añadir el siguiente amino C6 dA fosforamidito, el grupo S-tritilo se desprotege primero con nitrato de plata acuoso 50 mM y la resina se lava con agua. Entonces, la resina se trata normalmente con un agente reductor tal como DTT para eliminar disulfuros secundarios formados durante la escisión. La columna se lava entonces de nuevo con agua y con acetonitrilo, y el tiol libre se hace reaccionar bajo condiciones estándar con la amino C6 desoxiadenosina fosforamidito en presencia de tetrazol, formando así “ATA” con un enlace fosforotiolato conector S3'→P5' entre la desoxiadenosina terminal y la 3-mercapto-timidina. En el siguiente ciclo se añade desoxitimidina fosforamidito estándar. Finalmente, la adición de un reactivo de fosforilación química, seguido de métodos de escisión, desprotección y de purificación convencionales, completa la síntesis. El producto de tetranucleótido es un 5' fosfato (dT) (aminoC6-dA) (dT) (aminoC6-dA) 3'.

El enlace fosforotiolato es selectivamente escindible, por ejemplo, con AgCl, ácido o con yodoetanol (Mag et al., “Synthesis and selective cleavage of an oligodeoxynucleotide containing a bridged internucleotide 5'-phosphorothioate linkage”, Nucleic Acids Research, 19(7):1437-1441, 1991). Debido a que el enlace selectivamente escindible es entre la segunda y tercera nucleobases, el anclaje se diseña para conectar este enlace, y puede unirse a cualesquiera dos nucleobases (o cualesquiera dos puntos de unión del esqueleto primario) sobre cualquier lado del enlace selectivamente escindible. Los métodos para las químicas de conector cero y conector incluyen, por ejemplo, provisión de conectores con aminas primarias usadas en la síntesis del oligómero, como se describe en el Ejemplo 1. Los conectores modificados con amina sobre el oligómero “TATA” normalmente se protegen durante la síntesis de oligómeros y se desprotegen en el transcurso normal de completarse la síntesis de oligonucleótidos.

El anclaje de Xsonda para este ejemplo se construye entonces a partir de éter diglicídico de poli(etilenglicol) activado con bis-epóxido (SigmaAldrich, St. Louis MO, Producto N° 475696). La reacción de epóxido de las aminas

con los grupos terminales PEG activados se realiza en disolución diluida para minimizar cualquier reacción de concatenación de competición. Se obtiene reactividad similar con anhídridos mixtos o incluso cloruros de ácido, y pueden emplearse grupos de enlace heterobifuncionales de manera que orienten la unión del anclaje. Los productos de reacción anclados se separan por HPLC preparativa y se caracterizan por espectroscopía de masas. Se obtiene un producto del peso molecular aproximado para la construcción 4mera circularizada de PEG-sonda (aproximadamente 2,5 Kd). Con este método se obtiene una distribución de anclajes de PEG con aproximadamente  $M_n = 500$ . Esto se corresponde con un anclaje de aproximadamente 40 Angstroms (a aproximadamente 3,36 A/unidad de PEG).

### EJEMPLO 3.

#### 10 SÍNTESIS DE UNA XSONDA 3MERA "CTA" CON ENLACE 5'-3' FOSFODIÉSTER SELECTIVAMENTE ESCINDIBLE

También pueden sintetizarse Xsondas con un enlace fosfodiéster como el enlace selectivamente escindible. Para el siguiente ejemplo se describe la síntesis de un trinucleótido de 5' fosfato (aminoC6-dC) (aminoC6-dT) (dA) 3' con una modificación de fosforotioato no conectora. Los enlaces fosfodiéster son atacados por una variedad de nucleasas. Un enlace fosforotioato, con azufre no conector, se usa como enlace resistente a nucleasas en este ejemplo.

Para la síntesis automatizada en la dirección 3' a 5' se usa un soporte sólido de desoxiadenosina inmovilizado en CPG (5'-dimetoxitritil-N-benzoil-2'-desoxiadenosina, 3'-succinoil-alquilo de cadena larga-CPG 500). En el primer ciclo, el modificador de amino C6-desoxitimidina fosforamidito (5'-dimetoxitritil-5-[N-(trifluoroacetilaminohexil)-3-acrilimido]-2'-desoxiuridina, 3'-[(2-cianoetil)-(N,N-diisopropil)]-fosforamidito) se acopla. Antes del encapuchado, la dA inmovilizada se hace reaccionar con reactivo sulfurizante (Glen Research, Sterling VA; Cat N° 40-4036), también conocido como reactivo de Beaucage, siguiendo el protocolo del fabricante. El reactivo se añade generalmente mediante un puerto separado en el sintetizador. Tras la tiolación, la C6-desoxi-citidina modificada con amino (5'-dimetoxitritil-N-dimetilformamidina-5-[N-(trifluoroacetilaminohexil)-3-acrilimido]-2'-desoxicitidina, 3'-[(2-cianoetil)-(N,N-diisopropil)]-fosforamidito) se acopla. Antes del encapuchado, la dC inmovilizada se hace reaccionar con reactivo sulfurizante (Glen Research, Sterling VA; Cat N° 40-4036), también conocido como reactivo de Beaucage, siguiendo el protocolo del fabricante. La "CTA" 3mera resultante tiene un enlace fosforotioato entre la T y la A, y un enlace fosfodiéster entre la C y la T. La adición de un reactivo de fosforilación química, seguido de métodos de escisión, desprotección y de purificación estándar, completa la síntesis.

La resistencia de los enlaces fosforotioato al ataque de nucleasas está bien caracterizado, por ejemplo, por Matsukura et al ("Phosphorothioate analogs of oligodeoxynucleotides: inhibitors of replication and cytopathic effects of human immunodeficiency virus", PNAS 84:7706-10, 1987), por Agrawal et al ("Oligodeoxynucleoside phosphoramidates and phosphorothioates as inhibitors of human immunodeficiency virus", PNAS 85:7079-83, 1988), y en la patente de EE.UU. N° 5770713. Tanto C y T están modificados con conector o con conector cero, sirviendo la derivatización para la unión de un miembro de anclaje.

### EJEMPLO 4.

#### SÍNTESIS DE UNA XSONDA 3MERA "ATA" CON ENLACE 5' N-P-O 3' FOSFORAMIDATO SELECTIVAMENTE ESCINDIBLE

El modificador de amino C6-desoxiadenosina se une primero a un soporte universal usando 5'-dimetoxitritil-N6-benzoil-N8-[6-(trifluoroacetilamino)-hex-1-il]-8-amino-2'-desoxiadenosina-3'-[(2-cianoetil)-(N,N-diisopropil)]-fosforamidito, seguido en el siguiente ciclo de la adición de 5'amino-dT bloqueado con MMT (5'monometoxitritilamino-2'-desoxitimidina). Después del desbloqueo, el extremo 5'-amino se hace reaccionar con C6-desoxiadenosina fosforamidito modificado con amino bajo condiciones estándar. La adición de un reactivo de fosforilación química, seguido de métodos de escisión, desprotección y de purificación estándar completa la síntesis. El trinucleótido 5' (aminoC6-dA) (O-P-N) (dT) (O-P-O) (aminoC6-dA) 3' tiene un enlace fosforamidato entre la 5' aminoC6-dA y la penúltima dT.

Este enlace fosforamidato es selectivamente escindible en condiciones en las que enlaces fosfodiéster siguen intactos tratando el oligómero con 80 % de ácido acético como se describe por Mag et al. ("Synthesis and selective cleavage of oligodeoxynucleotides containing non-chiral internucleotide phosphoramidate linkages", Nucl. Acids Res. 17: 5973-5988, 1989). Uniendo un anclaje para conectar el enlace fosforamidato 5'N-P-O, un Xpandómero que contiene dímeros de este tipo puede expandirse por escisión selectiva de los enlaces fosforamidato del esqueleto primario.

### EJEMPLO 5.

#### SÍNTESIS DE UNA XSONDA 6MERA "CACCAC" CON UN ENLACE FOTOESCINDIBLE INTERNO

Después de la síntesis estándar de 3' a 5' con desoxicitosina y desoxiadenosina fosforamiditos sin modificar, se acopla C6 dC modificada con amina (5'-dimetoxitritil-N-dimetilformamidina-5-[N-(trifluoroacetilaminohexil)-3-

acrilimido]-2'-desoxicidina,3'-[(2-cianoetil)-(N,N-diisopropil)]-fosforamidito; Glen Research; Cat N° 10-1019), formando así un trímero CAC. Para el siguiente ciclo, se acopla un conector fotoescindible (3-(4,4'-dimetoxitritil)-1-(2-nitrofenil)-propan-1-il-[(2-cianoetil)-(N,N-diisopropil)]-fosforamidito; Glen Research; Cat N° 10-4920). En el siguiente ciclo, se añade una segunda dC modificada con amino, seguido de dos rondas finales de adición estándar de dA y dC fosforamiditos, respectivamente. El producto resultante, "CAC-pc-CAC" contiene conectores de amino en la tercera y cuarta posiciones de base, y puede modificarse mediante la adición de un anclaje que conecta el enlace selectivamente escindible formado por la construcción de nitrobeneno fotoescindible entre las dos bases modificadas con amino. La escisión selectiva de un esqueleto de fosfodiéster modificado con conector fotoescindible se desvela, por ejemplo, por Sauer et al. ("MALDI mass spectrometry analysis of single nucleotide polymorphisms by photocleavage and charge-tagging", *Nucleic Acids Research* 31,11 e63, 2003), Vallone et al. ("Genotyping SNPs using a UV-photocleavable oligonucleotide in MALDI-TOF MS", *Methods Mol. Bio.* 297:169-78, 2005) y Ordoukhanian et al. ("Design and synthesis of a versatile photocleavable DNA building block, application to phototriggered hybridization", *J. Am. Chem. Soc.* 117, 9570-9571, 1995).

#### EJEMPLO 6.

#### 15 SÍNTESIS DE UNA CONSTRUCCIÓN DE SUSTRATO XMERA

Las construcciones de sustrato Xmeras están estrechamente relacionadas en diseño y composición con Xsondas. Se sintetiza una biblioteca Xmera, por ejemplo, por 5'-pirofosforilación de Xsondas. Procedimientos establecidos para el tratamiento con pirofosfato de 5'-monofosfatos incluyen, por ejemplo, Abramova et al. ("A facile and effective synthesis of dinucleotide 5'-triphosphates", *Bioorganic Medicinal Chemistry* 15: 6549-55, 2007). En este método, el monofosfato terminal del oligómero se activa según sea necesario para la reacción posterior con pirofosfato haciendo reaccionar primero el fosfato terminal como una sal de cetiltrimetilamonio con cantidades equimolares de trifetilfosfina (Ph<sub>3</sub>P) y disulfuro de 2,2'-dipiridilo (PyS)<sub>2</sub> en DMF/DMSO, usando DMAP (4-dimetilaminopiridina) o 1-Melm (1-metilimidazol) como catalizador nucleófilo. El producto se precipita con LiClO<sub>4</sub> en acetona y se purifica por cromatografía de intercambio aniónico.

25 Puede considerarse una variedad de otros métodos para la síntesis robusta de Xmeros de 5'-trifosfato. Como se describe por Burgess y Cook (*Chem Rev* 100(6):2047-2060), estos métodos incluyen, pero no se limitan a, reacciones usando nucleósido fosforamiditos, síntesis mediante ataque nucleófilo de pirofosfato sobre nucleósido monofosfatos activados, síntesis mediante ataque nucleófilo de fosfato sobre nucleósido pirofosfato activado, síntesis mediante ataque nucleófilo de difosfato sobre fosfato sintona activada, síntesis que implica fosfitos o fosforamiditos activados derivados de nucleósidos, síntesis que implica el desplazamiento directo de grupos salientes 5'-O- por nucleófilos de trifosfato, y métodos biocatalíticos. Un método específico que produjo sustratos de dinucleótido compatibles con polimerasa usa N-metilimidazol para activar el grupo 5'-monofosfato; la posterior reacción con pirofosfato (sal de tributilamonio) produce el trifosfato (Bogachev, 1996). En otro procedimiento, los trinucleótido fosforamidatos se han sintetizado por Kayushin (Kayushin AL et al. 1996. A convenient approach to the synthesis of trinucleotide phosphoramidites. *Nucl Acids Res* 24:3748-55).

#### EJEMPLO 7.

#### SÍNTESIS DE XPANDÓMEROS CON ESCISIÓN DE FOSFORAMIDATO POR POLIMERASA

La síntesis de un Xpandómero se realiza usando una construcción de sustrato preparada con extremos 5'-trifosfato y 3'-OH. La patente de EE.UU. N° 7.060.440 a Kless describe el uso de polimerasas para polimerizar oligómeros de trifosfato, y el método está adaptado aquí para la síntesis de Xpandómeros. La construcción de sustrato consiste en miembro de sonda 2mera "pppCA" con enlace fosforamidato inter-nucleótido 5' N-P-O selectivamente escindible y una construcción de bucle de enlace de PEG. Se sintetiza una hebra molde y cebador acompañante y se purifica antes de uso. La secuencia: "TGTGTGTGTGTGTGTGTGTGATCTACCGTCCGTCCC" se usa como molde. La secuencia "5'GGGACGGACGGTAGAT" se usa como cebador. Se usa una HEX del extremo 5' (5'hexafluoro fluoresceína) sobre el cebador como marca. La hibridación del cebador y el molde forma un molde con cebador de dúplex y nucleótido protuberante monocatenario de los extremos 3'-OH y 5' libres de veinte bases en longitud. Entonces se añaden las construcciones de sustrato y la ADN polimerasa T7 recombinante de la marca Sequenase™ (US Biochemicals Corp., Cleveland, OH) y la polimerización continúa durante 30 min en condiciones ajustadas para la polimerización óptima. Se mezcla una muestra de la reacción de polimerización con tampón de carga de gel y se somete a electroforesis sobre un gel de 20 % de TBE-acrilamida (Invitrogen, EE.UU.) junto con un control negativo no de polimerasa y un marcador de MW para confirmar la polimerización de Xmeros.

El producto intermedio de Xpandómero se trata con 80 % de ácido acético durante 5 h a temperatura ambiente según el procedimiento de Mag et al. ("Synthesis and selective cleavage of oligodeoxyribonucleotides containing non-chiral internucleotide phosphoramidate linkages", *Nucl. Acids Res.*, 17: 5973-88, 1989) para escindir selectivamente los enlaces fosforamidato, que asimismo se confirma por electroforesis.

EJEMPLO 8.

## SÍNTESIS DE XPANDÓMEROS CON ESCISIÓN DE FOSFOROTIOATO POR POLIMERASA

Se realiza la síntesis de un Xpandómero usando una construcción de sustrato preparada con extremos 5'-trifosfato y 3'-OH. La patente de EE.UU. N° 7.060.440 a Kless describe el uso de polimerasas para polimerizar oligómeros de trifosfato. La construcción de sustrato es una "pppCA". La construcción de sustrato se diseña con un enlace de esqueleto de fosforotiolato inter-nucleótido selectivamente escindible y una construcción de bucle de anclaje de PEG 2000. Se sintetiza una hebra molde y cebador acompañante y se purifica antes de uso. La secuencia: "TGTGTGTGTGTGTGTGTGTGTGATCTACCGTCCGTC" se usa como molde diana. La secuencia "5'GGGACGGACGGTAGAT" se usa como cebador. Se usa una HEX del extremo 5' (5'hexacloro-fluoresceína) sobre el cebador como marca. La hibridación del cebador y el molde forma un molde con cebador de dúplex y nucleótido protuberante monocatenario de los extremos 3'-OH y 5' libres de veinte bases en longitud. Entonces se añaden las construcciones de sustrato y Thermanator™ ADN polimerasa (New England Biolabs, EE.UU.) con tampón y sales optimizadas para la polimerización. La polimerización continuó durante 60 min en condiciones ajustadas para la polimerización óptima. Se mezcla una muestra de la reacción de polimerización con tampón de carga de gel y se somete a electroforesis sobre un gel de 20 % de TBE-acrilamida (Invitrogen, EE.UU.) junto con un control negativo no de polimerasa y un marcador de MW para confirmar la polimerización de Xmeros.

Los enlaces fosforotiolato del producto intermedio de Xpandómero son selectivamente escindibles, por ejemplo, con AgCl, ácido o con yodoetanol (Mag et al., "Synthesis and selective cleavage of an oligodeoxynucleotide containing a bridged internucleotide 5'-phosphorothioate linkage", *Nucleic Acids Research*, 19(7):1437-1441, 1991). La escisión se confirma por electroforesis en gel.

EJEMPLO 9.

## SÍNTESIS DE XPANDÓMEROS QUIMÉRICOS CON LIGASA

Se realiza la síntesis de un Xpandómero usando una construcción de sustrato preparada con extremos 5'-monofosfato y 3'-OH. La construcción de sustrato es un "5' p dC rA~dC dA 3'" 4mero quimérico en el que la penúltima adenosina de 5' es un ribonucleótido y el resto del sustrato es desoxirribonucleótido. La construcción de sustrato se diseña con un enlace ribosil 5'-3' fosfodiéster inter-nucleótido selectivamente escindible (como se muestra por "~") y una construcción de bucle de anclaje de PEG 2000, en la que el anclaje está unido a la "C" y "A" terminal del 4mero. Se sintetiza una hebra molde y cebador acompañante y se purifica antes de uso. La secuencia: "TGTGTGTGTGTGTGTGTGTGATCTACCGTCCGTC" se usa como molde diana. La secuencia "5'GGGACGGACGGTAGAT" se usa como cebador. Se usa una HEX del extremo 5' (5'hexacloro-fluoresceína) sobre el cebador como marca. La hibridación del cebador y el molde forma un molde con cebador de dúplex y nucleótido protuberante monocatenario de los extremos 3'-OH y 5' libres de veinte bases en longitud. Entonces se añaden las construcciones de sustrato y ADN ligasa T4 (Promega Corp, Madison, WI, EE.UU.; Cat N° M1801) con temperatura, tampón y sales optimizadas para la hibridación y ligación de sondas transitoria. La ligación continúa durante 6 horas. Una muestra de la reacción de ligación se mezcla con tampón de carga de gel y se somete a electroforesis sobre un gel de 20 % de TBE-acrilamida (Invitrogen, Carlsbad, CA, EE.UU.) junto con un control negativo no de ligasa y un marcador de MW para confirmar la ligación de Xsondas.

En una segunda etapa, el producto intermedio de Xpandómero se trata con ribonucleasa H para escindir el enlace fosfodiéster 5'-3' lábil a RNasa para producir un producto de Xpandómero que se confirma asimismo por electroforesis en gel.

EJEMPLO 10.

## PREPARACIÓN DE UNA CONSTRUCCIÓN DE CONECTOR DE ALFA-FOSFATO

También pueden unirse conectores de anclaje al diéster de fosforotioato S o a una N-amida de fosforamido, como se trata por Agrawal ("Site specific functionalization of oligonucleotides for attaching two different reporter groups", *Nuc. Acids Res.* 18:5419-23, 1990). Se desvela un método de funcionalización de dos enlaces de esqueleto inter-nucleótido diferentes: conector de fosforamido de aminohexilo (N-1 amino alquilo) y fosforotioato. La amina C6 preparada como se describe por Agrawal se usa como conector para la síntesis de una construcción de anclaje inter-nucleobase de la presente invención. También se ha informado de la derivatización de enlaces N3'-P5' (Sinyakov et al., "Functionalization of the oligonucleotides containing an internucleotide phosphoramidate bond", *Russian J Bioorganic Chem*, 29:100-102, 2003).

EJEMPLO 11.

## VÍAS HETEROBIFUNCIONALES PARA LAS CONSTRUCCIONES DE ANCLAJE

La síntesis de oligómeros modificados que contienen un enlace selectivamente escindible se describen en los Ejemplos 1-5. Aquí, un 4mero con una base modificada con C6-amino en la segunda posición y base modificada con conector de 4-formilbenzoato en la tercera posición del 4mero se prepara por química sintética de oligómeros

estándar. La segunda y tercera bases se separan por un enlace selectivamente escindible seleccionado de enlace ribosil 5'-3' fosfodiéster, enlace desoxirribosil 5'-3' fosfodiéster, enlace fosfortiolato (5' O-P-S 3' o 5' S-P-O 3'), enlace fosforamidato (5' O-P-N 3' o 5' N-P-O 3'), o un enlace fotoescindible. La amina del conector C6 se hace reaccionar con sulfo-EGS para formar un grupo NHS de éster activo. La síntesis de bucles de anclaje heterobifuncionales se lleva entonces a cabo haciendo reaccionar el miembro de sonda modificado con un anclaje funcionalizado con grupos terminales del extremo amina y hidrazida. Se obtiene un producto final circularizado en el anclaje.

#### EJEMPLO 12.

##### CONSTRUCCIÓN DE ANCLAJE MARCADO CON COLORANTE

Se prepara una construcción de sustrato de Xsonda con un único indicador. Se sintetiza primero un 2mero con C6 aminas sobre la primera y segunda bases usando métodos convencionales. La primera y segunda bases se separan por un enlace selectivamente escindible seleccionado de enlace ribosil 5'-3' fosfodiéster, enlace desoxirribosil 5'-3' fosfodiéster, enlace fosfortiolato (5' O-P-S 3' o 5' S-P-O 3'), enlace fosforamidato (5' O-P-N 3' o 5' N-P-O 3'), o un enlace fotoescindible. En este ejemplo, el anclaje es una molécula de PEG 2000 funcionalizada en los extremos específica de especie con un único grupo conector interno, que en este caso es un anclaje localizado centrado de conector funcionalizado con maleimido. El indicador es un dendrímero marcado con colorante unido con el grupo conector de maleimido sobre el anclaje mediante un resto sulfhidrilo sobre el dendrímero. La construcción de indicador resultante contiene así un indicador sobre el anclaje. Un dendrímero de cistamina (Dendritic Nanotechnologies, Mt Pleasant, MI, USA; Cat N° DNT-294 G3) con un diámetro de 5,4 nm y 16 aminas superficiales por medio dendrímero) se usa como indicador. Uniendo colorantes específicos, o combinaciones de colorantes, a los grupos amina sobre el dendrímero, la especie de construcción de sustrato está únicamente marcada para la identificación.

#### EJEMPLO 13.

##### CONSTRUCCIÓN DE ANCLAJE MARCADO CON PÉPTIDO

Se prepara una construcción de sustrato de Xsonda con un único indicador. Se sintetiza primero un 2mero con C6 aminas sobre la primera y segunda bases usando métodos convencionales. La primera y segunda bases se separan por un enlace selectivamente escindible seleccionado de enlace ribosil 5'-3' fosfodiéster, enlace desoxirribosil 5'-3' fosfodiéster, enlace fosfortiolato (5' O-P-S 3' o 5' S-P-O 3'), enlace fosforamidato (5' O-P-N 3' o 5' N-P-O 3'), o un enlace fotoescindible. En este ejemplo, el anclaje es una molécula de PEG 2000 funcionalizada en los extremos específica de especie con un único grupo conector interno, que en este caso es un anclaje localizado centrado de conector funcionalizado con maleimido. El indicador es un dendrímero marcado con colorante unido con el grupo conector de maleimido sobre el anclaje mediante un resto sulfhidrilo sobre el dendrímero. La construcción de indicador resultante contiene así un indicador sobre el anclaje. Un dendrímero de cistamina (Dendritic Nanotechnologies, Mt Pleasant, MI, USA; Cat N° DNT-294 G3) con un diámetro de 5,4 nm y 16 aminas superficiales por medio dendrímero) se usa como indicador. Uniendo péptidos específicos con el dendrímero, que está funcionalizado con amina, la especie de construcción de sustrato se marca para la posterior identificación. Las dimensiones y la carga de los péptidos unidos se usan como características de detección en un aparato de detección.

#### EJEMPLO 14.

##### VÍA HETEROBIFUNCIONAL PARA CONSTRUCCIONES DE ANCLAJE DE INDICADOR

Se preparan construcciones de sustrato con múltiples indicadores. Una biblioteca de 2meros con base modificada con C6-amino en la primera posición y base modificada con 4'-formilbenzoato (4FB) en la segunda posición del 2mero se prepara por química orgánica estándar. La primera y segunda bases se separan por un enlace selectivamente escindible seleccionado de enlace ribosil 5'-3' fosfodiéster, enlace desoxirribosil 5'-3' fosfodiéster, enlace fosfortiolato (5' O-P-S 3' o 5' S-P-O 3'), enlace fosforamidato (5' O-P-N 3' o 5' N-P-O 3'), o un enlace fotoescindible. La amina se hace reaccionar con sulfo-EGS para formar un grupo NHS de éster activo. En una segunda etapa, un segmento de indicador de amina bifuncional específico de especie (segmento 1) se hace reaccionar con el grupo NHS activo y un segmento de indicador de hidrazida bifuncional específico de especie (segmento 4) se hace reaccionar con 4FB. A continuación, la amina libre sobre el segmento 1 se hace reaccionar con sulfo-EGS para formar un NHS de éster activo. A continuación se hace reaccionar un encapuchado heterobifuncional específico de especie que consiste en un par de segmentos de indicador (segmentos 2 y 3) con grupos terminales amina y 4FB con la construcción, cerrando el bucle de anclaje.

La construcción de indicador resultante contiene así cuatro indicadores sobre la construcción de anclaje. En este ejemplo, los indicadores son dendrones de poliamina con un conector de cistamina para unir covalentemente a cada segmento de polímero por enlaces tioéter (Dendritic Nanotechnologies, Mt Pleasant MI, EE.UU.; Cat N° DNT-294: G4, 4,5 nm de diámetro, 32 aminas superficiales por medio dendrímero) y los miembros de polímero son PEG 2000 de extremos funcionalizados, cada uno con un grupo conector interno. Cada segmento de anclaje comprende un

único indicador. Con cuatro segmentos direccionalmente acoplados, están disponibles  $2^4$  combinaciones posibles de código de indicador.

#### EJEMPLO 15.

##### VÍA HETEROBIFUNCIONAL PARA CONSTRUCCIONES DE INDICADOR CON MARCADO POST-SÍNTESIS

5 Se preparan construcciones de sustrato con múltiples indicadores. Se prepara un 4mero con base modificada con C6-amino en la primera posición y base modificada con 4'-formilbenzoato (4FB) en la segunda posición del 2mero por química orgánica estándar. La segunda y tercer bases se separan por un enlace selectivamente escindible seleccionado de enlace ribosil 5'-3' fosfodiéster, enlace desoxirribosil 5'-3' fosfodiéster, enlace fosforotiolato (5' O-P-S 3' o 5' S-P-O 3'), enlace fosforamidato (5' O-P-N 3' o 5' N-P-O 3'), o un enlace fotoescindible. La amina se hace reaccionar con sulfo-EGS para formar un grupo NHS de éster activo. En una segunda etapa, un segmento de  
10 indicador de amina bifuncional específico de especie (segmento 1) se hace reaccionar con el grupo NHS activo y se hace reaccionar un segmento de indicador de hidrazida bifuncional específico de especie (segmento 4) con 4FB. A continuación, la amina libre en el segmento 1 se hace reaccionar con sulfo-EGS para formar un NHS de éster activo. Entonces, un encapsulado heterobifuncional específico de especie que consiste en un par de segmentos de  
15 indicador (segmentos 2 y 3) con grupos terminales amina y 4FB se hace reaccionar con la construcción, cerrando el bucle de anclaje.

La construcción de indicador resultante contiene así cuatro indicadores sobre la construcción de anclaje. En este ejemplo se describe el post-marcado de la construcción de anclaje. Covalentemente unido a cada segmento de anclaje está un oligómero 16mero que se usa para la unión del indicador. Cada segmento de anclaje está  
20 compuesto en parte de una molécula de PEG funcionalizada. El indicador es dendrímero de poliamina marcado con colorante con un conector de cistamina (Dendritic Nanotechnologies, Mt Pleasant MI, EE.UU.; Cat N° DNT-294: G5, 4,5 nm de diámetro, 64 aminas superficiales por medio dendrímero) para el acoplamiento con la sonda de oligómero 16mera. Tras el ensamblaje del Xpandómero, los dendrímeros marcados con colorante se hibridan con los segmentos de anclaje oligoméricos. Este enfoque de marcado es análogo al método descrito por DeMattei et al. ("Designed Dendrimer Syntheses by Self-Assembly of Single-Site, ssDNA Functionalized Dendrons", Nano Letters,  
25 4:771-77, 2004).

#### EJEMPLO 16.

##### ELEMENTOS INDICADORES MARCADOS CON COLORANTE

30 Con referencia a los Ejemplos 14 y 15, las aminas superficiales de los elementos indicadores dendrímeros se marcan por química de éster activo. Están disponibles Alexa Fluor 488 (verde) y Alexa Fluor 680 (rojo) para la unión de una etapa como ésteres activos de sNHS de Molecular Probes (Eugene OR). La densidad y relación de los colorantes se varían para producir una marca molecular distintiva sobre cada elemento indicador.

#### EJEMPLO 17.

##### ELEMENTOS INDICADORES MULTI-ESTADO

35 Se seleccionan diversas paletas de colorantes por técnicas similares a aquellas usadas en M-FISH y SKY, como se conoce para aquellos expertos en la materia, y se conjugan con un indicador dendrímero. Así, el elemento indicador de cada anclaje constituye una "dirección espectral", por lo que una construcción dendrímica única con una multiplicidad de sitios de unión del colorante puede crear una pluralidad de códigos de indicador. Con referencia a la construcción de anclaje descrita en los Ejemplos 14 y 15, una dirección espectral de 5 estados produce 625  
40 combinaciones de código de indicador.

#### EJEMPLO 18.

##### ANCLAJE DE ESPACIADOR PEG-5000

45 Se construyen segmentos de anclaje de un polímero duradero soluble en disolvente acuoso/orgánico que posee de poca a ninguna afinidad de unión por reactantes de SBX. Se usa PEG 5000 modificado para los espaciadores de anclaje flexibles que flanquean un indicador de poli-lisina 5000. Los extremos de PEG libres están funcionalizados para la unión al miembro de sonda. El polímero se circulariza por reticulación a puntos de unión de sonda usando química de conectores heterobifuncionales.

#### EJEMPLO 19.

##### PREPARACIÓN DE UNA COMPOSICIÓN DE INDICADOR DE MARCA DE MASA

50 Se sintetiza un anclaje del siguiente modo: Marcas de masa escindibles se acoplan covalentemente a dendrímeros G5 funcionalizados con el péptido poli-l-lisina (Dendritic Nanotechnologies, Mt Pleasant MI) usando química de acoplamiento de amina estándar. Los dendrímeros G5 tienen ~5,7 nm de diámetro y proporcionan 128 grupos superficiales reactivos. Una tira de diez dendrímeros G5 funcionalizados con péptidos de polilisina de 10.000 de

5 peso molecular proporciona aproximadamente 100.000 sitios de unión de indicador conjuntamente en un segmento de dendrímero de ~ 57 nm. Están disponibles un total de aproximadamente 10.000 marcas de masa para la detección sobre el segmento completamente ensamblado, suponiendo solo el 10 % de ocupación de los sitios de unión disponibles. Usando el método de codificación de 3 indicadores de marca de masa descritos previamente (Figura 37), aproximadamente 3.300 copias de cada marca de masa están disponibles para la medición. Como alternativa, un único dendrímero G9 (con 2048 grupos reactivos) funcionalizado con poli-lisina de peso molecular 10.000 tiene disponibles aproximadamente 170.000 sitios de unión de marcas de masa en un segmento de 12 nm de la construcción de indicador.

10 Para detectar una secuencia con marcas de masa, se usa un método de liberación controlada de los indicadores de marca de masa en el punto de medición por uso de conectores fotoescindibles. La fragmentación secuencial del anclaje no es necesaria. Los indicadores de marca de masa asociados a cada subunidad del polímero Xpandómero se miden en una etapa. Por ejemplo, un conjunto de 13 indicadores de marca de masa que oscilan de 350 Dalton a 710 Dalton (es decir, una escala de 30 Dalton de indicadores de marca de masa) tiene 286 combinaciones de tres marcas de masa cada uno. De este modo, uno cualquiera de los 256 4meros diferentes está asociado a solo una combinación particular de 3 indicadores de marca de masa. La información de secuencias codificada del Xpandómero se detecta fácilmente por espectroscopía de masas de las subunidades. Debido a que las subunidades del Xpandómero están espacialmente bien separadas, la tecnología de manipulación y detección del Xpandómero no necesita ser altamente sofisticada.

#### EJEMPLO 20.

#### 20 ANÁLISIS DIRECTO DE CONSTRUCCIONES DE SUSTRATO SIN MARCAR

Se sintetiza una biblioteca de construcciones de sustrato; los anclajes no contienen indicadores. Tras la preparación de un producto de Xpandómero, las bases individuales del producto de Xpandómero se analizan por espectroscopía túnel electrónica.

#### EJEMPLO 21.

#### 25 ANÁLISIS ASISTIDO POR HIBRIDACIÓN DE CONSTRUCCIONES DE SUSTRATO SIN MARCAR

Se sintetiza una biblioteca de construcciones de sustrato oligomérico; los anclajes no contienen indicadores. Tras la preparación de un producto de Xpandómero, un conjunto completo de sondas marcadas se hibridan entonces con el producto de Xpandómero y las sondas duplexadas se analizan secuencialmente.

#### EJEMPLO 22.

#### 30 SÍNTESIS DE UN TRIFOSFATO DE DESOXIADENOSINA CON ANCLAJE DE LYS-CYS-PEG-POLIGLUTAMATO-PEG-CYS-COOH

35 Se inmoviliza lisina con una cadena lateral de amino protegida con BOC sobre una resina y se hace reaccionar con un residuo de cisteína usando métodos de síntesis de péptidos estándar. La cadena lateral de amino de la lisina será el grupo funcional reactivo épsilon del anclaje de RT-NTP (véase clase VI, VII). La cisteína formará una primera mitad de un enlace disulfuro intra-anclaje. La amina desprotegida sobre la cisteína se modifica con SANH (Pierce-Thermo Fisher, USA; Cat N° 22400: Bioconjugate Toolkit) para formar una hidrazida.

40 Por separado, se prepara un segmento de espaciador a partir de bis-amino PEG 2000 (Creative PEGWorks, Winston Salem NC; Cat N° PSB 330) por funcionalización de las aminas libres con C6-SFB (Pierce-Thermo Fisher, EE.UU.; Cat N° 22400: Bioconjugate Toolkit), formando un segmento de espaciador de bis-4FB PEG; el producto se purifica.

El segmento de espaciador de bis-4FB PEG se acopla entonces al conector de hidrazida sobre la cisteína, dejando un grupo 4FB como grupo reactivo terminal, y la resina se lava.

45 Por separado, se prepara un segmento de poliglutamato (cada glutamato derivatizado en el gamma-carboxilo con 5 unidades de PEO de PEG encapuchado con metilo). El extremo C se convierte en una amina con el agente de acoplamiento EDC y diaminohexano. Se usa SANH para formar un segmento de poliglutamato terminado con dihidrazida, y el producto se purifica. El segmento de poliglutamato terminado con dihidrazida se hace reaccionar con el grupo 4FB terminal sobre la resina, formando una hidrazida terminal, y la resina se lava.

50 Por separado, se hace reaccionar un segmento de espaciador de PEG-2000 (con extremos amina y carboxilo, Creative PEGWorks; PHB-930) con SFB para generar un grupo terminal 4FB. Este segmento de espaciador se hace reaccionar con el grupo hidrazida sobre la resina, formando una cadena terminada con carboxilo. La resina se lava de nuevo.

Se acopla un residuo de cisteína al carboxilo libre usando síntesis de péptidos reactivos estándar. El carboxilo terminal de la cisteína está protegido con O-bencilo. El producto resultante se lava de nuevo y a continuación se escinde de la resina. El carboxilo libre generado por escisión se modifica entonces con EDC, aminohexilo y SANH

para formar una hidrazida reactiva.

- 5 Por separado, una desoxiadenosina trifosfato modificada con C6 amina (N6-(6-amino)hexil-dATP, Jena Bioscience, Jena DE; Cat N° NU-835) se trata con SFB (Pierce Bioconjugate Toolkit, Cat N° 22419) para formar un grupo funcional 4FB. Combinando la base modificada con la hidrazida reactiva de las etapas precedentes, se ensambla una construcción de sustrato de anclaje-sonda. BOC de la cadena lateral de lisina se elimina antes de uso. Bajo condiciones generalmente oxidantes, las cisteínas se asocian para formar un enlace disulfuro intra-anclaje, estabilizando el anclaje en una forma compacta limitada.

#### EJEMPLO 23.

##### SÍNTESIS DE UNA BIBLIOTECA DE TRIFOSFATO DE RT-NTP

- 10 Las bases de nucleótido trifosfato ancladas A, T, C y G con enlace -S-S- intra-anclaje se preparan como se describe en el Ejemplo 22, pero la carga y los parámetros físicos de los segmentos de glutamato PEGilado usados para cada base se seleccionan para proporcionar una característica de indicador distinta.

#### EJEMPLO 24.

##### SÍNTESIS DE XPANDÓMEROS POR SBE USANDO RT-NTP ANCLADOS CON POLIGLUTAMATO

- 15 Se preparan nucleótido trifosfatos de adenosina y guanosina RT-NTP modificados con enlaces disulfuro intra-anclaje. Las bases se modifican adicionalmente de manera que se bloqueen reversiblemente en la posición 3'. La química de bloqueo reversible basada en alilo es como se describe por Ruparel ("Design and synthesis of a 3'-O-allyl photocleavable fluorescent nucleotide as a reversible terminator for DNA sequencing by synthesis" PNAS, 102:5932-37, 2005). Los anclajes de las bases modificadas se construyen con grupo funcional delta y grupo funcional épsilon generalmente como se muestra en la Figura 61. El grupo funcional delta es un carboxilo de una cisteína lateral del anclaje y el grupo funcional épsilon es una cadena lateral amina de una lisina cerca de la unión del anclaje con las purinas. Los anclajes se modifican adicionalmente de manera que contengan segmentos de poliglutamato modificados específicos de nucleobase.

- 25 La secuencia TCTCTCTCTCTCTCATCTACCGTCCGTCCC" se usa como molde. La secuencia "5'GGGACGGACGGTAGAT" se usa como cebador. Se usa una HEX del extremo 5' (5'hexafluoro-fluoresceína) sobre el cebador como marca. El método de síntesis de Xpandómeros es esencialmente como se describe para la Figura 61. En un primer ciclo de cebado de SBE, la nucleobase modificada se añade con polimerasa Klenow en condiciones adaptadas para la polimerización y se añade una única base con la hebra hija naciente sobre el extremo 3'-OH del cebador. Debido a que la construcción de sustrato está bloqueada en la posición 3', no se produce polimerización adicional.

El grupo conector de la cadena lateral de amino (épsilon) sobre el primer RT-NTP añadido está encapuchado y seguirá así durante la reacción de SBE. El grupo carboxilo terminal del anclaje se desprotege y el 3' OH sobre el sustrato se desbloquea; el complejo se lava antes de la siguiente ronda de SBE.

- 35 En un segundo ciclo de SBE, otra nucleobase se polimeriza con el producto intermedio de Xpandómero naciente. El enlace  $\chi$  se forma entre la amina libre del grupo conector épsilon sobre la primera nucleobase y el grupo conector de carboxilo sobre el anclaje de la segunda nucleobase usando EDC y sulfo-NHS como agente de reticulación (Pierce Cat N° 22980 y 24510). El carboxilo sobre el grupo conector delta del anclaje se desprotege y el 3' OH sobre el sustrato se desbloquea; el complejo se lava antes de una siguiente ronda de SBE.

- 40 El ciclo de SBE puede repetirse múltiples veces, formando así un producto intermedio de Xpandómero en la configuración limitada. Cada anclaje en la cadena creciente de nucleobases unidas por  $\chi$  está en la configuración de Xpandómero limitado.

#### EJEMPLO 25.

##### ESCISIÓN POR NUCLEASA Y TCEP PARA FORMAR XPANDÓMERO DE CLASE X

- 45 El producto intermedio de Xpandómero del Ejemplo 24 se escinde con nucleasa, formando un producto de Xpandómero compuesto de nucleobases individuales enlazadas por segmentos de anclaje y enlaces  $\chi$ . La nucleasa también degrada el molde y cualquier cebador asociado, liberando el producto. Los enlaces disulfuro intra-anclaje se escinden mediante la adición de un agente reductor (TCEP, Pierce Cat. N°. 20490).

- 50 El producto de Xpandómero se filtra y se purifica para eliminar sintonas truncadas y productos secundarios de la digestión con nucleasa. La detección y el análisis del Xpandómero linealizado pueden hacerse usando una amplia variedad de métodos de generación existentes y futuros.



EJEMPLO 26.

## SÍNTESIS DE UNA DESOXIADENOSINA TRIFOSFATO CON UN CONECTOR FOTOESCINDIBLE INTRA-ANCLAJE

5 Se inmoviliza glicina sobre una resina y se hizo reaccionar con una cisteína. A continuación, el grupo amino de cisteína se desprotege y se hace reaccionar con un glutamato, el glutamato con una cadena lateral modificada con un conector fotolábil que termina en un carboxilo protegido con OBencil, tal como un conector de 2-nitroveratrilamina adaptado a partir del descrito por Holmes et al. ("Reagents for combinatorial organic synthesis: development of a new O-nitrobenzyl photolabile linker for solid phase synthesis", J Org Chem, 60:2318-19, 1995). La cisteína será el "grupo funcional épsilon" del anclaje de RT-NTP. La amina desprotegida sobre el glutamato se modifica con SANH (Pierce-Thermo Fisher, EE.UU.; Cat N° 22400: Bioconjugate Toolkit) para formar una hidrazida. La cadena lateral del glutamato formará un conector intra-anclaje fotoescindible tras la síntesis del anclaje.

15 Por separado, se prepara un segmento de espaciador a partir de bis-amino PEG 2000 (Creative PEGWorks, Winston Salem, NC, USA; Cat N° PSB 330) por funcionalización de las aminas libres con C6-SFB (Pierce Bioconjugate Toolkit, Cat N° 22423), formando un segmento de espaciador de bis-4FB PEG, y el producto se purifica. El segmento de espaciador de bis-4FB PEG se acopla entonces al conector de hidrazida sobre el glutamato, dejando un grupo 4FB como el grupo reactivo terminal, y la resina se lava.

20 Por separado, se prepara un segmento de poliglutamato (con cadenas laterales protegidas con t-butilo). El extremo C se convierte en una amina con el agente de acoplamiento EDC y diaminohexano. Se usa SANH para formar un segmento de poliglutamato terminado con di-hidrazida, y el producto se purifica. El segmento de poliglutamato terminado con dihidrazida se hace reaccionar con el grupo 4FB terminal sobre la resina, formando una hidrazida terminal sobre la resina, y la resina se lava.

25 Se modifica un segmento de espaciador de PEG-2000 (con amino libre y extremos amino protegidos con FMOC; Cat N° PHB-0982, Creative PEGWorks) con C6 SFB para formar un 4FB y segmento de espaciador de PEG modificado con FMOC-amino. El extremo de 4FB se hace reaccionar con el grupo hidrazida sobre la resina, formando una cadena terminada con FMOC-amino. La resina se lava de nuevo.

Se acopla un residuo de lisina a la amina libre del espaciador por un enlace peptídico. El residuo de lisina se protege sobre la cadena lateral por BOC y la alfa-amino de la lisina se protege por FMOC. A continuación se desprotegen el carboxilo terminal de OBencilo del conector fotoescindible y la cadena lateral protegida con BOC de la lisina y se reticulan con EDC/sulfo-NHS para circularizar el anclaje.

30 El producto resultante se lava de nuevo y a continuación se escinde de la resina. El carboxilo de glicina libre generado por la escisión se modifica entonces con EDC, aminohexilo y SANH para formar una hidrazida reactiva.

35 Por separado, una desoxiadenosina trifosfato modificada con C6 amina (N6-(6-amino)hexil-dATP, Jena Bioscience, Jena DE; Cat N° NU-835) se trata con SFB (Pierce Bioconjugate Toolkit, Cat N° 22419) para formar un grupo funcional 4FB. Combinando la base modificada con la hidrazida reactiva de las etapas precedentes se ensambla una construcción de sustrato de anclaje-sonda. El conector intra-anclaje fotoescindible estabiliza el anclaje en una forma compacta limitada. Entonces se elimina FMOC y la amina terminal libre se hace reaccionar con sulfo-EMCS (Pierce; Cat N° 22307) para introducir un grupo funcional maleimido terminal.

EJEMPLO 27.

## SÍNTESIS DE XPANDÓMEROS POR SBE USANDO RT-NTP FOTOESCINDIBLE

40 Se preparan nucleótido trifosfatos de adenosina y guanosina de RT-NTP modificados con enlaces intra-anclaje fotoescindibles. Las bases se modifican adicionalmente de manera que se bloqueen reversiblemente en la posición 3'. La química de bloqueo reversible basada en alilo es como se describe por Ruparel ("Design and synthesis of a 3'-O-allyl photocleavable fluorescent nucleotide as a reversible terminator for DNA sequencing by synthesis", PNAS, 102:5932-37, 2005). Los anclajes de las bases modificadas se construyen con grupo funcional delta y grupo funcional épsilon generalmente como se muestra en la Figura 61. El grupo conector delta es una amina de una lisina lateral terminal del anclaje y el grupo conector épsilon es un sulfhidrilo de una cisteína cerca del punto de unión del anclaje. Los anclajes se modifican adicionalmente de manera que contengan segmentos de poliglutamato modificados específicos de especie.

50 La secuencia TCTCTCTCTCTCTCATCTACCGTCCGTCCC" se usa como molde. La secuencia "5'GGGACGGACGGTAGAT" se usa como cebador. Se usa una HEX del extremo 5' (5'hexafluoro-fluoresceína) sobre el cebador como marca. El método de síntesis de Xpandómeros es esencialmente como se ha descrito para la Figura 61. En un primer ciclo de cebado de SBE, la nucleobase modificada (A) se añade con polimerasa de Klenow en condiciones adaptadas para la polimerización y se añade una única base con la hebra hija naciente sobre el extremo 3'-OH del cebador. Debido a que la construcción de sustrato se bloquea en la posición 3', no se produce polimerización adicional. El complejo de cebador-molde inmovilizado se lava entonces para eliminar el sustrato sin reaccionar.

El grupo conector de la cadena lateral de sulfhidrilo (épsilon) sobre el primer RT-NTP añadido está encapuchado y seguirá así durante la reacción de SBE. El grupo amino terminal del anclaje se desprotege y el 3' OH sobre el sustrato se desbloquea; el complejo se lava antes de la siguiente ronda de SBE.

5 En un segundo ciclo de SBE, otra nucleobase (G) se polimeriza con el producto intermedio de Xpandómero naciente. El enlace  $\chi$  se forma entre el amino (grupo conector delta) sobre la primera nucleobase y el sulfhidrilo (grupo conector épsilon) sobre el anclaje de la segunda base usando el reactivo de reticulación GMBS (Pierce; Cat N° 22309). El grupo conector amino delta sobre el segundo RT-NTP se desprotege y el 3'-OH del sustrato se desbloquea; el complejo se lava antes de la siguiente ronda de SBE.

10 El ciclo de SBE puede repetirse múltiples veces, formando así un producto intermedio de Xpandómero en la configuración limitada. Cada anclaje en la cadena creciente de nucleobases unidas por  $\chi$  está en la configuración de Xpandómero limitado.

#### EJEMPLO 28.

##### NUCLEASA Y FOTOESCISIÓN PARA FORMAR XPANDÓMERO DE CLASE X

15 El producto intermedio de Xpandómero del Ejemplo 27 se escinde con nucleasa, formando un producto de Xpandómero compuesto de nucleobases individuales enlazadas por segmentos de anclaje y enlaces  $\chi$ . La nucleasa también degrada el molde y cualquier cebador asociado, liberando el producto. Los enlaces fotoescindibles intra-anclaje se escinden por exposición a una luz UV.

20 El producto de Xpandómero se filtra y se purifica para eliminar sintonas truncadas y productos secundarios de la digestión con nucleasa. La detección y el análisis del Xpandómero linealizado puede hacerse usando una amplia variedad de métodos de generación existentes y futuros.

#### EJEMPLO 29.

##### SÍNTESIS DE UN ANCLAJE DE RT-NTP IN SITU

25 Usando métodos de síntesis de péptidos estándar sobre un soporte sólido, se prepara un péptido que tiene la estructura (Resina-C')-Glu-Cys-(Gly-Ala)<sub>10</sub>-Pro-Ser-Gly-Ser-Pro-(Ala-Gly)<sub>10</sub>-Cys-Lys. La amina terminal se hace reaccionar con SANH (Pierce, Cat N° 22400) para crear un conector de hidrazida.

Por separado, una desoxiadenosina trifosfato modificada con C6 amina (N6-(6-amino)hexil-dATP, Jena Bioscience, Jena DE; Cat N° NU-835) se trata con SFB (Pierce Bioconjugate Toolkit, Cat N° 22419) para formar un grupo funcional 4FB. Combinando la base modificada con la hidrazida reactiva de las etapas precedentes, se ensambla una construcción de sustrato de anclaje-sonda.

30 La construcción se escinde a continuación de la resina. Tras la desprotección y bajo condiciones generalmente oxidantes, las cisteínas se asocian para formar un enlace disulfuro intra-anclaje, estabilizando la horquilla beta, que contiene un carboxilo libre terminal (un grupo conector delta lateral sobre el anclaje) y una lisina cerca del punto de unión del anclaje (la amina de la cadena lateral y grupo conector épsilon).

35 El disulfuro es representativo de la estabilización intra-anclaje representada en las construcciones de sustrato de clases II, III, VI, VII y VIII (véanse las Figuras 8 y 9), aunque ilustradas aquí con referencia más específica a construcciones de sustrato monomérico de clases VI, VII y VIII. La longitud del anclaje sin plegar, suponiendo una longitud del enlace peptídico C-C del residuo de 3,8 Å, es aproximadamente 10 nm, pero supone una forma compacta debido al enlace de hidrógeno en la horquilla beta.

40 Como se describe por Gellman ("Foldamers, a manifesto", Acc Chem Res 31:173-80, 1998), una amplia variedad de polímeros, no simplemente péptidos, pueden plegarse en formas compactas. Tales polímeros incluyen oligopiridinas, poliisocianuros, poliisocianatos, metacrilatos de poli(triarilmetilo), polialdehídos, poliprolina, ARN, oligopirrolinonas y oligoureas, todos los cuales han presentado la capacidad de plegarse en estructuras secundarias compactas y expandirse bajo condiciones adecuadas. Por tanto, los ejemplos de péptidos presentados aquí son representativos de una clase mucho mayor de químicas de anclaje, en la que las limitaciones al anclaje sin expandir pueden incluir enlace de hidrógeno e interacciones hidrófobas, por ejemplo, además de reticulaciones intra-anclaje.

#### EJEMPLO 30.

##### SÍNTESIS DE XPANDÓMEROS USANDO XSONDAS

50 En una realización de SBX, una biblioteca de Xsondas de 256 Xsondas 4meras se presenta a una diana de ADN monocatenario anclada y alargada para la hibridación. La etapa de hibridación continúa bajo un ciclado térmico preciso rutinario para promover cadenas de Xsondas largas. Se eliminan dúplex de sonda-diana no específicos débilmente unidos por una simple etapa de lavado, de nuevo bajo control térmico preciso. La ligación enzimática se realiza para enlazar cualquier cadena de Xsondas a lo largo del ADN diana, seguido de un segundo lavado. Repitiendo el ciclo de hibridación / lavado / ligación / lavado, secuencias enlazadas más largas crecen en múltiples

loci a lo largo del ADN diana hasta que la replicación del molde diana sea generalmente completa.

Los huecos sin llenar a lo largo del ADN diana se llenan usando una ADN polimerasa bien establecida y proceso de llenado de huecos basado en ligasa (Lee, "Ligase Chain Reaction", Biologicals, 24(3): 197-199, 1996). Los nucleótidos incorporados en los huecos tienen un código de indicador único para indicar un nucleótido de hueco. El producto intermedio de Xpandómero completado, que está compuesto por la diana de ADN original con Xsondas complementariamente duplexadas y ligadas con cargas de 1, 2 o 3 huecos de nucleótido ocasionales, se escinde para producir un Xpandómero. El conector escindible para este ejemplo es una modificación de esqueleto del sustrato 3' OP-N 5'. La escisión selectiva se cataliza mediante la adición de ácido acético a temperatura ambiente.

El Xpandómero se filtra y se purifica para eliminar productos truncados y posteriormente se alarga para formar una estructura lineal de códigos de indicador enlazados. La detección y análisis del producto de Xpandómero puede hacerse usando una amplia variedad de métodos existentes.

#### EJEMPLO 31.

##### SÍNTESIS DE XPANDÓMEROS XNTP USANDO POLIMERASA

La síntesis de un Xpandómero de clase X se realiza usando una construcción de sustrato modificada 8-[(6-amino)hexil]-amino-desoxiadenosina trifosfato que tiene un esqueleto mixto que consiste en un fosfonato de 2-aminoetilo no conector y un fosforotiolato conector (3' O-P- S 5') en el fosfato alfa. Un anclaje intra-nucleótido está unido al conector de fosfonato de 2-aminoetilo y a un conector de C6 amino sobre 8-[(6-amino)hexil]-amino-desoxiATP. La secuencia TTTTTTTTTTTTTTTTTTTTATCTACCGTCCGTCCTCC" se usa como molde. La secuencia "5'GGGACGGACGGTAGAT" se usa como cebador. Se usa una HEX del extremo 5' (5'hexafluoro-fluoresceína) sobre el cebador como marca. La hibridación del cebador y el molde forma un molde con cebador de dúplex y nucleótido protuberante monocatenario de los extremos 3'-OH y 5' libres de veinte bases en longitud. Entonces se añaden las construcciones de sustrato y polimerasa y la polimerización continúa durante 60 min en condiciones ajustadas para la polimerización óptima. Se mezcla una muestra de la reacción de polimerización con tampón de carga de gel y se somete a electroforesis sobre un gel de 20 % de TBE-acrilamida (Invitrogen, EE.UU.) junto con un control negativo no de polimerasa y un marcador de MW para confirmar la polimerización de XNTP.

El producto intermedio de Xpandómero se trata con catión divalente (véase Mag et al. 1991. "Synthesis and selective cleavage of an oligodeoxynucleotide containing a bridged internucleotide 5'-phosphorothioate linkage", Nucleic Acids Research, 19(7):1437-1441) para escindir selectivamente los enlaces fosforotiolato entre la unión del anclaje y la desoxirribosa, que se confirma por electroforesis.

#### EJEMPLO 32.

##### SÍNTESIS DE XPANDÓMEROS XNTP USANDO POLIMERASA

La síntesis de un Xpandómero de clase X se realiza usando una construcción de sustrato modificada N<sup>6</sup>-(6-amino)hexil-desoxiadenosina trifosfato que tiene un esqueleto mixto que consiste en un (N-1-aminoalquil)fosforamidato no conector y un fosforotiolato conector (3' O-P-S 5') en el fosfato alfa. Un anclaje intra-nucleótido está unido al grupo N-1-aminoalquilo y a un conector de C6 amino sobre N<sup>6</sup>-(6-amino)hexil-desoxiATP. La secuencia TTTTTTTTTTTTTTTTTTTTATCTACCGTCCGTCCTCC" se usa como molde. La secuencia "5'GGGACGGACGGTAGAT" se usa como cebador. Se usa una HEX del extremo 5' (5'hexafluoro-fluoresceína) sobre el cebador como marca. La hibridación del cebador y el molde forma un molde con cebador de dúplex y nucleótido protuberante monocatenario de los extremos 3'-OH y 5' libres de veinte bases en longitud. Entonces se añaden las construcciones de sustrato y polimerasa y la polimerización continúa durante 60 min en condiciones ajustadas para la polimerización óptima. Se mezcla una muestra de la reacción de polimerización con tampón de carga de gel y se somete a electroforesis sobre un gel de 20 % de TBE-acrilamida (Invitrogen, EE.UU.) junto con un control negativo no de polimerasa y un marcador de MW para confirmar la polimerización de XNTP.

El producto intermedio de Xpandómero se trata con yodoetanol (véase Gish et al ("DNA and RNA sequence determination based on phosphorothioate chemistry", Science, 240(4858): 1520-1522, 1988) o por escisión con cationes metálicos divalentes como se describe por Vyle et al ("Sequence- and strand-specific cleavage in oligodeoxyribonucleotides and DNA containing 3'-thiothymidine". Biochemistry 31(11): 3012-8, 1992) para escindir selectivamente los enlaces fosforotiolato entre la unión del anclaje y la desoxirribosa, que se confirma por electroforesis.

#### EJEMPLO 33.

##### SÍNTESIS DE XPANDÓMEROS XNTP USANDO LIGASA

La síntesis de un Xpandómero de clase X se realiza usando una construcción de sustrato modificada 8-[(6-amino)hexil]-amino-desoxiadenosina monofosfato que tiene un esqueleto mixto que consiste en un fosfonato de 2-aminoetilo no conector y un fosforotiolato conector (3' O-P- S 5') en el fosfato alfa. Un anclaje intra-nucleótido está unido al conector de fosfonato de 2-aminoetilo y a un conector de C6 amino sobre la 8-[(6-amino)hexil]-amino-

5 desoxiAMP. La secuencia TTTTTTTTTTTTTTTTTTTTATCTACCGTCCGTCCC" se usa como molde. La secuencia "5'GGGACGGACGGTAGAT" se usa como cebador. Se usa una HEX del extremo 5' (5'hexacloro-fluoresceína) sobre el cebador como marca. La hibridación del cebador y el molde forma un molde con cebador de dúplex y nucleótido protuberante monocatenario de los extremos 3'-OH y 5' libres de veinte bases en longitud. Entonces se añaden las construcciones de sustrato y ligasa y la ligación continúa durante 5 horas en condiciones ajustadas para la ligación. Se mezcla una muestra de la reacción de ligación con tampón de carga de gel y se somete a electroforesis sobre un gel de 20 % de TBE-acrilamida (Invitrogen, EE.UU.) junto con un control negativo no de polimerasa y un marcador de MW para confirmar la ligación de XNTP.

10 El producto intermedio de Xpandómero se trata con 80 % de ácido acético durante 5 h a temperatura ambiente según el procedimiento de Mag et al ("Synthesis and selective cleavage of oligodeoxyribonucleotides containing non-chiral internucleotide phosphoramidate linkages", Nucl. Acids Res. 17: 5973-88, 1989) para escindir selectivamente los enlaces fosforamidato entre el punto de unión del anclaje y la desoxirribosa, que se confirma por electroforesis.

Las diversas realizaciones descritas anteriormente pueden combinarse para proporcionar realizaciones adicionales.

**REIVINDICACIONES**

1. Un método para la secuenciación de un ácido nucleico diana, que comprende:

- 5 a) proporcionar una hebra hija producida por una síntesis dirigida por molde, comprendiendo la hebra hija una pluralidad de subunidades acopladas en una secuencia correspondiente a una secuencia de nucleótidos contigua de toda o una porción del ácido nucleico diana, en la que las subunidades individuales comprenden un anclaje, al menos una sonda o residuo de nucleobase, y al menos un enlace selectivamente escindible;
- 10 b) escindir el al menos un enlace selectivamente escindible para dar un Xpandómero de una longitud mayor que la pluralidad de las subunidades de la hebra hija, comprendiendo el Xpandómero los anclajes y elementos indicadores para analizar la información genética en una secuencia correspondiente a la secuencia de nucleótidos contigua de toda o una porción del ácido nucleico diana; y
- c) detectar los elementos indicadores del Xpandómero.

2. El método de la reivindicación 1, en el que los elementos indicadores para analizar la información genética están asociados con:

- 15 los anclajes del Xpandómero;
- la hebra hija antes de la escisión del al menos un enlace selectivamente escindible; o
- el Xpandómero después de la escisión del al menos un enlace selectivamente escindible.

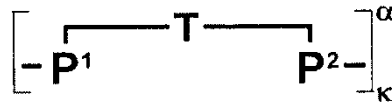
3. El método de la reivindicación 1, en el que el Xpandómero comprende además toda o una porción de la al menos una sonda o residuo de nucleobase.

4. El método de la reivindicación 3, en el que los elementos indicadores para analizar la información genética son o están asociados con la al menos una sonda o residuo de nucleobase.

5. El método de la reivindicación 1, en el que el al menos un enlace selectivamente escindible es:

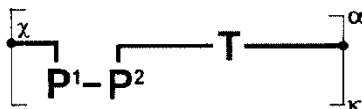
- 25 un enlace covalente;
- un enlace intra-anclaje;
- un enlace entre o dentro de las sondas o residuos de nucleobase de la hebra hija; o
- un enlace entre las sondas o residuos de nucleobase de la hebra hija y un molde diana.

6. El método de la reivindicación 1, en el que el Xpandómero comprende la siguiente estructura:



en la que

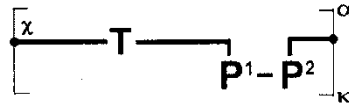
- 30 T representa el anclaje;
- P<sup>1</sup> representa un primer resto de sonda;
- P<sup>2</sup> representa un segundo resto de sonda;
- κ representa la subunidad k<sup>n</sup> en una cadena de m subunidades, en la que m es un número entero superior a tres; y
- 35 α representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos contigua de una porción del ácido nucleico diana;



en la que

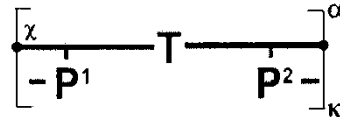
- 40 T representa el anclaje;
- P<sup>1</sup> representa un primer resto de sonda;
- P<sup>2</sup> representa un segundo resto de sonda;
- κ representa la subunidad k<sup>n</sup> en una cadena de m subunidades, en la que m es un número entero superior a tres;
- 45 α representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos

contigua de una porción del ácido nucleico diana; y  
 $\chi$  representa un enlace con el anclaje de una subunidad adyacente;



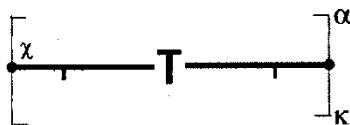
en la que

- 5 T representa el anclaje;  
 $P^1$  representa un primer resto de sonda;  
 $P^2$  representa un segundo resto de sonda;  
 $\kappa$  representa la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $m$  es un número entero superior a tres;
- 10  $\alpha$  representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y  
 $\chi$  representa un enlace con el anclaje de una subunidad adyacente;



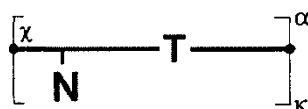
15 en la que

- T representa el anclaje;  
 $P^1$  representa un primer resto de sonda;  
 $P^2$  representa un segundo resto de sonda;  
 $\kappa$  representa la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $m$  es un número entero superior a tres;
- 20  $\alpha$  representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y  
 $\chi$  representa un enlace con el anclaje de una subunidad adyacente;



25 en la que

- T representa el anclaje;  
 $\kappa$  representa la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $m$  es un número entero superior a tres;
- 30  $\alpha$  representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y  
 $\chi$  representa un enlace con el anclaje de una subunidad adyacente;



35 en la que

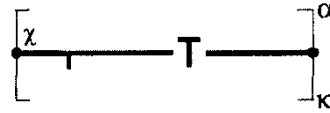
- T representa el anclaje;  
 $N$  representa un residuo de nucleobase;

$\kappa$  representa la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $m$  es un número entero superior a diez;

$\alpha$  representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y

5

$\chi$  representa un enlace con el anclaje de una subunidad adyacente;



en la que

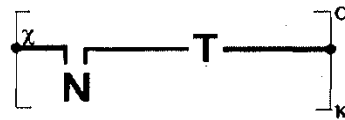
T representa el anclaje;

10  $\kappa$  representa la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $m$  es un número entero superior a diez;

$\alpha$  representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y

15

$\chi$  representa un enlace con el anclaje de una subunidad adyacente;



en la que

T representa el anclaje;

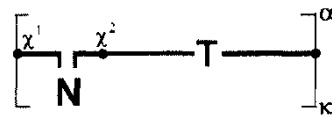
N representa un residuo de nucleobase;

20  $\kappa$  representa la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $m$  es un número entero superior a diez;

$\alpha$  representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y

25

$\chi$  representa un enlace con el anclaje de una subunidad adyacente;



en la que

T representa el anclaje;

N representa un residuo de nucleobase;

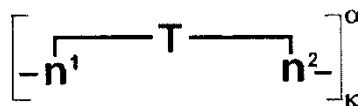
30  $\kappa$  representa la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $m$  es un número entero superior a diez;

$\alpha$  representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos contigua de una porción del ácido nucleico diana;

35

$\chi^1$  representa un enlace con el anclaje de una subunidad adyacente; y

$\chi^2$  representa un enlace inter-anclaje; o



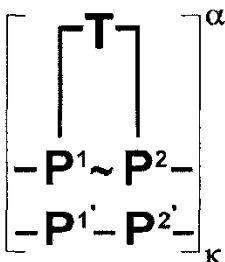
en la que

T representa el anclaje;

$n^1$  y  $n^2$  representa una primera porción y una segunda porción, respectivamente, de un residuo de nucleobase;  
 $k$  representa la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $m$  es un número entero superior a diez; y

5  $\alpha$  representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies comprende información de secuencias de la secuencia de nucleótidos contigua de una porción del ácido nucleico diana.

7. El método de la reivindicación 6, en el que, antes de la escisión del al menos un enlace selectivamente escindible, la hebra hija comprende un dúplex de molde-hebra hija que tiene la siguiente estructura:



10

en la que

T representa el anclaje;

$P^1$  representa un primer resto de sonda;

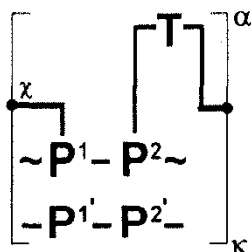
$P^2$  representa un segundo resto de sonda;

15  $\sim$  representa el al menos un enlace selectivamente escindible;  $P^{1'}$  representa una secuencia de nucleótidos contigua de al menos un residuo de nucleótido de la hebra molde a la que  $P^1$  es complementaria;

$P^{2'}$  representa una secuencia de nucleótidos contigua de al menos un residuo de nucleótido de la hebra molde a la que  $P^2$  es complementaria;

20  $k$  representa la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $m$  es un número entero superior a tres; y

$\alpha$  representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies es complementaria a la secuencia de nucleótidos contigua de una porción del ácido nucleico diana;



25 en la que

T representa el anclaje;

$P^1$  representa un primer resto de sonda;

$P^2$  representa un segundo resto de sonda;

30  $\sim$  representa el al menos un enlace selectivamente escindible;  $P^{1'}$  representa una secuencia de nucleótidos contigua de al menos un residuo de nucleótido de la hebra molde a la que  $P^1$  es complementaria;

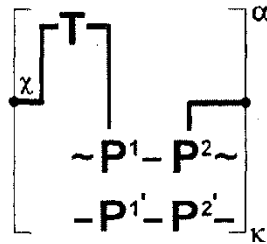
$P^{2'}$  representa una secuencia de nucleótidos contigua de al menos un residuo de nucleótido de la hebra molde a la que  $P^2$  es complementaria;

$k$  representa la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $m$  es un número entero superior a tres;

35  $\alpha$  representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies es complementaria a la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y

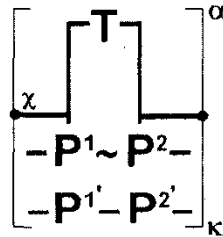
$\chi$  representa un enlace con el anclaje de una subunidad adyacente;





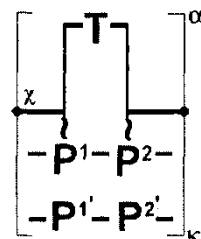
en la que

- T representa el anclaje;
- P<sup>1</sup> representa un primer resto de sonda;
- P<sup>2</sup> representa un segundo resto de sonda;
- ~ representa el al menos un enlace selectivamente escindible;
- P<sup>1</sup> representa una secuencia de nucleótidos contigua de al menos un residuo de nucleótido de la hebra molde a la que P<sup>1</sup> es complementaria;
- P<sup>2</sup> representa una secuencia de nucleótidos contigua de al menos un residuo de nucleótido de la hebra molde a la que P<sup>2</sup> es complementaria;
- k representa la subunidad k<sup>n</sup> en una cadena de m subunidades, en la que m es un número entero superior a tres;
- α representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies es complementaria a la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y
- χ representa un enlace con el anclaje de una subunidad adyacente;



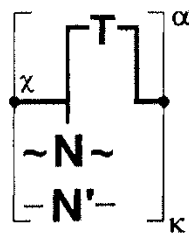
en la que

- T representa el anclaje;
- P<sup>1</sup> representa un primer resto de sonda;
- P<sup>2</sup> representa un segundo resto de sonda;
- ~ representa el al menos un enlace selectivamente escindible; P<sup>1</sup> representa una secuencia de nucleótidos contigua de al menos un residuo de nucleótido de la hebra molde a la que P<sup>1</sup> es complementaria;
- P<sup>2</sup> representa una secuencia de nucleótidos contigua de al menos un residuo de nucleótido de la hebra molde a la que P<sup>2</sup> es complementaria;
- k representa la subunidad k<sup>n</sup> en una cadena de m subunidades, en la que m es un número entero superior a tres;
- α representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies es complementaria a la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y
- χ representa un enlace con el anclaje de una subunidad adyacente;



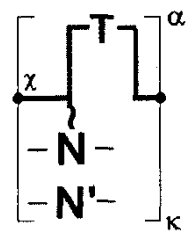
en la que

- T representa el anclaje;
- P<sup>1</sup> representa un primer resto de sonda;
- P<sup>2</sup> representa un segundo resto de sonda;
- 5 ~ representa el al menos un enlace selectivamente escindible;
- P<sup>1</sup> representa una secuencia de nucleótidos contigua de al menos un residuo de nucleótido de la hebra molde a la que P<sup>1</sup> es complementaria;
- P<sup>2</sup> representa una secuencia de nucleótidos contigua de al menos un residuo de nucleótido de la hebra molde a la que P<sup>2</sup> es complementaria;
- 10 κ representa la subunidad k<sup>n</sup> en una cadena de m subunidades, en la que m es un número entero superior a tres;
- α representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies es complementaria a la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y
- 15 χ representa un enlace con el anclaje de una subunidad adyacente;



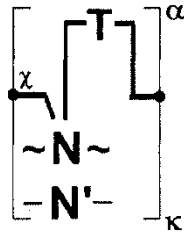
en la que

- T representa el anclaje;
- N representa un residuo de nucleobase;
- 20 N' representa un residuo de nucleótido de la hebra molde al que N es complementario;
- ~ representa el al menos un enlace selectivamente escindible;
- κ representa la subunidad k<sup>n</sup> en una cadena de m subunidades, en la que m es un número entero superior a diez;
- α representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies es complementaria a la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y
- 25 χ representa un enlace con el anclaje de una subunidad adyacente;



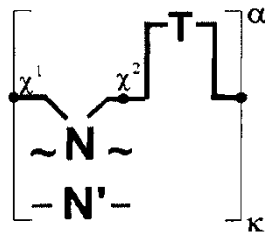
en la que

- 30 T representa el anclaje;
- N representa un residuo de nucleobase;
- N' representa un residuo de nucleótido de la hebra molde al que N es complementario;
- ~ representa el al menos un enlace selectivamente escindible;
- κ representa la subunidad k<sup>n</sup> en una cadena de m subunidades, en la que m es un número entero superior a diez;
- 35 α representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies es complementaria a la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y
- χ representa un enlace con el anclaje de una subunidad adyacente;



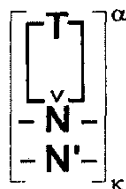
en la que

- T representa el anclaje;
- N representa un residuo de nucleobase;
- 5 N' representa un residuo de nucleótido de la hebra molde al que N es complementario;
- ~ representa el al menos un enlace selectivamente escindible;
- kappa representa la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $m$  es un número entero superior a diez;
- 10 alpha representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies es complementaria a la secuencia de nucleótidos contigua de una porción del ácido nucleico diana; y
- chi representa un enlace con el anclaje de una subunidad adyacente;



en la que

- 15 T representa el anclaje;
- N representa un residuo de nucleobase;
- N' representa un residuo de nucleótido de la hebra molde al que N es complementario;
- ~ representa el al menos un enlace selectivamente escindible;
- 20 kappa representa la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $m$  es un número entero superior a diez;
- alpha representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad, en la que cada una de las especies es complementaria a la secuencia de nucleótidos contigua de una porción del ácido nucleico diana;
- 25  $\chi^1$  representa un enlace con el anclaje de una subunidad adyacente; y
- $\chi^2$  representa un enlace inter-anclaje; o

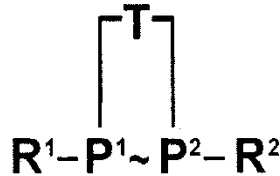


en la que

- T representa el anclaje;
- N representa un residuo de nucleobase;
- 30 N' representa un residuo de nucleótido de la hebra molde al que N es complementario;
- V representa un sitio de escisión interno del residuo de nucleobase;
- kappa representa la subunidad  $k^n$  en una cadena de  $m$  subunidades, en la que  $m$  es un número entero superior a diez; y
- alpha representa una especie de un motivo de subunidad seleccionado de una biblioteca de motivos de subunidad,

en la que cada una de las especies es complementaria a la secuencia de nucleótidos contigua de una porción del ácido nucleico diana.

8. El método de la reivindicación 6, en la que la hebra hija se forma a partir de una pluralidad de construcciones de sustrato de oligómero que tienen la siguiente estructura:



5

en la que

T representa el anclaje;

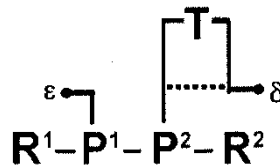
P<sup>1</sup> representa un primer resto de sonda;

P<sup>2</sup> representa un segundo resto de sonda;

~ representa el al menos un enlace selectivamente escindible; y

R<sup>1</sup> y R<sup>2</sup> representan los mismos grupos terminales o diferentes para la síntesis dirigida por molde de la hebra hija;

10



en la que

T representa el anclaje;

P<sup>1</sup> representa un primer resto de sonda;

P<sup>2</sup> representa un segundo resto de sonda;

R<sup>1</sup> y R<sup>2</sup> representan los mismos grupos terminales o diferentes para la síntesis dirigida por molde de la hebra hija;

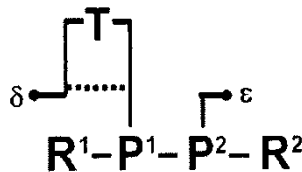
ε representa un primer grupo conector;

δ representa un segundo grupo conector; y

"- - -" representa una reticulación intra-anclaje escindible;

15

20



en la que

T representa el anclaje;

P<sup>1</sup> representa un primer resto de sonda;

P<sup>2</sup> representa un segundo resto de sonda; R<sup>1</sup> y R<sup>2</sup> representan los mismos grupos terminales o diferentes para la síntesis dirigida por molde de la hebra hija;

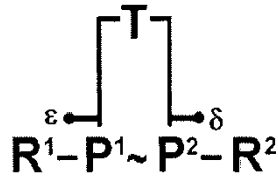
ε representa un primer grupo conector;

δ representa un segundo grupo conector; y

"- - -" representa una reticulación intra-anclaje escindible;

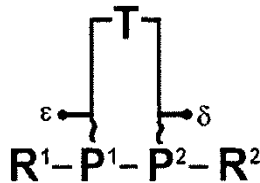
25

30



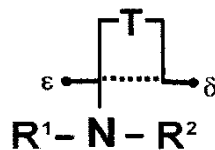
en la que

- 5 T representa el anclaje;
- P<sup>1</sup> representa un primer resto de sonda;
- P<sup>2</sup> representa un segundo resto de sonda;
- ~ representa el al menos un enlace selectivamente escindible;
- R<sup>1</sup> y R<sup>2</sup> representan los mismos grupos terminales o diferentes para la síntesis dirigida por molde de la hebra hija;
- ε representa un primer grupo conector; y
- 10 δ representa un segundo grupo conector;



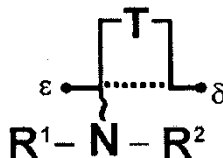
en la que

- 15 T representa el anclaje;
- P<sup>1</sup> representa un primer resto de sonda;
- P<sup>2</sup> representa un segundo resto de sonda;
- ~ representa el al menos un enlace selectivamente escindible;
- R<sup>1</sup> y R<sup>2</sup> representan los mismos grupos terminales o diferentes para la síntesis dirigida por molde de la hebra hija;
- ε representa un primer grupo conector; y
- 20 δ representa un segundo grupo conector;



en la que

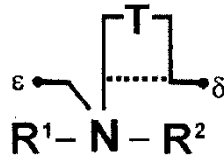
- 25 T representa el anclaje;
- N representa un residuo de nucleobase;
- R<sup>1</sup> y R<sup>2</sup> representan los mismos grupos terminales o diferentes para la síntesis dirigida por molde de la hebra hija;
- ε representa un primer grupo conector;
- δ representa un segundo grupo conector; y
- "- - -" representa una reticulación intra-anclaje escindible;



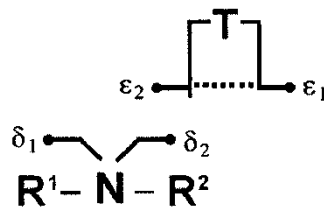
30

en la que

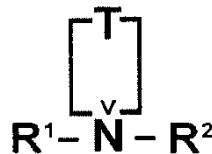
- 5 T representa el anclaje;  
 N representa un residuo de nucleobase;  
 $R^1$  y  $R^2$  representan los mismos grupos terminales o diferentes para la síntesis dirigida por molde de la hebra  
 hija;  
 ~ representa el al menos un enlace selectivamente escindible;  
 $\varepsilon$  representa un primer grupo conector;  
 $\delta$  representa un segundo grupo conector; y  
 “- - -” representa una reticulación intra-anclaje escindible;



- 10 en la que  
 T representa el anclaje;  
 N representa un residuo de nucleobase;  
 $R^1$  y  $R^2$  representan los mismos grupos terminales o diferentes para la síntesis dirigida por molde de la hebra  
 hija;  
 $\varepsilon$  representa un primer grupo conector;  
 $\delta$  representa un segundo grupo conector; y  
 “- - -” representa una reticulación intra-anclaje escindible;
- 15

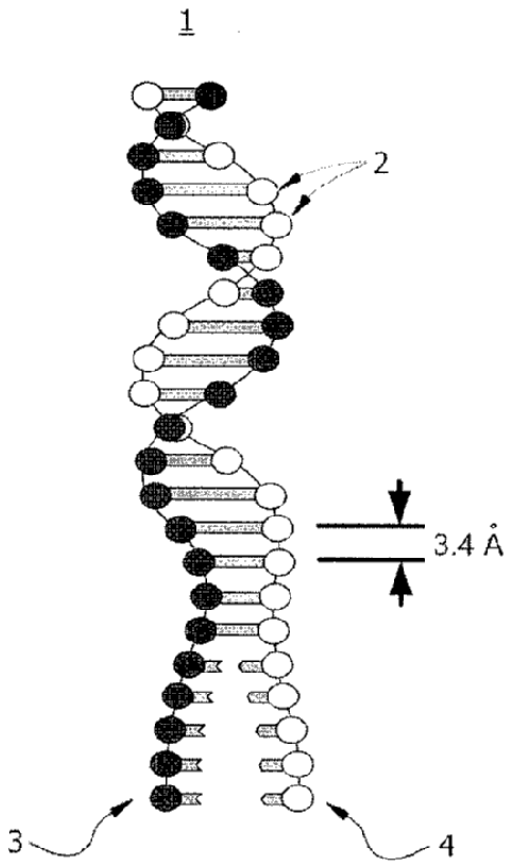


- 20 en la que  
 T representa el anclaje;  
 N representa un residuo de nucleobase;  
 $R^1$  y  $R^2$  representan los mismos grupos terminales o diferentes para la síntesis dirigida por molde de la hebra  
 hija;  
 $\varepsilon_1$  y  $\varepsilon_2$  representan los primeros grupos conectores iguales o diferentes;  
 $\delta_1$  y  $\delta_2$  representan los segundos grupos conectores iguales o diferentes; y  
 “- - -” representa una reticulación intra-anclaje escindible; o
- 25

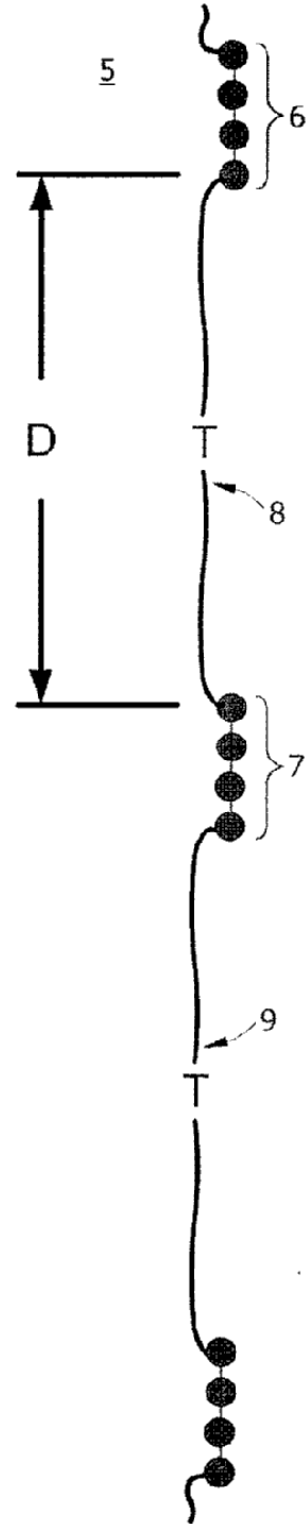


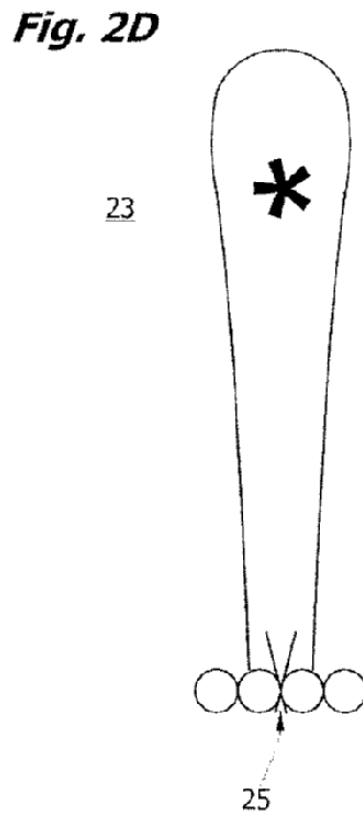
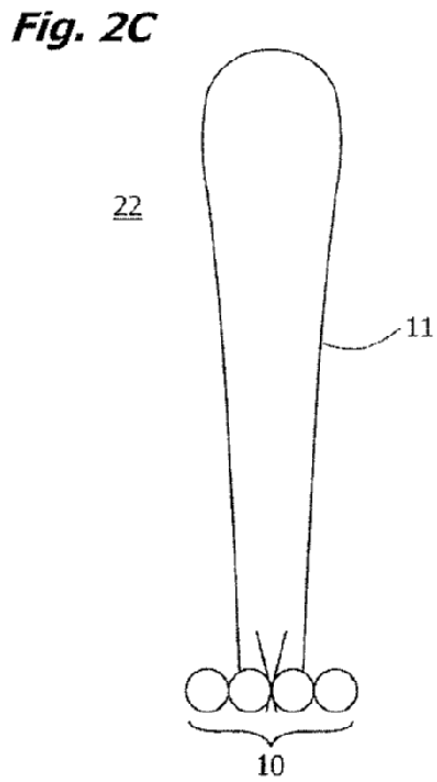
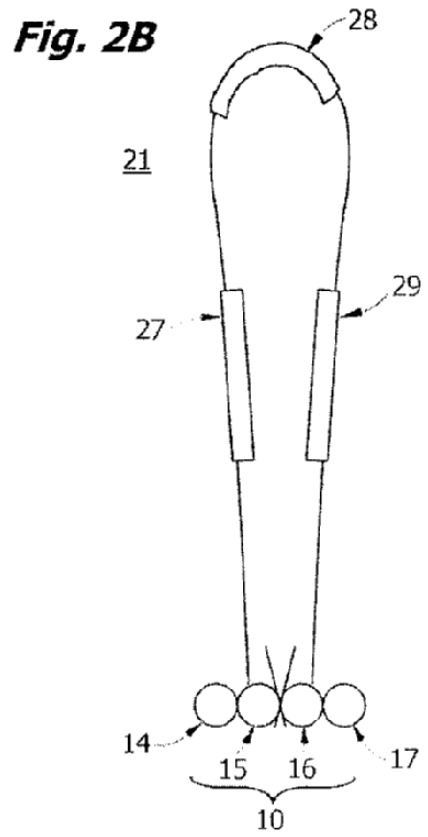
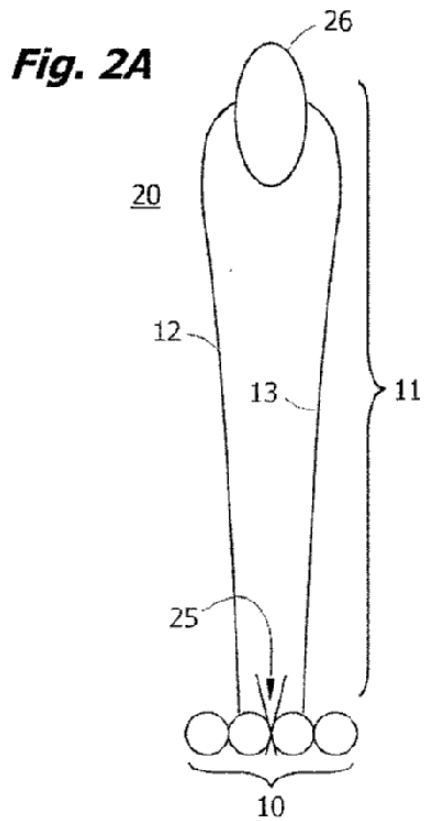
- 30 en la que  
 T representa el anclaje;  
 N representa un residuo de nucleobase;  
 V representa un sitio de escisión interno del residuo de nucleobase; y  
 $R^1$  y  $R^2$  representan los mismos grupos terminales o diferentes para la síntesis dirigida por molde de la hebra  
 hija.
- 35

**Fig. 1A**



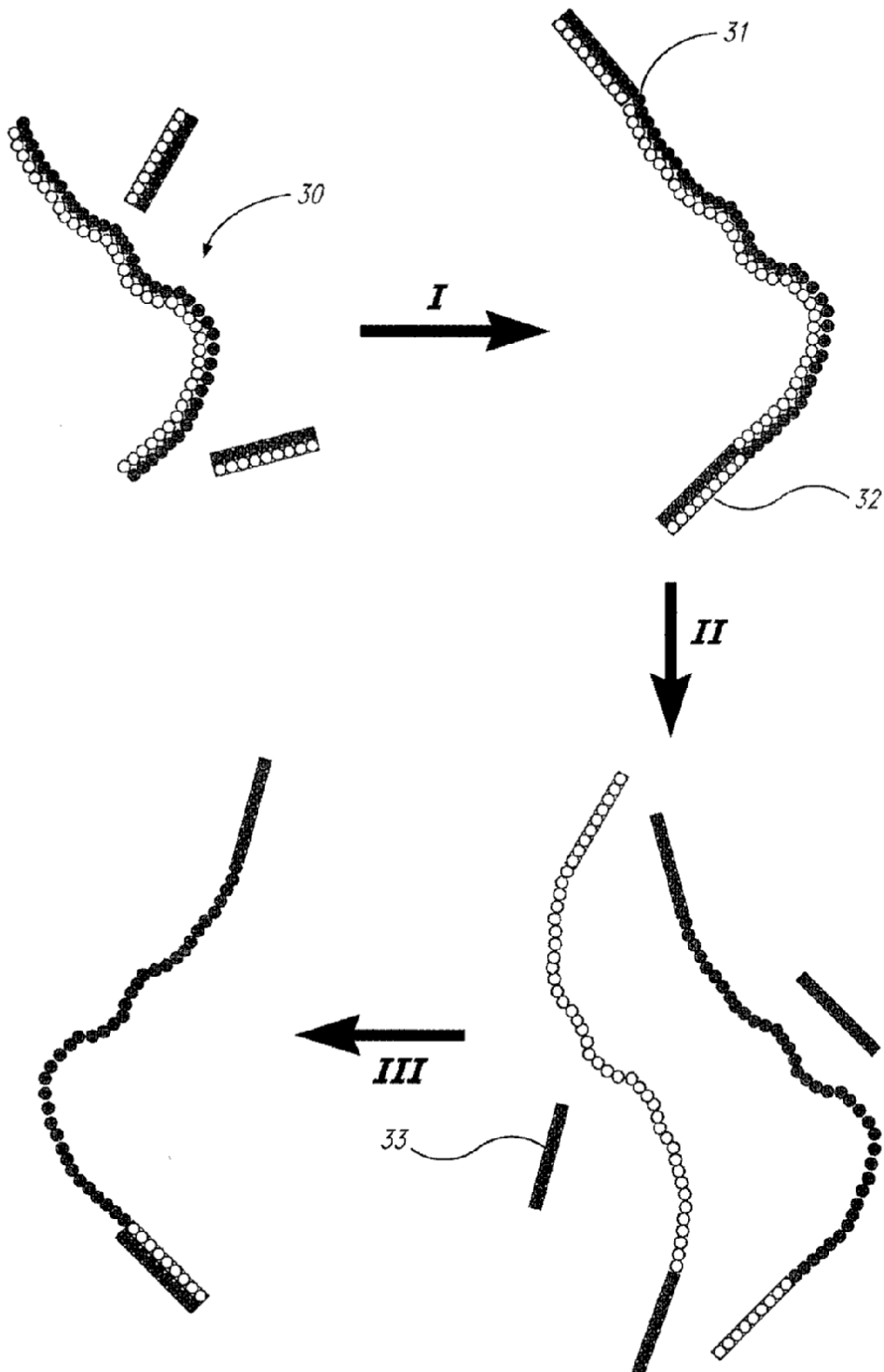
**Fig. 1B**



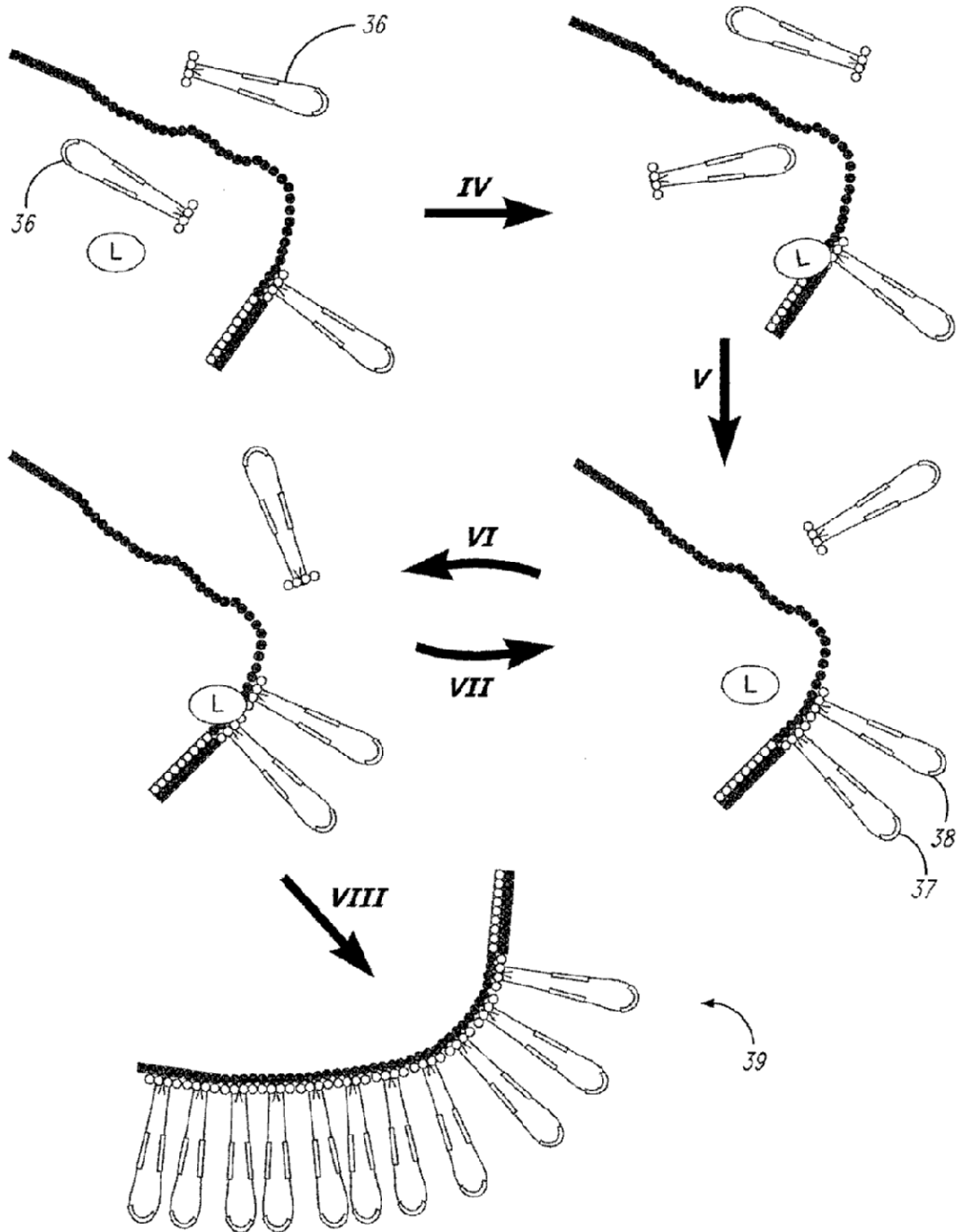




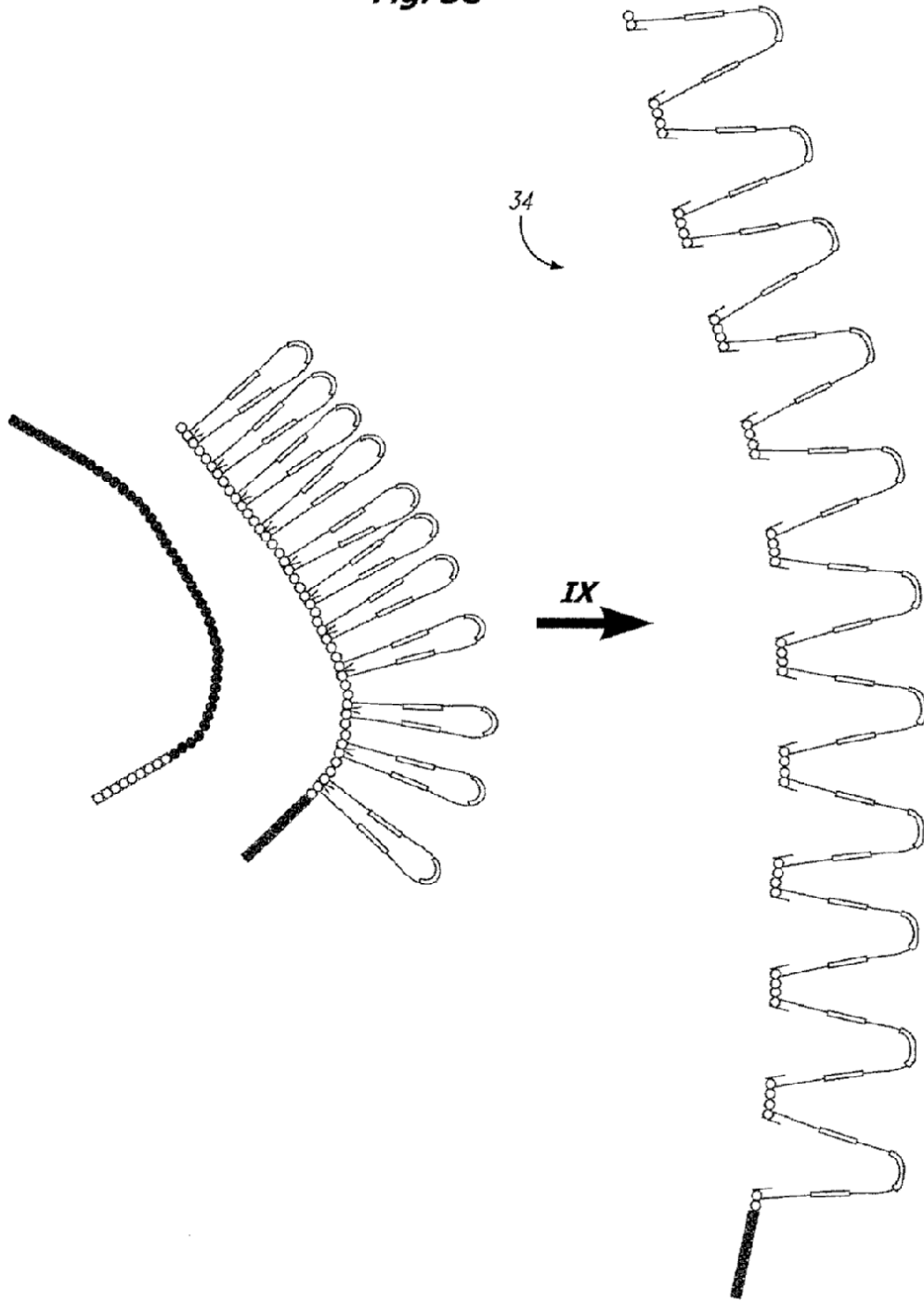
**Fig. 3A**



**Fig. 3B**



**Fig. 3C**



**Fig. 4**

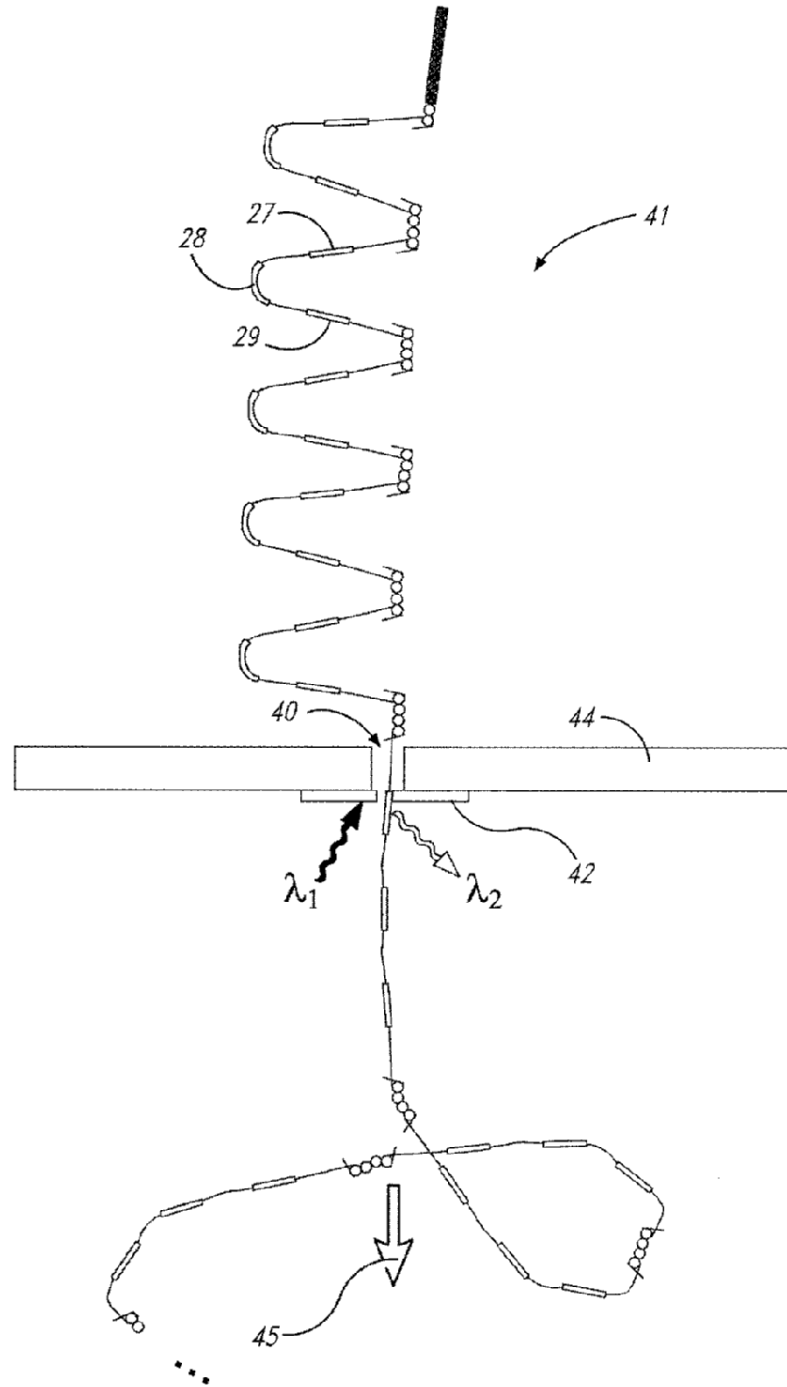
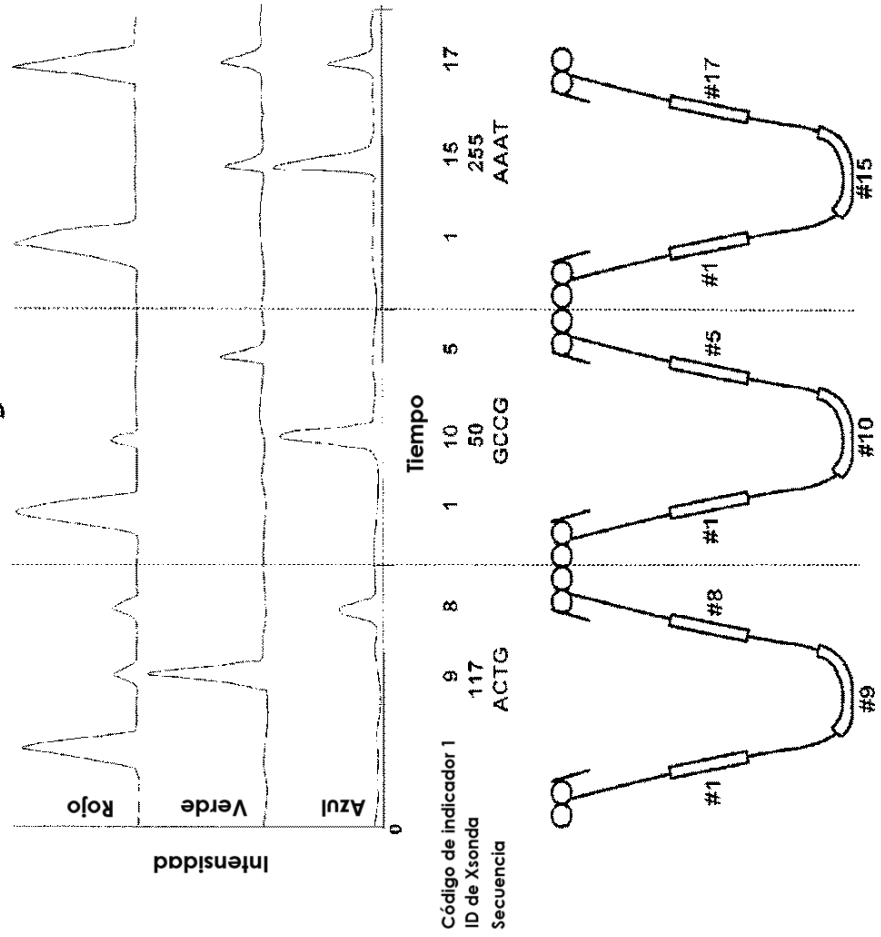
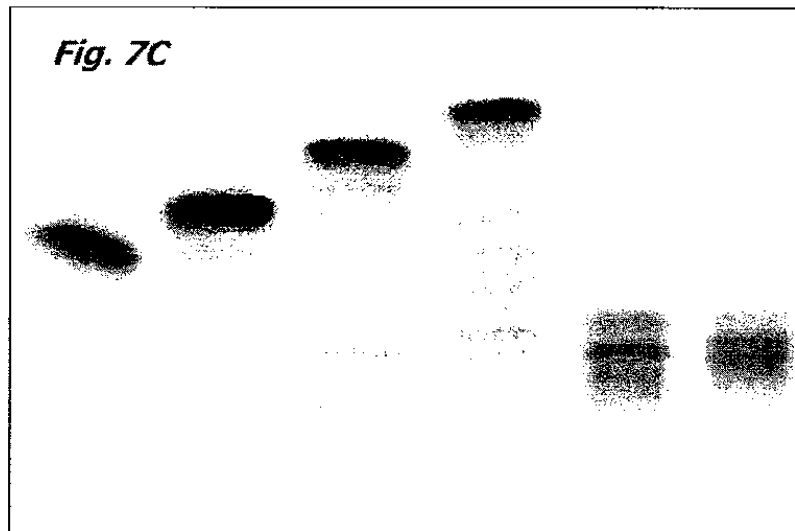
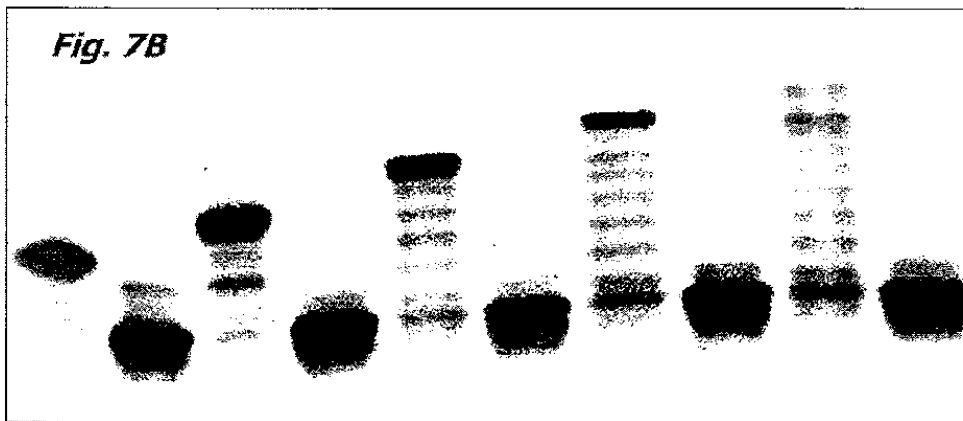
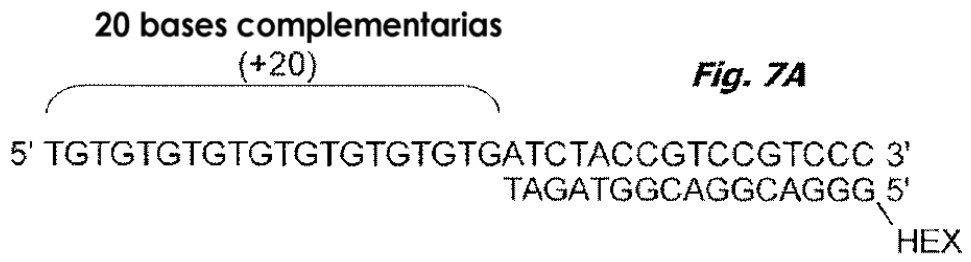


Fig. 5

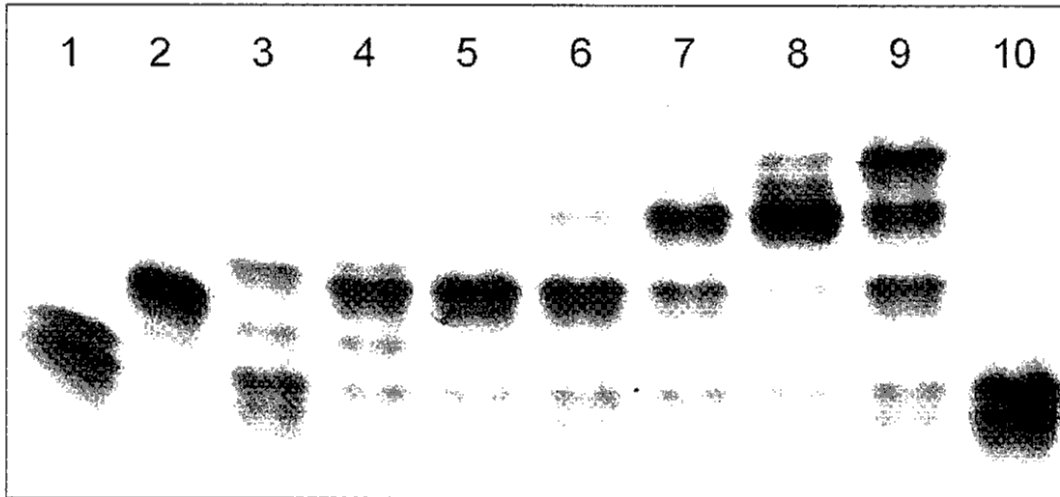


**FIG. 6**

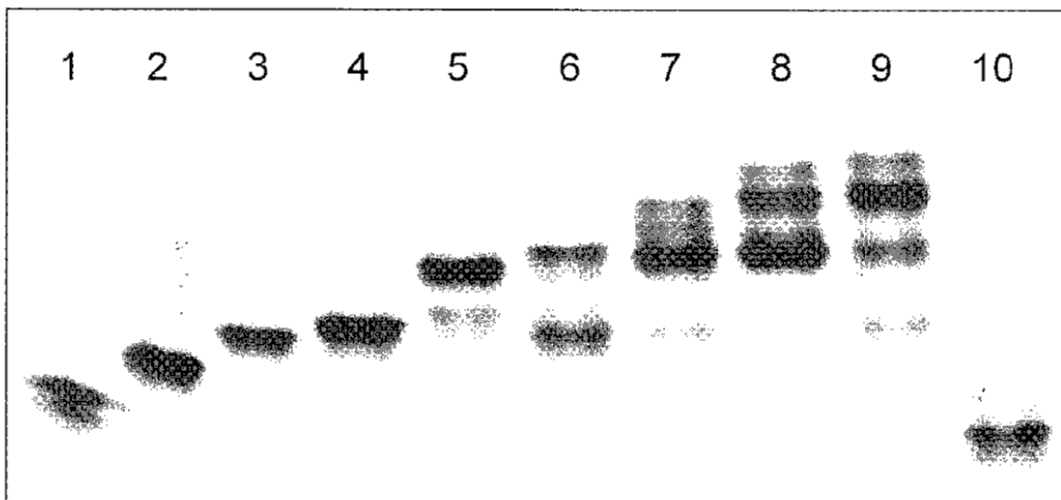
Código de Indicador	Fluorescencia (F)	Código de Indicador	Fluorescencia (F)
1	ROJA	12	ROJA-azul
2	VERDE	13	verde-azul
3	AZUL	14	VERDE-azul
4	roja	15	verde-Azul
5	verde	16	roja-verde-azul
6	azul	17	ROJA-verde-azul
7	roja-verde	18	roja-VERDE-azul
8	roja-azul	19	roja-verde-AZUL
9	roja-VERDE	20	ROJA-VERDE-azul
10	roja-AZUL	21	roja-VERDE-AZUL
11	ROJA-verde	22	ROJA-verde-AZUL



**Fig. 7D**



**Fig. 7E**





**Fig. 8**


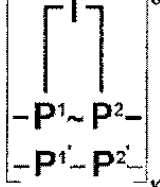
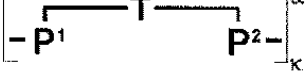
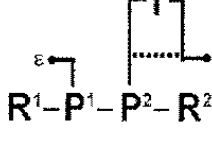
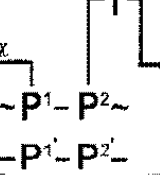
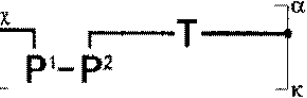
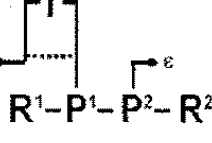
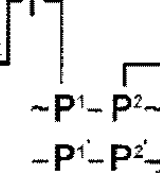
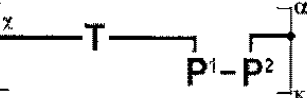
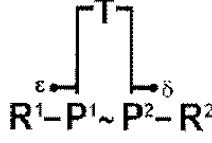
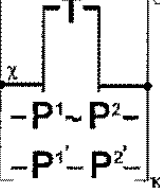
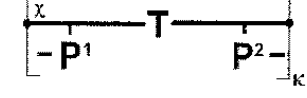
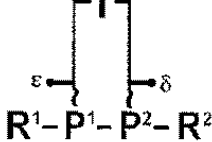
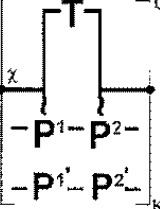
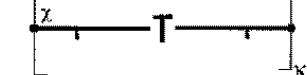
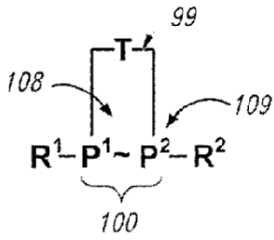
CLASE	CONCENTRACION DE SUSTRATO OLIGOMÉRICO	PRODUCTO INTERMEDIO DE XPANDÓMERO	XPANDÓMERO
I.	$R^1-P^1 \sim P^2-R^2$ 	$\left[ \begin{array}{c} \text{---} \text{T} \text{---} \\   \\ \text{---} P^1 \sim P^2 \text{---} \\   \\ \text{---} P^{1'} \sim P^{2'} \text{---} \end{array} \right]_{\alpha}^{\alpha}$ 	$\left[ \text{---} P^1 \text{---} T \text{---} P^2 \text{---} \right]_{\alpha}^{\alpha}$ 
II.	$R^1-P^1 \sim P^2-R^2$ 	$\left[ \begin{array}{c} \text{---} \text{T} \text{---} \\   \\ \text{---} P^1 \sim P^2 \text{---} \\   \\ \text{---} P^{1'} \sim P^{2'} \text{---} \end{array} \right]_{\alpha}^{\alpha}$ 	$\left[ \text{---} P^1 \text{---} T \text{---} P^2 \text{---} \right]_{\alpha}^{\alpha}$ 
III.	$R^1-P^1 \sim P^2-R^2$ 	$\left[ \begin{array}{c} \text{---} \text{T} \text{---} \\   \\ \text{---} P^1 \sim P^2 \text{---} \\   \\ \text{---} P^{1'} \sim P^{2'} \text{---} \end{array} \right]_{\alpha}^{\alpha}$ 	$\left[ \text{---} T \text{---} P^1 \text{---} P^2 \text{---} \right]_{\alpha}^{\alpha}$ 
IV.	$R^1-P^1 \sim P^2-R^2$ 	$\left[ \begin{array}{c} \text{---} \text{T} \text{---} \\   \\ \text{---} P^1 \sim P^2 \text{---} \\   \\ \text{---} P^{1'} \sim P^{2'} \text{---} \end{array} \right]_{\alpha}^{\alpha}$ 	$\left[ \text{---} P^1 \text{---} T \text{---} P^2 \text{---} \right]_{\alpha}^{\alpha}$ 
V.	$R^1-P^1 \sim P^2-R^2$ 	$\left[ \begin{array}{c} \text{---} \text{T} \text{---} \\   \\ \text{---} P^1 \sim P^2 \text{---} \\   \\ \text{---} P^{1'} \sim P^{2'} \text{---} \end{array} \right]_{\alpha}^{-\alpha}$ 	$\left[ \text{---} T \text{---} \right]_{\alpha}^{\alpha}$ 

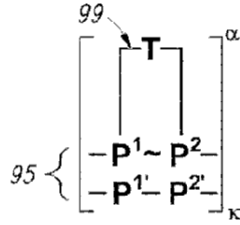
Fig. 9

CLASE	CONCENTRACION DE SUSTRATO OLIGOMÉRICO	PRODUCTO INTERMEDIO DE XPANDÓMERO	XPANDÓMERO
VI.	 $R^1-N-R^2$		
VII.	 $R^1-N-R^2$		
VIII.	 $R^1-N-R^2$		
IX.	 $R^1-N-R^2$		
X.	 $R^1-N-R^2$		

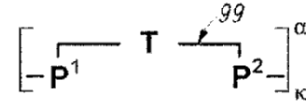
**Fig. 10A**



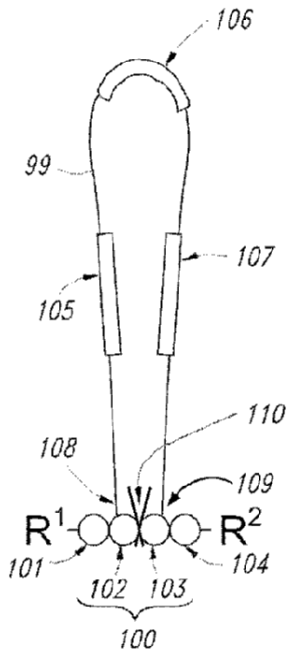
**Fig. 10B**



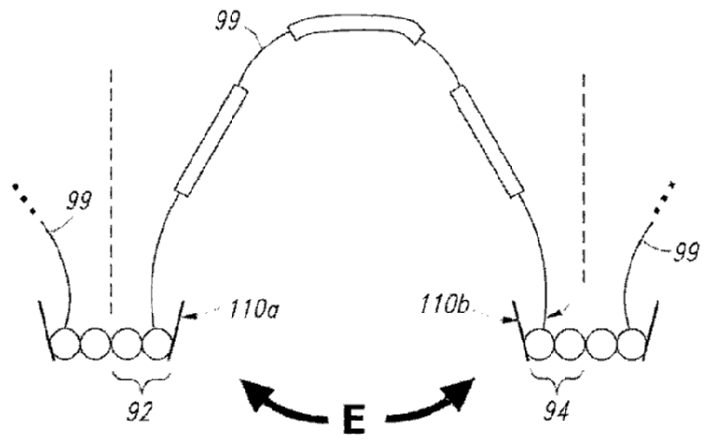
**Fig. 10C**



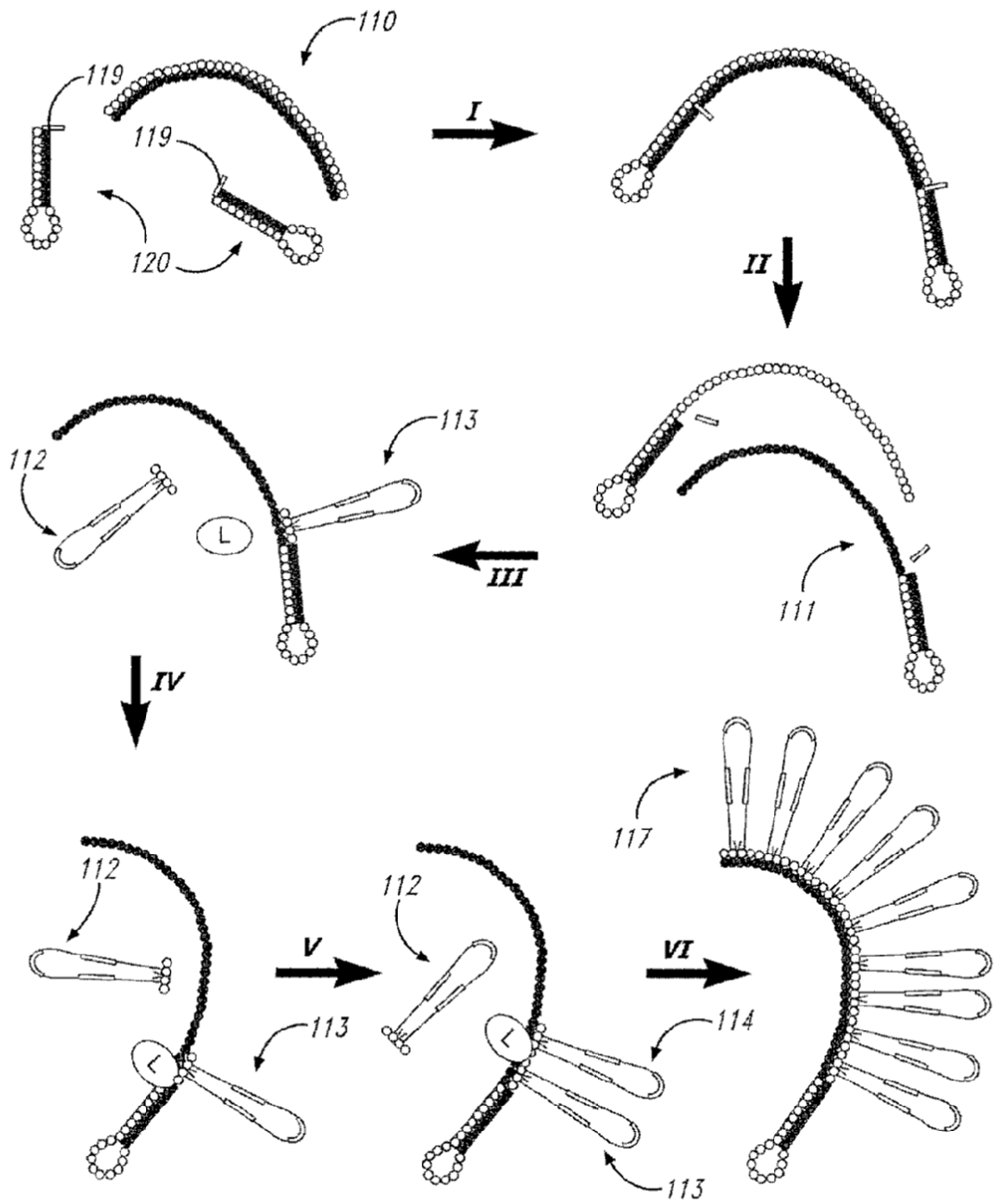
**Fig. 10D**



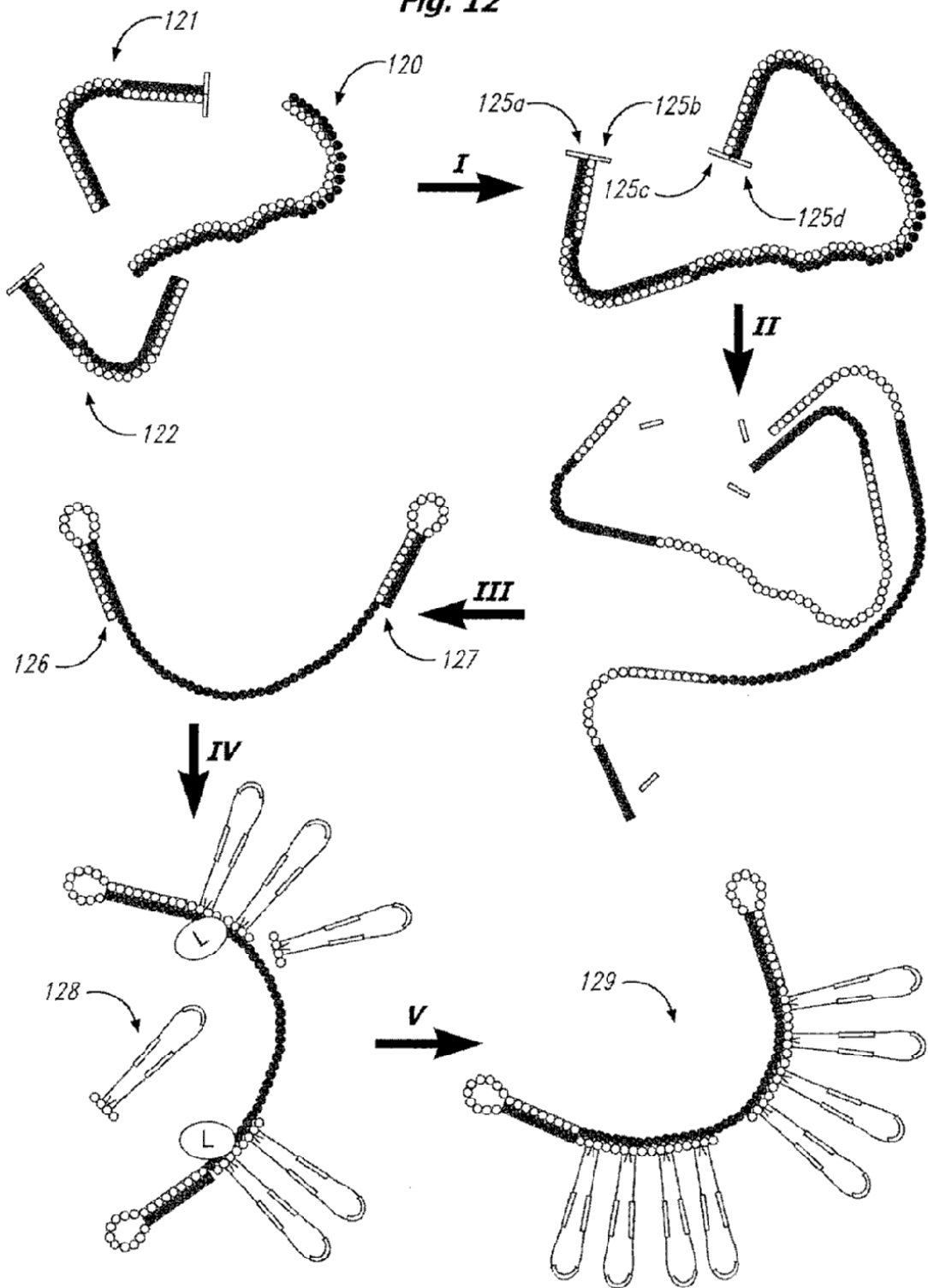
**Fig. 10E**



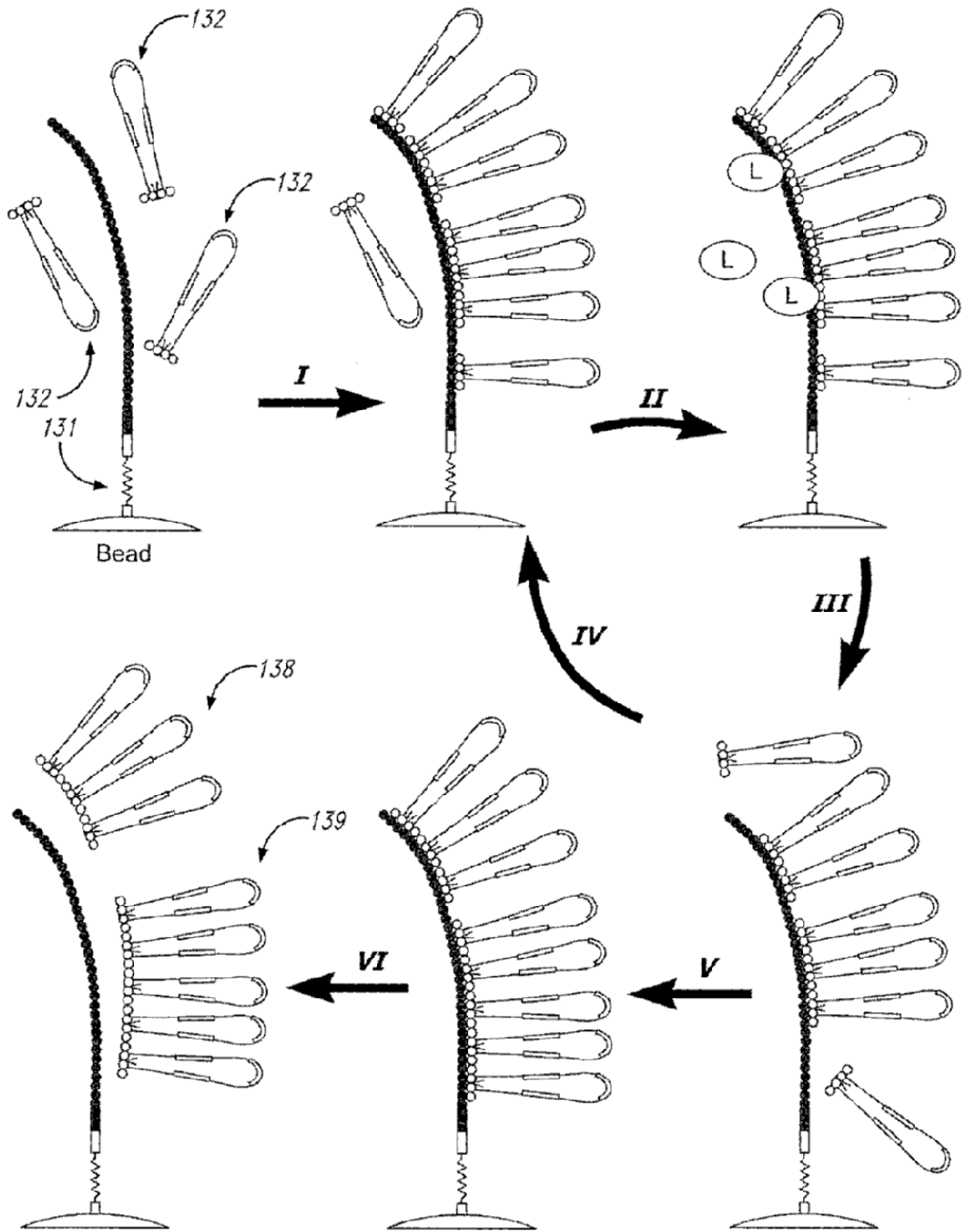
**Fig. 11**



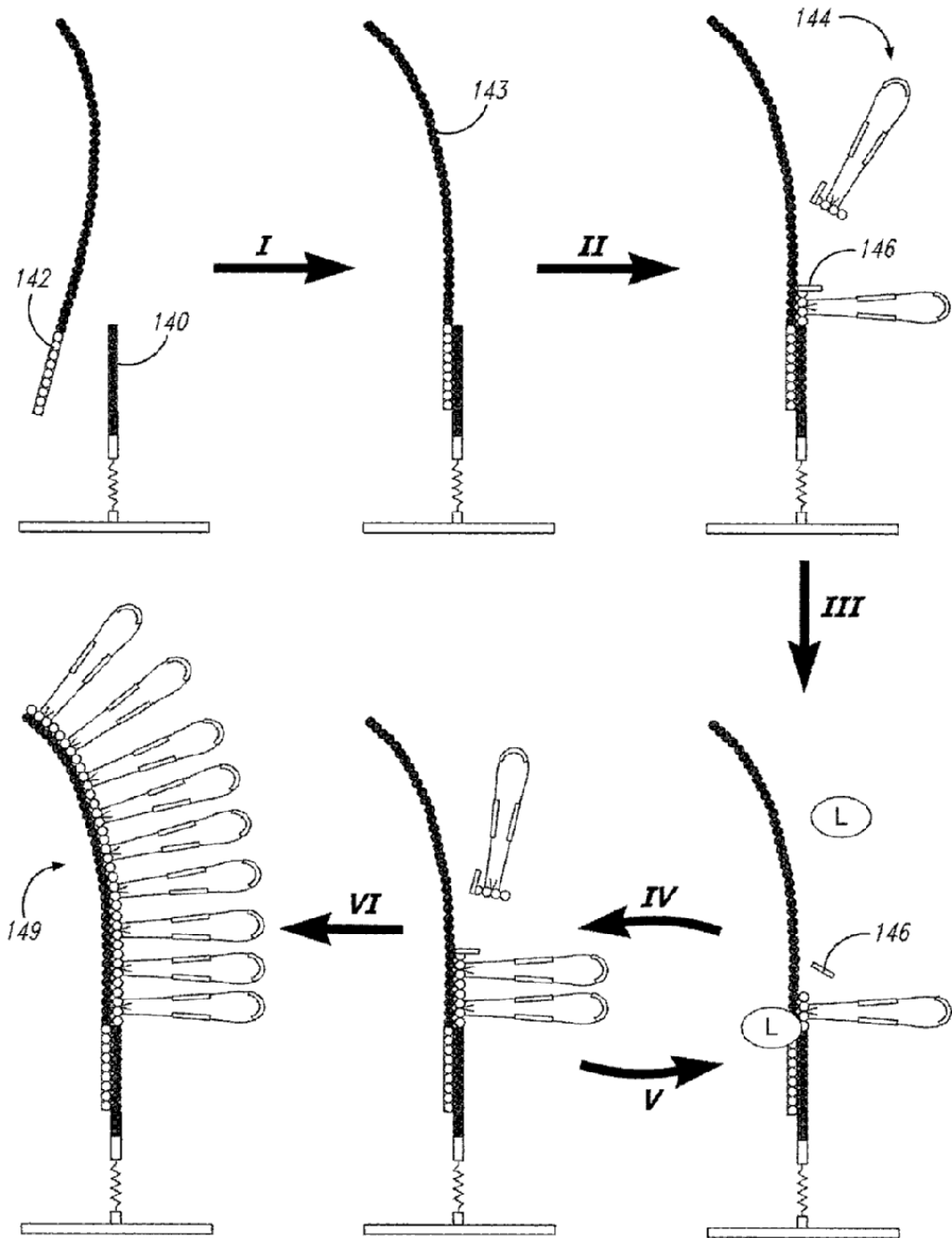
**Fig. 12**



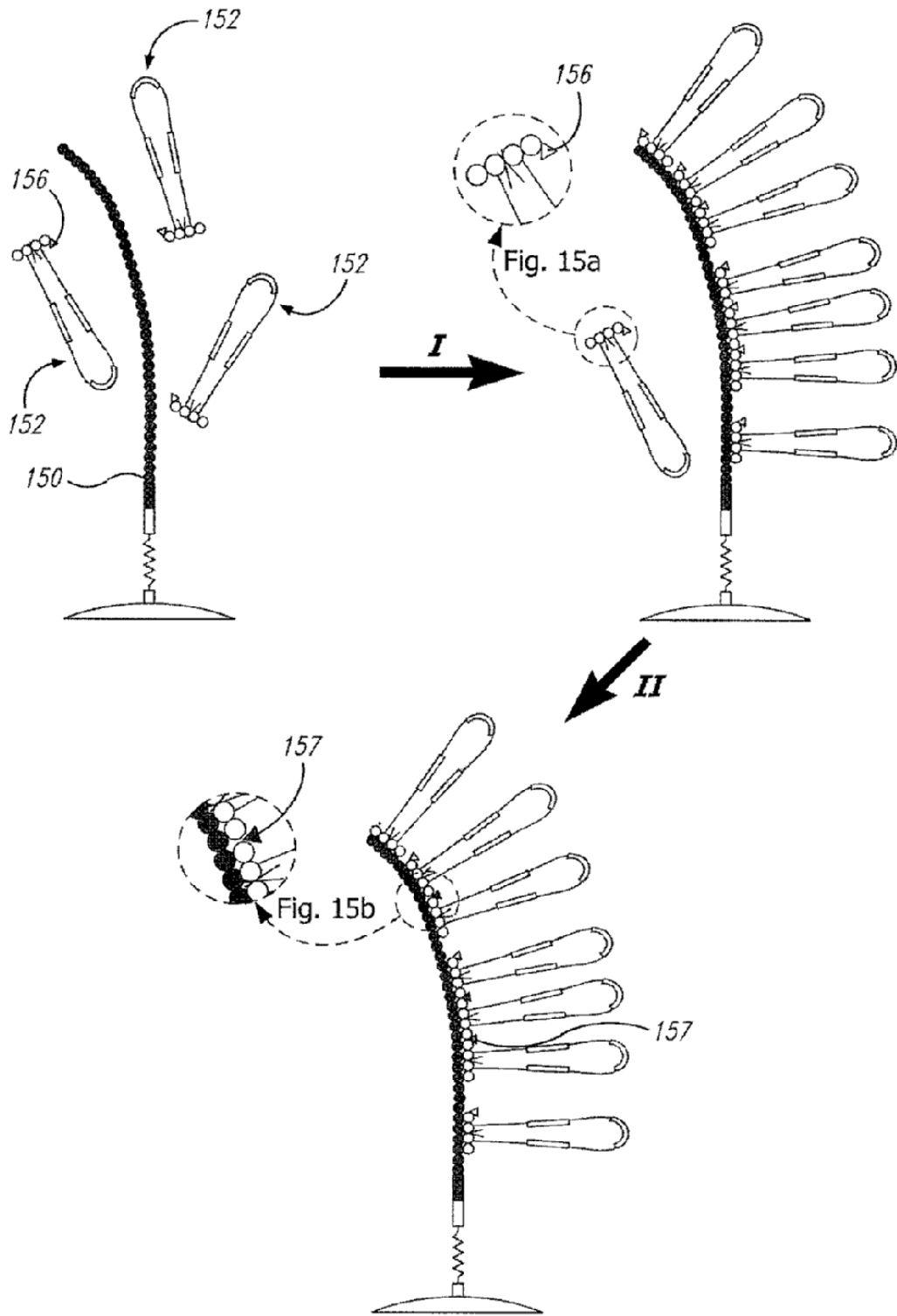
**Fig. 13**



**Fig. 14**

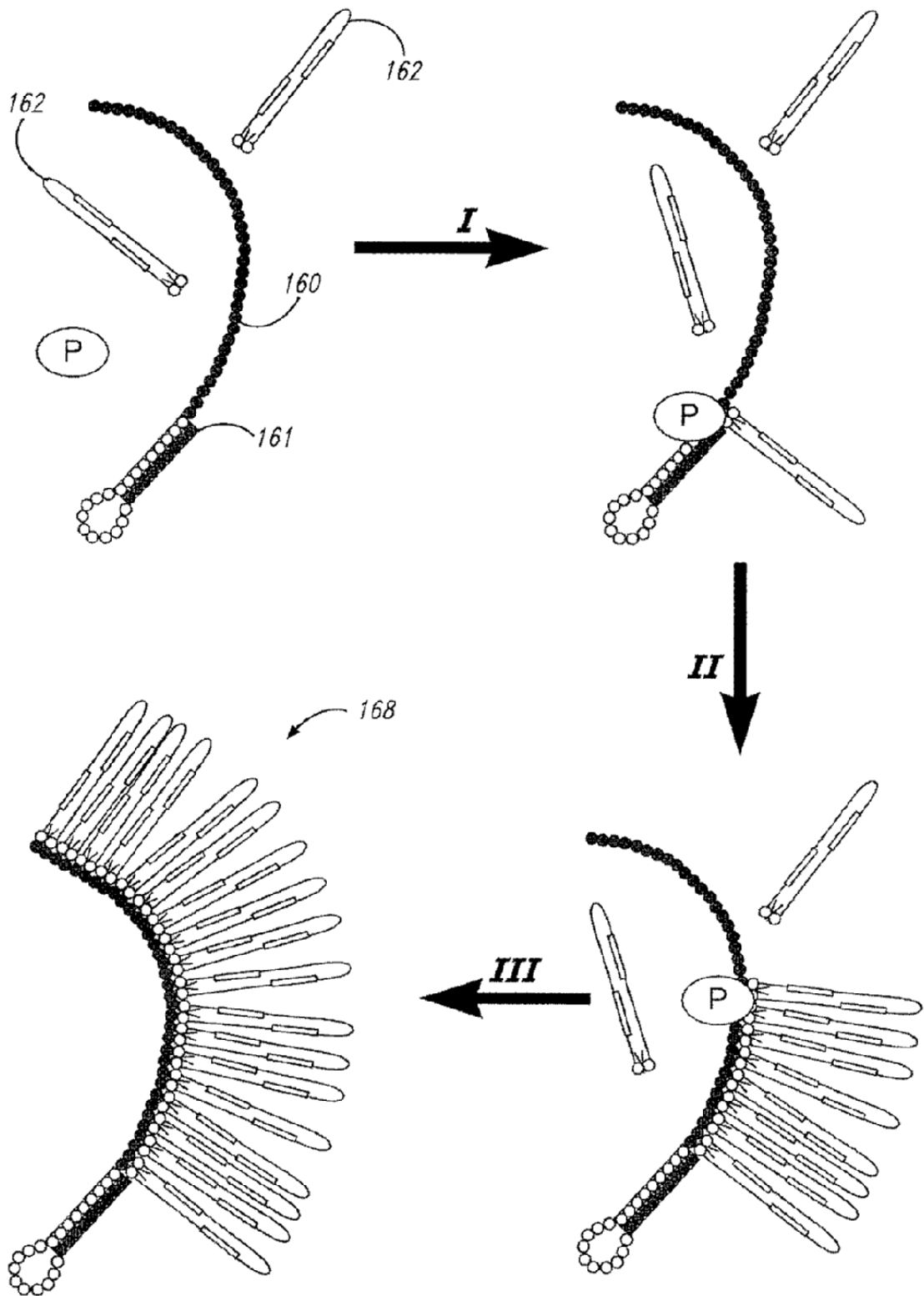


**Fig. 15**





**Fig. 16**



**Fig. 17**

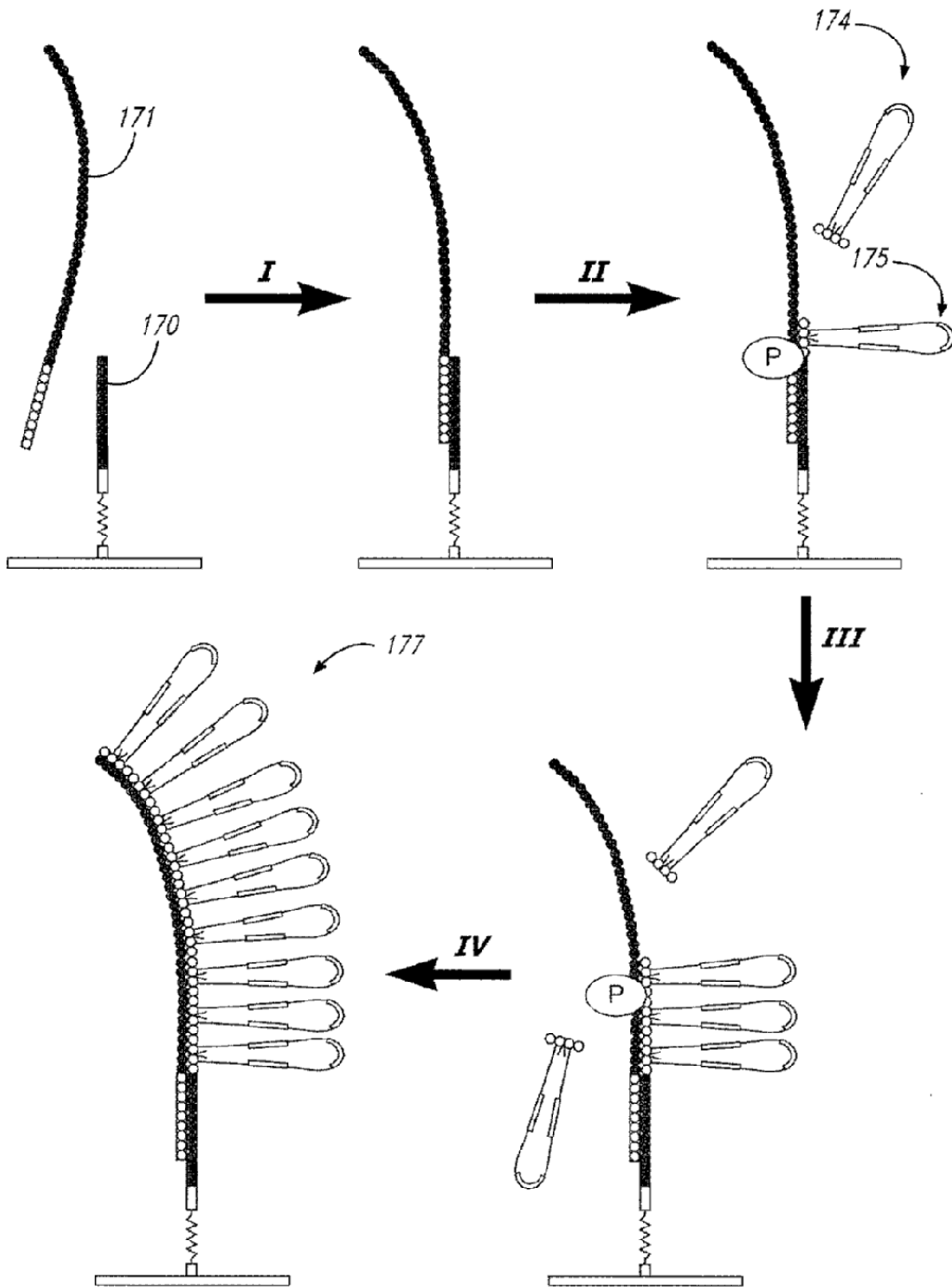




FIG. 19A

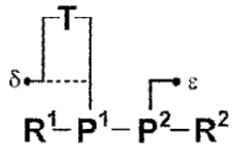


FIG. 19B

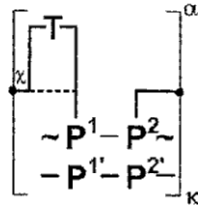


FIG. 19C

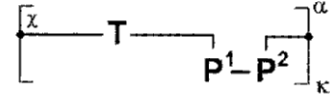


FIG. 19D

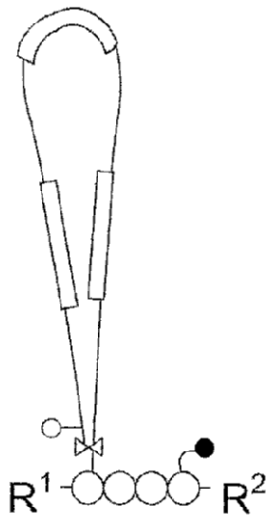
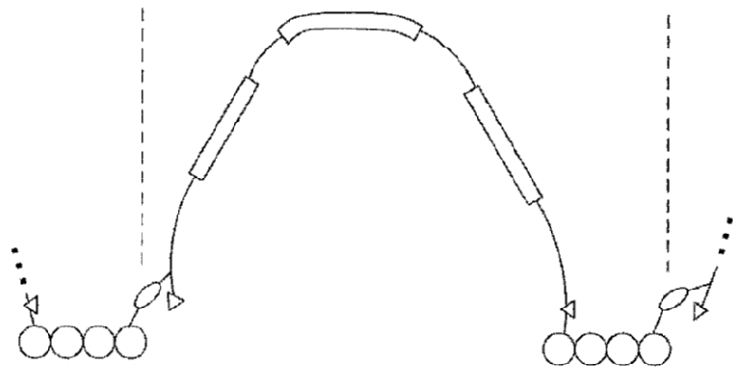
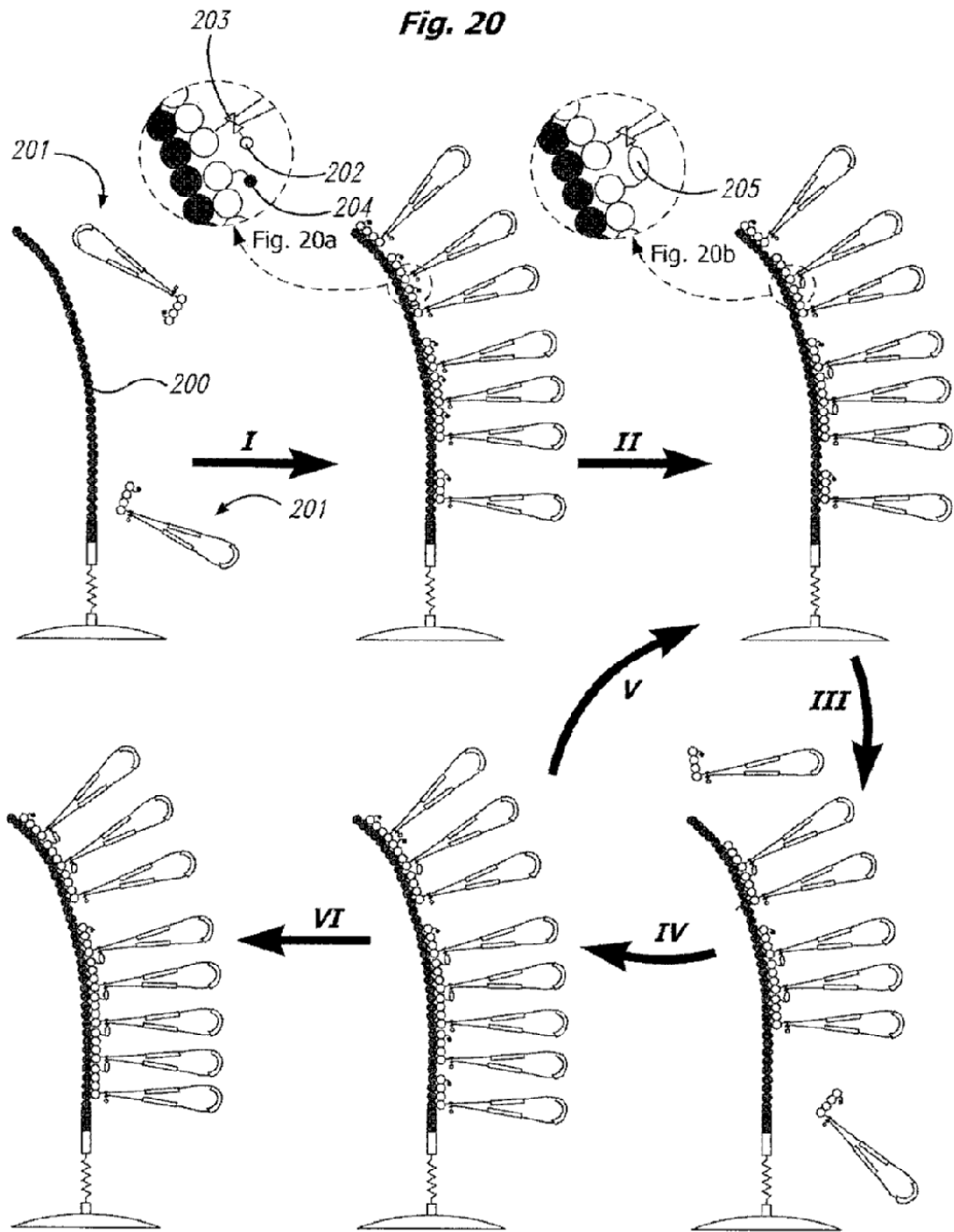
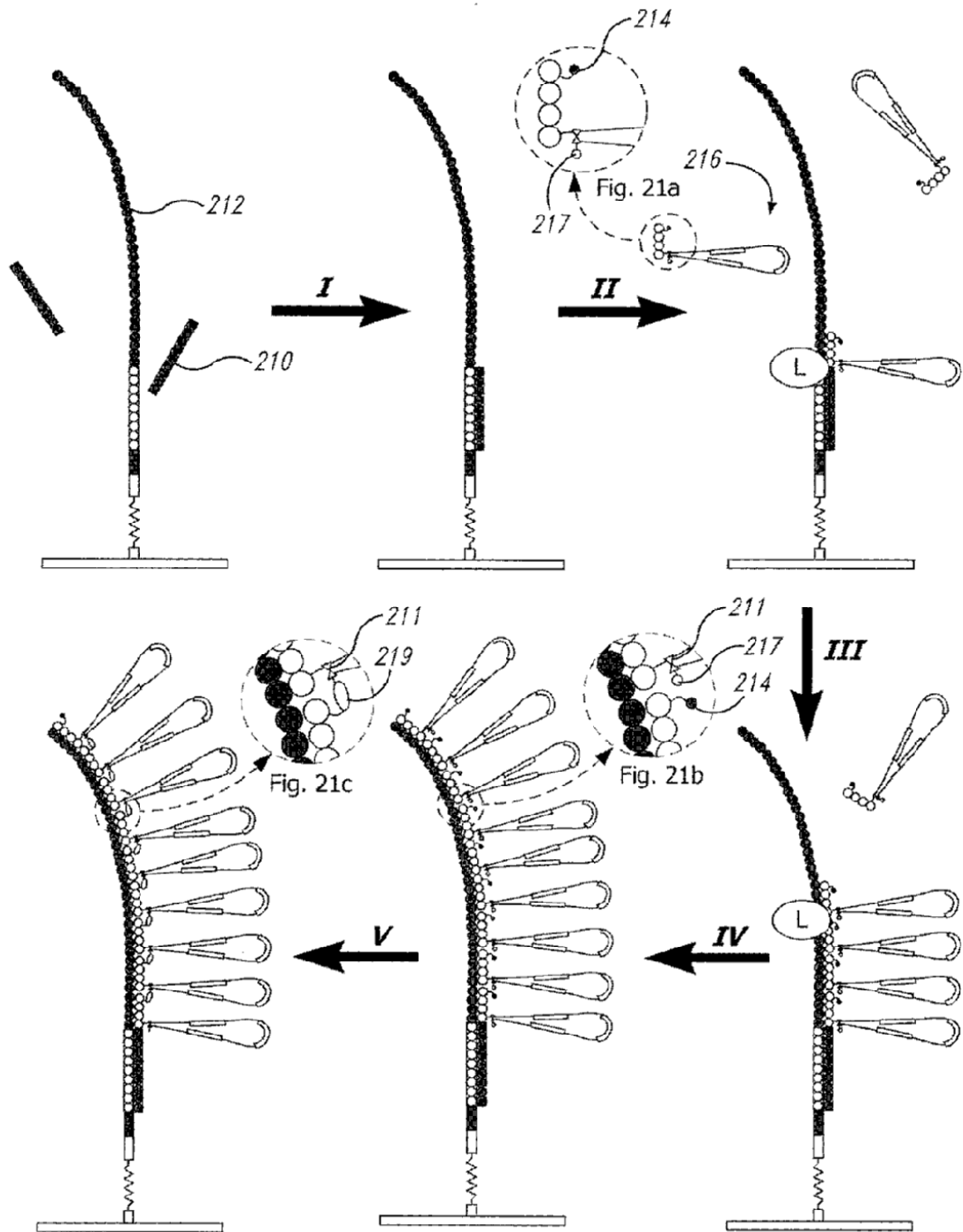


FIG. 19E

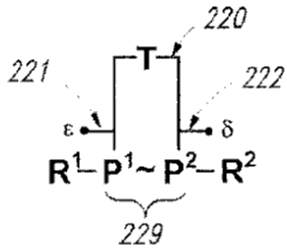




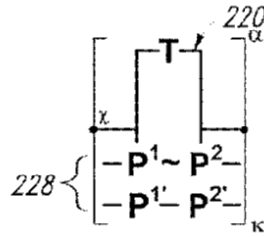
**Fig. 21**



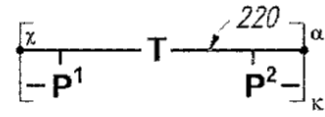
**Fig. 22A**



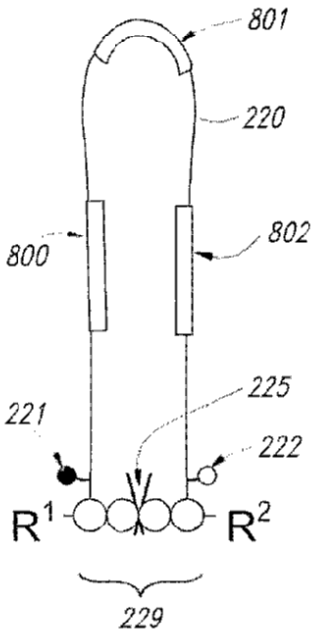
**Fig. 22B**



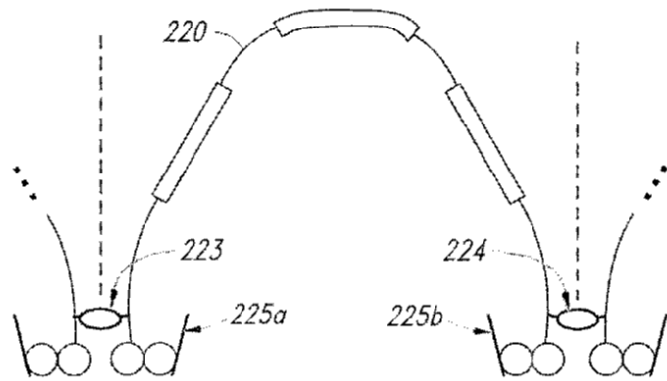
**Fig. 22C**



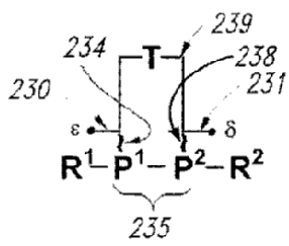
**Fig. 22D**



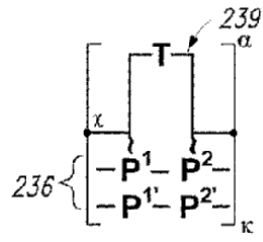
**Fig. 22E**



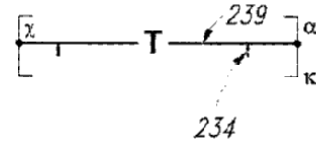
**Fig. 23A**



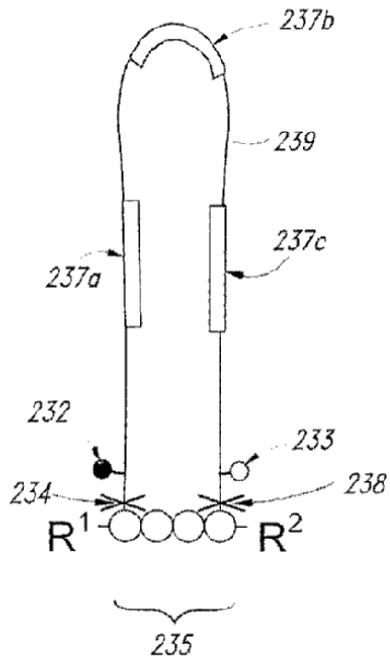
**Fig. 23B**



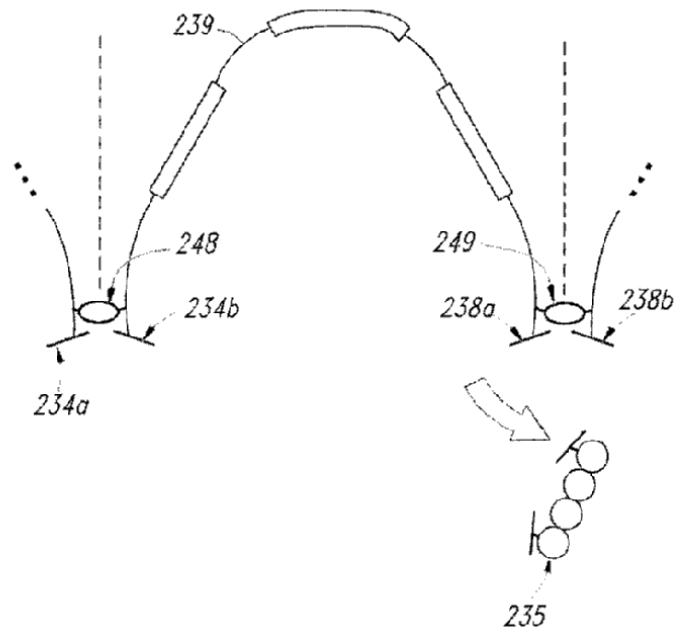
**Fig. 23C**



**Fig. 23D**



**Fig. 23E**





**Fig. 24**

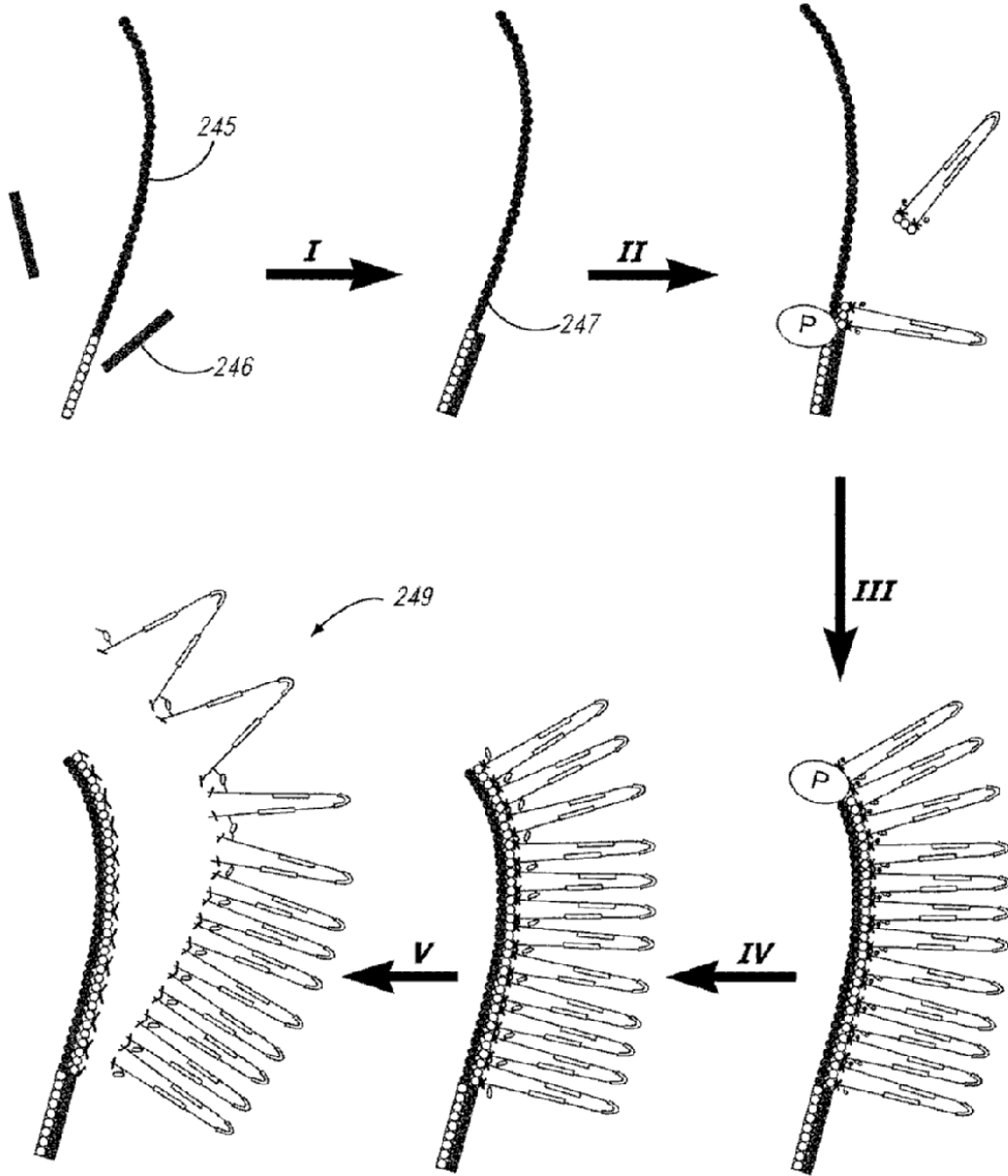
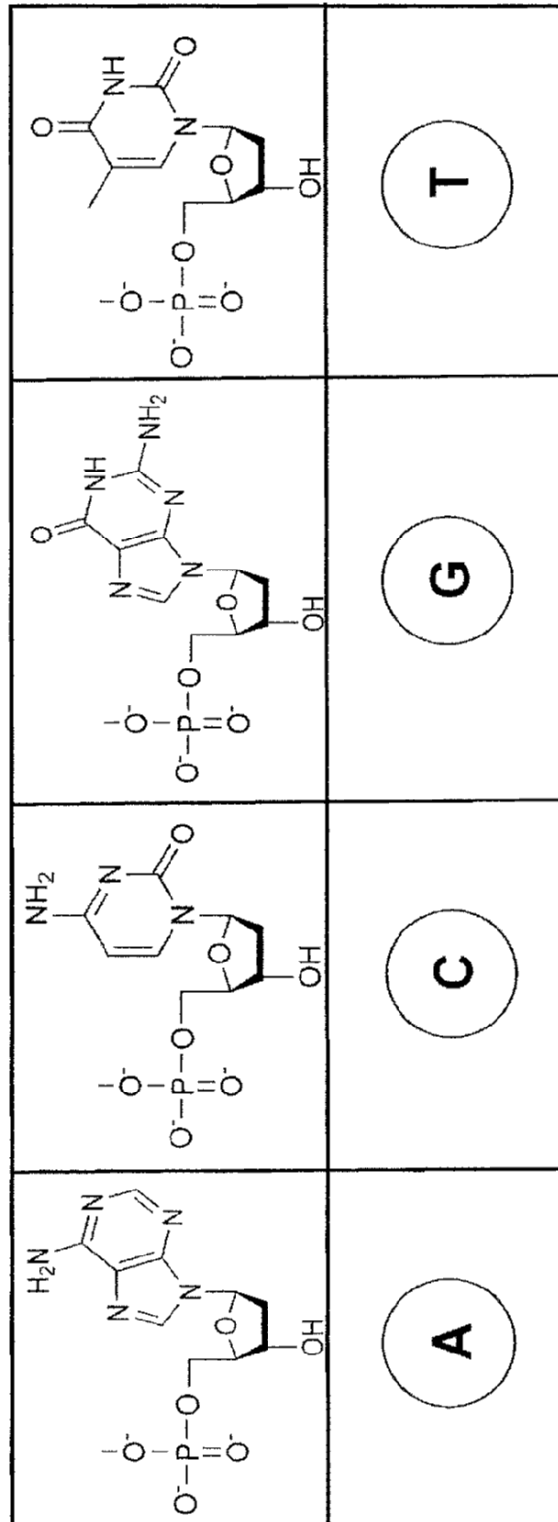
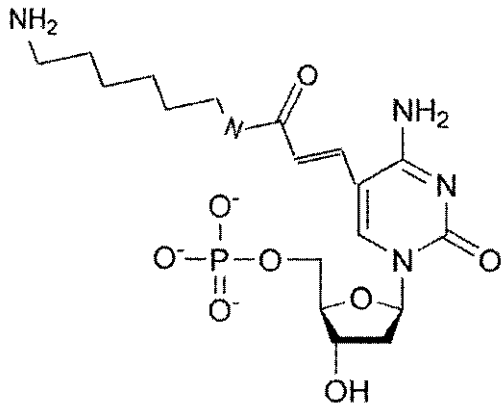


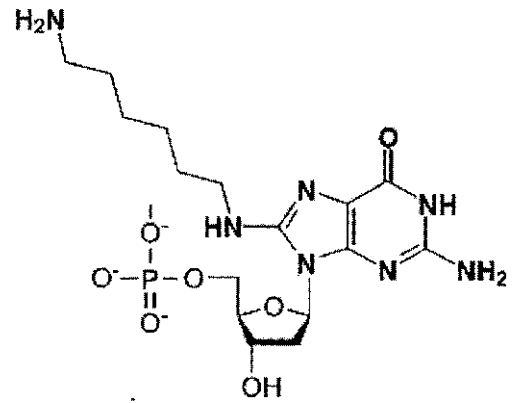
Fig. 25



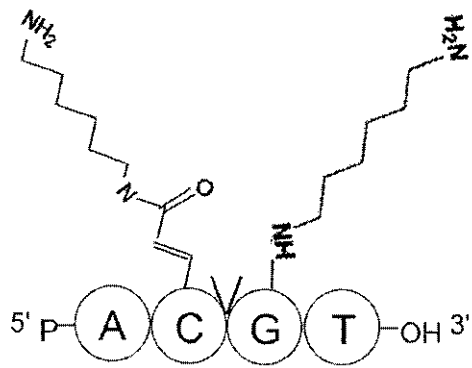
**Fig. 26A**



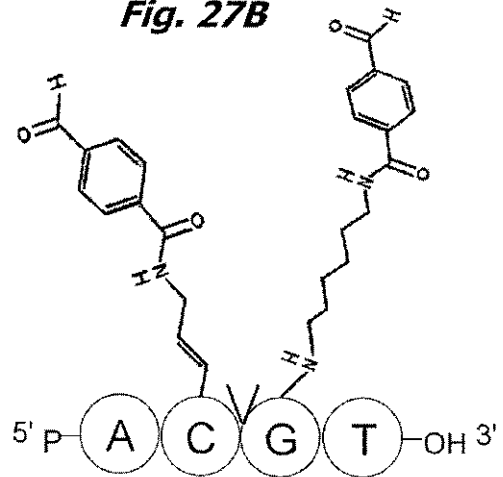
**Fig. 26B**



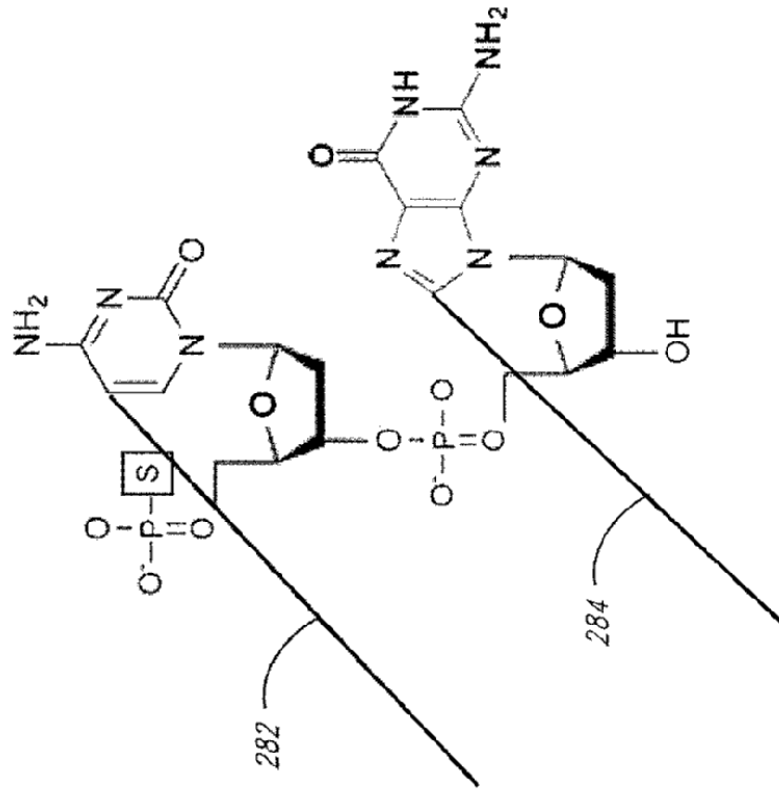
**Fig. 27A**



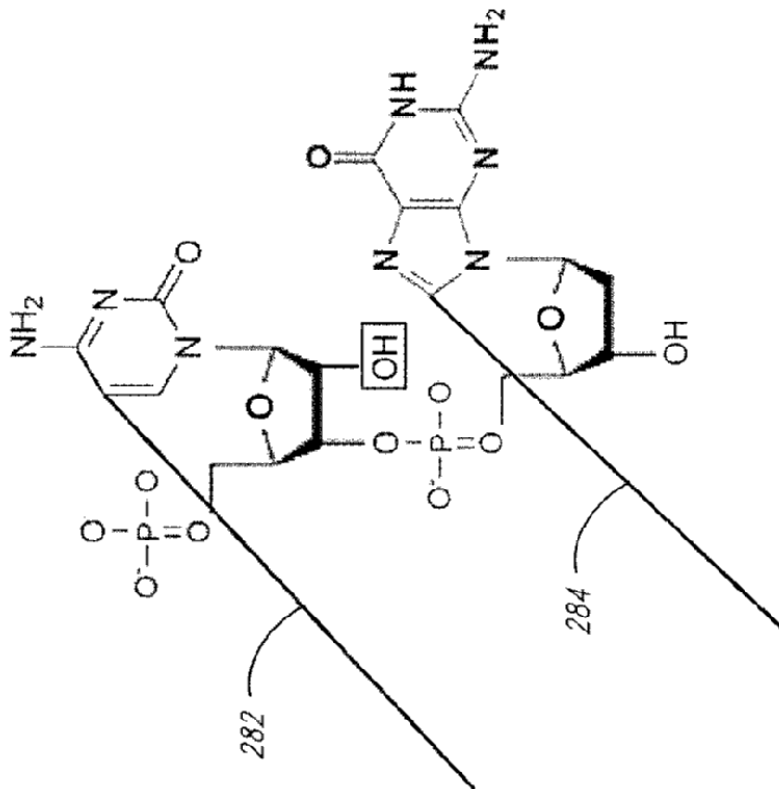
**Fig. 27B**



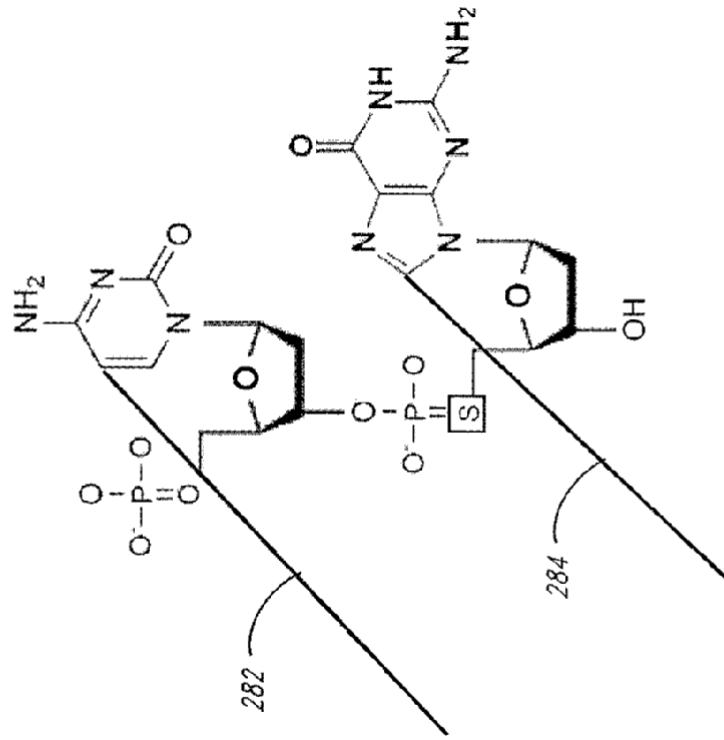
**Fig. 28B**



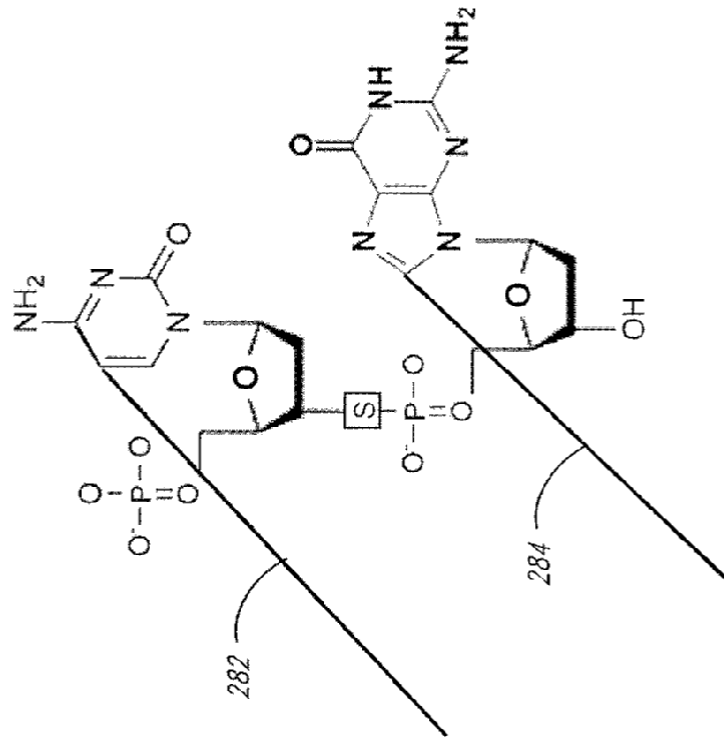
**Fig. 28A**



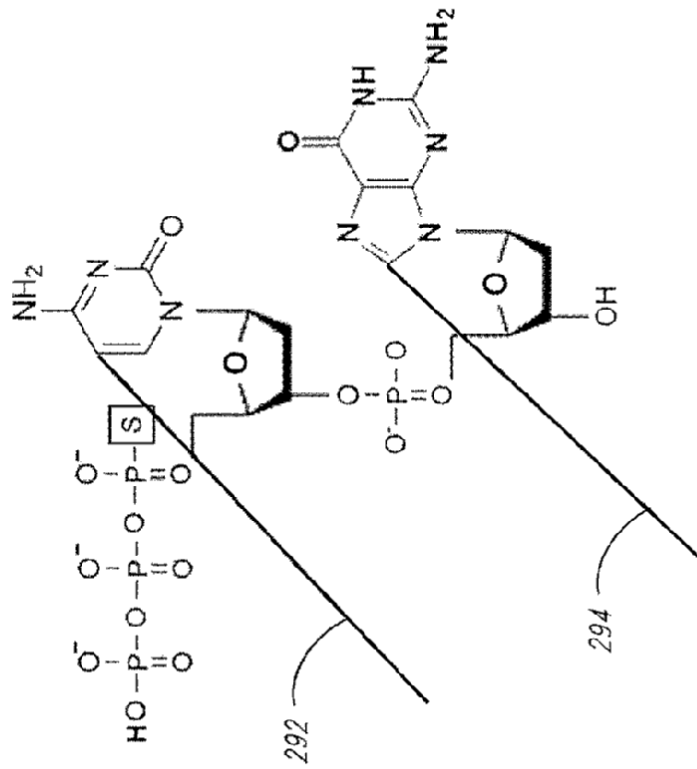
**Fig. 28D**



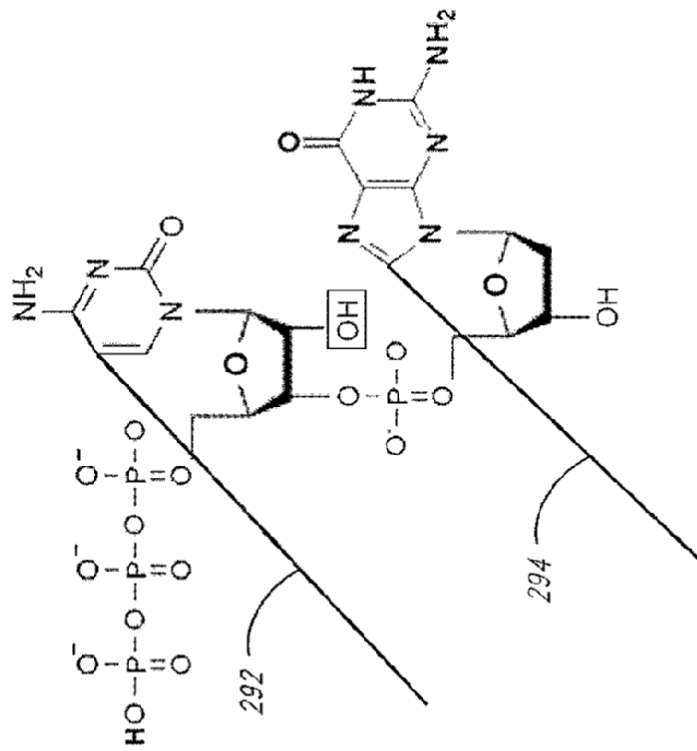
**Fig. 28C**



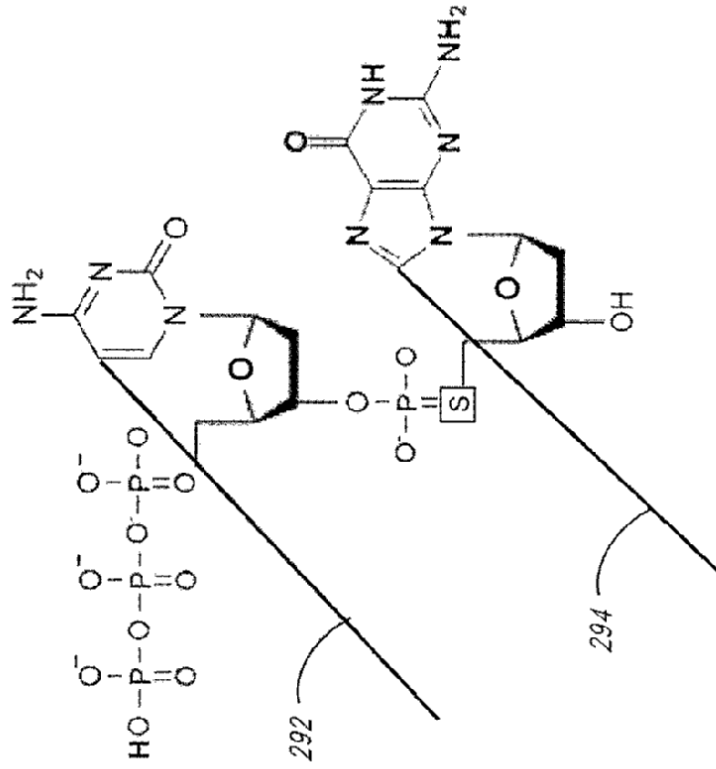
**Fig. 29B**



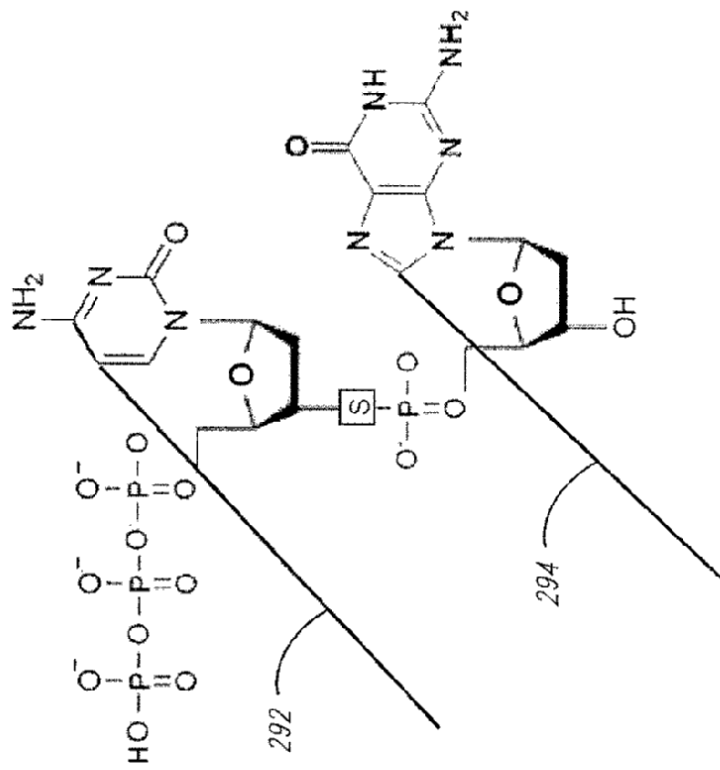
**Fig. 29A**



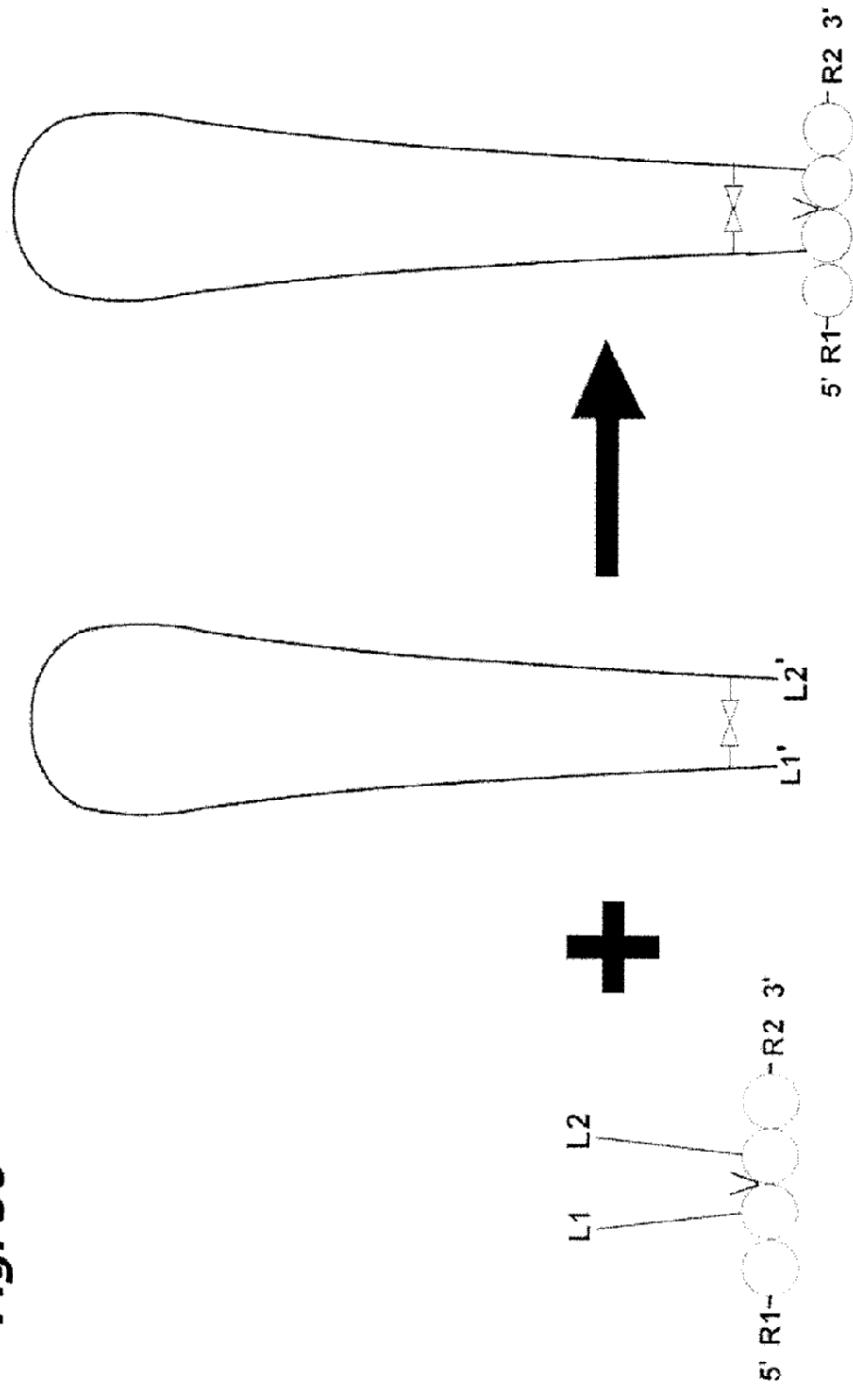
**Fig. 29D**



**Fig. 29C**

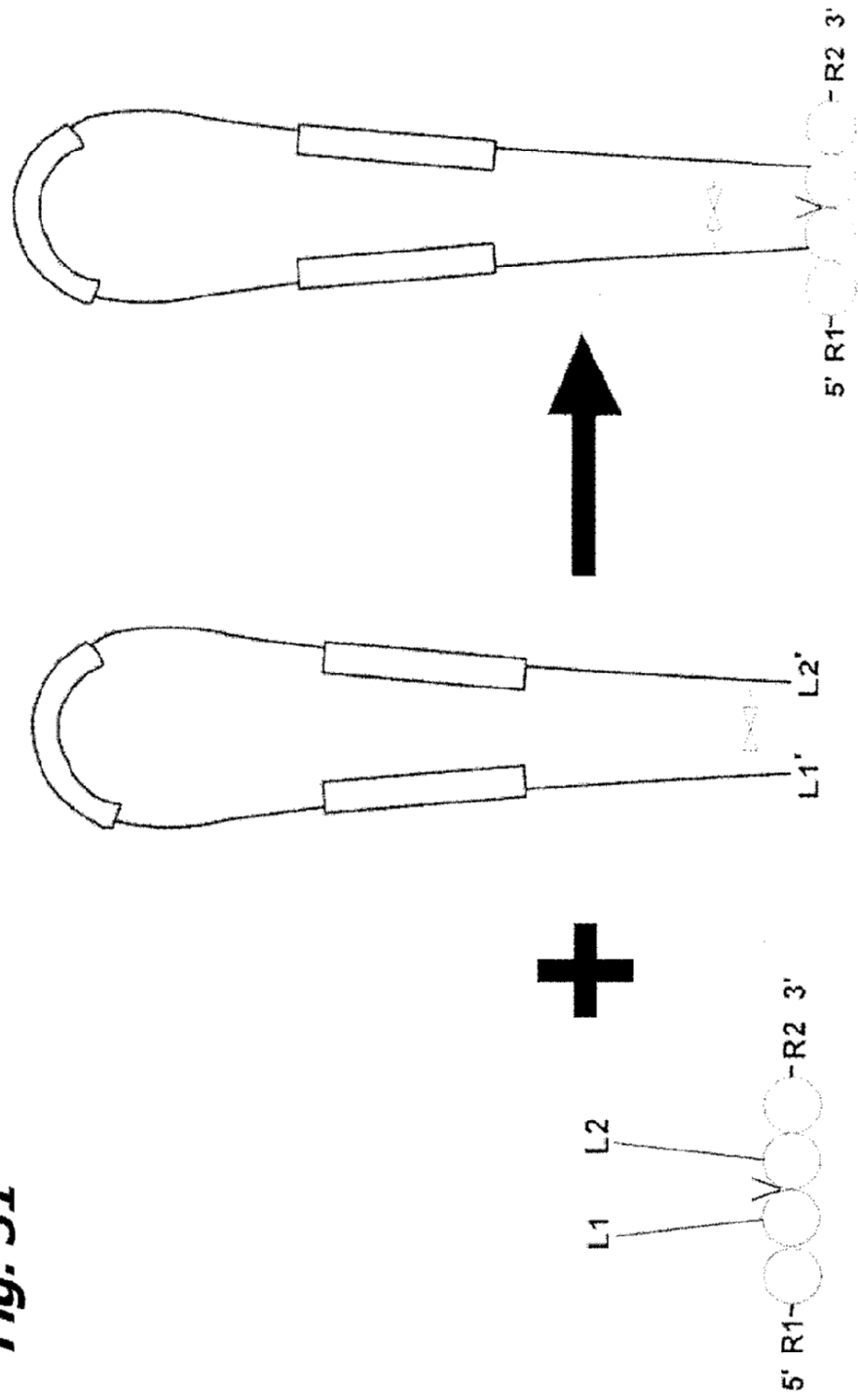


**Fig. 30**

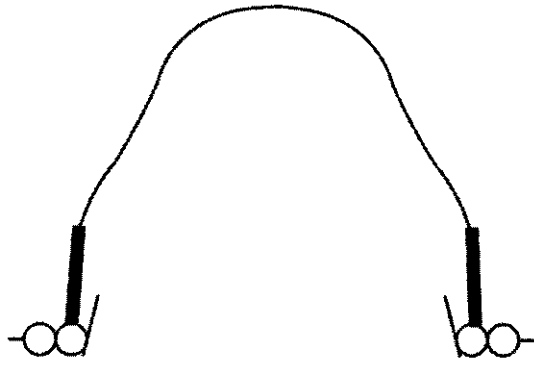




**Fig. 31**



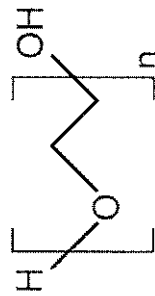
**Fig. 32C**



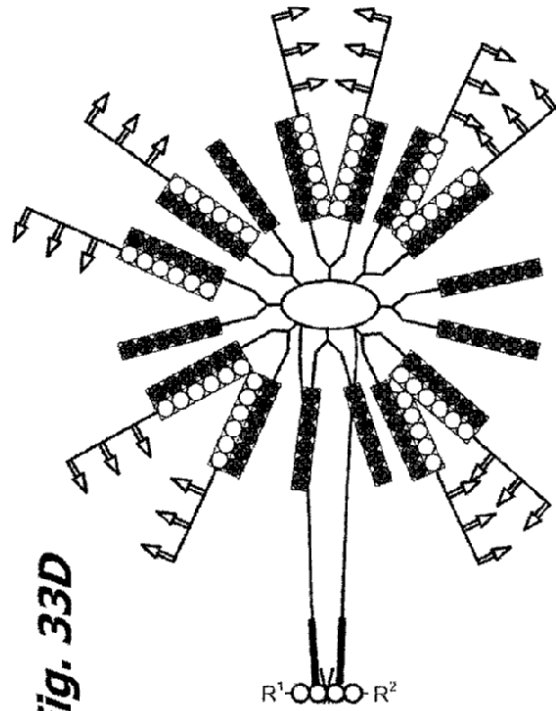
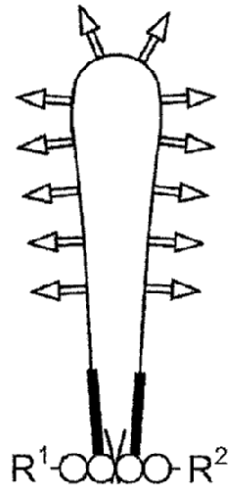
**Fig. 32B**



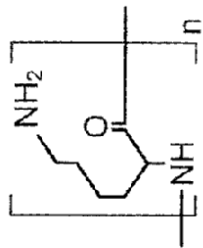
**Fig. 32A**



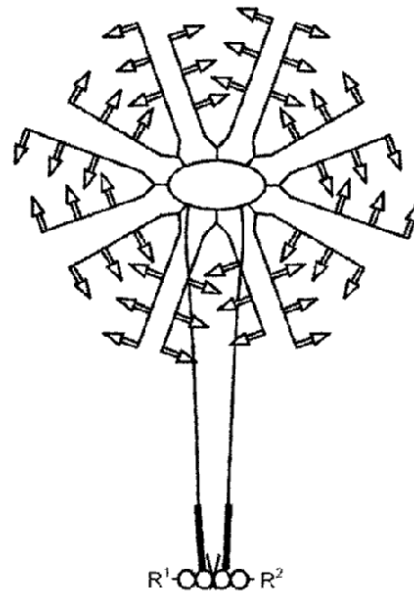
**Fig. 33B**

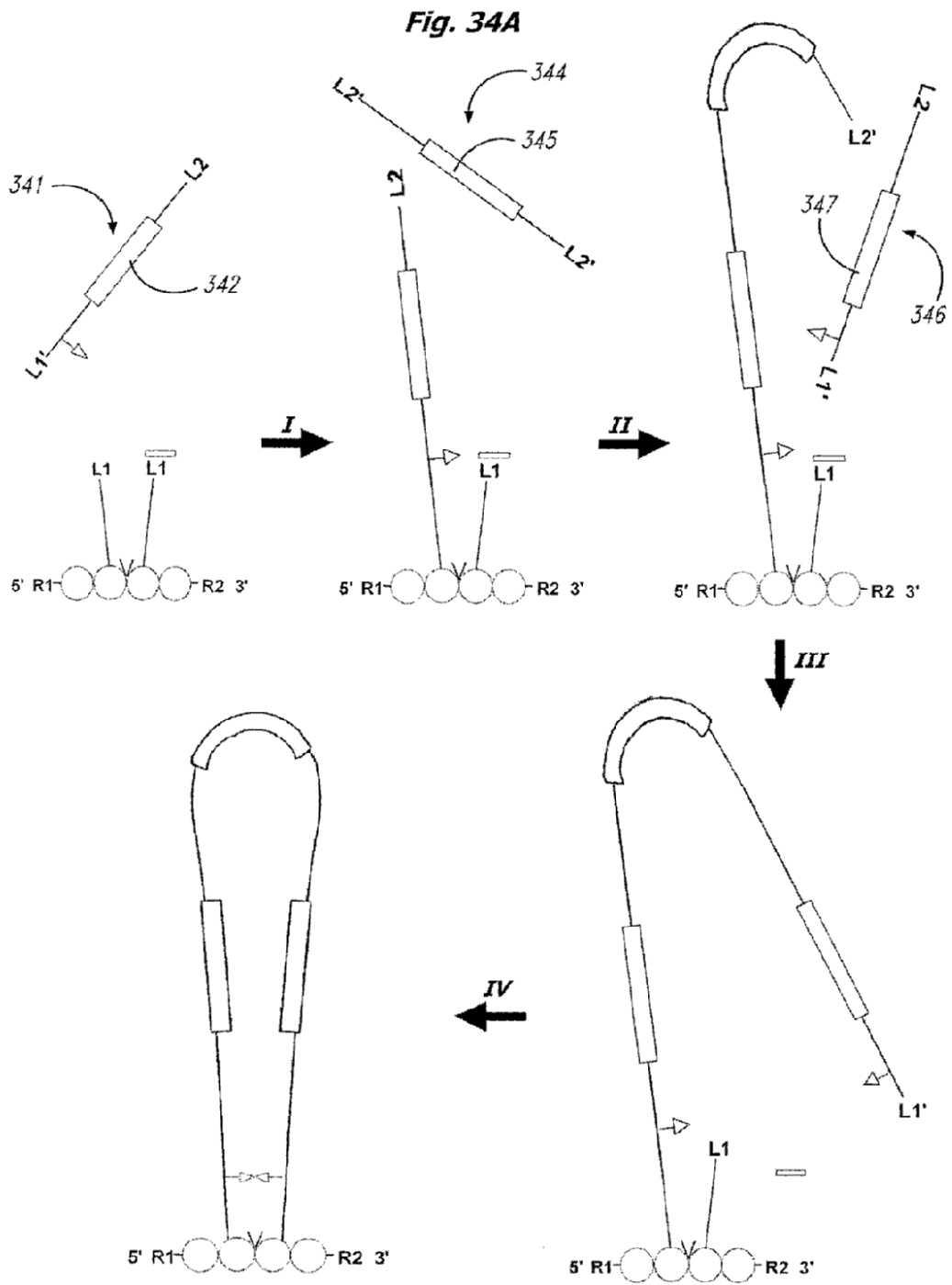


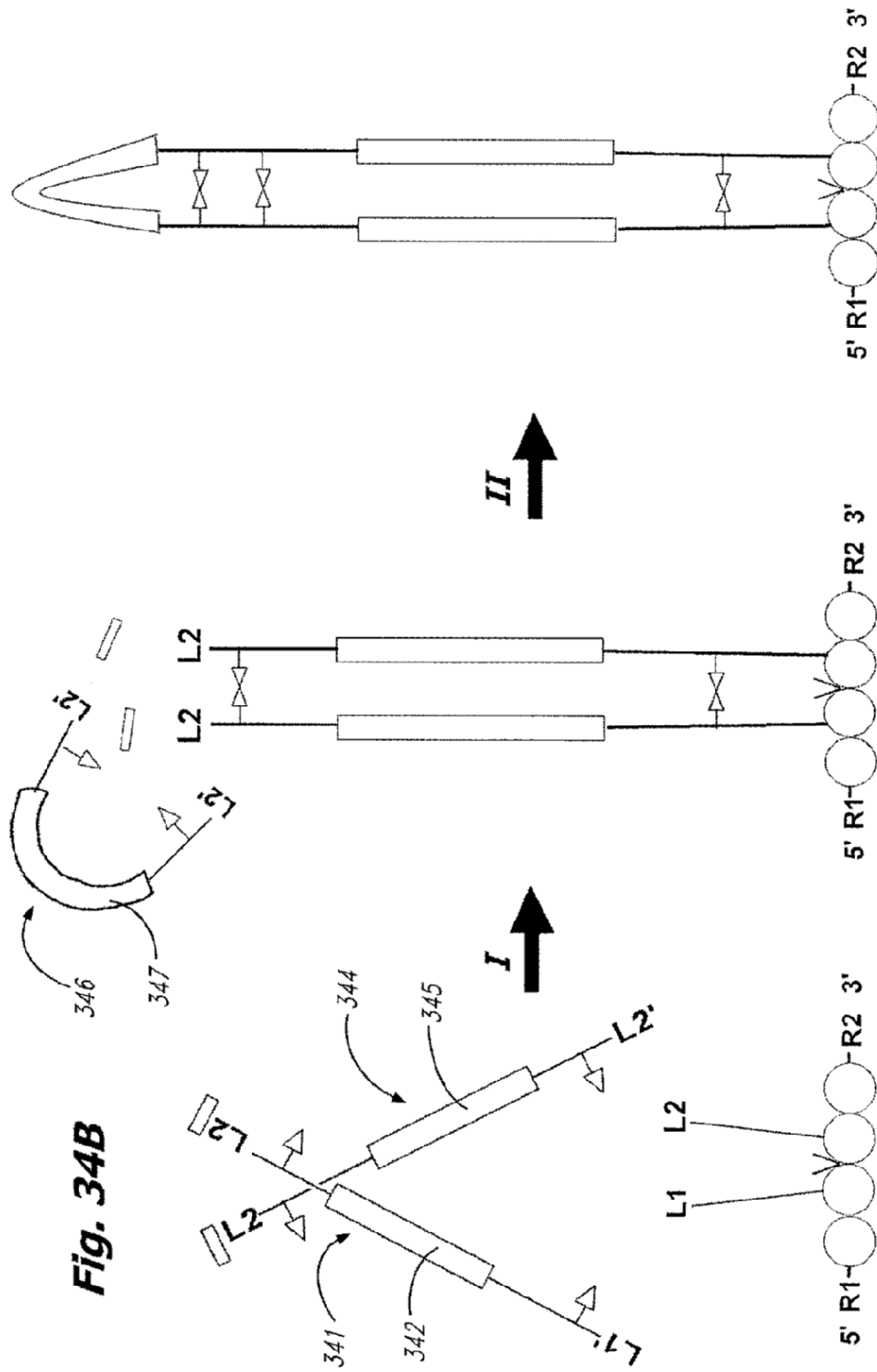
**Fig. 33A**



**Fig. 33C**







**Fig. 34B**

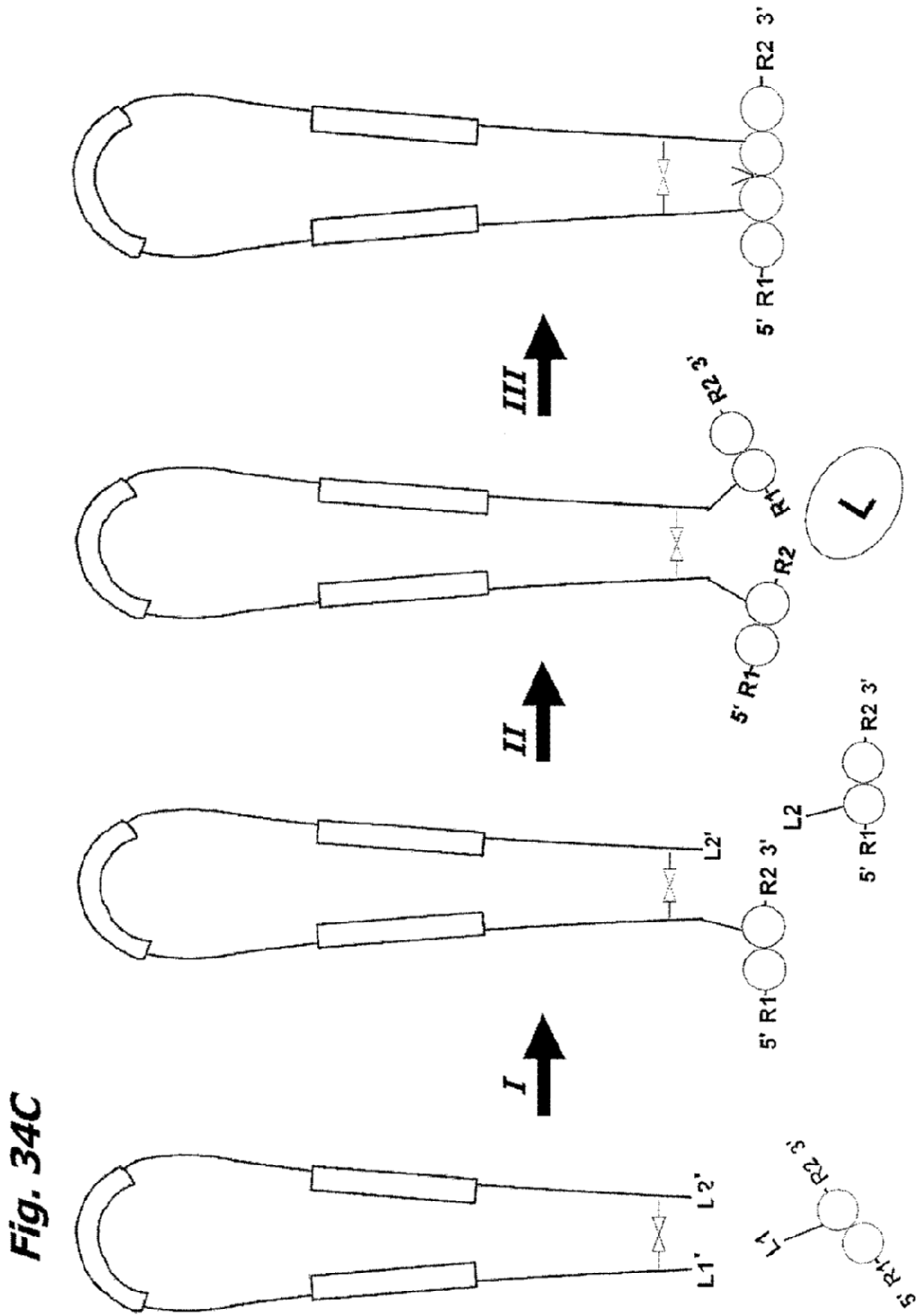
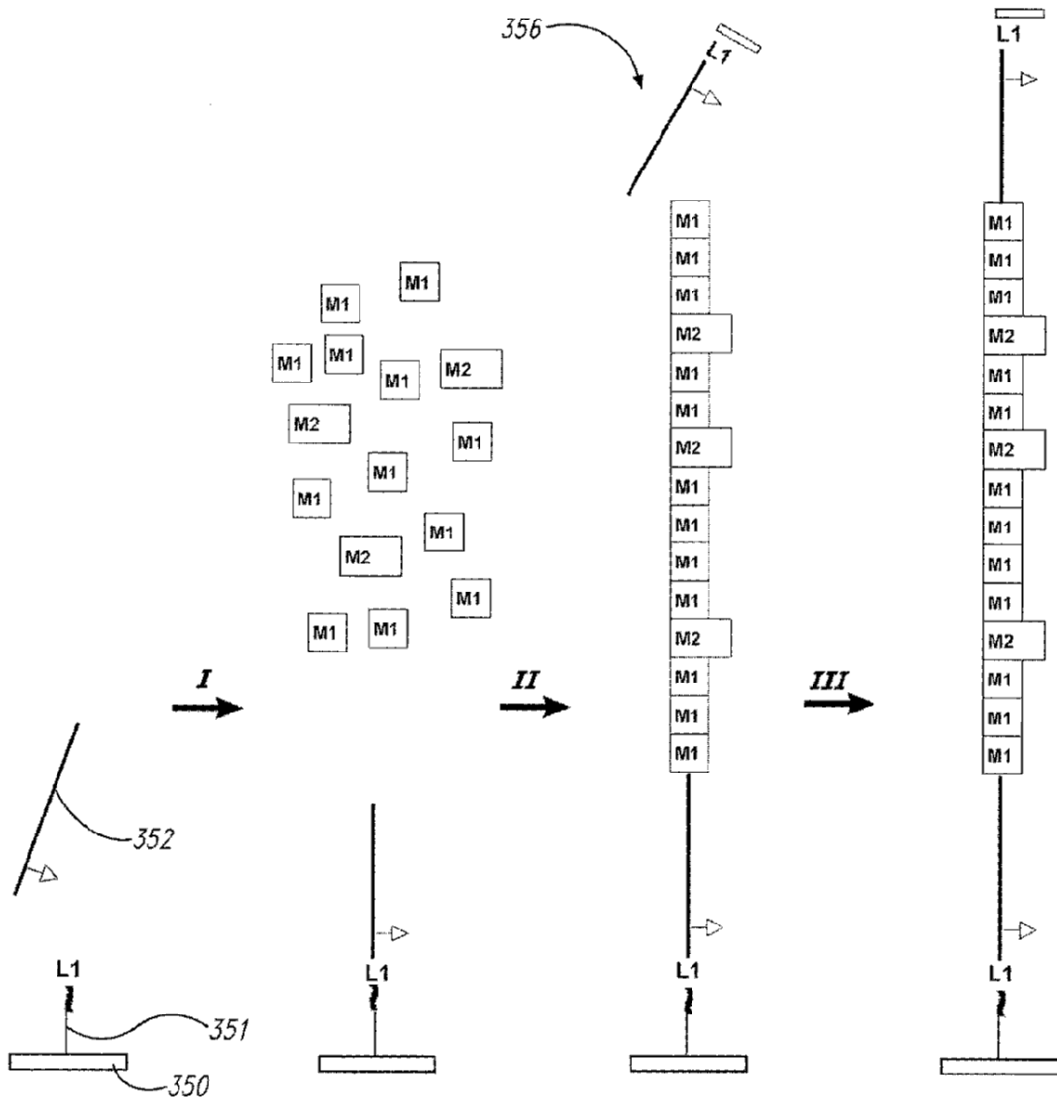
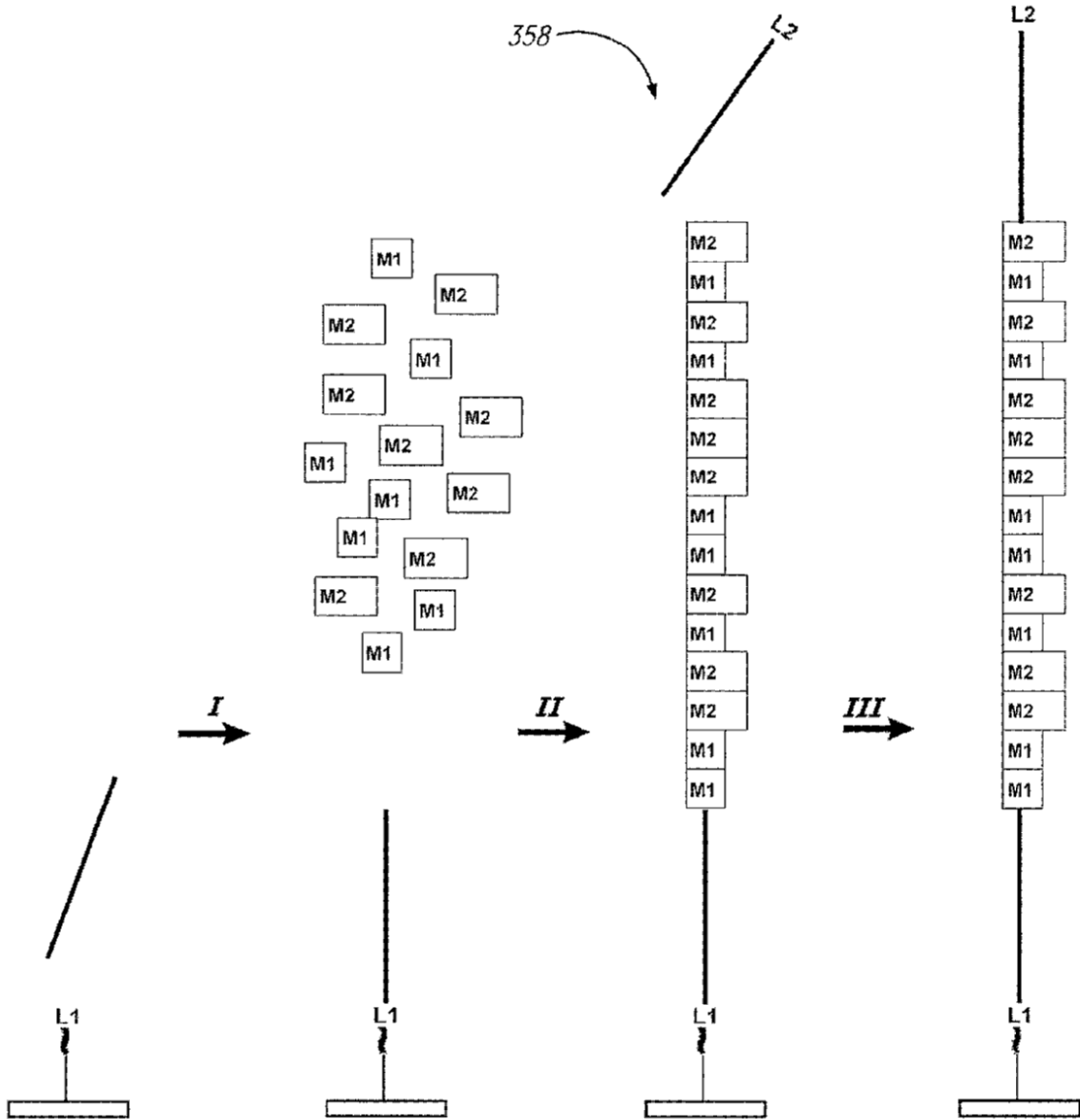


Fig. 35A



**Fig. 35B**





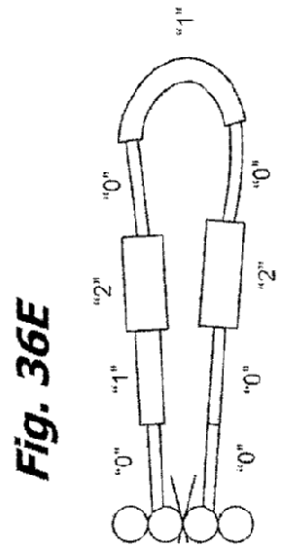
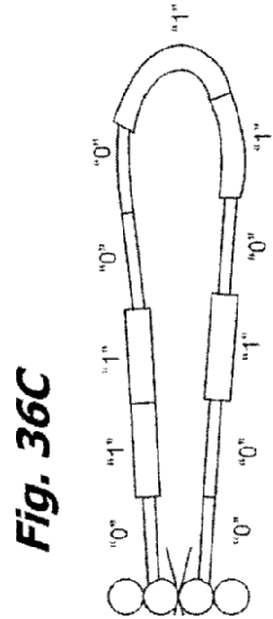
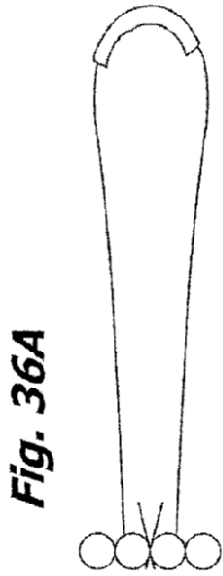
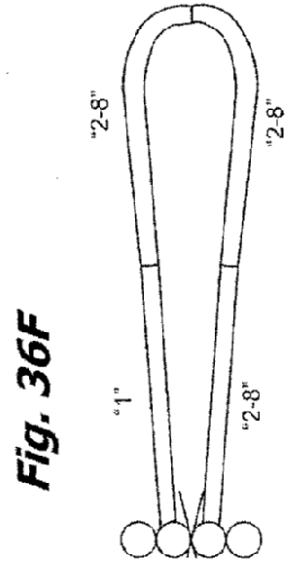
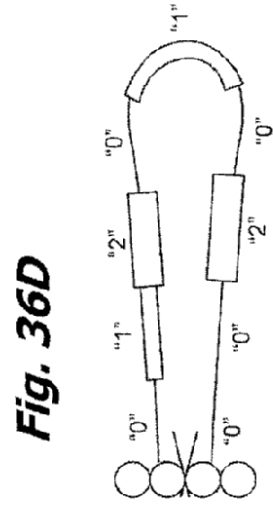
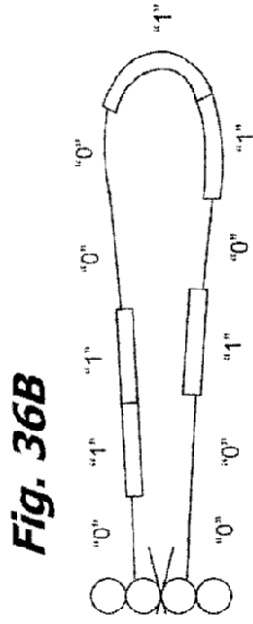




Fig. 37

Identidad de marca	Masa	Fluorescencia (F)	Conectores químicos
1	M1	ROJA	HIDARACINA -
2	M2	VERDE	AMINA -
3	M3	AZUL	CARBOXILO -
4	m1	Roja	hidracina -
5	m2	Verde	amina -
6	m3	Azul	carboxilo -
7	m1-m2	roja-verde	hidracina amina
8	m1-m3	roja-azul	hidracina -
9	m1-M2	roja-VERDE	hidracina AMINA
10	m1-M3	roja-AZUL	hidracina -
11	M1-m2	ROJA-verde	HIDRACINA amina
12	M1-m3	ROJA-azul	HIDRACINA -
13	m2-m3	verde-azul	carboxilo
14	M2-m3	VERDE-azul	carboxilo
15	m2-m3	verde-Azul	carboxilo
16	m1-m2-m3	roja-verde-azul	CARBOXILO
17	M1-m2-m3	ROJA-verde-azul	amina
18	m1-M2-m3	roja-VERDE-azul	amina
19	m1-m2-M3	roja-verde-AZUL	AMINA
20	M1-M2-m3	RED-VERDE-azul	amina
21	m1-M2-M3	roja-VERDE-AZUL	AMINA
22	M1-m2-M3	ROJA-verde-AZUL	amina

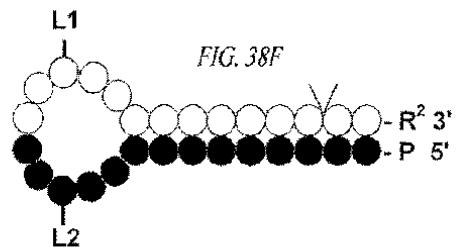
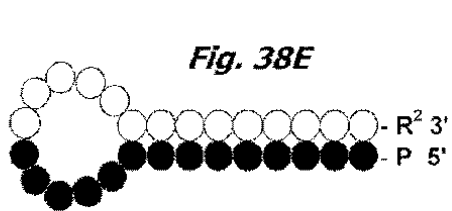
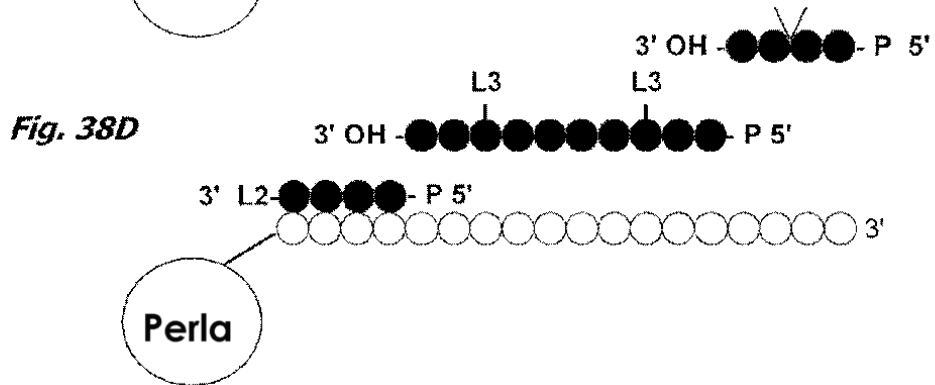
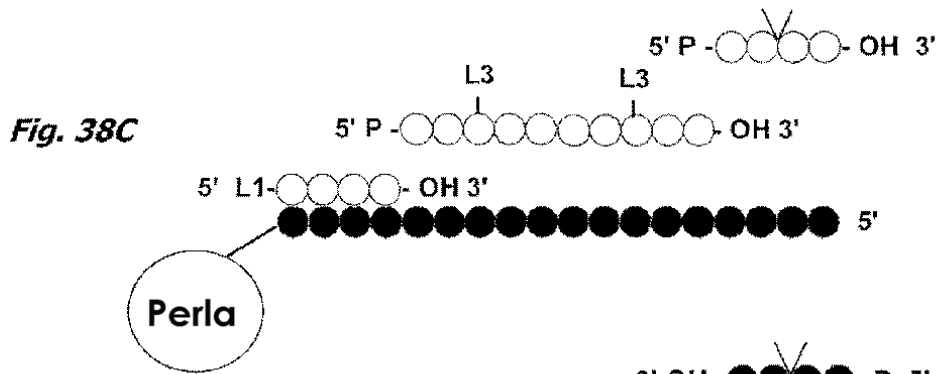
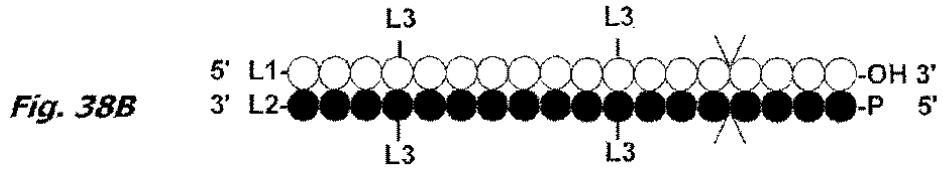


Fig. 39

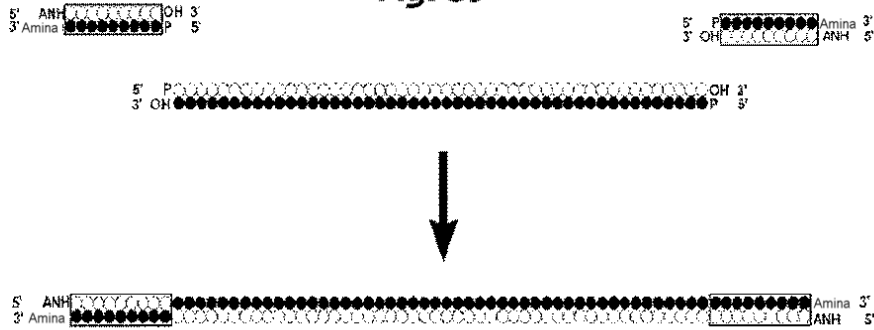


Fig. 40

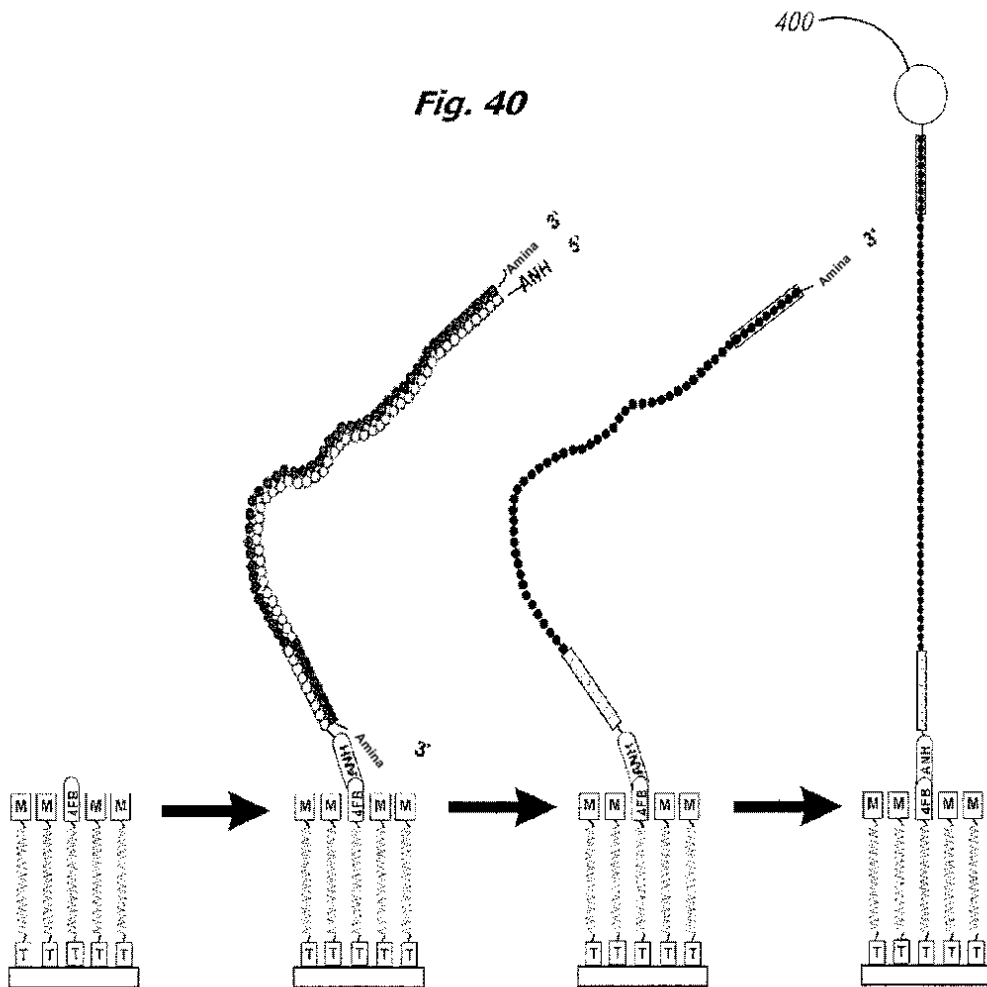


Fig. 41D

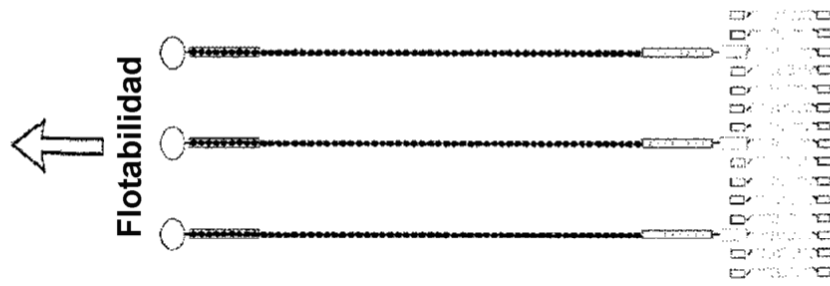


Fig. 41C

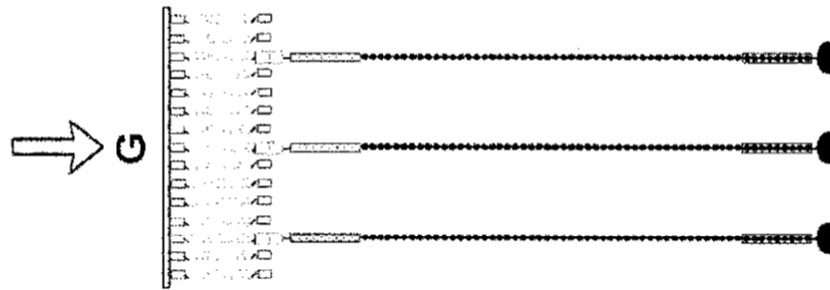


Fig. 41B

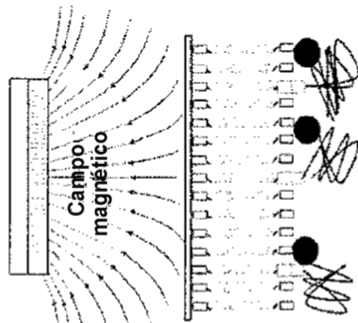
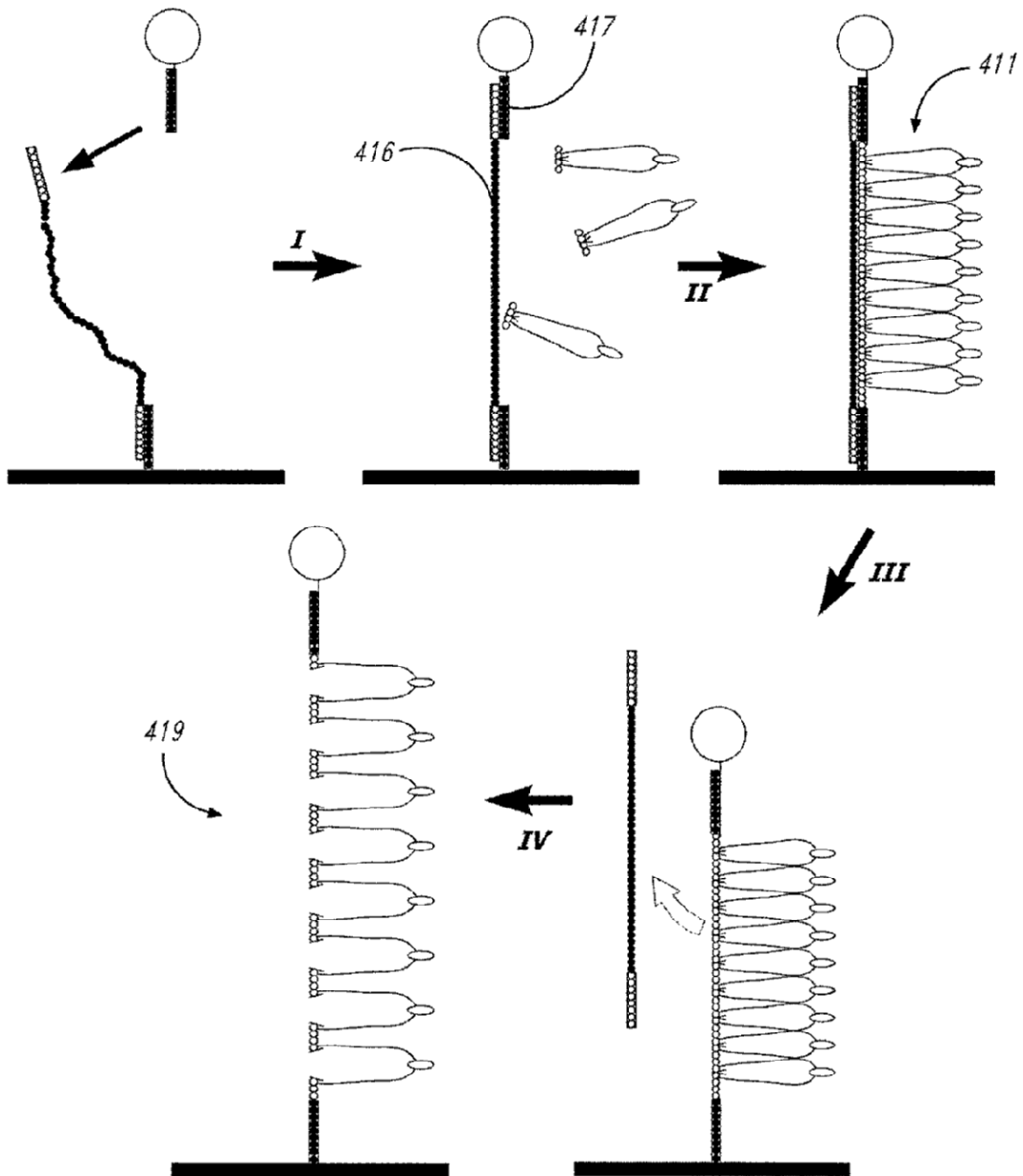


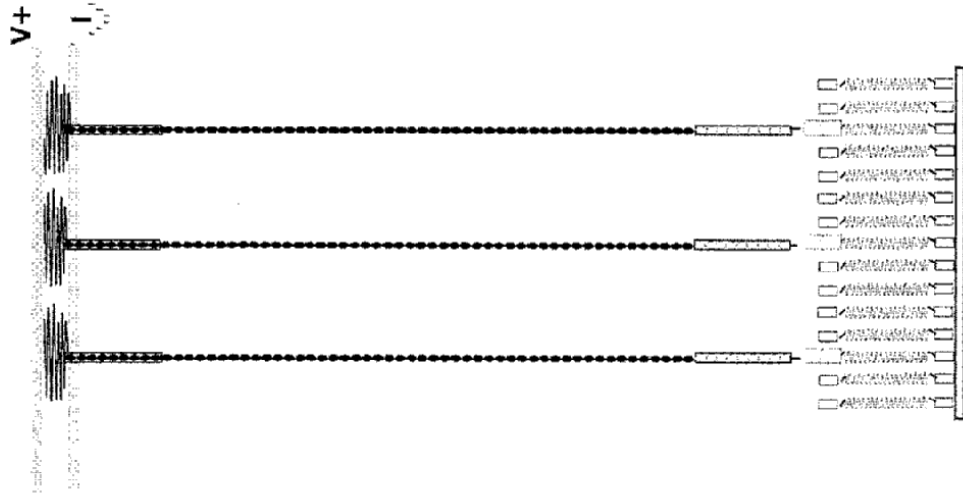
Fig. 41A



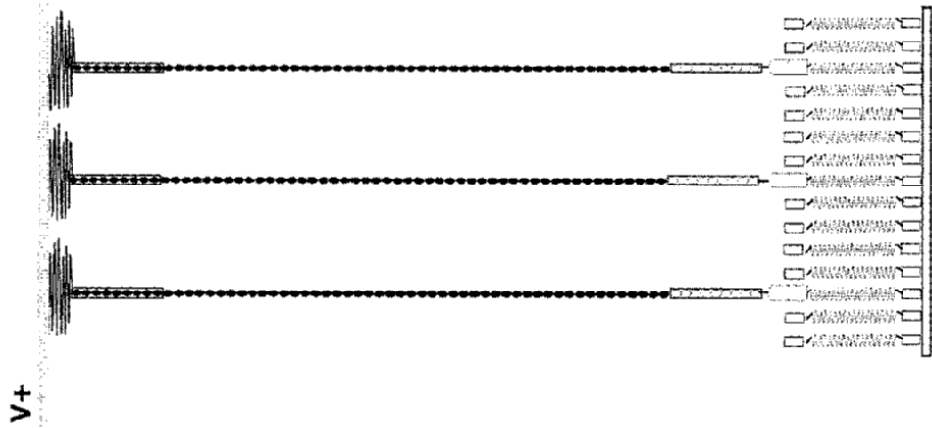
**Fig. 41E**



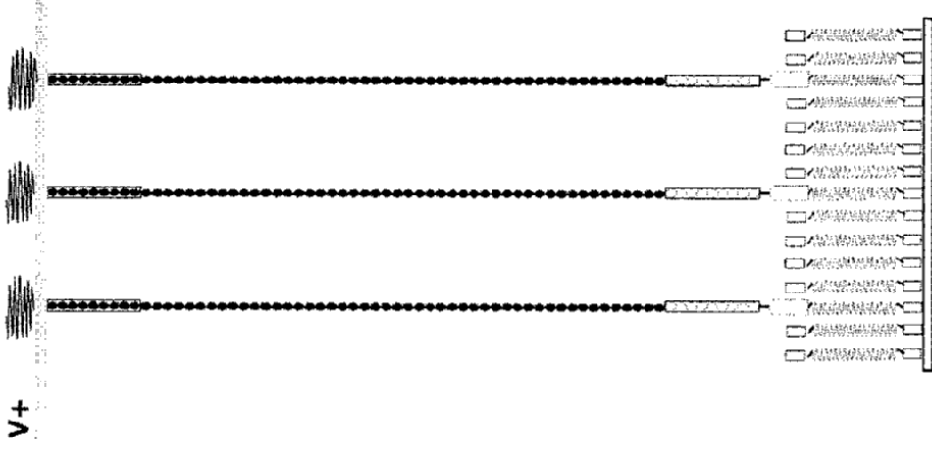
**Fig. 42C**



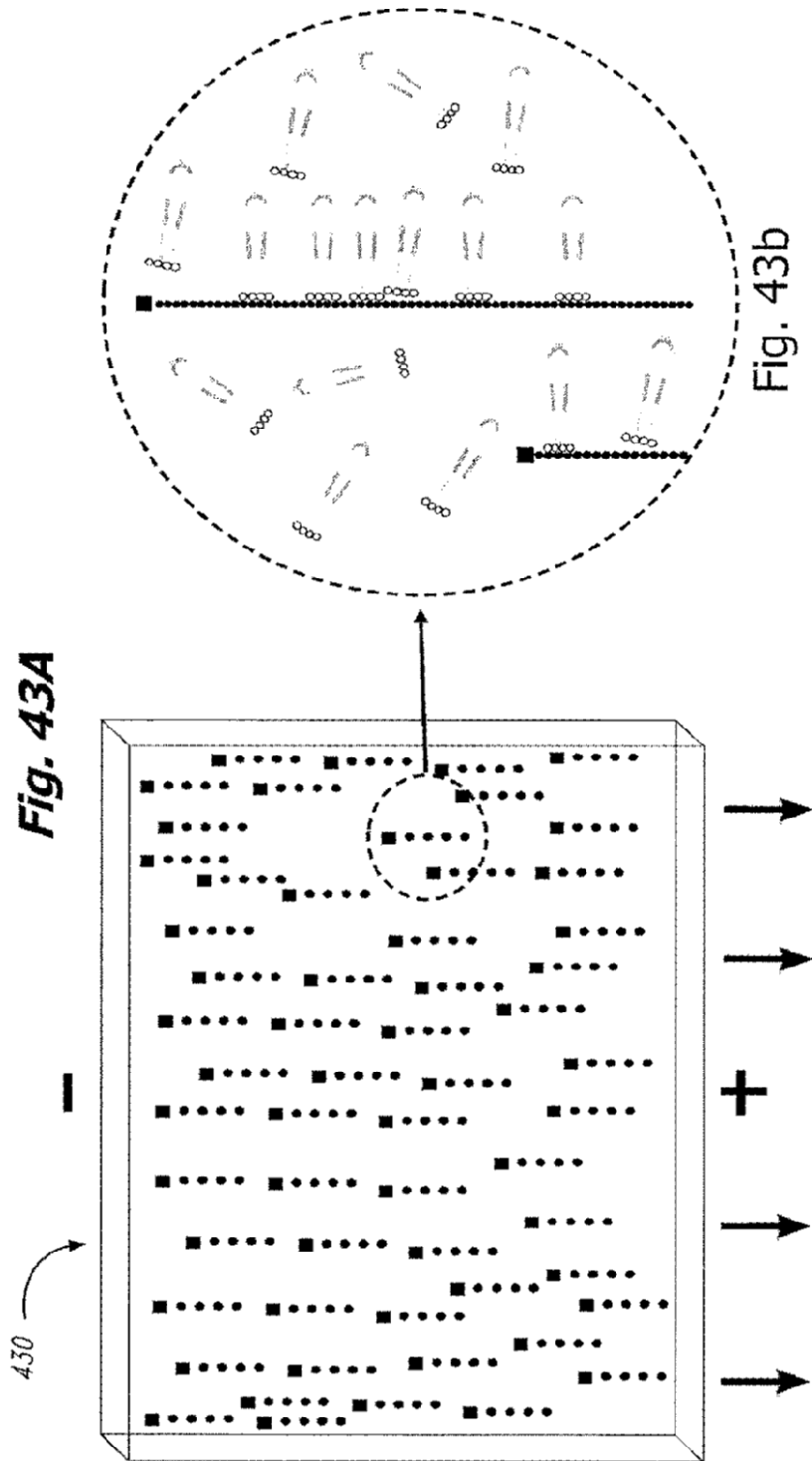
**Fig. 42B**



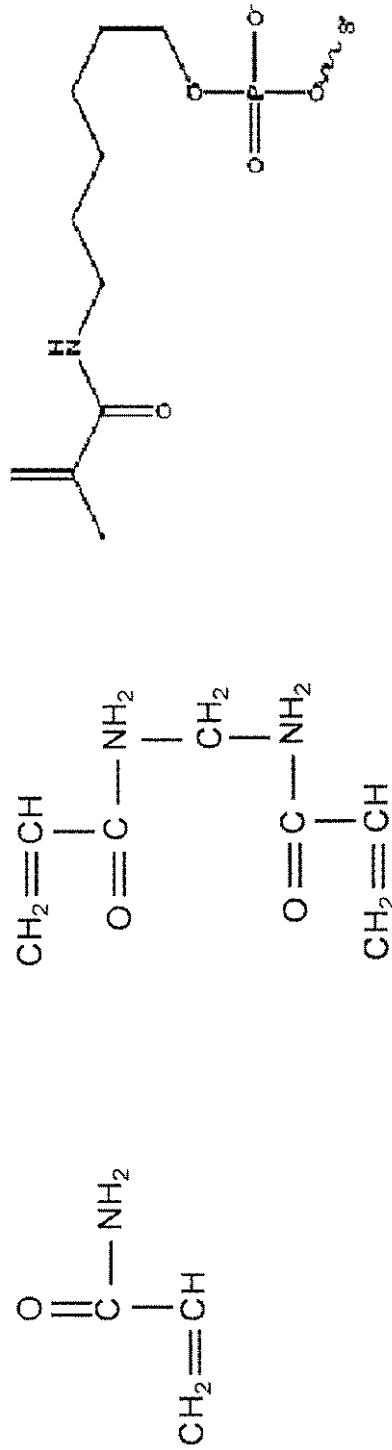
**Fig. 42A**





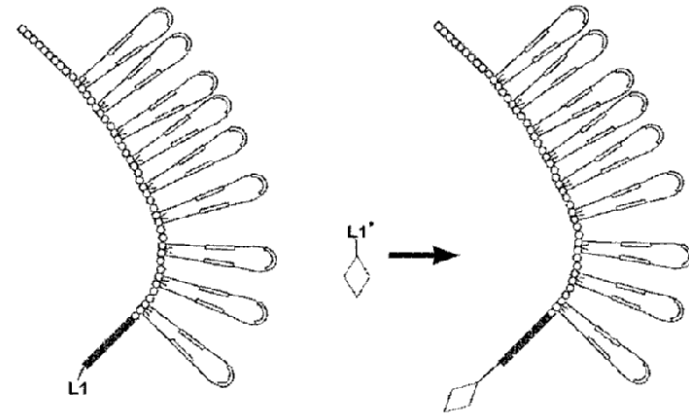


**Fig. 43C**

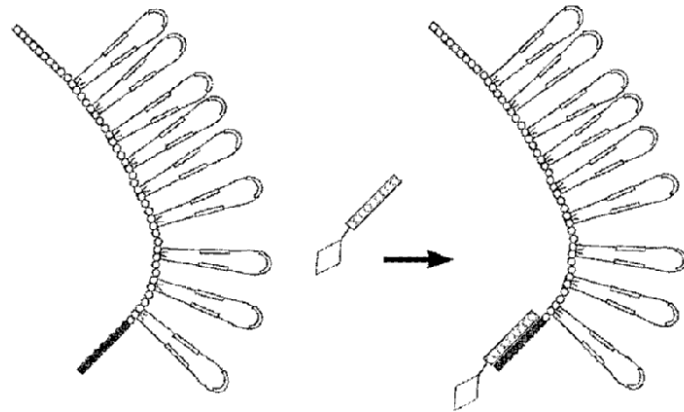




**Fig. 44A**



**Fig. 44B**



**Fig. 44C**

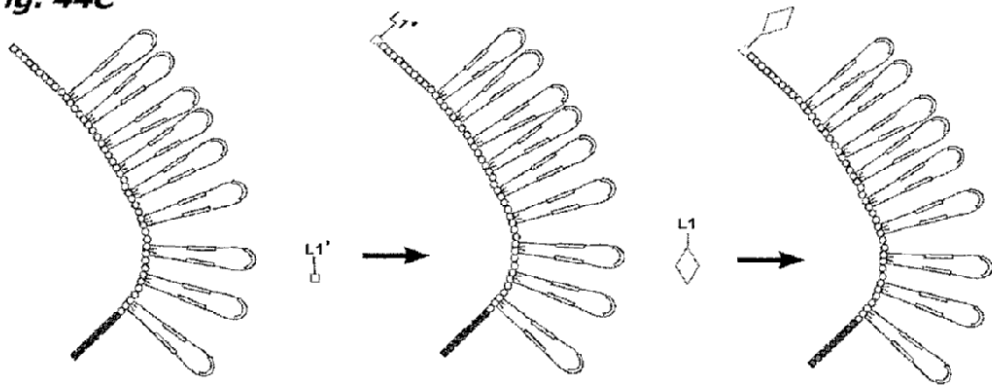
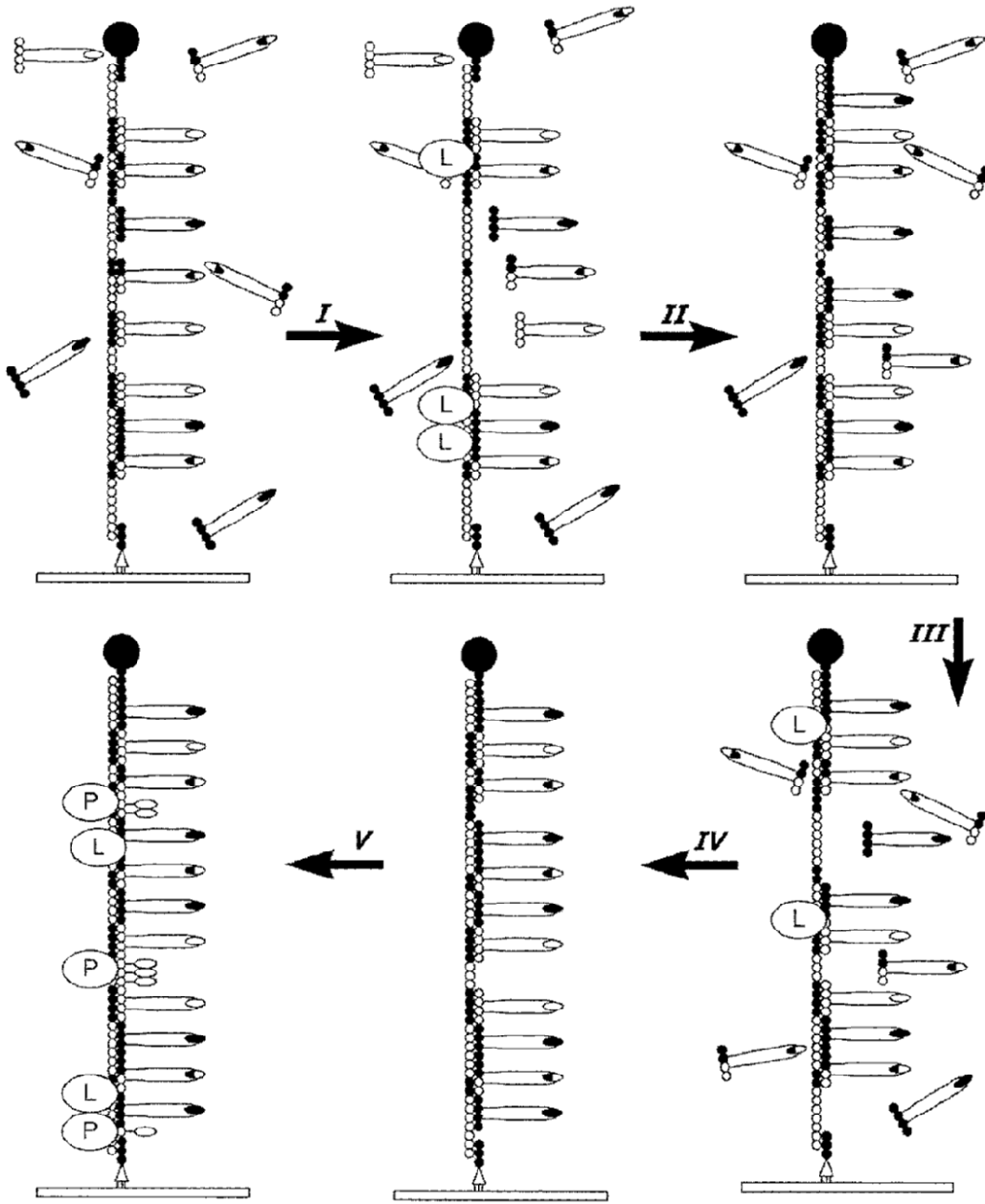
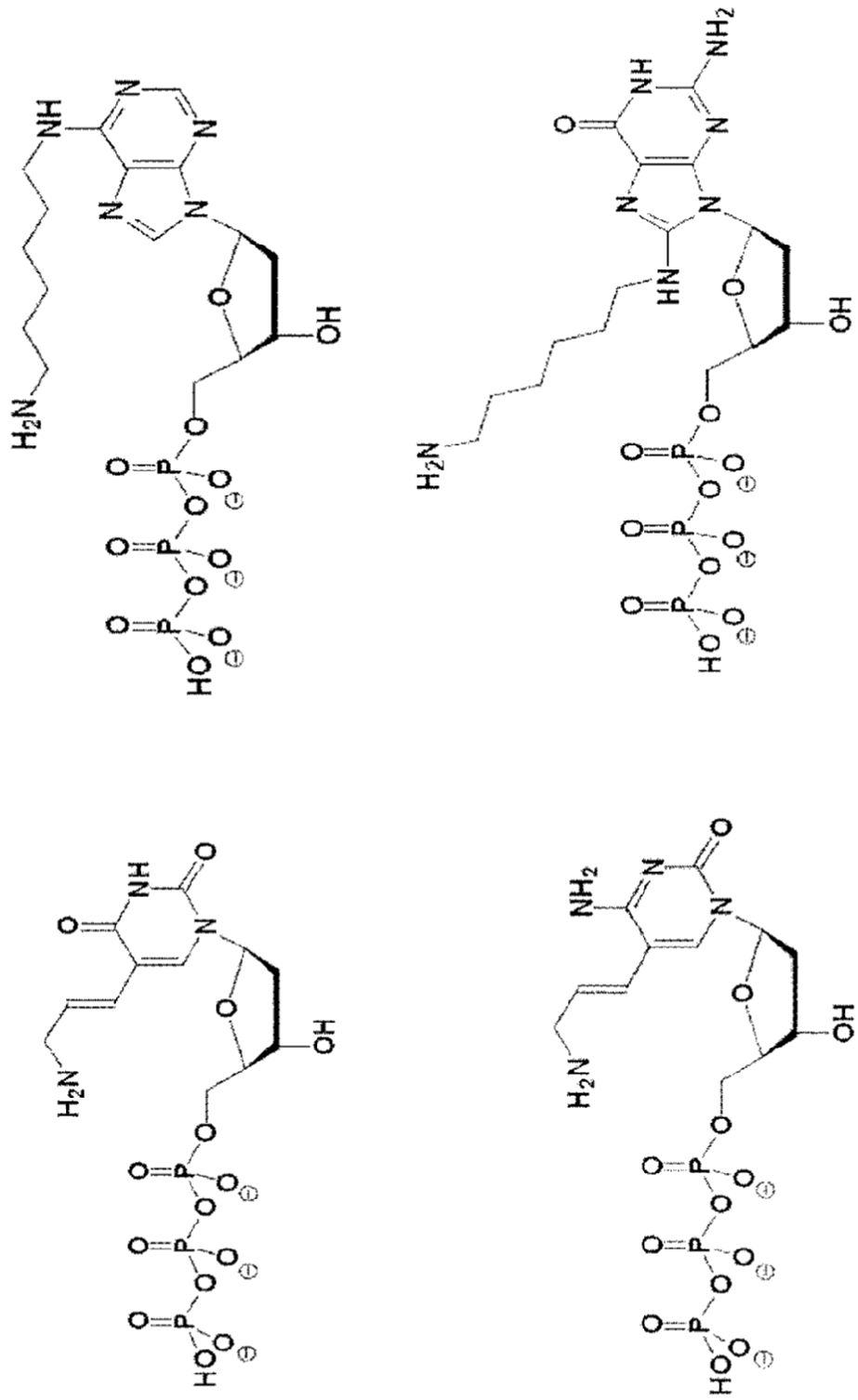


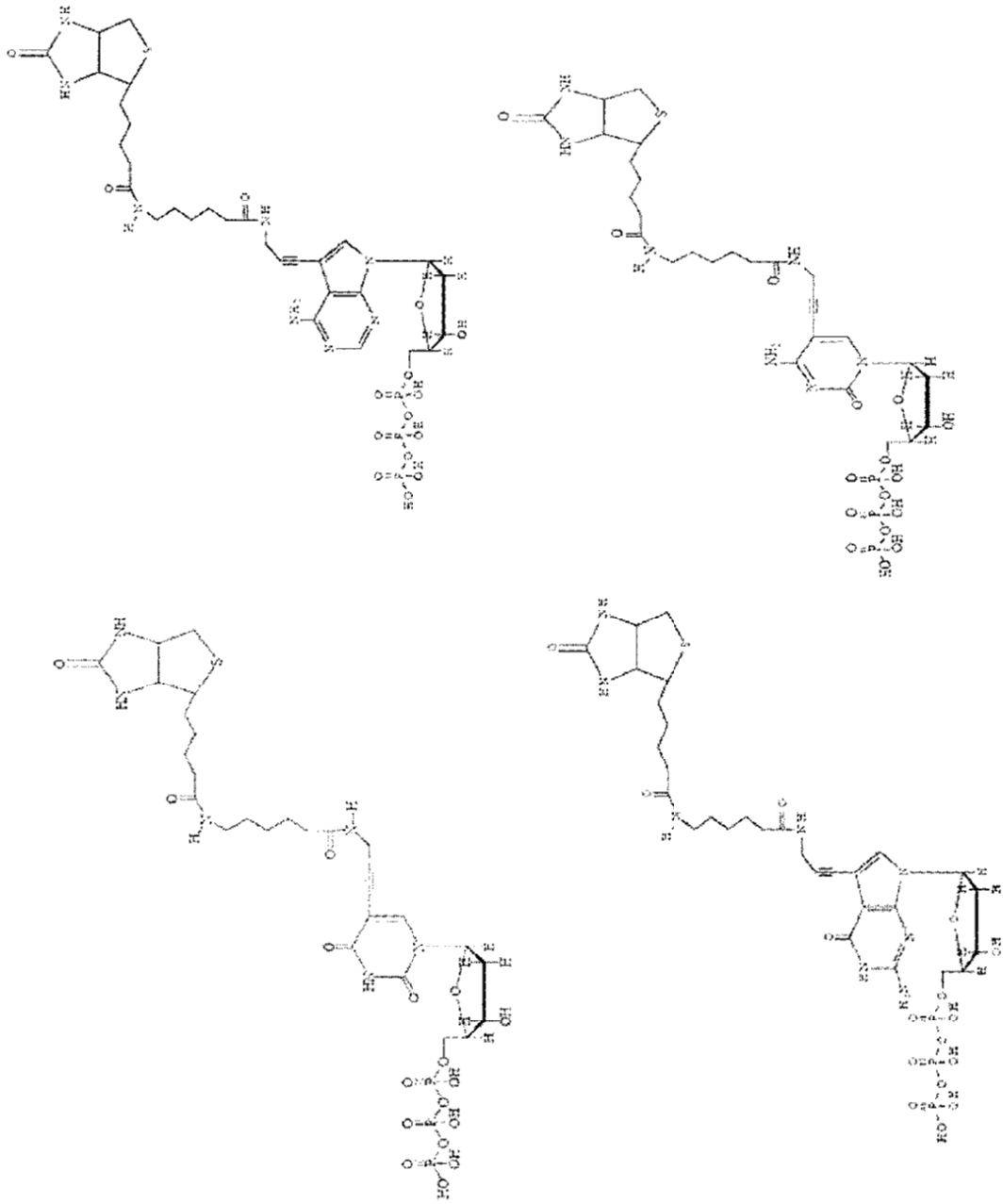
Fig. 45



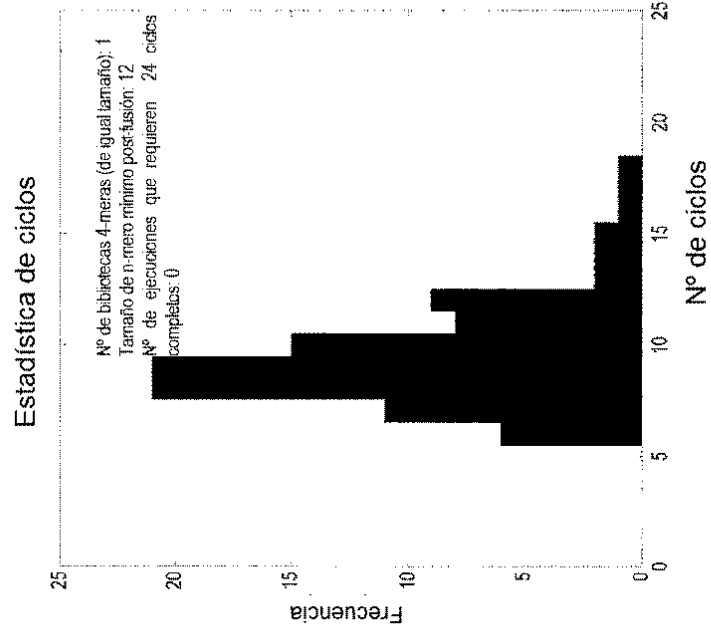
**Fig. 46A**



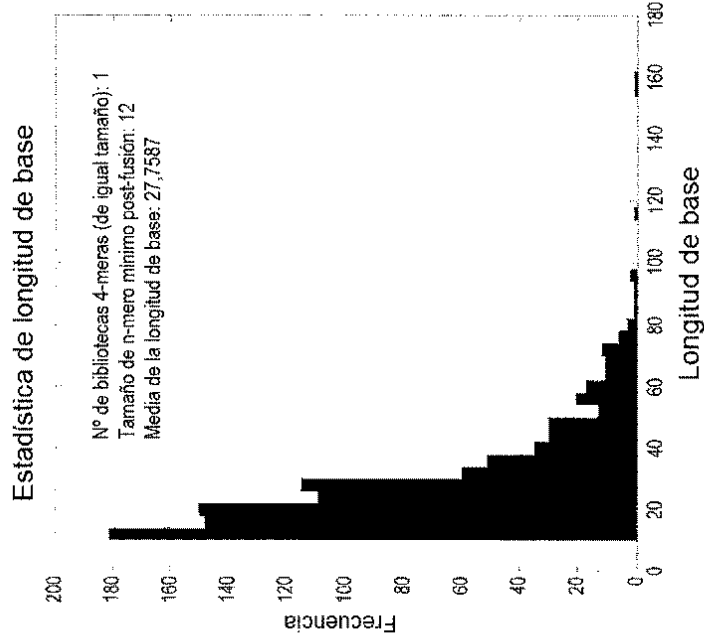
**Fig. 46B**



**Fig. 47B**

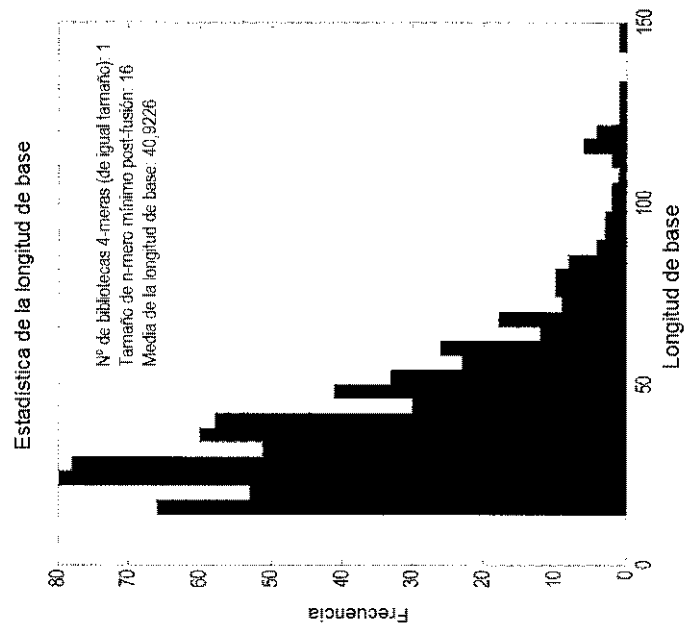


**Fig. 47A**

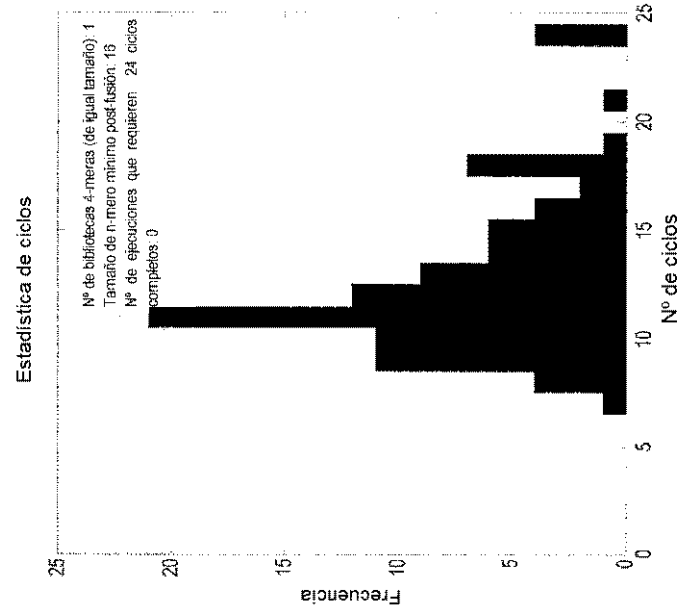




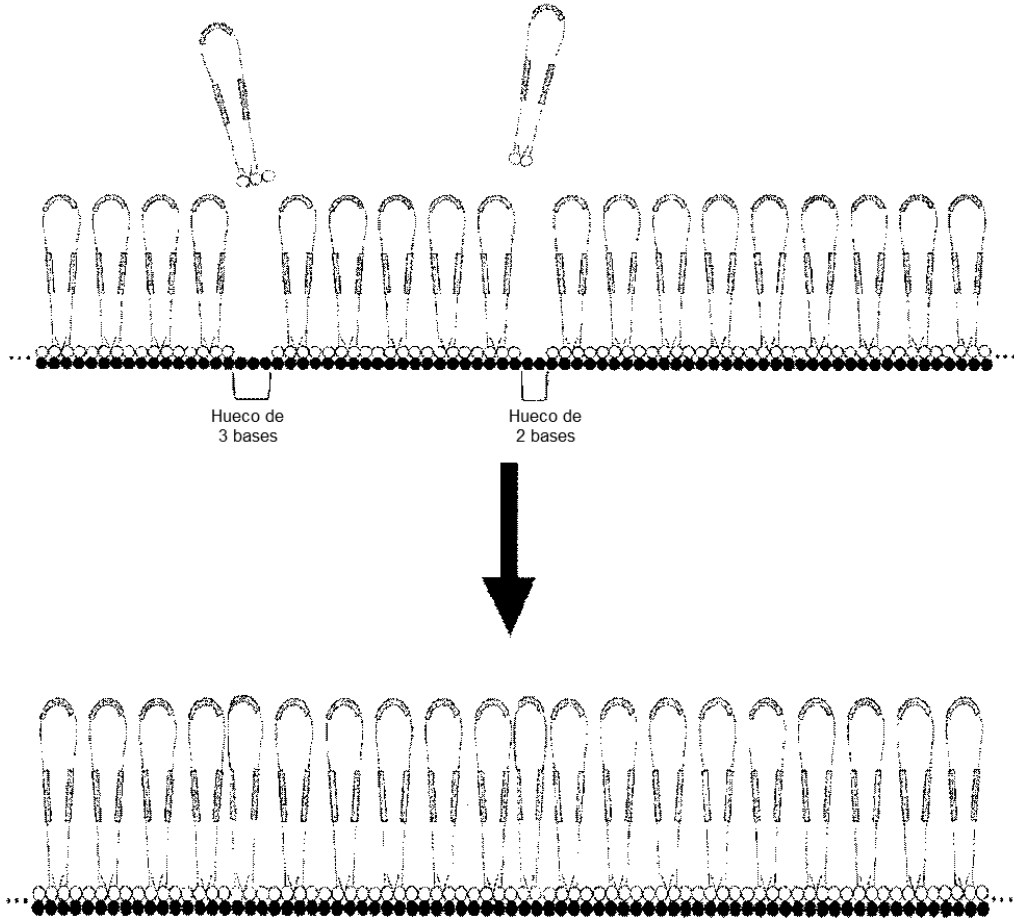
**Fig. 48A**



**Fig. 48B**

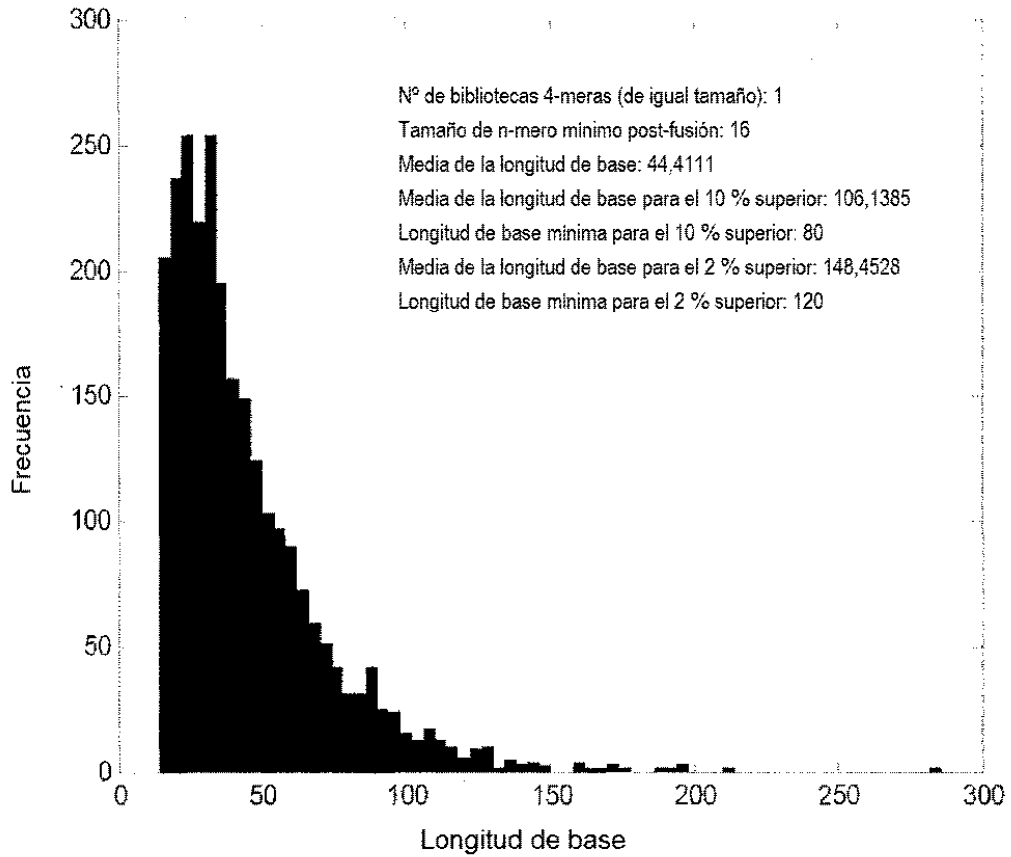


**Fig. 49**

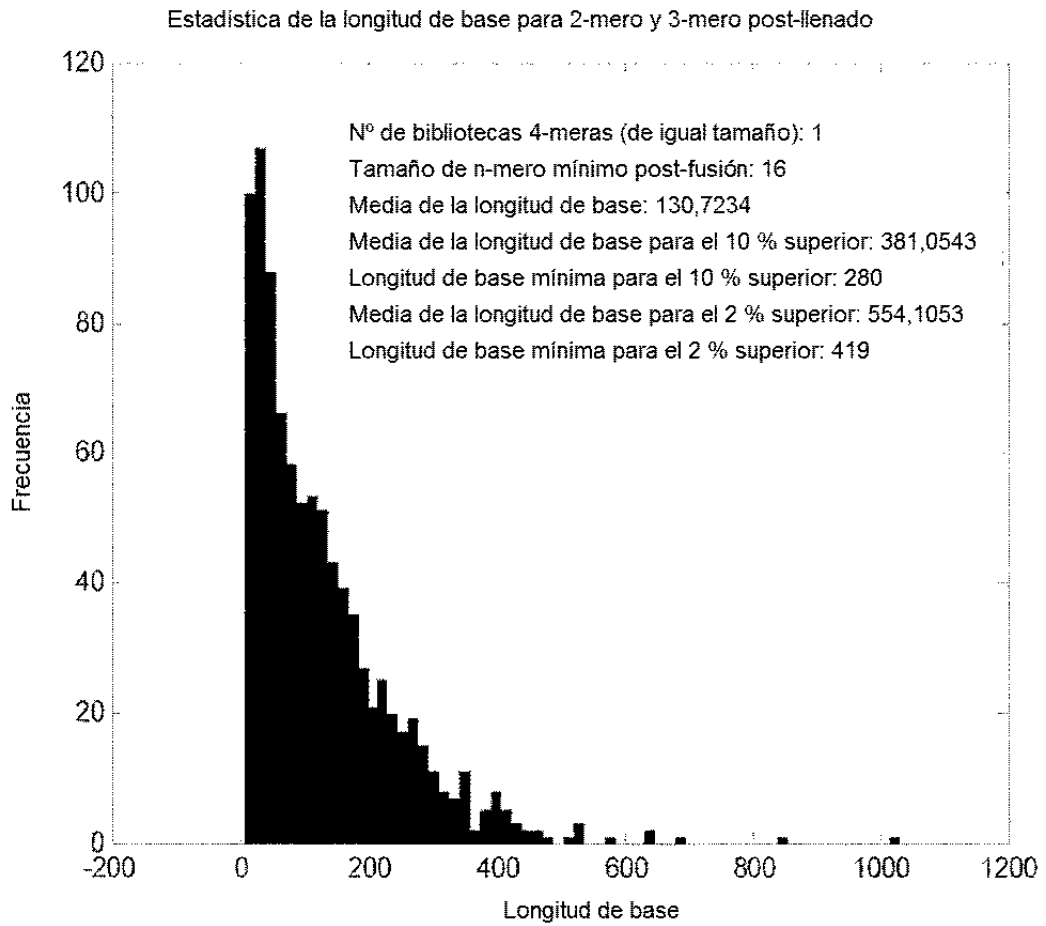


**Fig. 50A**

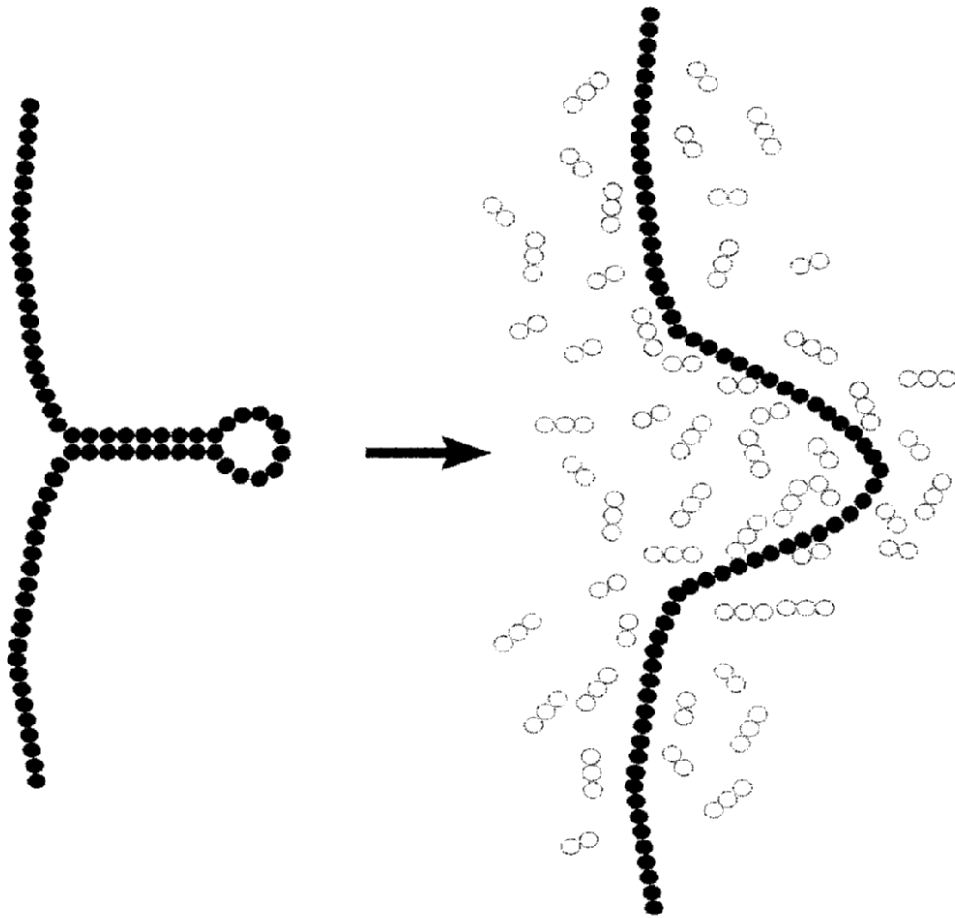
Estadística de la longitud de base



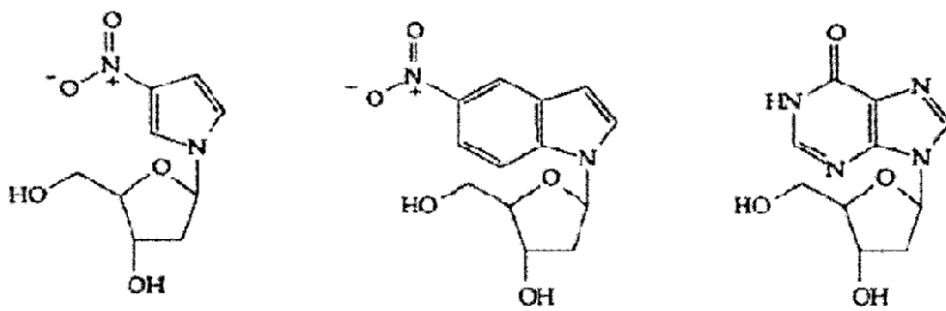
**Fig. 50B**



**Fig. 51**



**Fig. 52**



**Fig. 53**

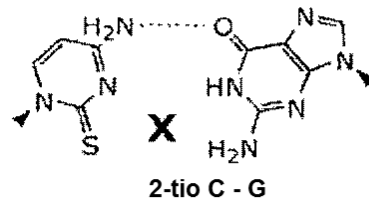
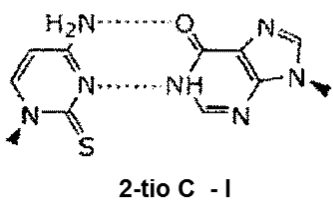
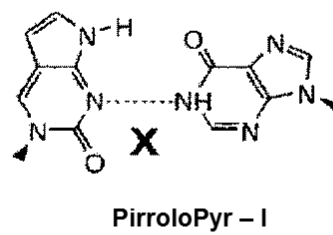
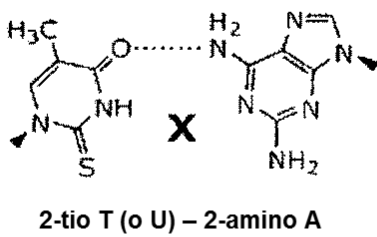
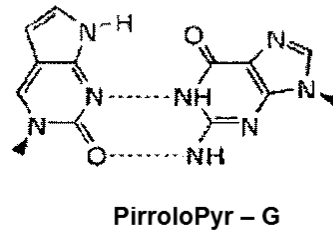
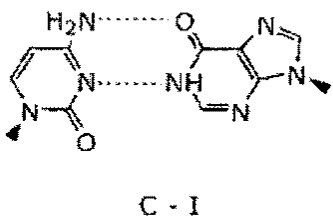
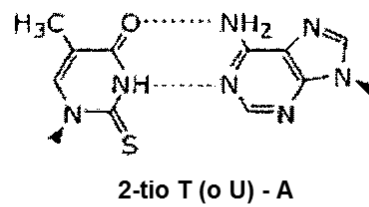
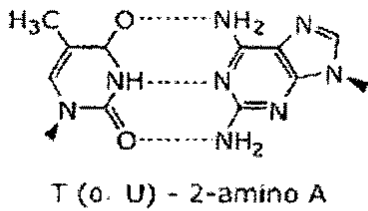
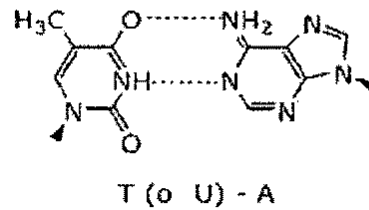
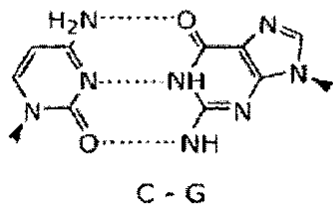
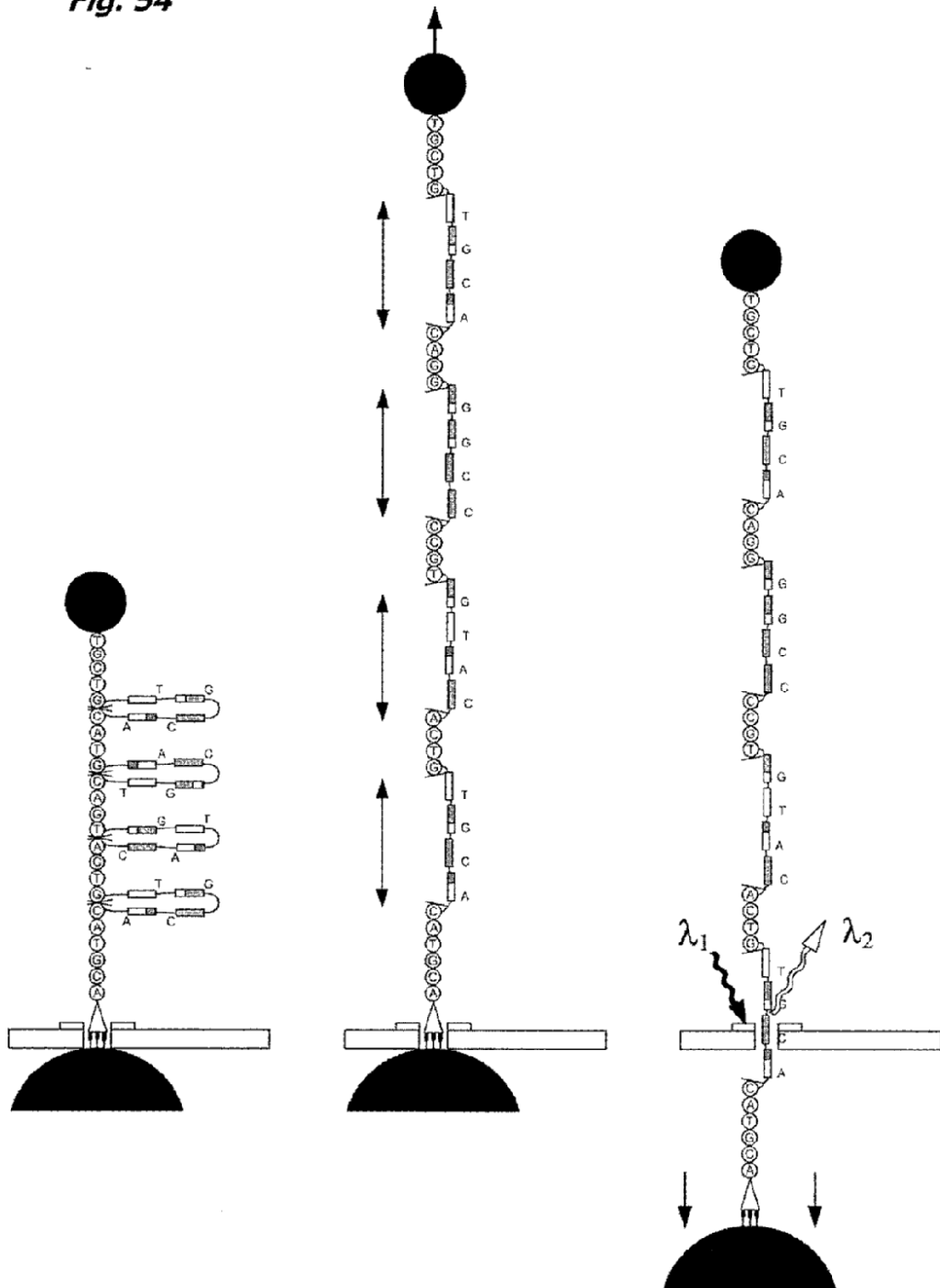
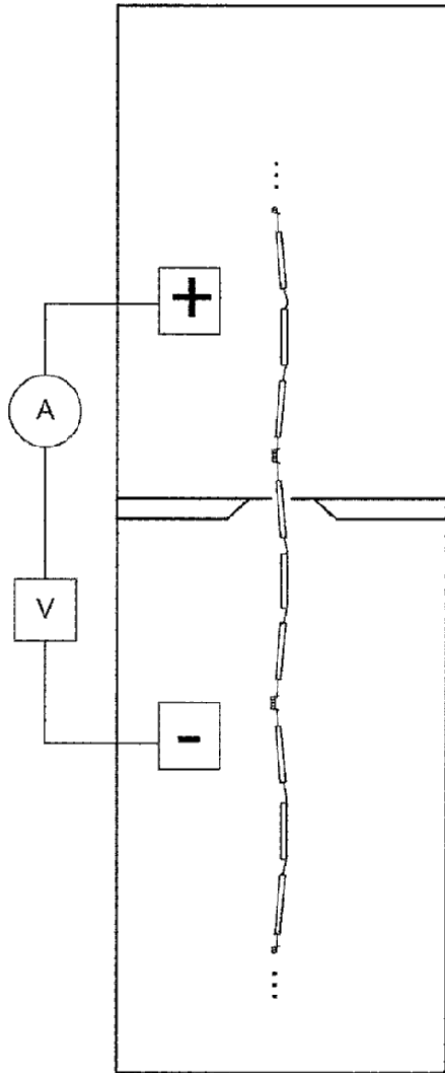


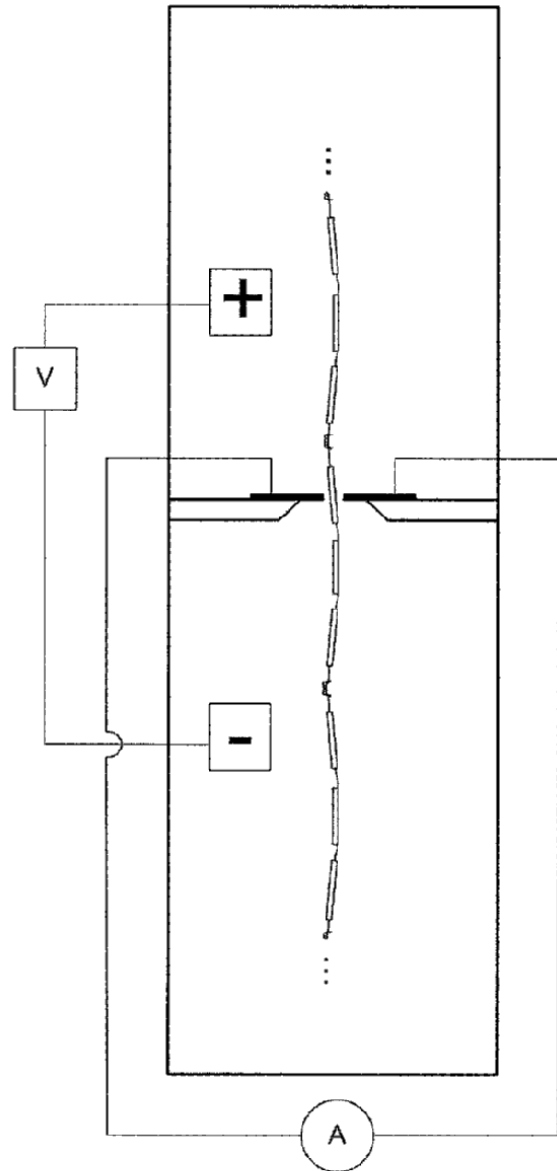
Fig. 54



**Fig. 55**

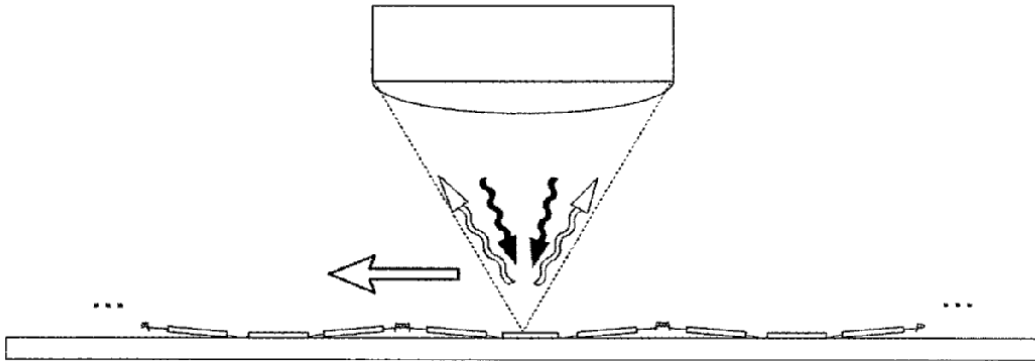


**Fig. 56**

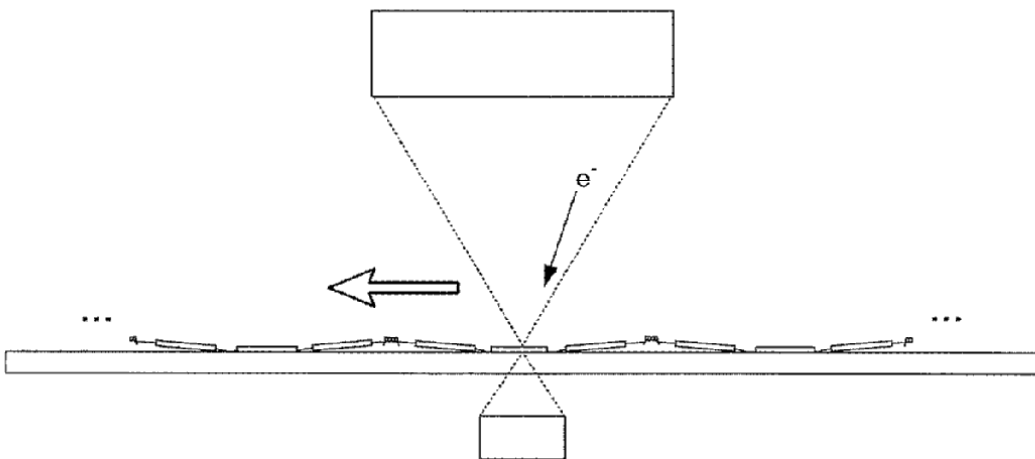




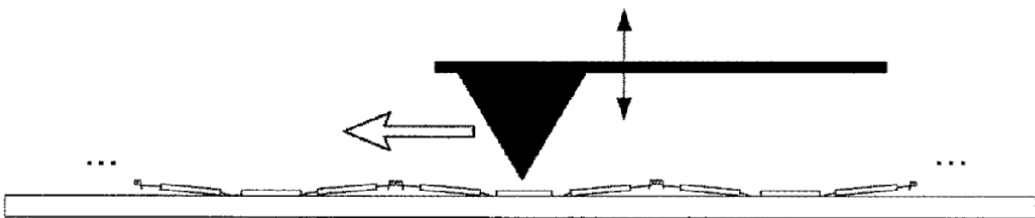
**Fig. 57**



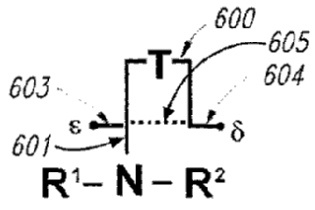
**Fig. 58**



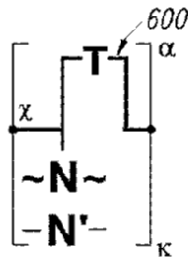
**Fig. 59**



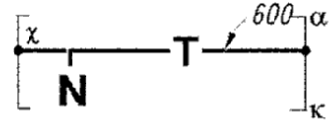
**Fig. 60A**



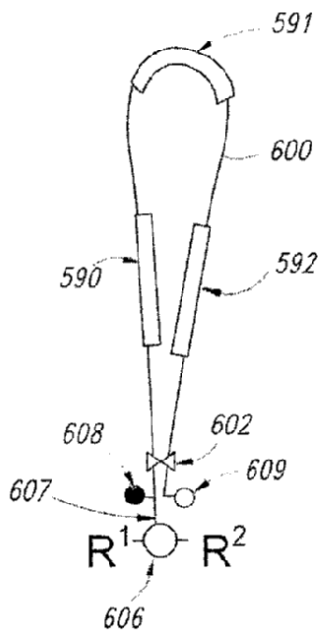
**Fig. 60B**



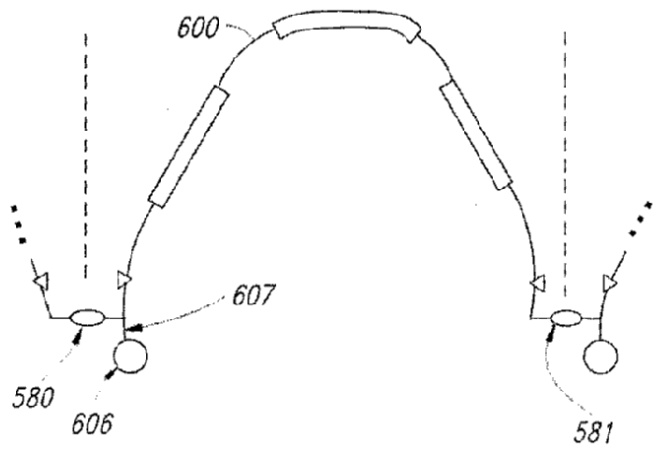
**Fig. 60C**

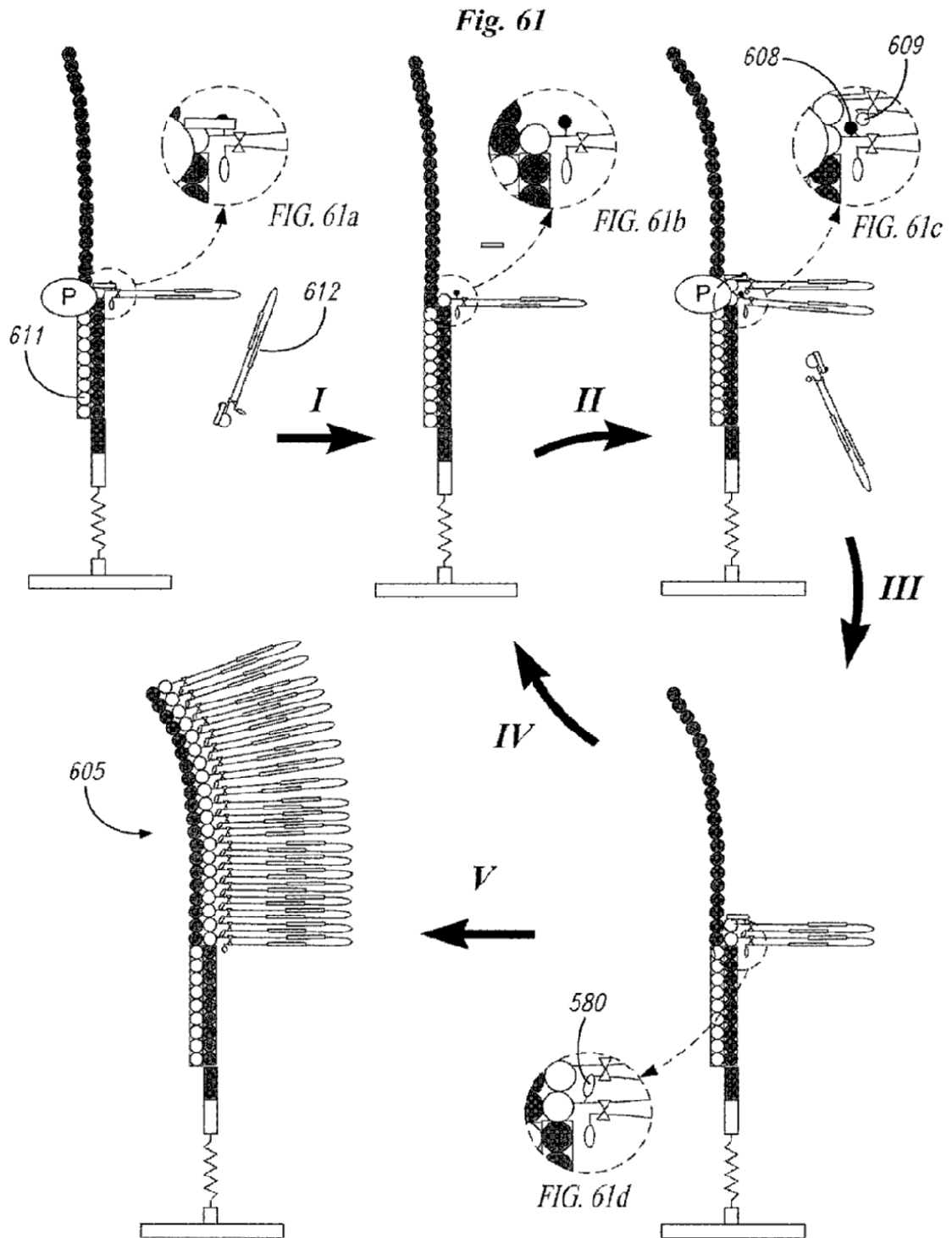


**Fig. 60D**

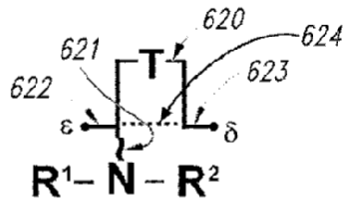


**Fig. 60E**

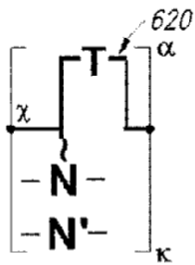




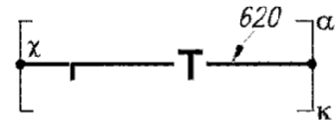
**Fig. 62A**



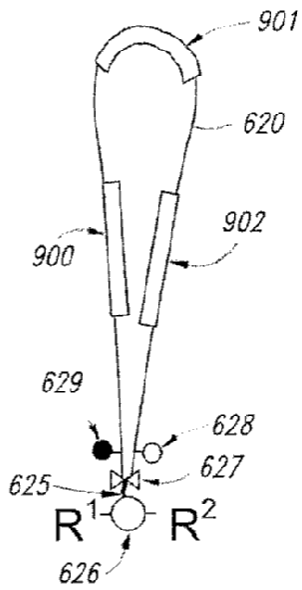
**Fig. 62B**



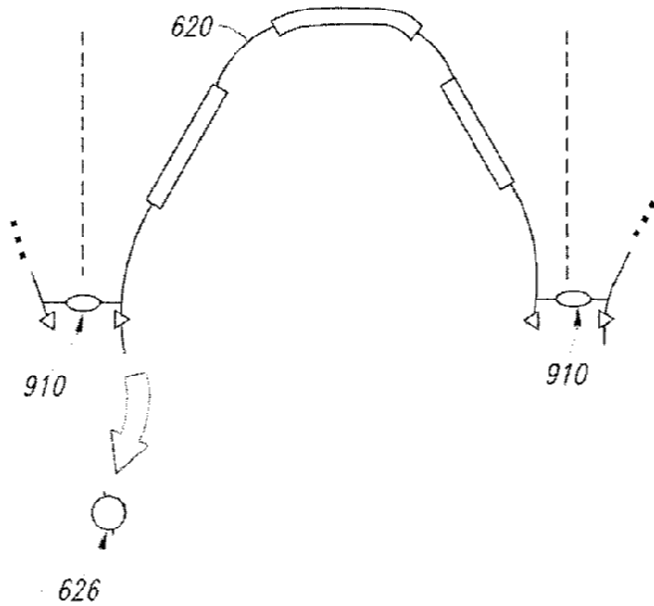
**Fig. 62C**



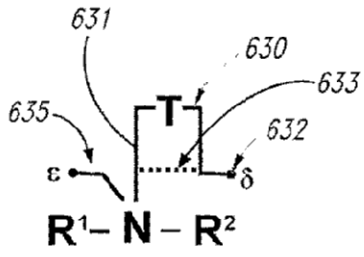
**Fig. 62D**



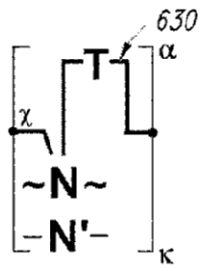
**Fig. 62E**



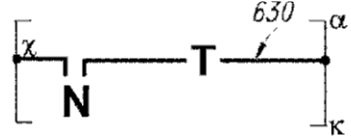
**Fig. 63A**



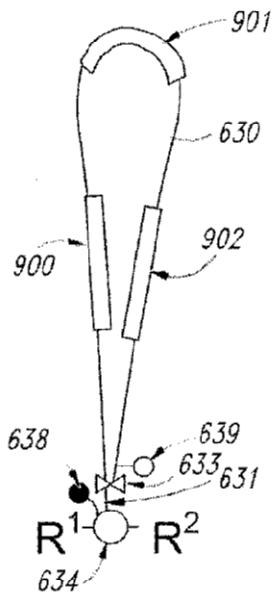
**Fig. 63B**



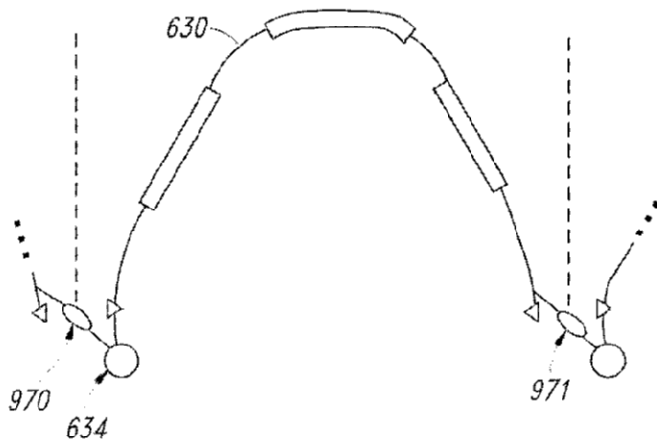
**Fig. 63C**



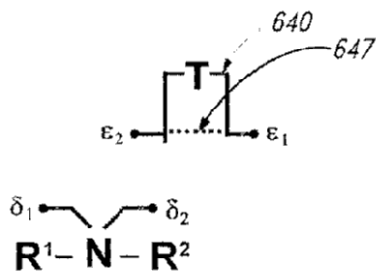
**Fig. 63D**



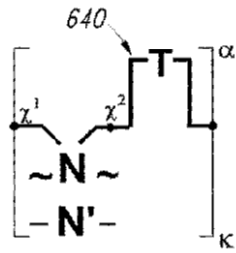
**Fig. 63E**



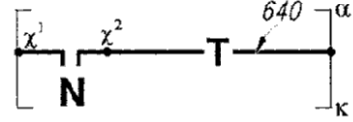
**Fig. 64A**



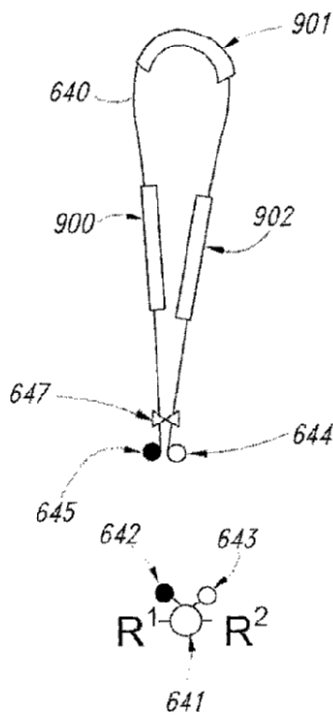
**Fig. 64B**



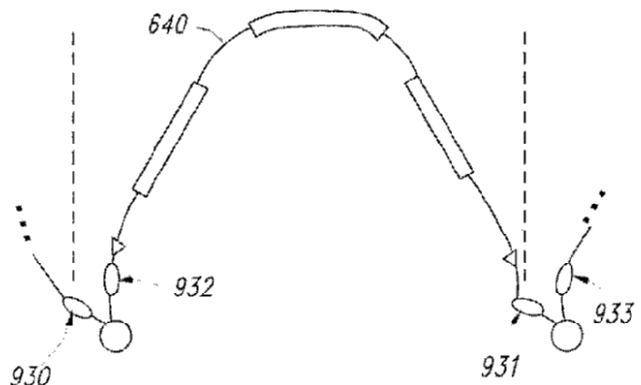
**Fig. 64C**



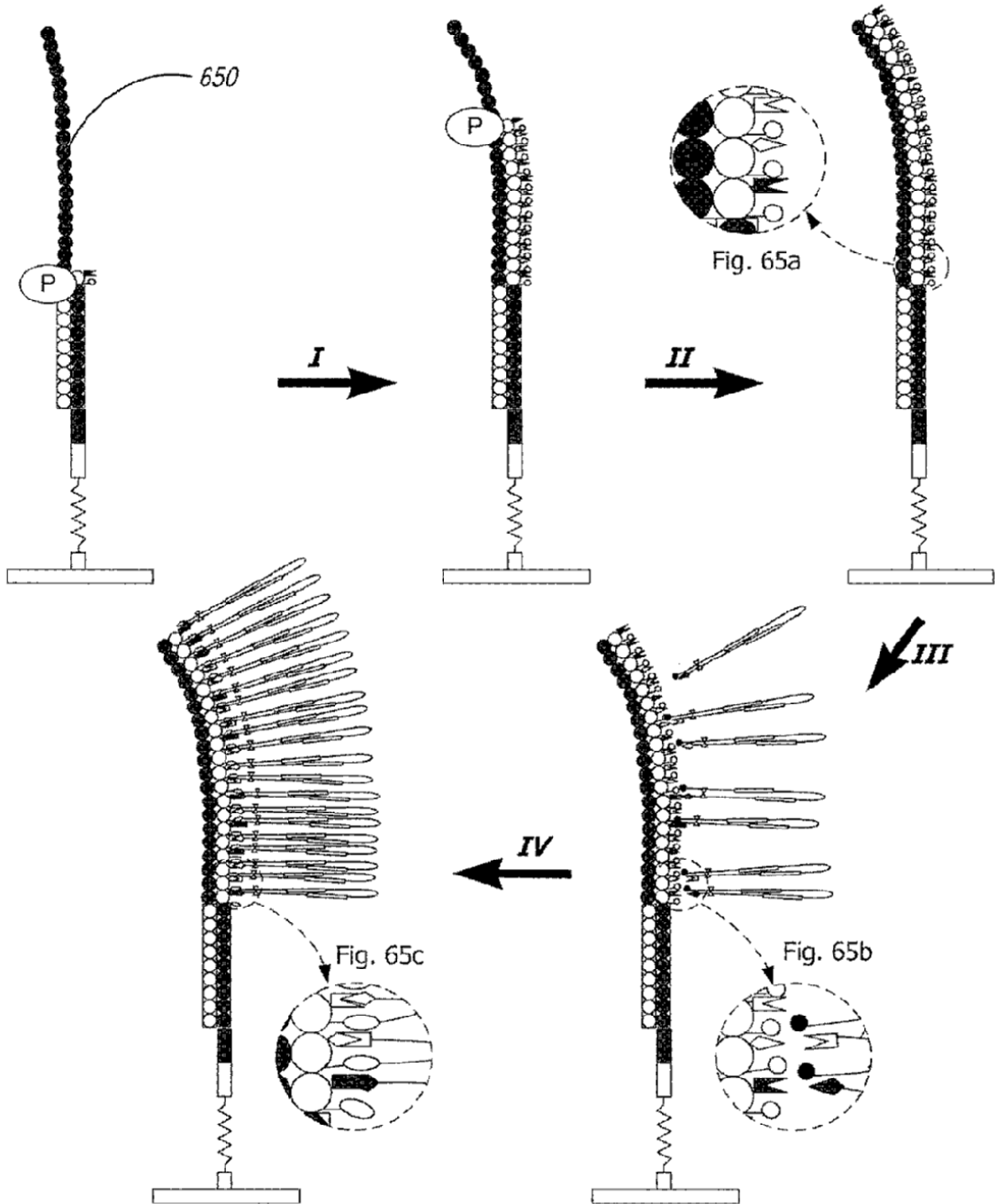
**Fig. 64D**



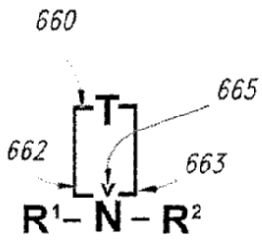
**Fig. 64E**



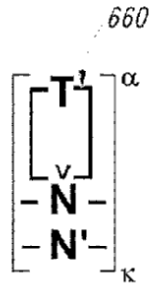
**Fig. 65**



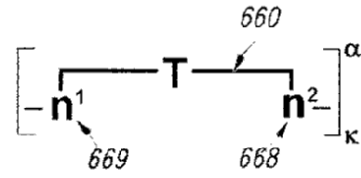
**Fig. 66A**



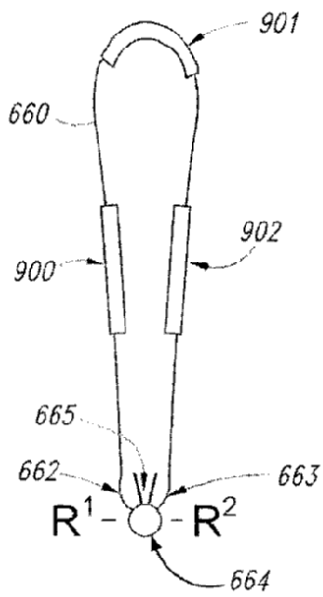
**Fig. 66B**



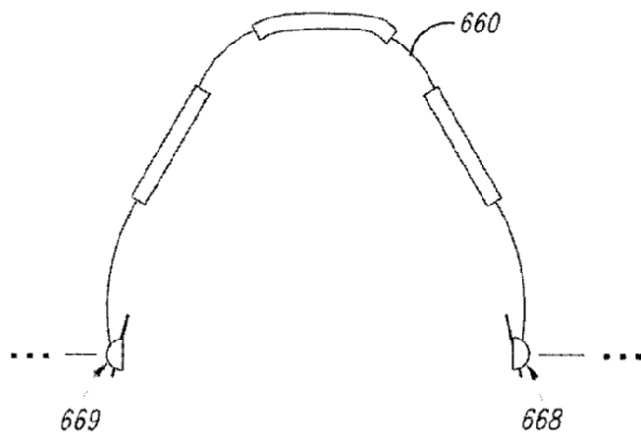
**Fig. 66C**



**Fig. 66D**



**Fig. 66E**





**Fig. 67**

