

(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS
ESPAÑA



(11) Número de publicación: **2 959 127**

(51) Int. Cl.:

C12Q 1/68 (2008.01)
C07H 21/02 (2006.01)
C07H 21/04 (2006.01)
C12Q 1/6806 (2008.01)
C12Q 1/6869 (2008.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Fecha de presentación y número de la solicitud europea: **19.06.2008 E 15164186 (7)**

(97) Fecha y número de publicación de la concesión europea: **05.07.2023 EP 2952587**

(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:
20.02.2024

(73) Titular/es:

F. HOFFMANN-LA ROCHE AG (100.0%)
Grenzacher Strasse 124
4070 Basel, CH

(72) Inventor/es:

KOKORIS, MARK STAMATIOS y
MCRUER, ROBERT N.

(74) Agente/Representante:

MILTENYI , Peter

Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 959 127 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 se refiere en general a la secuenciación de ácidos nucleicos, así como a productos 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 la vida. Por lo tanto, determinar dichas secuencias es una herramienta útil en la investigación pura sobre cómo y dónde viven los organismos, así como en las ciencias aplicadas, tal como el desarrollo de fármacos. En medicina, se pueden usar herramientas de secuenciación para el diagnóstico y para desarrollar tratamientos para diversas enfermedades, incluidas cáncer, enfermedades cardíacas, trastornos autoinmunitarios, 15 esclerosis múltiple u obesidad. En la industria, se puede usar la secuenciación para diseñar procedimientos enzimáticos mejorados u organismos sintéticos. En biología, dichas herramientas se pueden usar para estudiar la salud de los ecosistemas, por ejemplo, y, por tanto, tienen una amplia gama de utilidades.

- 20 La secuencia del ADN única de un individuo proporciona información valiosa sobre su predisposición a determinadas enfermedades. La secuencia proporcionará a los pacientes la oportunidad de someterse a un cribado para detección precoz y de recibir un tratamiento preventivo. Además, dado el modelo individual de un paciente, los médicos podrán administrar un tratamiento personalizado para maximizar la eficacia de un fármaco y para minimizar el riesgo de una respuesta adversa al fármaco. De forma similar, determinar el diseño de organismos patógenos puede dar lugar a nuevos tratamientos para enfermedades infecciosas y una vigilancia de patógenos más sólida. La secuenciación del ADN del 25 genoma completo proporcionará la base de la medicina moderna.

- 30 La secuenciación del ADN es el procedimiento de determinar el orden de los componentes químicos de un polímero de ADN dado. Estos componentes químicos, que se llaman nucleótidos, existen en el ADN en cuatro formas habituales: desoxiadenosina (A), desoxiguanosina (G), desoxicitidina (C) y desoxitimidina (T). La secuenciación de un genoma humano diploide requiere determinar el orden secuencial de aproximadamente 6 mil millones de nucleótidos.

- 35 Actualmente, la mayor parte de la secuenciación del ADN se realiza usando el procedimiento de terminación de cadena desarrollado por Frederick Sanger. Esta técnica, denominada secuenciación de Sanger, usa la terminación específica de la secuencia de la síntesis de ADN y sustratos indicadores de nucleótidos modificados con fluorescencia para obtener información de la secuencia. Este procedimiento secuencia una hebra de ácido nucleico diana, o una 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 base concretos (A, C, G o T) y las longitudes de las secuencias interrumpidas se determinan mediante electroforesis capilar en gel. A continuación, la longitud determina qué tipo de base está localizado en esa longitud. Se producen muchas longitudes de lectura superpuestas y sus 40 secuencias se superponen usando el procesamiento de datos para determinar el ajuste más fiable de los datos. Este procedimiento de producción de longitudes de lectura de secuencia es muy laborioso y costoso, y actualmente se está reemplazando por nuevos procedimientos que son más eficaces.

- 45 El procedimiento de Sanger se usó 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 tardó en completarse más de 10 años y costó casi 3 mil millones de dólares. Dadas estas importantes limitaciones de rendimiento y coste, está claro que las tecnologías de secuenciación de ADN deberán mejorar drásticamente para lograr los objetivos indicados establecidos por la comunidad científica. Con ese fin, una serie de tecnologías de segunda generación, que superan con creces las limitaciones de rendimiento y coste por base de la secuenciación de Sanger, están ganando una cuota cada vez mayor 50 en el mercado de la secuenciación. Aún así, estos procedimientos de "secuenciación por síntesis" no alcanzan los objetivos de rendimiento, coste y calidad requeridos por mercados tales como la secuenciación del genoma completo para la medicina personalizada.

- 55 Por ejemplo, 454 Life Sciences está produciendo instrumentos (por ejemplo, Genome Sequencer) que pueden 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, de cientos de bases de longitud, en la superficie de una microesfera. Este procedimiento se denomina PCR en emulsión. Se disponen a continuación cientos de miles de microesferas de este tipo en una "placa de picotitulación". La placa se prepara a continuación para una secuenciación adicional, con lo que cada tipo de base de ácido nucleico se lava secuencialmente sobre la placa. Las microesferas con diana que incorporan la base producen un subproducto de pirofosfato que se puede usar para catalizar una reacción que produce luz que a continuación se detecta con una cámara.

Illumina Inc. tiene un procedimiento similar que usa nucleótidos de terminación reversible y marcadores fluorescentes para realizar la secuenciación de ácidos nucleicos. La longitud de lectura promedio para el analizador 1G de Illumina es de menos de 40 nucleótidos. En lugar de usar la PCR en emulsión para amplificar las dianas de secuencia, Illumina tiene un enfoque para amplificar las colonias de PCR en una superficie de matriz. Tanto el enfoque de 454 como el de Illumina usan una amplificación de polimerasa complicada para incrementar la intensidad de señal, realizar mediciones de bases durante el ciclo de extensión de secuencia limitante de velocidad y tienen longitudes de lectura limitadas debido a errores de incorporación que degradan la señal de medición a ruido proporcionalmente a la longitud de lectura.

Applied Biosystems usa ligación de terminación reversible en lugar de secuenciación por síntesis para leer el ADN. Al igual que Genome Sequencer de 454, la tecnología usa PCR en emulsión basada en microesferas para amplificar la muestra. Dado que la mayoría de las microesferas no llevan productos de PCR, los investigadores usan a continuación una etapa de enriquecimiento para seleccionar microesferas recubiertas con ADN. Las microesferas recubiertas de biotina se extienden e inmovilizan en una matriz de portaobjetos de vidrio cubierta con estreptavidina. Las microesferas inmovilizadas se someten a continuación a un procedimiento de hibridación con sonda de octómeros (cada una marcada con cuatro tintes fluorescentes diferentes), ligación y escisión (entre las bases 5.^a y 6.^a para crear un sitio para la próxima ronda de ligación). Cada sonda explora dos bases, en las posiciones 4 y 5, usando un sistema de codificación de 2 bases, que es grabado por una cámara. De forma similar al enfoque de Illumina, la longitud de lectura promedio para la plataforma SOLiD de Applied Biosystems es de menos de 40 nucleótidos.

Se están desarrollando otros enfoques para evitar el tiempo y el gasto de la etapa de amplificación de la polimerasa midiendo moléculas individuales de ADN directamente. Visigen Biotechnologies, Inc. está midiendo bases marcadas con fluorescencia a medida que se secuencian incorporando un segundo fluoróforo en una polimerasa de ADN genomanipulado y usando la transferencia de energía de resonancia de Forster (FRET) para la identificación de nucleótidos. Esta técnica se enfrenta a los desafíos de separación de las señales de bases que se separan por menos de un nanómetro y por una acción de incorporación de polimerasa que tendrá una variación estadística muy grande.

Un procedimiento que está desarrollado LingVitae secuencia ADNc insertado en vectores plasmídicos inmovilizados. El procedimiento usa una enzima de restricción de clase IIS para escindir el ácido nucleico diana y ligar un oligómero a la diana. Típicamente, uno o dos nucleótidos en el saliente 5' o 3' terminal generado por la enzima de restricción determinan cuál de una colección de oligómeros en la mezcla de ligación se añadirá al extremo cortado y pegajoso de la diana. Cada oligómero contiene secuencias de "señal" que identifican de forma única el(s) nucleótido(s) que reemplaza. A continuación, se repite el procedimiento de escisión y ligación. La nueva molécula se secuencia a continuación usando marcas específicas para los diversos oligómeros. El producto de este procedimiento se denomina "polímero de diseño" y siempre consiste en un ácido nucleico más largo que el que reemplaza (por ejemplo, una secuencia diana de dinucleótido se reemplaza por una secuencia de polinucleótido "aumentada" de hasta 100 pares de bases). Una ventaja de este procedimiento es que la hebra del producto bicatenario se puede amplificar si se desea. Una desventaja es que el procedimiento es necesariamente cíclico y se perdería la continuidad del molde si se hicieran múltiples cortes de restricción simultáneos.

La patente de EE. UU. n.º 7.060.440 concedida a Kless describe un procedimiento de secuenciación que implica incorporar oligómeros por polimerización con una polimerasa. Se usa una modificación del procedimiento de Sanger, con oligómeros finalizados en los extremos como sustratos, para construir escaleras de secuenciación mediante electroforesis en gel o cromatografía capilar. Si bien el acoplamiento de oligómeros por ligación de extremos es bien conocido, el uso de una polimerasa para acoplar oligómeros en un procedimiento dirigido por molde se usó con nuevas ventajas.

Se prevé que las técnicas de polimerización aumenten su potencia a medida que polimerasas modificadas (y ligasas) estén disponibles a través de genomanipulación y bioprospección, y ya se conocen procedimientos para la eliminación de la actividad exonucleasa por modificación de la polimerasa. Por ejemplo, la solicitud de patente de EE. UU. publicada 2007/0048748 concedida a Williams describe el uso de polimerasas mutantes para incorporar nucleótidos marcados con tinte y otros modificados. Los sustratos para estas polimerasas también incluyen nucleótidos marcados con γ-fosfato. Tanto la velocidad incrementada de incorporación como la reducción de la tasa de error se descubrieron con polimerasas químéricas y mutantes.

Además, se ha hecho un gran esfuerzo por equipos académicos e industriales para secuenciar ADN natural usando procedimientos no sintéticos. Por ejemplo, Agilent Technologies, Inc. junto con colaboradores de universidades están desarrollando un procedimiento de detección de una sola molécula que enhebra el ADN a través de un nanoporo para hacer mediciones a medida que pasa por él. Al igual que con Visigen y LingVitae, este procedimiento debe superar el problema de obtener eficazmente y con exactitud señales claras de nucleobases individuales separadas por dimensiones subnanométricas, así como el problema de desarrollar tamaños de poro reproducibles de tamaño similar. Como tal, la secuenciación directa del ADN por detección de sus partes constituyentes aún no se ha logrado en un procedimiento de alto rendimiento debido al pequeño tamaño de los nucleótidos en la cadena (aproximadamente 4 Angstroms de centro a centro) y la señal correspondiente a limitaciones de resolución de señal y ruido en los mismos. La detección directa se complica además por la estructura secundaria inherente del ADN, que no se alarga fácilmente en un polímero perfectamente lineal.

Si bien se han hecho avances significativos en el campo de la secuenciación del ADN, continúa existiendo la necesidad

en la técnica de procedimientos nuevos y mejorados. La presente invención satisface estas necesidades y proporciona otras ventajas relacionadas.

5 El documento US 2002/028458 A1 divulga un procedimiento de secuenciación de la totalidad o parte de una molécula de ácido nucleico diana determinando la secuencia de una parte de dicha molécula de ácido nucleico además de información con respecto a la posición de dicha parte y, en particular, proporciona un nuevo procedimiento de secuenciación que implica la ampliación de una o más bases de dichas bases para ayudar en la identificación.

Breve sumario

10 En términos generales, se describen procedimientos y dispositivos y productos correspondientes que superan los desafíos de resolución espacial que presentan las 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 "Xpandómero") se forma por síntesis dirigida 15 por molde que conserva la información genética original del ácido nucleico diana, al mismo tiempo que incrementa también la separación lineal de los elementos individuales de los datos de secuencia.

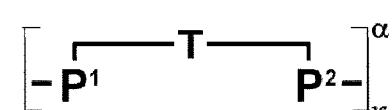
La invención se refiere a un sustrato monomérico como se define en la reivindicación 1, una hebra hija híbrida como se define en la reivindicación 2 y un Xpandómero como se define en la reivindicación 3. Se definen modos de realización 20 preferentes del sustrato monomérico, la hebra hija híbrida y el Xpandómero en las reivindicaciones dependientes.

Se divulga un procedimiento para secuenciar 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 todos o una parte del ácido nucleico diana, 25 en la que las subunidades individuales comprenden una atadura, al menos una sonda o residuo de nucleobase, y al menos un enlace escindible selectivamente; b) escindir el al menos un enlace escindible selectivamente para proporcionar un Xpandómero de una longitud mayor que la pluralidad de las subunidades de la hebra hija, comprendiendo el Xpandómero las ataduras y los elementos indicadores para analizar la información genética en una secuencia correspondiente a la secuencia de nucleótidos contigua de la totalidad o una parte del ácido nucleico diana; y c) detectar 30 los elementos indicadores del Xpandómero.

En aspectos más específicos, los elementos indicadores para analizar la información genética se pueden asociar con las ataduras del Xpandómero, con la hebra hija antes de la escisión del al menos un enlace escindible selectivamente, y/o con el Xpandómero después de la escisión del al menos un enlace escindible selectivamente. El Xpandómero puede 35 comprender además la totalidad o una parte de la al menos una sonda o residuo de nucleobase, y los elementos indicadores para analizar la información genética se pueden asociar con al menos una sonda o residuo de nucleobase o pueden ser la propia sonda o residuos de nucleobase. Además, el enlace escindible selectivamente puede ser un enlace covalente, un enlace intraatadura, un enlace entre o dentro de 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.

40 Los Xpandómeros divulgados en el presente documento tienen las siguientes estructuras (I) a (X):

Estructura (I) (no reivindicada)



45 en la que

T representa la atadura;

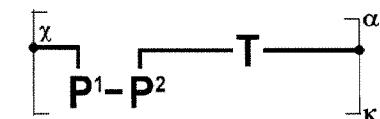
50 P¹ representa un primer resto de sonda;

P² representa un segundo resto de sonda;

55 κ representa la κésima subunidad en una cadena de m subunidades, donde m es un número entero mayor que tres; y

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

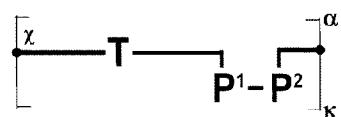
Estructura (II) (no reivindicada)



en la que

- 5 T representa la atadura;
 P¹ representa un primer resto de sonda;
 10 P² representa un segundo resto de sonda;
 kappa representa la $\kappa^{\text{ésima}}$ subunidad en una cadena de m subunidades, donde m es un número entero mayor que tres;
 15 alpha representa una especie de un motivo de subunidad seleccionado de una colección de motivos de subunidad, en el que cada una de las especies comprende información de secuencia de la secuencia de nucleótidos contigua de una parte del ácido nucleico diana; y
 x representa un enlace con la atadura de una subunidad adyacente.

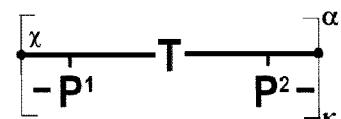
20 Estructura (III) (no reivindicada)



en la que

- 25 T representa la atadura;
 P¹ representa un primer resto de sonda;
 30 P² representa un segundo resto de sonda;
 kappa representa la $\kappa^{\text{ésima}}$ subunidad en una cadena de m subunidades, donde m es un número entero mayor que tres;
 35 alpha representa una especie de un motivo de subunidad seleccionado de una colección de motivos de subunidad, en el que cada una de las especies comprende información de secuencia de la secuencia de nucleótidos contigua de una parte del ácido nucleico diana; y
 x representa un enlace con la atadura de una subunidad adyacente.

40 Estructura (IV) (no reivindicada)



en la que

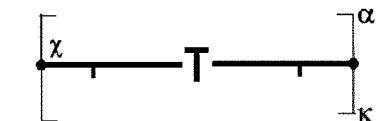
- 45 T representa la atadura;
 P¹ representa un primer resto de sonda;
 P² representa un segundo resto de sonda;
 50 kappa representa la $\kappa^{\text{ésima}}$ subunidad en una cadena de m subunidades, donde m es un número entero mayor que tres;

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

5

χ representa un enlace con la atadura de una subunidad adyacente.

Estructura (V) (no reivindicada)



10

en la que

T representa la atadura;

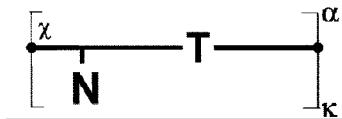
15 κ representa la κ ésmia subunidad en una cadena de m subunidades, donde m es un número entero mayor que tres;

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

20

χ representa un enlace con la atadura de una subunidad adyacente.

Estructura (VI) (no reivindicada)



25

en la que

T representa la atadura;

30 N representa un residuo de nucleobase;

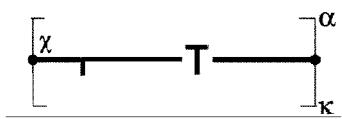
κ representa la κ ésmia subunidad en una cadena de m subunidades, donde m es un número entero mayor que diez;

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

χ representa un enlace con la atadura de una subunidad adyacente.

40

Estructura (VII) (no reivindicada)



en la que

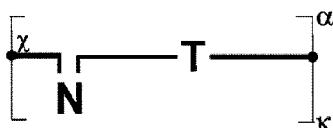
45 T representa la atadura;

κ representa la κ ésmia subunidad en una cadena de m subunidades, donde m es un número entero mayor que diez;

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

- 5 χ representa un enlace con la atadura de una subunidad adyacente.

Estructura (VIII) (no reivindicada)



- 10 en la que

T representa la atadura;

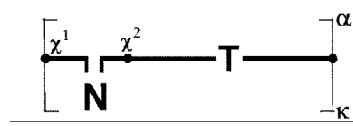
N representa un residuo de nucleobase;

- 15 κ representa la κ ^{ésima} subunidad en una cadena de m subunidades, donde m es un número entero mayor que diez;

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

χ representa un enlace con la atadura de una subunidad adyacente.

Estructura (IX) (no reivindicada)



- 25 en la que

T representa la atadura;

- 30 N representa un residuo de nucleobase;

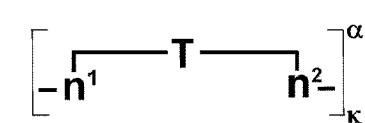
κ representa la κ ^{ésima} subunidad en una cadena de m subunidades, donde m es un número entero mayor que diez;

35 α representa una especie de un motivo de subunidad seleccionado de una colección de motivos de subunidad, en el que cada una de las especies comprende información de secuencia de la secuencia de nucleótidos contigua de una parte del ácido nucleico diana;

40 χ^1 representa un enlace con la atadura de una subunidad contigua; y

χ^2 representa un enlace entre ataduras.

Estructura (X)



- 45 en la que

- T representa la atadura;
- n¹ y n² representan una primera parte y una segunda parte, respectivamente, de un residuo de nucleobase;
- 5 κ representa la κ ^{ésima} subunidad en una cadena de m subunidades, donde m es un número entero mayor que diez; y α representa una especie de un motivo de subunidad seleccionado de una colección de motivos de subunidad, en el que cada una de las especies comprende información de secuencia de la secuencia de nucleótidos contigua de una parte del ácido nucleico diana.
- 10 Además, se divultan construcciones de sustrato oligoméricas para su uso en una síntesis dirigida por molde para secuenciar un ácido nucleico diana. Las construcciones de sustrato oligoméricas comprenden un primer resto de sonda unido a un segundo resto de sonda, teniendo cada uno de los restos de sonda primero y segundo un grupo terminal adecuado para la síntesis dirigida por molde, y una atadura que tiene un primer extremo y un segundo extremo con al menos el primer extremo de la atadura unido a al menos uno de los restos de sonda primero y segundo, en las que la construcción de sustrato oligomérica cuando se usa en la síntesis dirigida por molde puede formar una cadena hija que comprende un Xpandómero restringido y que tiene una pluralidad de subunidades acopladas en una secuencia correspondiente a la secuencia de nucleótidos contigua de la totalidad o una parte del ácido nucleico diana, en las que las subunidades individuales comprenden una atadura, los restos de sonda primero y segundo y al menos un enlace escindible selectivamente.
- 15
- Además, se divultan construcciones de sustrato monoméricas para su uso en una síntesis dirigida por molde para secuenciar un ácido nucleico diana. Las construcciones de sustrato monomérico comprenden un residuo de nucleobase con grupos terminales adecuados para la síntesis dirigida por molde, y una atadura que tiene un primer extremo y un segundo extremo con al menos el primer extremo de la atadura unido al residuo de nucleobase, en las que la construcción de sustrato monomérico cuando se usa en la síntesis dirigida por molde puede formar una hebra hija que comprende un Xpandómero restringido y que tiene una pluralidad de subunidades acopladas en una secuencia correspondiente a la secuencia de nucleótidos contigua de la totalidad o una parte del ácido nucleico diana, en las que las subunidades individuales comprenden una atadura, el residuo de nucleobase y al menos un enlace escindible selectivamente.
- 20
- Además, se divultan híbridos de molde-hebra hija que comprenden una cadena hija hibridada con una hebra molde, así como procedimientos para formar los mismos a partir de la hebra molde y las construcciones de sustrato oligoméricas o monoméricas.
- 25
- 30 Estos y otros aspectos de la invención serán evidentes por referencia a los dibujos adjuntos y a la siguiente descripción detallada. Con este fin, se exponen en el presente documento diversas referencias que describen con mayor detalle determinados procedimientos, compuestos y/o composiciones.
- 35
- Breve descripción de los dibujos**
- 40 En las figuras, números de referencia idénticos identifican elementos similares. Los tamaños y las posiciones relativas de los elementos de las figuras no están necesariamente dibujados a escala y algunos de estos elementos están arbitrariamente agrandados y situados para mejorar la legibilidad de las figuras. Además, las formas particulares de los elementos como se dibujan no pretenden transmitir ninguna información con respecto a la forma real de los elementos particulares, y se han seleccionado únicamente para facilitar su reconocimiento en las figuras.
- 45
- Las figuras 1A y 1B ilustran la separación limitada entre nucleobases que se debe resolver para determinar la secuencia de nucleótidos en un ácido nucleico diana.
- 50 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 a partir de un ácido nucleico diana.
- 55 La figura 4 es un modelo simple que ilustra un dispositivo de tipo nanoporo por FRET para secuenciar un Xpandómero.
- La figura 5 es un gráfico con canales para emisiones de fluorescencia roja, verde y azul e ilustra cómo las señales analógicas se pueden decodificar en información digital que corresponde a la información de la secuencia genética codificada en un Xpandómero. La tabla adjunta (figura 6) muestra cómo se decodifican los datos. Empleando tres fluoróforos multiestado, la secuencia de bases se puede leer con alta resolución en forma digital a partir de una sola molécula de manera ultrarrápida a medida que el Xpandómero se enrolla a través del nanoporo.
- 60
- La figura 6 es una tabla de consulta de la que se derivan los datos de la figura 5.
- 65 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.
- 5 Las figuras 10A a 10E representan Xpandómeros de clase I, intermedios y precursores en una forma simbólica y gráfica. Estos precursores se denominan Xsondas si son monofosfatos o Xmeros si son trifosfatos.
- La figura 11 es un esquema condensado de un procedimiento para la síntesis de un Xpandómero por ligación en solución usando un cebador de horquilla finalizado en el extremo y construcciones de sustrato de clase I.
- 10 La figura 12 es un esquema condensado de un procedimiento para la síntesis de un Xpandómero por ligación en solución usando un cebador de horquilla de dos extremos y construcciones de sustrato de clase I.
- La figura 13 es un esquema condensado de un procedimiento para la síntesis de un Xpandómero por ligación en un molde inmovilizado, sin cebadores, usando construcciones de sustrato de clase I.
- 15 La figura 14 es un esquema condensado de un procedimiento para la síntesis de un Xpandómero por ligación gradual cíclica usando construcciones de sustrato de clase I finalizadas reversiblemente en moldes hibridados con cebadores inmovilizados.
- 20 La figura 15 es un esquema condensado de un procedimiento para la síntesis de un Xpandómero por ensamblaje indiscriminado y acoplamiento químico usando construcciones de sustrato de clase I sin cebadores.
- 25 La figura 16 es un esquema condensado de un procedimiento para la síntesis de un Xpandómero por polimerización en solución usando un cebador de horquilla y construcciones de sustrato de trifosfato de clase I.
- La figura 17 es un esquema condensado de un procedimiento para la síntesis de un Xpandómero en un molde inmovilizado usando construcciones de sustrato de trifosfato de clase I y una polimerasa.
- 30 Las figuras 18A a 18E representan un Xpandómero de clase II, un intermedio de Xpandómero y una construcción de sustrato en un lenguaje simbólico y gráfico.
- Las figuras 19A a 19E representan un Xpandómero de clase III, un intermedio de Xpandómero y una construcción de sustrato en un lenguaje simbólico y gráfico.
- 35 La figura 20 es un esquema condensado de un procedimiento para la síntesis de un Xpandómero en 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 procedimiento para la síntesis de un Xpandómero en un molde inmovilizado usando un cebador, construcciones de sustrato de clase II y una ligasa.
- 40 Las figuras 22A a 22E representan un Xpandómero de clase IV, un intermedio de Xpandómero y una construcción de sustrato en forma simbólica y gráfica.
- 45 Las figuras 23A a 23E representan un Xpandómero de clase V, un intermedio de Xpandómero y una construcción de sustrato en forma simbólica y gráfica.
- La figura 24 es un esquema condensado de un procedimiento para la síntesis de un Xpandómero por polimerización en solución usando un cebador de adaptador y construcciones de sustrato de trifosfato de clase V.
- 50 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.
- 55 Las figuras 27A y 27B ilustran miembros de sonda que incorporan nucleobases derivatizadas.
- Las figuras 28A a 28D ilustran con más detalle los sustratos de clase I-IV de la invención, mostrándose aquí ejemplos de sitios de escisión de enlaces escindibles selectivamente en la cadena principal de la sonda e indicándose enlaces de extremo de bucle que forman un puente entre los sitios de escisión.
- 60 Las figuras 29A a 29D ilustran con más detalle los sustratos de clase I-IV de la invención, mostrándose aquí ejemplos de sitios de escisión de enlaces escindibles selectivamente en la cadena principal de la sonda e indicándose enlaces de extremo de bucle que forman un puente entre los sitios de escisión.
- 65 La figura 30 ilustra un procedimiento de ensamblaje de una construcción "sonda-bucle", tal como una Xsonda o un Xmero.

- La figura 31 ilustra un procedimiento para ensamblar una construcción de sustrato de clase I, donde el bucle contiene construcciones de indicadores.
- Las figuras 32A a 32C ilustran el uso de PEG como atadura polimérica.
- 5 Las figuras 33A a 33D ilustran la polilisina como una atadura polimérica y construcciones dendriméricas derivadas de estructuras de polilisina.
- 10 Las figuras 34A a 34C ilustran procedimientos seleccionados de cierre de bucle de atadura integrados con ensamblaje de construcción de indicadores a partir de segmentos.
- Las figuras 35A y 35B ilustran un procedimiento para la síntesis de un segmento indicador individual por polimerización aleatorizada de bloques de precursores.
- 15 Las figuras 36A a 36I ilustran construcciones de indicadores.
- La figura 37 es una tabla que muestra la composición y los procedimientos químicos para ensamblar construcciones de indicadores y sus correspondientes códigos de indicadores.
- 20 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 adaptador para introducir un grupo funcional ANH terminal en un molde de ADN bicatenario.
- 25 La figura 40 es un esquema para inmovilizar y preparar un molde para síntesis de un Xpandómero.
- Las figuras 41A a 41E ilustran procedimientos de estiramiento físico seleccionados.
- 30 Las figuras 42A a 42C ilustran procedimientos de electroestiramiento seleccionados.
- Las figuras 43A a 43D ilustran procedimientos, reactivos y adaptadores para estirar en matrices de gel.
- 35 Las figuras 44A a 44C describen la construcción y el uso de "marcas de arrastre".
- La figura 45 describe un procedimiento basado en hibridación indiscriminada/ligación para sintetizar un Xpandómero.
- Las figuras 46A y B describen nucleobases usadas para llenar huecos.
- 40 Las figuras 47A y B describen simulaciones de existencia de huecos.
- La figura 49 ilustra cómo se llenan huecos con dímeros y trímeros.
- 45 Las figuras 50A y B describen simulaciones de llenado de huecos usando combinaciones de dímeros y trímeros.
- La figura 51 ilustra el uso de adyuvantes diméricos y triméricos para romper una estructura secundaria.
- 50 La figura 52 describe bases útiles como adyuvantes.
- La figura 53 describe sustituciones de nucleótidos usadas para reducir una estructura secundaria.
- La figura 54 describe un modelo de detección de nanoporos con transporte de microesfera magnéticas.
- 55 La figura 55 ilustra un procedimiento convencional de detección de nanoporos.
- La figura 56 ilustra un procedimiento de detección de nanoporos con electrodos transversales.
- 60 La figura 57 ilustra un procedimiento de detección microscópica.
- La figura 58 ilustra la detección por microscopía electrónica.
- La figura 59 ilustra la detección usando microscopía de fuerza atómica.
- 65 Las figuras 60A a 60E representan un Xpandómero de clase VI, un intermedio de Xpandómero y una construcción de sustrato en un lenguaje simbólico y gráfico. Estos precursores se denominan RT-NTP.

La figura 61 es un esquema condensado de un procedimiento para la síntesis de un Xpandómero en un molde inmovilizado usando construcciones de sustrato de trifosfato de clase VI finalizadas reversiblemente y una polimerasa.

- 5 Las figuras 62A a 62E representan un Xpandómero de clase VII, un intermedio de Xpandómero y una construcción de sustrato en un lenguaje simbólico y gráfico. Estos precursores se denominan RT-NTP.

Las figuras 63A a 63E representan un Xpandómero de clase VIII, un intermedio de Xpandómero y una construcción de sustrato en un lenguaje simbólico y gráfico. Estos precursores se denominan RT-NTP.

- 10 10 Las figuras 64A a 64E representan un Xpandómero de clase IX, un intermedio de Xpandómero y una construcción de sustrato en un lenguaje simbólico y gráfico. Estos precursores se denominan RT-NTP.

15 La figura 65 es un esquema condensado de un procedimiento para la síntesis de un Xpandómero en un molde inmovilizado usando construcciones de sustrato de trifosfato de clase IX y una polimerasa.

Las figuras 66A a 66E representan un Xpandómero de clase X, un intermedio de Xpandómero y una construcción de sustrato en un lenguaje simbólico y gráfico. Estos precursores se denominan XNTP.

- 20 20 La figura 67 es un esquema condensado de un procedimiento para la síntesis de un Xpandómero por polimerización en solución usando un cebador de horquilla y construcciones de sustrato de trifosfato de clase X.

Descripción detallada

- 25 En la siguiente descripción, se exponen determinados detalles específicos para proporcionar una comprensión completa de diversos aspectos y modos de realización. Sin embargo, un experto en la técnica comprenderá que la invención se puede poner en práctica sin estos detalles. En otros ejemplos, no se han mostrado o descrito en detalle estructuras bien conocidas para evitar eclipsar descripciones de los modos de realización. A menos que el contexto lo requiera de otro modo, a lo largo de la memoria descriptiva y las reivindicaciones que siguen, la palabra "comprender" y las variaciones 30 de la misma, tales como "comprende" y "que comprende" se deben interpretar en un sentido abierto e inclusivo, es decir, como "incluyendo, pero sin limitaciones". Además, los encabezados proporcionados en el presente documento son solo por comodidad y no interpretan el alcance o el significado de la invención reivindicada.

35 La referencia a lo largo de esta memoria descriptiva a "un modo de realización" significa que un rasgo característico, estructura o característica particular descrito en relación con el modo de realización se incluye en al menos un modo de realización. Por tanto, las apariciones de la expresión "en un modo de realización" a lo largo de esta memoria descriptiva pueden, pero no necesariamente, hacer referencia al mismo modo de realización. Además, los rasgos característicos, estructuras, funcionamientos o características particulares se pueden combinar de cualquier manera adecuada en uno o más modos de realización. Además, como se usa en esta memoria descriptiva y en las reivindicaciones adjuntas, las 40 formas en singular "un", "uno/a" y "el/la" incluyen las referencias al plural a menos que el contenido lo indique claramente de otro modo. Cabe destacar también que el término "o" se emplea en general en su sentido que incluye "y/o" a menos que el contenido lo indique claramente lo contrario.

Definiciones

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

50 "Nucleobase" es una base heterocíclica tal como adenina, guanina, citosina, timina, uracilo, inosina, xantina, hipoxantina o un derivado heterocíclico, análogo o tautómero de las mismas. Una nucleobase puede ser 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-(C3-C6)-alquinilcitosina, 5-fluorouracilo, 5-bromouracilo, tiouracilo, pseudouracilo, isocitosina, isoguanina, inosina, 7,8-dimetilaloxazina, 6-dihidrotimina, 5,6-dihidrouracilo, 4-metil-indol, etenoadenina y las nucleobases no naturales descritas en las patentes de EE. UU. n.º 5.432.272 y 6.150.510 y las solicitudes PCT WO92/002258, WO93/10820, WO94/22892 y WO94/24144, y Fasman ("Practical Handbook of Biochemistry and Molecular Biology", pp. 385-394, 1989, CRC Press, Boca Raton, LO).

- 60 "Residuo de nucleobase" incluye nucleótidos, nucleósidos, fragmentos de los mismos y moléculas relacionadas que tienen la propiedad de unirse a un nucleótido complementario. Los 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.

65 Los "polinucleótidos", también llamados ácidos nucleicos, son series de nucleótidos enlazados covalentemente en los que la posición 3' de la pentosa de un nucleótido está unida por un grupo fosfodiéster a la posición 5' del siguiente. El ADN

(ácido desoxirribonucleico) y el ARN (ácido ribonucleico) son polinucleótidos biológicos en los que los residuos de nucleótido están enlazados en una secuencia específica por enlaces fosfodiéster. Como se usa en el presente documento, los términos "polinucleótido" u "oligonucleótido" engloban cualquier compuesto polimérico que tenga una cadena principal lineal de nucleótidos. Los oligonucleótidos, también denominados oligómeros, son, en general, polinucleótidos de cadena más corta.

"Complementario" se refiere, en general, a la hibridación de nucleótidos específicos para formar pares de bases de Watson-Crick canónicos, como entienden los expertos en la técnica. Sin embargo, complementario como se denomina en el presente documento también incluye emparejamiento de bases de análogos de nucleótido, que incluyen, pero sin limitaciones, 2'-desoxiinosina y 5-nitroindol-2'-desoxirribósido, que pueden emparejar bases universales con nucleótidos de A, T, G o C y ácidos nucleicos bloqueados, que potencian la estabilidad térmica de los híbridos. Un experto en la técnica reconocerá que la restricción de la hibridación es un factor determinante en el grado de emparejamiento o emparejamiento erróneo en el híbrido formado por hibridación.

15 "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, en general, "ácidos nucleicos diana" o "secuencia diana" si son seleccionados para secuenciación. Los ácidos nucleicos pueden ser mezclas o grupos de moléculas seleccionados para secuenciación.

20 "Sonda" es una hebra corta de residuos de nucleobase, que se refiere, en general, a dos o más residuos de nucleobase contiguos que, en general, son monocatenarios y complementarios a una secuencia diana de un ácido nucleico. Como se abarca en "miembros de sustrato" y "construcciones de sustrato", las sondas pueden tener una longitud de hasta 20 residuos de nucleobase. Las sondas pueden incluir residuos de nucleobase modificados y enlaces intranucleobase modificados en cualquier combinación. Las estructuras principales de las sondas se pueden enlazar entre sí por cualquiera 25 de una serie de tipos de enlaces covalentes, que incluyen, pero sin limitaciones, éster, fosfodiéster, fosforamida, fosfonato, fosforotioato, fosforotiolato, enlace amida y cualquier combinación de los mismos. La sonda también puede tener enlaces en los extremos 5' y 3' que incluyen, pero sin limitaciones, los siguientes restos: monofosfato, trifosfato, hidroxilo, hidrógeno, éster, éter, glicol, amina, amida y tioéster.

30 "Hibridación selectiva" se refiere a una unión complementaria específica. Los polinucleótidos, oligonucleótidos, sondas, residuos de nucleobase y fragmentos de los mismos hibridan selectivamente con hebras de ácido nucleico diana, en condiciones de hibridación y lavado que minimizan una unión inespecífica. Como es conocido en la técnica, se pueden usar condiciones de alta restricción para lograr condiciones de hibridación selectiva que favorezcan un emparejamiento perfecto. Las condiciones para la hibridación, tal como concentración de sal, temperatura, detergentes, PEG y agentes 35 neutralizantes de GC tal como betaína, se pueden variar para incrementar la restricción de la hibridación, es decir, el requisito de emparejamientos exactos de C con un par de bases con G, y A con un par de bases con T o U, a lo largo de una hebra contigua de un ácido nucleico híbrido.

40 "Síntesis dirigida por molde", "ensamblaje dirigido por molde", "hibridación dirigida por molde", "unión dirigida por molde" y cualquier otro procedimiento dirigido por molde, se refieren a un procedimiento con el que los residuos de nucleobase o las sondas se unen selectivamente a un ácido nucleico diana complementario, y se incorporan a una hebra hija emergente. Una hebra hija producida por una síntesis dirigida por molde es complementaria a la diana monocatenaria a partir del cual se sintetiza. Cabe señalar que la secuencia correspondiente de una hebra diana se puede inferir de la secuencia de su 45 hebra hija, si es conocida. La "polimerización dirigida por molde" y la "ligación dirigida por molde" son casos especiales de síntesis dirigida por molde, con lo que la hebra hija resultante se polimeriza o liga, respectivamente.

"Contiguo/a" indica que una secuencia continúa sin interrupción o nucleobase perdida. Se dice que la secuencia contigua de nucleótidos de la hebra molde es complementaria a la secuencia contigua de la hebra hija.

50 Los "sustratos" o "miembros de sustrato" son oligómeros, sondas o residuos de nucleobase que tienen especificidad de unión por el molde diana. En general, los sustratos se combinan con ataduras para formar construcciones de sustrato. Los sustratos de construcciones de sustrato que forman la cadena principal primaria de la hebra hija también son sustratos o miembros de sustrato de la hebra hija.

55 Las "construcciones de sustrato" son reactivos para síntesis dirigida por molde de cadenas hijas, y en general se proporcionan en forma de colecciones. Las construcciones de sustrato contienen, en general, un miembro de sustrato para unión complementaria a un molde diana y un miembro de atadura o sitios de unión de atadura a los que se puede unir una atadura. 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 denominadas "construcciones de sustrato de sonda") como "construcciones de sustrato monomérico" (también denominadas "construcciones de sustrato de nucleobase").

60 "Motivo de subunidad" o "motivo" se refiere a una subunidad repetitiva de una cadena principal de polímero, en la que la subunidad tiene una forma general característica de las subunidades repetitivas, pero también tiene elementos específicos de especie que codifican información genética. Los motivos de residuos de nucleobase complementarios se representan en colecciones de construcciones de sustrato de acuerdo con el número de combinaciones posibles de los elementos de

nucleobase de unión a secuencias complementarias básicas en cada motivo. Si los elementos de unión a nucleobase son cuatro (por ejemplo, A, C, G y T), el número de posibles motivos de combinaciones de cuatro elementos es 4^x , donde x es el número de residuos de nucleobase en el motivo. Sin embargo, otros motivos basados en bases de emparejamiento degeneradas, en la sustitución de timidina por uracilo en residuos de ribonucleobase u otros conjuntos de residuos de nucleobase, pueden dar lugar a colecciones más grandes (o colecciones más pequeñas) de construcciones de sustrato que portan motivos. Los motivos también están representados por construcciones de indicadores específicos de especie, tal como los indicadores que componen una atadura de indicadores. En general, existe una correlación unívoca entre el motivo de la construcción de indicadores que identifica una especie de sustrato particular y la complementariedad de unión y especificidad del motivo.

10 El "intermedio de Xpandómero" es un producto intermedio (también denominado en el presente documento "cadena hija") ensamblado a partir de construcciones de sustrato, y está formado 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 colindantes que pueden incluir polimerización o ligación de los sustratos, enlaces de atadura a atadura o enlaces de atadura a sustrato. El intermedio de Xpandómero contiene dos estructuras; a saber, el Xpandómero restringido y la cadena principal primaria. El Xpandómero restringido comprende todas las ataduras en la hebra hija, pero puede comprender la totalidad, una parte o nada del sustrato según se requiera por el procedimiento. La cadena principal primaria comprende todos los sustratos colindantes. Durante la etapa del procedimiento en la que la cadena principal primaria se fragmenta o disocia, el Xpandómero restringido ya no está restringido y es el producto Xpandómero que se extiende a medida que se estiran las ataduras. "Cadena hija híbrida" se refiere a un intermedio de Xpandómero que hibrida con el molde diana.

15 "Cadena principal primaria" se refiere a una cadena principal contigua o segmentada de sustratos de la hebra hija. Una cadena principal primaria que se encuentra comúnmente es la cadena principal de ribosil 5'-3' fosfodiéster de un polinucleótido natural. Sin embargo, la cadena principal primaria de una hebra hija puede contener análogos de nucleobases y análogos de oligómeros no unidos por enlaces fosfodiéster o unidos por una mezcla de enlaces fosfodiéster y otros enlaces de la cadena principal, que incluyen, pero sin limitaciones, los siguientes enlaces: fosforotioato, fosforotiolato, fosfonato, fosforamidato y enlaces de cadena principal "PNA" de ácido peptidonucleico que incluyen fosfono-PNA, serina-PNA, hidroxiprolina-PNA y combinaciones de los mismos. Cuando la hebra hija está en su forma 20 híbrida (es decir, hebra hija híbrida), y los sustratos no están unidos covalentemente entre las subunidades, los sustratos son, no obstante, contiguos y forman la cadena principal primaria de la hebra hija.

25 "Xpandómero restringido" es un Xpandómero en una configuración antes de que se haya expandido. El Xpandómero restringido comprende todos los miembros de atadura de la hebra hija. Está restringido para expandirse por al menos una unión o enlace por atadura que se une a la cadena principal primaria. Durante el procedimiento de expansión, la cadena 30 principal primaria de la hebra hija se fragmenta o disocia para transformar el Xpandómero restringido en un Xpandómero.

30 "Cadena principal del Xpandómero restringido" se refiere a la cadena principal del Xpandómero restringido. Es una cadena principal covalente sintética coensamblada junto con la cadena principal primaria en la formación de la hebra hija. En algunos casos, ambas cadenas principales pueden no ser discretas, pero ambas pueden tener el mismo sustrato o porciones del sustrato en su composición. La cadena principal del Xpandómero restringido siempre comprende las ataduras, mientras que la cadena principal primaria no comprende miembros de atadura.

35 "Xpandómero" o "producto Xpandómero" es una construcción molecular sintética producida por expansión de un Xpandómero restringido, que a su vez se sintetiza por ensamblaje dirigido por molde de construcciones de sustrato. El Xpandómero se alarga en relación con el molde diana a partir de la cual se produjo. Se compone de una concatenación de subunidades, cada subunidad un motivo, cada motivo un miembro de una colección, que comprende información de secuencia, una atadura y, opcionalmente, una parte o la totalidad del sustrato, todas las cuales se obtienen de la construcción de sustrato formativa. El Xpandómero está diseñado para expandirse para ser más largo que el molde diana; de este modo, se reduce la densidad lineal de la información de la secuencia del molde diana a lo largo de su longitud. Además, el Xpandómero proporciona opcionalmente una plataforma para incrementar el tamaño y la abundancia de indicadores, lo que a su vez mejora la relación señal-ruido para la detección. Una densidad de información lineal menor y señales más fuertes incrementan la resolución y reducen los requisitos de sensibilidad para detectar y descodificar la secuencia de la hebra molde.

40 45 "Enlace escindible selectivamente" se refiere a un enlace que se puede romper en condiciones controladas tales como, por ejemplo, condiciones para escisión selectiva de un enlace fosforotiolato, un enlace fotoescindible, un enlace fosforamida, un enlace 3'-OBD-ribofuranosil-2', un enlace tioéter, un enlace selenoéter, un enlace sulfóxido, un enlace disulfuro, un enlace desoxiribosil-5'-3' fosfodiéster o un enlace ribosil-5'-3' fosfodiéster, así como otros enlaces escindibles conocidos en la técnica. Un enlace escindible selectivamente puede ser un enlace intraatadura 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 escindibles selectivamente no se limitan a enlaces covalentes, y pueden ser enlaces no covalentes o asociaciones, tales como los basados en enlaces de hidrógeno, enlaces hidrófobos, enlaces iónicos, interacciones de apilamiento de anillos de enlace pi, interacciones de Van der Waals y similares.

50 55 60 65 "Resto" es una de dos o más partes en las que se puede dividir algo, tal como, por ejemplo, las diversas partes de una

atadura, una molécula o una sonda.

- "Atadura" o "miembro de atadura" se refiere a un polímero o construcción molecular que tiene una dimensión en general lineal y con un resto de extremo en cada uno de los dos extremos opuestos. Una atadura se une a un sustrato con un enlace en al menos un resto de extremo para formar una construcción de sustrato. Los restos de extremo de la atadura se pueden conectar a enlaces escindibles con el sustrato o enlaces escindibles intraatadura que sirven para restringir la atadura en una "configuración restringida". Después de que se sintetiza la hebra hija, cada resto de extremo tiene un enlace de extremo que se acopla directa o indirectamente a otras ataduras. Las ataduras acopladas comprenden el Xpandómero restringido que comprende además la hebra hija. Las ataduras tienen una "configuración restringida" y una "configuración expandida". La configuración restringida se encuentra en construcciones de sustrato y en la hebra hija. La configuración restringida de la atadura es la precursora de la configuración expandida, como se encuentra en los productos Xpandómeros. La transición de la configuración restringida a la configuración expandida da como resultado la escisión de enlaces escindibles selectivamente que pueden estar dentro de la cadena principal primaria de la hebra hija o enlaces intraatadura. También se usa una atadura en una configuración restringida donde se añade una atadura para formar la hebra hija después del ensamblaje de la "cadena principal primaria". Las ataduras pueden comprender opcionalmente uno o más indicadores o construcciones de indicadores a lo largo de su longitud que pueden codificar información de secuencia de sustratos. La atadura proporciona un medio para expandir la longitud del Xpandómero, de modo que reduce la densidad lineal de información de la secuencia.
- "Construcciones de ataduras" son ataduras o precursores de ataduras compuestas por uno o más segmentos de atadura u otros componentes arquitectónicos para ensamblar ataduras tales como construcciones de indicadores o precursores de indicadores, incluyendo 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 atadura" o "segmento de atadura" es un polímero que tiene una dimensión en general lineal con dos extremos terminales, donde los extremos forman enlaces de extremo para concatenar los elementos de atadura. Los elementos de atadura pueden ser segmentos de construcciones de atadura. Dichos polímeros pueden incluir, pero sin limitaciones: polietilenglicoles, poliglicoles, polipirimidinas, poliisocianuros, poliisocianatos, poli(triarylmetil)metacrilatos, polialdehídos, polipirrolinonas, poliureas, fosfodiésteres de poliglicol, poliacrilatos, polimetacrilatos, poliacrilamidas, ésteres de polivinilo, poliestirenos, poliamidas, poliuretanos, policarbonatos, polibutiratos, polibutadienos, polibutirolactonas, polipirrolidinonas, polivinilfosfonatos, poliacetamidas, polisacáridos, polihialuranatos, poliamidas, poliimidás, poliésteres, polietilenos, polipropilenos, poliestirenos, policarbonatos, politereftalatos, polisilanos, poliuretanos, poliéteres, poliaminoácidos, poliglicinas, poliprolinas, polilisina N-sustituida, polipéptidos, péptidos N-sustituidos de cadena lateral, glicina poli-N-sustituida, peptoides, péptidos sustituidos con carboxilo de cadena lateral, homopéptidos, oligonucleótidos, oligonucleótidos de ácido ribonucleico, oligonucleótidos de ácido desoxinucleico, oligonucleótidos modificados para evitar el emparejamiento de bases de Watson-Crick, análogos de oligonucleótido, ácido policitidílico, ácido poliadenílico, ácido poliuridílico, politimidina, polifosfato, polinucleótidos, polirribonucleótidos, fosfodiésteres de polietilenglicol, análogos de polinucleótidos peptídicos, análogos de treosil-polinucleótido, análogos de glicol-polinucleótido, análogos de morfolino-polinucleótido, análogos de oligómero de nucleótido bloqueado, análogos de polipéptido, polímeros ramificados, polímeros de peine, polímeros de estrella, polímeros dendríticos, copolímeros aleatorios, de gradiente y de bloque, polímeros aniónicos, polímeros catiónicos, polímeros que forman tallo-bucles, segmentos rígidos y segmentos flexibles.
- "Ácido peptidonucleico" o "PNA" es un análogo de ácido nucleico que tiene residuos de nucleobase adecuados para hibridación con un ácido nucleico, pero con una cadena principal que comprende aminoácidos o derivados o análogos de los mismos.
- "Ácido fosfono-peptidonucleico" o "pPNA" es un ácido peptidonucleico en el que la cadena principal comprende análogos de aminoácido, tales como N-(2-hidroxietil)fosfonoglicina o N-(2-aminoetil)fosfonoglicina, y los enlaces entre unidades de nucleobase son a través de enlaces fosfonoéster o fosfonoamida.
- "Ácido nucleico de serina" o "SerNA" es un ácido peptidonucleico en el que la cadena principal comprende residuos de serina. Dichos residuos se pueden unir a través de enlaces amida o éster.
- "Ácido nucleico de hidroxiprolina" o "HypNA" es un ácido peptidonucleico en el que la cadena principal comprende residuos de 4-hidroxiprolina. Dichos residuos se pueden unir a través de enlaces amida o éster.
- "Elemento indicador" es un elemento de señalización, complejo molecular, compuesto, molécula o átomo que también está compuesto de una "característica de detección de indicador" asociada. Otros elementos indicadores incluyen, pero sin limitaciones, donante o acceptor resonante por FRET, colorante, punto cuántico, microesfera, dendrímero, fluoróforo de conversión ascendente, partícula magnética, dispersor de electrones (por ejemplo, boro), masa, microesfera de oro, resonancia magnética, grupo ionizable, grupo polar, grupo hidrófobo. Todavía otros son marcadores fluorescentes, tales como, pero sin limitaciones, 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, ficolcianina, aloficocianina, 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

5 Detection Systems, Inc.), eritrosina, ácido cumárico, umbelifera, texas red rodaina, tetrametil rodamina, Rox, 7-nitrobenzo-1-oxa-1-diazol (NBD), oxazol, tiazol, pireno, fluoresceína o lantamidas; también radioisótopos (tales como ^{33}P , ^3H , ^{14}C , ^{35}S , ^{125}I , ^{32}P o ^{131}I), etidio, europio, ruteno 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 masa pueden ser, por ejemplo, halógeno, éter o poliéster, alquilo, éster o poliéster, o del tipo general XR, en el que X es un grupo de enlace y R es un grupo modificador de masa, marcadores quimioluminiscentes, marcadores de espín, enzimas (tales como peroxidases, fosfatases alcalinas, beta-galactosidasas y oxidasa), fragmentos de anticuerpos y ligandos de afinidad (tal como un oligómero, haptene y aptámero). La asociación del elemento indicador con la atadura puede ser covalente o no covalente, y directa o indirecta. Las asociaciones covalentes representativas incluyen conector y enlaces sin conector. Se incluyen enlaces a la cadena principal de la atadura o a un elemento unido a atadura, tal como un dendrímero o una cadena lateral. Los enlaces no covalentes representativos incluyen enlaces de hidrógeno, enlaces hidrófobos, enlaces iónicos, apilamiento de anillos de enlaces pi, interacciones de Van der Waals y similares. Los ligandos, por ejemplo, se asocian por unión específica por afinidad con sitios de unión en el elemento indicador. La asociación directa puede tener lugar en el momento de la síntesis de la atadura, después de la síntesis de la atadura y antes o después de la síntesis del Xpandómero.

10 Un "indicador" se compone de uno o más elementos indicadores. Los indicadores incluyen lo que es conocido como "marcas" y "marcadores". La sonda o residuo de nucleobase del Xpandómero se puede considerar un indicador. Los indicadores sirven para analizar la información genética del ácido nucleico diana.

15 20 "Construcción de indicadores" comprende uno o más indicadores que pueden producir una señal(es) detectable(s), en la que la(s) señal(es) detectable(s) contiene(n), en general, información de secuencia. Esta información de señal se denomina "código de indicador" y posteriormente se descodifica en datos de secuencia genética. Una construcción indicadora también puede comprender segmentos de atadura u otros componentes arquitectónicos, incluyendo polímeros, copolímeros de injerto, copolímeros de bloque, ligandos de afinidad, oligómeros, haptenos, aptámeros, dendrímeros, grupos de unión o grupo de unión por afinidad (por ejemplo, biotina).

25 30 35 La "característica de detección de indicador", denominada "señal", describe todos los posibles elementos, propiedades o características medibles o detectables usados para comunicar la información de secuencia genética de un indicador directa o indirectamente a un dispositivo de medición. Estos incluyen, pero sin limitaciones, fluorescencia, fluorescencia de longitud de onda múltiple, extinción de fluorescencia del espectro de emisión, FRET, emisión, absorbancia, reflectancia, emisión de tinte, emisión de puntos cuánticos, imagen de microesfera, imagen de complejo molecular, 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, desintegración nuclear, resonancia o complementariedad.

40 45 "Código de indicador" es la información genética de una señal medida de una construcción de indicadores. El código de indicador se descodifica para proporcionar datos de información genética específicos de secuencia.

50 "Xsonda" es una construcción de sustrato oligomérico expandible. Cada Xsonda tiene un miembro de sonda y un miembro de atadura. El miembro de atadura tiene, en general, una o más construcciones de indicadores. Las Xsondas con modificaciones de 5'-monofosfato son compatibles con procedimientos basados en ligación enzimática para síntesis de Xpandómeros. Las Xsondas con modificaciones de conector en 5' y 3' son compatibles con procedimientos basados en ligación química para síntesis de Xpandómeros.

55 60 "Xmero" es una construcción de sustrato oligomérico expandible. Cada Xmero tiene un miembro de sustrato oligomérico y un miembro de atadura, teniendo, en general, el miembro de atadura una o más construcciones de indicadores. Los Xmeros son 5'-trifosfatos compatibles con procedimientos basados en polimerasa para sintetizar Xpandómeros.

65 "RT-NTP" es una construcción de sustrato de nucleótido modificado con 5' trifosfato expandible ("sustrato monomérico") compatible con polimerización enzimática dependiente de molde. Un RT-NTP tiene un desoxirribonucleótido trifosfato modificado ("DNTP"), un ribonucleótido trifosfato ("RNTP") o un sustrato análogo funcionalmente equivalente, denominados conjuntamente sustrato de nucleótido trifosfato ("NTPS"). Un RT-NTP tiene dos componentes funcionales distintos; a saber, una nucleobase 5'-trifosfato y una atadura o precursor de atadura. Después de la formación de la hebra hija, la atadura se une entre cada nucleótido en posiciones que permiten la expansión controlada de RT. En una clase de RT-NTP (por ejemplo, clase IX), la atadura se une después de la polimerización de RT-NTP. En algunos casos, el RT-NTP tiene un terminador de extremo reversible y una atadura que se reticula selectivamente directamente con ataduras contiguas. Cada atadura se puede codificar de forma exclusiva con indicadores que identifican específicamente el nucleótido al que está atado.

65 "XNTP" es un sustrato de nucleótido modificado con 5' trifosfato expandible compatible con polimerización enzimática dependiente de molde. Un XNTP tiene dos componentes funcionales distintivos; a saber, una nucleobase 5'-trifosfato y una atadura o precursor de atadura que se une dentro de cada nucleótido en posiciones que permiten la expansión controlada de RT por escisión intranucleotídica.

65 "Procesivo" se refiere a un procedimiento de acoplamiento de sustratos que, en general, es continuo y prosigue con

direccionalidad. Si bien no está vinculado por la teoría, tanto las ligasas como las polimerasas, por ejemplo, presentan un comportamiento procesivo si se añaden sustratos a una hebra hija emergente de forma incremental sin interrupción. Las etapas de hibridación y ligación, o hibridación y polimerización, no se consideran etapas independientes si el efecto neto es un crecimiento procesivo de la hebra hija emergente. Algunos pero no todos los procedimientos dependientes de cebador son procesivos.

5 "Indiscriminado" se refiere a un procedimiento de acoplamiento de sustratos que prosigue a partir de múltiples puntos en un molde a la vez, y no depende de cebador, e indica que la extensión de la cadena se produce en paralelo (simultáneamente) desde más de un punto de origen.

10 "Extensión de base única" se refiere a un procedimiento gradual cíclico en el que se añaden sustratos monoméricos uno a uno. En general, la reacción de acoplamiento se restringe para que proceda más allá de la extensión de un solo sustrato en cualquier etapa mediante el uso de grupos de bloqueo reversibles.

15 "Extensión de sonda única" se refiere a un procedimiento gradual cíclico en el que se añaden sustratos oligoméricos uno a uno. En general, la reacción de acoplamiento se restringe para que proceda más allá de la extensión de un solo sustrato en cualquier etapa mediante el uso de grupos de bloqueo reversibles.

20 "Corresponde a" o "correspondiente" se usa aquí en referencia a una secuencia monocatenaria contigua de una sonda, oligonucleótido, análogo de oligonucleótido o cadena hija que es complementario y, por tanto, "corresponde a" la totalidad o una parte de una secuencia de ácido nucleico diana. Se puede decir que la secuencia complementaria de una sonda corresponde a su diana. A menos que se indique de otro modo, tanto la secuencia complementaria de la sonda como la secuencia complementaria de la diana son secuencias contiguas individualmente.

25 "Resistente a nucleasa" se refiere a un enlace que es resistente a una enzima nucleasa en condiciones donde, en general, se escindirá un enlace fosfodiéster de ADN o ARN. Las enzimas nucleasas incluyen, pero sin limitaciones, DNase I, exonucleasa III, Mung Bean Nuclease, RNase I y RNase H. Un experto en esta técnica puede evaluar fácilmente la resistencia relativa a nucleasa de un enlace dado.

30 "Ligasa" es una enzima, en general, para unir nucleótidos de 3'-OH5'-monofosfato, oligómeros y sus análogos. Las ligasas incluyen, pero sin limitaciones, ligasas dependientes de NAD⁺ que incluyen ARN 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 de tipo VanC, ADN ligasa 9°N, ADN ligasa Tsp y ligasas novedosas descubiertas por bioprospección. Las ligasas también incluyen, pero sin limitaciones, ligasas dependientes de ATP que incluyen ARN ligasa T4, ADN ligasa T4, ADN ligasa T7, ADN ligasa Pfu, ADN ligasa ADN I, ADN ligasa ADN III, ADN ligasa IV y ligasas novedosas descubiertas por bioprospección. Estas ligasas incluyen isoformas naturales, mutantes y variantes genomanipuladas.

40 "Polimerasa" es una enzima, en general, para unir nucleótidos de 3'-OH5'-trifosfato, oligómeros y sus análogos. Las polimerasas incluyen, pero sin limitaciones, 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 Klenow, ADN polimerasa de *Thermophilus aquaticus*, ADN polimerasa Tth, ADN polimerasa VentR® (New England Biolabs), ADN polimerasa Deep VentR® (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 eucariota beta, telomerasa, polimerasa Therminator™ (New England Biolabs), ADN polimerasa KOD HiFi™ (Novagen), ADN polimerasa KOD1, Q-beta replicasa, transferasa terminal, retrotranscriptasa de AMV, retrotranscriptasa de M-MLV, retrotranscriptasa de Phi6, retrotranscriptasa de HIV-1, polimerasas novedosas descubiertas por bioprospección y polimerasas citadas en los documentos US 50 2007/0048748, US 6.329.178, US 6.602.695 y US 6.395.524. Estas polimerasas incluyen isoformas naturales, mutantes y variantes genomanipuladas.

55 "Codificar" o "analizar" son verbos que se refieren a la transferencia de un formato a otro, y se refieren a transferir información genética de la secuencia de base del molde diana a una disposición de indicadores.

"Extragenético" se refiere a cualquier estructura en la hebra hija que no sea parte de la cadena principal primaria; por ejemplo, un indicador extragenético no es la nucleobase en sí misma que se encuentra en la cadena principal primaria.

60 "Heterocopolímero" es un material formado combinando diferentes unidades (por ejemplo, especies de subunidades monoméricas) en cadenas de un "copolímero". Los heterocopolímeros se construyen a partir de construcciones de "subunidades" diferenciadas. Una "subunidad" es una región de un polímero compuesta por un motivo bien definido, donde cada motivo es una especie y porta información genética. El término heterocopolímero también se usa aquí para describir un polímero en el que todos los bloques son bloques construidos a partir de motivos repetitivos, teniendo cada motivo elementos específicos de especie. La hebra hija y el Xpandómero son ambos heterocopolímeros, con lo que cada motivo de subunidad codifica 1 o más bases de la secuencia de molde diana y la secuencia diana completa 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, níquel, polímeros, plásticos, cerámica o metales. Un soporte sólido también puede ser una membrana, tal como una membrana de níquel, de nitrocelulosa o polimérica, o una placa o plato y puede estar compuesto de vidrio, cerámica, metales o plásticos, tal como, por ejemplo, poliestireno, polipropileno, policarbonato, o polialímero. Un soporte sólido también puede ser una microesfera, una resina o una partícula de cualquier conformación. Dichas partículas o microesferas pueden estar compuestas de cualquier material adecuado, tal como vidrio o cerámica, y/o uno o más polímeros, tal como, por ejemplo, níquel, politetrafluoroetileno, TEFLON™, poliestireno, poliacrilamida, sepharose, agarosa, celulosa, celulosa derivados, o dextrano, y/o pueden contener metales, en particular metales paramagnéticos, tal como hierro.

"Bloqueo reversible" o "terminador" se refiere a un grupo químico que, cuando se une a un segundo grupo químico en un resto, evita que el segundo grupo químico entre en reacciones químicas particulares. Es conocida una amplia gama de grupos protectores en química orgánica y bioorgánica sintética que son adecuados para grupos químicos particulares y son compatibles con procedimientos químicos particulares, lo que significa que protegerán grupos particulares durante esos procedimientos y se pueden eliminar o modificar posteriormente (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 espacio entre las dos moléculas o restos, de modo que puedan funcionar de la manera prevista. Por ejemplo, un conector puede comprender una cadena hidrocarbonada de diamina que está unida covalentemente a través de un grupo reactivo en un extremo a una molécula de análogo de oligonucleótido y a través de un grupo reactivo en el otro extremo a un soporte sólido, tal como, por ejemplo, una superficie de microesferas. El acoplamiento de conectores a nucleótidos y construcciones de sustrato de interés se puede lograr a través del uso de reactivos de acoplamiento que son conocidos en la técnica (véase, por ejemplo, Efimov *et al.*, Nucleic Acids Res. 27: 4416-4426, 1999). Los procedimientos para derivatizar y acoplar moléculas orgánicas son bien conocidos en las técnicas de la química orgánica y bioorgánica. Un conector también puede ser escindible o reversible.

30 Aspectos generales

En términos generales, se describen procedimientos y dispositivos y productos correspondientes para replicar ácidos nucleicos diana de una sola molécula. Dichos procedimientos usan "Xpandómeros" que permiten la secuenciación del ácido nucleico diana con rendimiento y exactitud incrementados. Un Xpandómero codifica (analiza) los datos de la secuencia de nucleótidos del ácido nucleico diana en un formato expandido linealmente, mejorando de este modo la resolución espacial, opcionalmente con amplificación de la intensidad de señal. Estos procedimientos se denominan en el presente documento "secuenciación por expansión" o "SBX".

Como se muestra en la figura 1A, los ácidos nucleicos híbridos naturales tienen una densidad de datos lineal extremadamente compacta; una separación de centro a centro de aproximadamente 3,4 Å entre las bases apiladas secuenciales (2) de cada hebra de la doble hélice (1) y, por lo tanto, son tremadamente difíciles de visualizar o secuenciar directamente con exactitud y velocidad. Cuando la forma bicatenaria se desnaturaliza para formar polinucleótidos monocatenarios (3,4), las distancias de separación de base a base resultantes son similares, pero el problema se complica por los dominios de estructura secundaria.

Como se muestra en la figura 1B, el Xpandómero (5), aquí ilustrado como una concatenación de oligómeros cortos (6,7) que se mantienen unidos entre sí por ataduras extragenéticas T (8,9), es un reemplazo sintético o "sustituto" que se secuencie el ácido nucleico diana. Las bases complementarias al molde se incorporan en el Xpandómero, pero las ataduras regularmente espaciadas sirven para incrementar la distancia entre los oligómeros cortos (aquí mostrado cada uno con cuatro nucleobases representadas por círculos). El Xpandómero se elabora por un procedimiento en el que primero se forma un intermedio híbrido sintético replicando una hebra molde. La hebra hija es única en que tiene tanto una cadena principal lineal formado por los oligómeros como una cadena principal de Xpandómero restringido compuesto de ataduras plegadas. A continuación, las ataduras se abren o "expanden" para transformar el producto en una cadena de ataduras alargadas. En sentido figurado, se puede considerar que la hebra hija tiene dos cadenas principales superpuestas: una lineal (cadena principal primaria) y la otra con pliegues en "acordeón" (Xpandómero restringido). La escisión selectiva de enlaces en la hebra hija permite que los pliegues en acordeón se expandan para producir un producto Xpandómero. Este procedimiento se explicará con más detalle a continuación, pero cabe destacar que la elección de cuatro nucleobases por oligómero y los detalles de la atadura, como se muestra en la figura 1B, es solo con fines ilustrativos y de ninguna manera se debe interpretar que limita la invención.

La distancia de separación "D" entre oligómeros vecinos en el Xpandómero es ahora una variable dependiente del procedimiento y está determinada por la longitud de la atadura T. Como se mostrará, la longitud de la atadura T está diseñada en las construcciones de sustrato, los componentes básicos a partir de los cuales se elabora el Xpandómero. La distancia de separación D se puede seleccionar para que sea mayor de 0,5 nm, mayor de 2 nm, mayor de 5 nm, mayor de 10 nm o mayor de 50 nm, por ejemplo. A medida que se incrementa la distancia de separación, el procedimiento de discriminación o "resolución" de los oligómeros individuales se vuelve progresivamente más fácil. Esto también sería cierto

si, en lugar de oligómeros, se enhebran nucleobases individuales de otra especie de Xpandómero en una cadena de ataduras.

Volviendo a la figura 1A, el ADN natural se replica mediante un procedimiento de replicación semiconservadora; cada 5 nueva molécula de ADN es un "híbrido" de una hebra molde (3) y una hebra hija natural (4). La información de secuencia pasa del molde a la hebra hija natural mediante un procedimiento de "síntesis dirigida por molde" que conserva la información genética inherente a la secuencia de los pares de bases. La hebra hija natural, a su vez, se convierte en un molde para una hebra hija natural de siguiente generación, etc. Los Xpandómeros se forman por un procedimiento similar de síntesis dirigida por molde, que puede ser enzimático o un procedimiento de acoplamiento químico. Sin embargo, a 10 diferencia del ADN natural, una vez formados, los Xpandómeros no se pueden replicar por un procedimiento biológico de replicación semiconservadora y no son adecuados para amplificación por procedimientos tal como PCR. El producto Xpandómero está diseñado para limitar una estructura secundaria no deseada.

Las figuras 2A a 2D muestran sustratos de Xpandómero de clase I representativos (20, 21, 22, 23). Estos son los 15 componentes básicos a partir de los cuales se sintetizan Xpandómeros. Otros sustratos de Xpandómero (en el presente documento se divultan diez clases) se abordan en las secciones posteriores. Las construcciones de sustrato de Xpandómero que se muestran aquí tienen dos componentes funcionales; a saber, un miembro de sonda (10) y un miembro de "atadura" (11) en una configuración de bucle. El bucle forma la atadura alargada "T" del producto final. Únicamente por 20 comodidad en la explicación, el miembro de sonda se representa nuevamente con cuatro residuos de nucleobase (14, 15, 16, 17) como se muestra en la figura 2B.

Estas construcciones de sustrato se pueden modificar en los extremos con grupos R, por ejemplo, como un 5'-monofosfato, 3'-OH adecuado para su uso con una ligasa (denominada en el presente documento una "Xsonda") o como 25 un 5'-trifosfato, 3'-OH adecuado para su uso con una polimerasa (denominada en el presente documento "Xmero"). Otros grupos R pueden ser útiles en diversos protocolos. En el primer ejemplo que se muestra en la figura 3B, los autores de la invención presentan el uso de Xsondas en la síntesis de un Xpandómero a partir de una hebra molde de un ácido nucleico diana por un procedimiento dependiente de ligasa.

Los cuatro restos de nucleobase (14, 15, 16, 17) del miembro de sonda (10) se seleccionan para que sean 30 complementarios a una secuencia contigua de cuatro nucleótidos del molde. Cada "sonda" está, por tanto, diseñada para hibridar con el molde en una secuencia complementaria de cuatro nucleótidos. Suministrando una colección de muchas secuencias de sonda de este tipo, se puede formar una réplica complementaria contigua del molde. Esta hebra hija se denomina "intermedio de Xpandómero". Los intermedios de Xpandómero tienen formas híbridas o monocatenarias.

El bucle de atadura se une al miembro de sonda (10) en los residuos de nucleobase segundo y tercero (15, 16). Los 35 residuos de nucleobase segundo y tercero (15, 16) también se unen entre sí por un "enlace escindible selectivamente" (25) representado por una "V". La escisión de este enlace escindible permite que el bucle de atadura se expanda. Se puede decir que la atadura linealizada "forma un puente" con el sitio de enlace escindible selectivamente de la cadena principal primaria de polinucleótido de una hebra hija. La escisión de estos enlaces rompe la cadena principal primaria y 40 forma el Xpandómero más largo.

La escisión selectiva de los enlaces escindibles selectivamente (25) se puede hacer en una variedad de maneras, incluyendo, pero sin limitaciones, escisión química de enlaces fosforotiolato, digestión con ribonucleasas de enlaces ribosil 5'-3' fosfodiéster, escisión de enlaces fotoescindibles y similares, como se analiza con mayor detalle a continuación.

La construcción de sustrato (20) que se muestra en la figura 2A tiene un solo segmento de atadura, representado aquí 45 por una elipse (26), para la unión de elementos indicadores. Este segmento está flanqueado por segmentos de atadura de espaciador (12, 13), todos los cuales forman conjuntamente la construcción de atadura. Se pueden usar de uno a muchos dendrímeros, polímeros, polímeros ramificados o combinaciones de los mismos, por ejemplo, para construir el segmento de atadura. Para la construcción de sustrato (21) de la figura 2B, la construcción de atadura se compone de 50 tres segmentos de atadura para unión de elementos indicadores (27, 28, 29), cada uno de los cuales está flanqueado por un segmento de atadura espaciador. La combinación de elementos indicadores forman conjuntamente una "construcción de indicadores" para producir un código de indicador digital exclusivo (para identificación de la secuencia de la sonda). Estos elementos indicadores incluyen, pero sin limitaciones, fluoróforos, marcas de FRET, microesferas, ligandos, 55 aptámeros, péptidos, haptenos, oligómeros, polinucleótidos, dendrímeros, estructuras tallo-bucle, marcas de afinidad, marcas de masa y similares. El bucle de atadura (11) de la construcción de sustrato (22) en la figura 2C está "desnudo". La información genética codificada en esta construcción no está codificada en la atadura, sino que se asocia con 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 el asterisco (*), la información de secuencia de la sonda se codifica o "analiza" en la construcción 60 de sustrato en una forma modificada que se detecta más fácilmente en un protocolo de secuenciación. Debido a que los datos de secuencia se resuelven mejor físicamente después de la escisión del enlace escindible selectivamente (25) para formar el polímero Xpandómero alargado linealmente, el asterisco (*) representa cualquier forma de información genética codificada para la cual esto es un beneficio. El elemento o elementos bioinformáticos (*) de la construcción de sustrato, cualquiera que sea su forma, pueden ser detectables directamente o pueden ser precursores a los que se añaden 65 elementos detectables en una etapa de marcado posterior al ensamblaje. En algunos casos, la información genética se codifica en una propiedad molecular de la propia construcción del sustrato, por ejemplo, una marca de masa multiestado.

En otros casos, la información genética está codificada por uno o más fluoróforos de pares donante:ceptor 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 técnica. Diversos aspectos y modos de realización se analizarán con más detalle a continuación.

- 5 La atadura tiene, en general, una serie de funciones: (1) enlazarse secuencialmente, directa o indirectamente, a ataduras contiguas que forman el intermedio de Xpandómero; (2) estirarse y expandirse para formar una cadena alargada de ataduras tras la escisión de enlaces seleccionados en la cadena principal primaria o dentro de la atadura (véase la figura 1B); y/o (3) proporcionar una construcción molecular para incorporar elementos indicadores, también denominados "marcas" o "marcadores", que codifican la información de secuencia de residuos de nucleobase de su sustrato asociado.
- 10 10 La atadura se puede diseñar para optimizar la función de codificación ajustando separaciones espaciales, abundancia, densidad de información e intensidad de señal de sus elementos indicadores constituyentes. Una amplia gama de propiedades del indicador son útiles para amplificar la intensidad de señal de la información genética codificada dentro de la construcción de sustrato. La bibliografía relativa a indicadores, códigos de barras moleculares, unión por afinidad, marcado molecular y otros elementos de los indicadores son bien conocidos por los expertos en este campo.
- 15 15 Se puede observar que, si cada sustrato de una construcción de sustrato contiene x nucleobases, entonces una colección que represente todas las posibles combinaciones secuenciales de x nucleobases contendría 4^x sondas (cuando se seleccionan las nucleobases entre A, T, C o G). Se pueden necesitar menos o más combinaciones si se usan otras bases. Estas colecciones de sustratos están diseñadas de modo que cada construcción de sustrato contenga (1) una sonda (o al menos un residuo de nucleobase) complementaria a cualquiera de las posibles secuencias diana del ácido nucleico que se va a secuenciar y (2) una construcción de indicadores exclusiva que codifique la identidad de la secuencia diana a la que esa sonda particular (o nucleobase) sea complementaria. Una colección de sondas que contenga dos nucleobases tendría 16 miembros únicos; una colección de sondas que contenga tres nucleobases tendría 64 miembros únicos, etc. Una colección representativa tendría las cuatro nucleobases individuales, pero configuradas para adaptarse a un medio de atadura.

La síntesis de un Xpandómero se ilustra en las figuras 3A a 3C. El sustrato representado aquí es una Xsonda y el procedimiento se puede describir como hibridación con ligación procesiva dependiente de cebador en solución libre.

- 30 30 Muchos protocolos de biología molecular bien conocidos, tal como protocolos para fragmentar el ADN diana y ligar adaptadores de extremo, se pueden adaptar para su uso en procedimientos de secuenciación y se usan aquí para preparar el ADN diana (30) para secuenciación. Los autores de la invención ilustran en el presente documento, en términos amplios que resultarán familiares para los expertos en la técnica, procedimientos 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 desnaturiza e hibrida con cebadores adecuados (33) complementarios a los adaptadores.

- 40 40 En la figura 3B, la hebra molde cebada de la etapa III se pone en contacto con una colección de construcciones de sustrato (36) y ligasa (L), y en la etapa IV, las condiciones se ajustan para favorecer la hibridación seguida de la ligación en un 3'-OH libre de un híbrido cebador-molde. Opcionalmente, en la etapa V, la ligasa se disocia, y en las etapas VI y VII, se puede reconocer que el procedimiento de hibridación y ligación da como resultado una extensión por adición acumulativa de sustratos (37, 38) al extremo del cebador. Aunque el cebado se puede producir a partir de adaptadores en ambos extremos de una molde monocatenario, aquí se muestra que el crecimiento de una cadena hija emergente de Xpandómero prosigue a partir de un solo cebador, únicamente por simplicidad. La extensión de la hebra hija se representa en las etapas VI y VII, que se repiten continuamente (de forma incremental, sin interrupción). Estas reacciones se producen en solución libre y prosiguen hasta que se ha sintetizado una cantidad suficiente de producto. En la etapa VIII, se muestra la formación de un intermedio de Xpandómero completo (39).

- 50 50 Longitudes relativamente largas de secuencias de nucleótidos contiguas se pueden replicar eficazmente de esta manera para formar intermedios de Xpandómero. Se puede observar que con esta tecnología se pueden lograr longitudes de lectura continuas ("cóntigos") correspondientes a fragmentos de hebra molde largos. Será evidente para un experto en la técnica que miles de millones de estas reacciones SBX de una sola molécula se pueden hacer simultáneamente en un procedimiento por lotes eficiente en un solo tubo. Posteriormente, los productos indiscriminados de estas síntesis se pueden secuenciar.

- 55 55 En la figura 3C, se representan las siguientes etapas del procedimiento SBX. La etapa IX muestra la desnaturización del intermedio del Xpandómero híbrido seguida de la escisión de los enlaces escindibles selectivamente en la cadena principal, con los enlaces escindibles selectivamente diseñados de modo que los bucles de atadura "se abran", formando el producto Xpandómero alargado linealmente (34). Dicha escisión selectiva se puede lograr por cualquier serie de técnicas conocidas para los expertos en la técnica, que incluyen, pero sin limitaciones escisión de fosforotiolato con cationes metálicos como se divulga 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 de fosforamidato catalizada por ácido como se divulga 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 selectiva de enlaces fosfodiéster por nucleasa como se divulga 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 de la cadena principal de fosfodiéster modificado con conector fotoescindible como se divulga 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 procedimiento básico, tales como etapas de lavado y ajuste de las condiciones de restricción, están dentro de la técnica de un biólogo molecular experimentado. Las variantes de este procedimiento 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 posterior a la expansión, funcionalización de extremos y las alternativas a ligasa para unir los sustratos se analizarán 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 todo tipo. El procedimiento es un procedimiento para "expandir" o "alargar" la longitud de elementos de la cadena principal (o subunidades) que codifican información de secuencia (expandida en relación con las pequeñas distancias de nucleótido a nucleótido de ácidos nucleicos naturales) y, opcionalmente, también sirve para incrementar la intensidad de señal (en relación con las señales de baja intensidad casi indistinguibles observadas para los nucleótidos naturales). De este modo, los elementos indicadores incorporados en la cadena principal sintética expandida de un Xpandómero se pueden detectar y procesar usando una variedad de procedimientos de detección, incluyendo procedimientos de detección bien conocidos en la técnica (por ejemplo, una cámara CCD, un microscopio de fuerza atómica o un espectrómetro de masas activado), así como por procedimientos tales como una matriz de sensores de nanoporos masivamente paralelos, o una combinación de procedimientos. Las técnicas de detección se seleccionan sobre la base de una relación entre señal y ruido, rendimiento, coste y factores similares óptimos.

Volviendo a la figura 4, se muestra un modelo simple de una tecnología de detección; a saber, un nanoporo (40) con donante FRET (42) en una membrana (44), que es excitado por luz de longitud de onda λ_1 . A medida que el producto Xpandómero (41) se alarga y se transporta a través del nanoporo (40) en el sentido de la flecha (45), se detectan ráfagas en serie de emisión de longitud de onda λ_2 desde fluoróforos excitados, en la proximidad del poro. Las longitudes de onda de emisión (λ_2) están espaciadas temporalmente en función de la longitud de la atadura y la velocidad del Xpandómero que pasa a través del nanoporo. Capturando estas señales analógicas y procesándolas digitalmente, la información de la secuencia se puede leer directamente desde el Xpandómero. Cabe destacar que, en este procedimiento de detección, el nanoporo y la membrana pueden tener muchas vías a través de las cuales el Xpandómero se puede translocar. La detección por FRET requiere que haya al menos un donante de FRET excitado a lo largo de cada vía. Por el contrario, un nanoporo basado en un contador Coulter solo puede tener un orificio de translocación adicional a costa de la relación señal-ruido.

En la técnica de detección basada en nanoporos de la figura 4, que representa cadenas de Xsondas de la estructura que se muestra en la figura 2B, las construcciones de atadura contienen construcciones de indicadores de elementos múltiples, como se indica por los miembros indicadores en forma de caja (27, 28, 29) dispuestos a lo largo de la atadura. La tecnología de secuenciación por nanoporos pertinente se divulga, por ejemplo, en 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 indicadores de elementos múltiples, aquí compuestas por fluoróforos aceptores de FRET, aparecen ante un detector situado en la puerta de FRET. En este gráfico de emisiones multicanal se puede observar que se generan señales analógicas en intervalos de tiempo en general regulares y se pueden analizar como un tipo de código digital (aquí denominado código de indicador y, para este ejemplo, una ID de Xsonda) que revela la identidad y orden de las subunidades de Xsonda y, por tanto, la secuencia genética del Xpandómero ilustrado. Se pueden usar diversas combinaciones de indicadores para crear una colección de códigos de indicadores 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 indicadores. De esta manera, se observa que la secuencia ACTG va seguida de GCCG; seguida de AAAT. Las líneas de puntos colocadas verticalmente separan datos de fluorimetría y las subunidades correspondientes del Xpandómero (que se muestra esquemáticamente). Un algoritmo interpretativo inmediatamente debajo de la curva muestra cómo las señales analógicas espaciadas regularmente se transforman en una secuencia genética legible.

La construcción de sustrato de Xsonda ilustrada en la figura 5, que usa una construcción de atadura de elementos múltiples compuesta de tres segmentos marcados de indicador, cada uno de los cuales está flanqueado por segmentos de atadura de espaciador, para codificar la identidad de secuencia del sustrato, se elabora aún más. El primer segmento de atadura es el código de indicador n.º 1 (que se lee de izquierda a derecha) y se lee como una señal alta en el canal rojo. El segundo segmento de atadura 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 atadura es el código de indicador n.º 8, y se lee como una señal azul baja y una señal roja baja. El código de indicador n.º 1 se usa como reloj o señal de sincronización; el código de indicador n.º 9 codifica el primer resto de sonda "AC"; el código de indicador número 8 codifica el segundo resto "TG" de la sonda. En conjunto, el

código de indicador secuencial de "1-9-8" corresponde a una especie particular de Xsonda (Xsonda ID117), que a su vez corresponde por diseño al fragmento de secuencia "ACTG". Tres ID de Xsonda codifican la secuencia contigua completa que se muestra en la curva, "ACTGGCCGAAAT". Las emisiones de fluoróforo, la tabla para descodificar códigos de indicadores y fragmentos de secuencia, y las representaciones físicas correspondientes de las construcciones de indicadores se separan por las líneas punteadas de las figuras de acuerdo con subunidades estructurales del Xpandómero, de modo que se pueda observar fácilmente cómo la información de secuencia se descodifica y digitaliza.

La figura 6 es una tabla de marcadores de fluoróforo a partir de la cual se preparó el ejemplo de la figura 5. Esto ilustra de forma más general el uso de combinaciones de códigos de indicadores multiestado para analizar información en forma de señales detectables. Se usan fluoróforos que tienen veintidós estados de emisión posibles para formar las construcciones de indicadores de este ejemplo. Tres marcadores de fluoróforo por oligómero son más que adecuadas para codificar todas las posibles combinaciones tetraméricas de A, T, C y G. Incrementando la longitud de la atadura, se mejora la resolución entre las emisiones de los marcadores de fluoróforo, lo que beneficia la exactitud de la etapa de detección, un principio que es en general aplicable.

Los indicadores que se pueden usar con construcciones de atadura de este tipo son de muchos tipos, no meramente fluoróforos, y se pueden medir usando una amplia gama correspondiente de tecnologías de detección de alto rendimiento y exactas, tecnologías que de otro modo podrían no ser útiles para secuenciar ácidos nucleicos nativos debido a una resolución limitada. Los procedimientos de detección del estado de la técnica y masivamente paralelos, tales como matrices de sensores de nanoporos, se ven facilitados por las características más medibles de los Xpandómeros. Las ineficiencias en los procedimientos de detección de secuenciación se pueden reducir purificando previamente lotes de Xpandómeros para eliminar productos de reacción incompletos o cortos. Se proporcionan procedimientos para modificar extremos de Xpandómeros sintetizados que se pueden usar para purificación y también como medio de facilitar la presentación de Xpandómeros al detector. Además, el procedimiento de lectura no está restringido por limitación de protección, desprotección, extensión de nucleótidos, marcado u otros procedimientos de procesamiento simultáneos.

La figura 7A describe un molde híbrido parcial diseñado con un nucleótido protuberante en 5' de veinte bases para demostrar ligación procesiva de sustratos y ligación dirigida por molde iniciada por cebador en solución libre. La figura 7B es una fotografía de un gel que demuestra 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 CA3' se hibridan con el molde en presencia de un cebador y ADN ligasa T4. El nucleótido protuberante de extremo no hibridado (si lo hay) se digiere a continuación con nucleasa y los productos de ligación se separan en un gel de acrilamida al 20 %. La ligación da como resultado polímeros de producto que contienen subunidades ligadas de forma demostrable. Como se indica por el patrón de bandas, las reacciones positivas de ligasa se agotan en los carriles 1, 3, 5, 7 y 9, que contienen moldes progresivamente más largos (4, 8, 12, 16 y 20 bases, respectivamente), lo que demuestra claramente ligación secuencial. de sustratos diméricos (longitudes incrementadas de híbridos protegidos frente a exonuclease). Los carriles 2, 4, 6, 8 y 10 son controles negativos que no contienen ligasa y muestran una digestión completa con exonuclease de productos no ligados.

La figura 7C es un segundo gel que muestra ligación de sustratos dirigida por molde. Se sometieron a ensayo cuatro moldes de control positivo progresivamente más largos, hibridados de nuevo 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 CA3' se hibridan con el molde en presencia de un cebador y ADN ligasa T4. El nucleótido protuberante de extremo no hibridado (si lo hay) se digiere a continuación con nucleasa y los productos de ligación se separan en un gel de acrilamida al 20 %. Se observa que los sustratos oligoméricos (nuevamente diméricos) se ligan al molde en los carriles 1, 2, 3 y 4, pero no en los carriles 5 y 6, donde las hebras molde contienen un emparejamiento erróneo con el 5' (fosfato) CA3' dinucleótido (carril 5 molde - 5' CGCG3'; carril 6 molde - 5' GGGG3').

Los resultados del gel que se muestran en la figura 7D demuestran múltiples ligaciones dirigidas por molde de una sonda de tetranucleótido Bis(aminomodificado). Los modificadores amino alifáticos eran 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 A3' se hibridó con una gama de moldes complementarios progresivamente más largos (hibridados con un cebador de extensión) en presencia de un cebador y ADN ligasa T4. El nucleótido protuberante de extremo no hibridado (si lo hay) se digirió a continuación con nucleasa y los productos de ligación se separan en un gel de acrilamida al 20 %. La ligación da como resultado polímeros de producto que contienen subunidades ligadas de forma demostrable. Los carriles 1 y 2 representan controles de tamaño hexadecámérico e icosámérico. 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 tetrámeros para reacciones de moldes más largas (carriles 6-9). El carril 10 muestra una inhibición de ligasa esencialmente completa debido a un emparejamiento erróneo molde-sonda (molde-5' CGCG3').

Los resultados del gel que se muestran en la figura 7E demuestran múltiples ligaciones dirigidas por molde de una sonda de hexanucleótido Bis(aminomodificado). Los modificadores amino alifáticos eran 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 A3' se hibridó con una gama de moldes complementarios progresivamente más largos (hibridados con un cebador de extensión) en presencia de un cebador y ADN ligasa T4. El nucleótido protuberante de extremo no hibridado

(si lo hay) se digirió a continuación con nucleasa y los productos de ligación se separan en un gel de acrilamida al 20 %. La ligación da como resultado polímeros de producto que contienen subunidades ligadas de forma demostrable. Los cariles 1 y 2 representan controles de tamaño hexadecamérico e icosamérico. Los cariles 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 tetrámeros para reacciones de moldes más largas (cariles 5-9). El caril 10 muestra una inhibición de ligasa casi completa debido a un emparejamiento erróneo molde-sonda (molde-5' CGCGCG 3').

Los sustratos incluyen tanto miembros de sonda (es decir, oligómeros como miembros de unión específica de molde para ensamblar el intermedio de Xpandómero), como monómeros (es decir, miembros de nucleobases individuales como elementos de unión específica de molde). Los autores de la invención denominaron los primeros sustratos "de tipo sonda" y los segundos sustratos "de tipo monómero". Como se ilustra en la figura 8, los Xpandómeros de tipo sonda tienen cinco subgéneros básicos, mientras que la figura 9 ilustra cinco subgéneros básicos de Xpandómeros de tipo monómero. Las tablas de las figuras 8 y 9 incluyen tres columnas: la primera que describe las construcciones de sustrato, la segunda los intermedios de Xpandómero y la tercera los productos Xpandómeros característicos del subgénero (por fila). Las tablas se proporcionan aquí como una visión general, divulgándose los procedimientos para elaborar y usar los mismos con mayor detalle a continuación. En las figuras 8 y 9, "P" se refiere a un miembro de sonda, "T" a un miembro de atadura (o atadura de bucle o precursor de brazo de atadura), "N" a un monómero (una nucleobase individual o residuo de nucleobase) y "R" a un grupo de extremo.

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 de P¹-P², donde P¹ es un primer resto de sonda y P² es un segundo resto de sonda;

T es una atadura;

Los paréntesis indican una subunidad de la cadena hija, en la que cada subunidad es un motivo de subunidad que tiene un miembro de sonda específico de especie, en la que además dichos miembros de sonda de dichos motivos de subunidad son complementarios en serie a la secuencia de nucleótidos contigua correspondiente de la hebra molde, indicada aquí como P¹-P², y forman una cadena principal primaria del intermedio de Xpandómero, y en la que los miembros de atadura, opcionalmente en combinación con los restos de sonda, forman una cadena principal de Xpandómero restringido. La escisión de uno o más enlaces escindibles selectivamente dentro del intermedio de Xpandómero permite la expansión de las subunidades para producir un producto Xpandómero, cuyas subunidades también se indican entre paréntesis;

α denota una especie de motivo de subunidad seleccionado de una colección de motivos de subunidad;

ϵ es un primer grupo conector unido a un primer extremo o resto de un miembro de sonda o atadura; en condiciones controladas, ϵ puede reaccionar selectivamente con, directamente o por medio de agentes de reticulación, el grupo conector δ de un extremo colindante de una subunidad contigua para formar enlaces covalentes o de duración equivalente;

δ es un segundo grupo conector unido a un primer extremo o resto de un miembro de sonda o atadura; en condiciones controladas, δ puede reaccionar selectivamente con, directamente o por medio de agentes de reticulación, el grupo conector ϵ de un extremo colindante de una subunidad contigua para formar enlaces covalentes o de duración equivalente;

χ representa un enlace con una subunidad contigua y es el enlace del producto de la reacción de los grupos conectores δ y ϵ ;

~ indica un enlace escindible selectivamente, que puede ser el mismo o diferente cuando están presentes múltiples enlaces escindibles selectivamente;

R¹ incluye, pero sin limitaciones, hidroxilo, hidrógeno, trifosfato, monofosfato, éster, éter, glicol, amina, amida y tioéster;

R² incluye, pero sin limitaciones, hidroxilo, hidrógeno, trifosfato, monofosfato, éster, éter, glicol, amina, amida y tioéster; y

κ indica la subunidad κ ^{ésima} en una cadena de m subunidades, donde $\kappa = 1, 2, \dots, m$, donde $m > 3$, y en general $m > 20$, y preferentemente $m > 50$, y más preferentemente $m > 1000$.

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 una atadura;

Los paréntesis indican una subunidad de la cadena hija, en la que cada subunidad es un motivo de subunidad que tiene

- un residuo de nucleobase específico de especie, en la que además dichos residuos de nucleobase de dichos motivos de subunidad son complementarios en serie a la secuencia de nucleótidos contigua correspondiente de la hebra molde, indicada aquí como N', y forman una cadena principal primaria del intermedio de Xpandómero, y en la que los miembros de atadura, opcionalmente en combinación con los residuos de nucleobase, forman una cadena principal de Xpandómero 5 restringido. La escisión de uno o más enlaces escindibles selectivamente dentro del intermedio de Xpandómero permite la expansión de las subunidades para producir un producto Xpandómero, cuyas subunidades también se indican entre paréntesis;
- n¹ es una primera porción de un residuo de nucleobase;
- n² es una segunda porción de un residuo de nucleobase;
- ε es un primer grupo conector unido a un primer extremo o resto de un miembro de sonda o atadura; en condiciones controladas, ε puede reaccionar selectivamente con, directamente o por medio de agentes de reticulación, el grupo conector δ de un extremo colindante de una subunidad contigua para formar enlaces covalentes o de duración equivalente;
- δ es un segundo grupo conector unido a un primer extremo o resto de un miembro de sonda o atadura; en condiciones controladas, δ puede reaccionar selectivamente con, directamente o por medio de agentes de reticulación, el grupo conector ε de un extremo colindante de una subunidad contigua para formar enlaces covalentes o de duración equivalente;
- x representa un enlace con una subunidad contigua y es el enlace del producto de la reacción de los grupos de enlace δ y ε;
- x¹ es el enlace de producto de la reacción de los grupos de enlace δ¹ y ε¹;
- x² es el enlace de producto de la reacción de los grupos de enlace δ² y ε²;
- ~ indica un enlace escindible selectivamente, que puede ser el mismo o diferente cuando están presentes múltiples enlaces escindibles selectivamente;
- R¹ incluye, pero sin limitaciones, hidroxilo, hidrógeno, trifosfato, monofosfato, éster, éter, glicol, amina, amida y tioéster;
- R² incluye, pero sin limitaciones, hidroxilo, hidrógeno, trifosfato, monofosfato, éster, éter, glicol, amina, amida y tioéster; y
- κ denota la κ^{ésima} subunidad en una cadena de m subunidades, donde κ = 1, 2, ..., a m, donde m > 10, y en general m > 50, y típicamente m > 500 o > 5000.

Construcciones oligoméricas

- Los precursores y construcciones de Xpandómeros se pueden dividir en dos categorías en base al sustrato (oligomérico o monomérico) usado para el ensamblaje dirigido por molde. La estructura, precursores y procedimientos de síntesis de los Xpandómeros para los basados en los sustratos de oligómero se analizan a continuación.
- Las construcciones de sustrato son reactivos precursores del Xpandómero y, en general, tienen un miembro de atadura y un sustrato. El sustrato que se analiza en el presente documento es un sustrato o sonda de oligómero, constituido en general por una pluralidad de residuos de nucleobase. Generando colecciones de tipo combinatorio de dos a veinte residuos de nucleobase por sonda, en general de 2 a 10 y típicamente 2, 3, 4, 5 o 6 residuos de nucleobase por sonda, se generan colecciones de sondas útiles como reactivos en la síntesis de precursores de Xpandómeros (construcciones de sustrato).
- La sonda se describe en general a continuación como que tiene dos restos de sonda, P¹ y P². Estos restos de sonda se representan en general en las figuras como dinucleótidos, pero en general P¹ y P² tienen cada uno al menos un residuo de nucleobase. En el ejemplo de una sonda con dos residuos de nucleobase, los restos de sonda P¹ y P² serían residuos de una sola nucleobase. El número de residuos de nucleobase para cada uno se selecciona, de forma apropiada, para el procedimiento de síntesis del Xpandómero y puede no ser igual en P¹ y P².
- Para las construcciones de sustrato donde se usan grupos conectores ε y δ para crear enlaces entre subunidades, se puede adaptar una amplia gama de químicas disponibles comercialmente (Pierce, Thermo Fisher Scientific, EE. UU.) con este fin. Las químicas de conector comunes incluyen, por ejemplo, ésteres de NHS con aminas, maleimidas con sulfhidrilos, imidoésteres con aminas, EDC con carboxilos para reacciones con aminas, disulfuros de piridilo con sulfhidrilos y similares. Otros aspectos implican el uso de grupos funcionales como hidrazida (HZ) y 4-formilbenzoato (4FB) que, a continuación, pueden reaccionar para formar enlaces. Más específicamente, están ampliamente disponibles (Pierce) una amplia gama de agentes de reticulación (hetero y homobifuncionales) que incluyen, pero sin limitaciones, Sulfo-SMCC (4-[N-maleimidometil] ciclohexano-1-carboxilato de sulfosuccinimido), SIA (yodoacetato de N-succinimido), Sulfo-EMCS (éster de [N-e-Maleimidocaproiloxy]sulfosuccinimida), Sulfo-GMBS (éster de N-[g-maleimidobutiriloxy]sulfosuccinimida), AMAS éster de N-(a-maleimidoacetoxi) succinimida), BMPS (éster de N EMCA

- (ácido ne-maleimidocaproico)-[β-maleimidopropiloxi] succinimida), EDC (clorhidrato de 1-etil-3-[3-dimetilaminopropil]carbodiimida), SAN PAH (N-succinimidil-6-[4'-azido -2'-nitrofenilamino]hexanoato), SADP (N-succinimidil(4-azidofenil)-1, 3'-ditiopropionato), PMPI (N-[p-maleimidofenil]isoci, BMPH (N-[ácido β-maleimidopropiónico]hidrazida, sal de ácido trifluoroacético)anato), EMCH ([ácido N-e-maleimidocaproico]hidrazida, sal de ácido trifluoroacético), SANH (4-hidrazinonicotinato acetona hidrazona de succinimidilo), SHTH (4-hidrazidotereftalato clorhidrato de succinimidilo) y C6-SFB (4-formilbenzoato de C6-succinimidilo). Además, el procedimiento divulgado por Letsinger *et al.* ("Phosphorothioate oligonucleotides having modified internucleoside linkages", patente de EE. UU. n.º 6.242.589) se puede adaptarse para formar enlaces de fosforotiolato.
- 10 Además, están ampliamente disponibles químicas de protección/desprotección bien establecidas para restos conectores comunes (Benoitton, "Chemistry of Peptide Synthesis", CRC Press, 2005). La protección de amino incluye, pero sin limitaciones, carbamato de 9-fluorenilmetilo (Fmoc-NRR'), carbamato de t-butilo (Boc-NRR'), carbamato de bencílico (Z-NRR', Cbz-NRR'), trifluoroacetamida de acetamida, ftalimida, bencilamina (Bn-NRR'), trifenilmetilamina (Tr-NRR') y bencildenamina p-toluenosulfonamida (Ts-NRR'). La protección de carboxilos incluye, pero sin limitaciones, éster metílico, éster t-butílico, éster bencílico, éster de S-t-butilo y 2-alquil-1,3-oxazolina. Los carbonilos incluyen, pero sin limitaciones, dimetilacetal 1,3-dioxano y 1,3-ditiano N,N-dimethylhidrazona. La protección de hidroxilos incluye, pero sin limitaciones, éter metoximetílico (MOM-OR), éter tetrahidropiranílico (THP-OR), éter t-butílico, éter alílico, éter bencílico (Bn-OR), éter t-butildimetsilsilílico (TBDMS-OR), éter t-butildifenilsilílico (TBDPS-OR), éster de ácido acético, éster de ácido piválico y éster de ácido benzoico.
- 20 Si bien la atadura a menudo se representa como una construcción de indicadores con tres grupos indicadores, se pueden organizar diversas configuraciones de indicadores en la atadura y pueden comprender indicadores individuales que identifican los constituyentes de sonda, indicadores individuales que identifican especies de sonda, códigos de barras moleculares que identifican especies de sonda o la atadura puede ser de polímero desnudo (sin indicadores). En el caso 25 del polímero desnudo, los indicadores pueden ser la propia sonda o pueden estar en una segunda atadura unida a la sonda. En algunos casos, uno o más precursores indicadores se disponen en la atadura, y los indicadores se unen por afinidad o se unen covalentemente después del ensamblaje del producto Xpandómero.
- 30 Como se ha analizado anteriormente, la figura 8 proporciona una visión general de las construcciones oligoméricas de la invención, distinguiéndose cinco clases: clases I, II, III, IV y V. Estas clases se aplican tanto a Xsondas como a Xmeros. Cada clase se analizará individualmente a continuación.
- ### Construcciones oligoméricas de clase I
- 35 Volviendo a la Figura 10, las construcciones oligoméricas de clase I se describen con más detalle. Las figuras 10A a 10C emplean una notación adaptada para mostrar estas moléculas como sustratos y como productos de heterocopolímero del procedimiento de SBX. Las figuras se leen de izquierda a derecha, mostrándose primero la construcción de sustrato de sonda (precursor oligomérico de Xpandómero), a continuación la hebra hija híbrida intermedia en el centro y, a la derecha, el producto Xpandómero preparado para secuenciación.
- 40 Como se muestra en la figura 10A, una construcción de sustrato de clase I tiene un miembro de sonda oligomérico (-P¹~P²-) (100) y un miembro de atadura, T (99). La atadura está unida por dos enlaces de extremo (108, 109) a los restos de sonda P¹ y P². Estas restricciones evitan que la atadura se alargue o se expanda y, por tanto, en una configuración restringida. En el ensamblaje dirigido por molde, los sustratos forman un híbrido con el molde diana, de modo que los sustratos son colindantes.
- 45 R¹ y R² son grupos de extremo configurados como es adecuado para el protocolo de síntesis en el que se usa la construcción de sustrato. Por ejemplo, R¹ = 5'-fosfato y R² = 3'-OH encontrarían uso en un protocolo de ligación, y R¹ = 5'-trifosfato y R² = 3'-OH para un protocolo de polimerasa. Opcionalmente, R² se puede configurar con un grupo de bloqueo reversible para la adición cíclica de un solo sustrato. De forma alternativa, R¹ y R² se pueden configurar con grupos de extremo conectores para acoplamiento químico o sin grupos conectores para un protocolo solo de hibridación. R¹ y R² pueden ser del tipo general XR, en el que X es un grupo de enlace y R es un grupo funcional.
- 50 La tilde (~) en las figuras 10A y 10B indica un enlace escindible selectivamente que separa dos restos del miembro de sonda. Los enlaces de extremo de la atadura se unen a los dos restos del miembro de sonda que se separan por el enlace escindible selectivamente. La atadura une el primer resto de sonda con el segundo resto de sonda, formando un bucle que forma un puente con el enlace escindible selectivamente. Cuando el miembro de la sonda está intacto (no escindido), el miembro de sonda se puede unir con alta fidelidad a la secuencia molde y la atadura se enlaza en la "configuración restringida". Cuando se escinde este enlace, el bucle de atadura se puede abrir y la atadura está en la "configuración expandida".
- 55 Las construcciones de sustrato son reactivos que se usan 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 híbrida, un heterocopolímero con subunidades repetitivas (que se muestran entre paréntesis). Se muestran la cadena principal primaria de la hebra hija (-P¹~P²-) y la hebra molde diana (-P¹-P²-) como un híbrido (95). Cada subunidad de la hebra hija es un motivo repetitivo compuesto de un miembro de sonda y un miembro de atadura, T (99), el miembro de atadura

- en configuración restringida. Los motivos tienen variabilidad específica de especie, indicada en el presente documento por el superíndice "α". Cada subunidad particular en la cadena hija se selecciona de una colección de motivos por un procedimiento dirigido por molde y su sonda se une a una secuencia correspondiente de nucleótidos complementarios en la hebra molde. De esta manera, la secuencia de residuos de nucleobase de las sondas forma una copia complementaria y contigua de la hebra molde diana.
- La hebra hija se compone de un precursor de Xpandómero llamado "Xpandómero restringido" que además se compone de ataduras en la "configuración restringida". Cuando las ataduras (99) se convierten a su "configuración expandida", el Xpandómero restringido se convierte en el producto Xpandómero.
- Se puede observar que la hebra hija tiene dos cadenas principales, un "cadena principal primario" y la cadena principal del "Xpandómero restringido". La cadena principal primaria está compuesta por sustratos de sonda colindantes. La "cadena principal del Xpandómero restringido" evita el enlace escindible selectivamente entre los restos de sonda P¹ y P² y está formada por restos de cadena principal enlazados, siendo cada resto de cadena principal un enlace lineal de P¹ a la atadura a P², y donde P² se puede unir con el P¹ del siguiente resto de cadena principal. Se puede observar que la cadena principal del Xpandómero restringido forma un puente o forma un bucle sobre los enlaces escindibles selectivamente de la cadena principal primaria, y permanecerá covalentemente intacto cuando estos enlaces escindibles selectivamente se escindan y la cadena principal primaria se fragmente.
- La figura 10C es una representación del producto Xpandómero de clase I después de la disociación de la hebra molde y después de la escisión de los enlaces escindibles selectivamente de la cadena principal primaria. Los procedimientos para la disociación de la hebra molde incluyen la desnaturización por calor o digestión selectiva con una nucleasa o degradación química. La cadena del producto Xpandómero contiene una pluralidad de subunidades κ, donde κ denota la κ^{ésima} subunidad en una cadena de m subunidades que componen la cadena hija, donde κ = 1, 2, 3 a m, donde m>3, y en general m>20, y preferentemente m>50, y más preferentemente m> 000. Cada subunidad está formada por una atadura (99) y restos de sonda P¹ y P². El miembro de atadura T, ahora en "configuración expandida", se observa estirado en toda su longitud entre los restos de sonda escindidos P¹ y P², que permanecen unidos covalentemente a las subunidades contiguas. Cada subunidad, un motivo de subunidad α, contiene información genética específica de especie establecida por ensamblaje dirigido por molde del intermedio de Xpandómero (hebra hija).
- La figura 10D muestra la construcción de sustrato de la figura 10A como un modelo molecular, donde el miembro de sonda (100) se representa arbitrariamente con cuatro residuos de nucleobase (101, 102, 103, 104), dos de los cuales (102, 103) están unidos a la atadura (99) por enlaces de extremo (108, 109). Entre los dos enlaces de los extremos de la atadura hay un enlace escindible selectivamente, que se muestra como la "V" (110) en el miembro de sonda (100). Este enlace une los restos de sonda P¹ y P² a los que se hace referencia en la figura 10A. El bucle de atadura que se muestra aquí tiene tres indicadores (105, 106, 107), que también pueden ser específicos de la especie del motivo.
- La figura 10E muestra el producto Xpandómero después de la escisión de los enlaces escindibles selectivamente en el sustrato. La escisión da como resultado la expansión del Xpandómero restringido y se indica por "E" (flechas oscuras). Los residuos (110a, 110b) del enlace escindible selectivamente marcan el evento de escisión. La subunidad se indica mediante líneas punteadas que enmarcan verticalmente la subunidad repetitiva, como se representa por corchetes en la figura 10C adjunta.
- En el producto Xpandómero (figura 10E), la cadena principal primaria está ahora fragmentado y no está covalentemente intacta porque los miembros de sonda se han escindido, separando cada P¹ (92) y P² (94). A través del procedimiento de escisión, el Xpandómero restringido se libera para convertirse en el producto Xpandómero. El Xpandómero incluye cada subunidad concatenada en secuencia. Unidos dentro de cada subunidad están el resto de sonda P¹, la atadura y el resto de sonda P². Los miembros de atadura (99) del Xpandómero, que estaban anteriormente en una configuración restringida, están ahora en una configuración expandida, funcionando de este modo para extender linealmente la información de secuencia de la diana del molde. La expansión de las ataduras reduce la densidad lineal de la información de secuencia a lo largo del Xpandómero y proporciona una plataforma para incrementar el tamaño y la abundancia de los indicadores, lo que a su vez mejora la relación señal-ruido para la detección y descodificación de la secuencia molde.
- La figura 11 representa un esquema resumido de un procedimiento para elaborar un Xpandómero de clase I; el procedimiento ilustra la elaboración y uso de sustratos y productos que se muestran en las figuras 10D y 10E. El procedimiento se puede realizar en solución libre y se describe usando una ligasa (L) para acoplar covalentemente Xsondas colindantes. Los procedimientos para aliviar la estructura secundaria en el molde se analizan en una sección posterior. Las condiciones adaptadas para hibridación y ligación son bien conocidas en la técnica y dichas condiciones se pueden optimizar fácilmente por un experto en este campo.
- Las ligasas incluyen, pero sin limitaciones, 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 Ampligase, ligasa de tipo VanC, ADN ligasa 9°N, ADN ligasa Tsp y ligasas novedosas descubiertas por bioprospección. Las ligasas también incluyen, pero sin limitaciones, ligasas dependientes de ATP que incluyen ARN ligasa T4, ADN ligasa T4, ADN ligasa T7, ADN ligasa Pfu, ADN ligasa ADN I, ADN ligasa ADN III, ADN ligasa IV y ligasas novedosas descubiertas por bioprospección. Estas ligasas incluyen isoformas naturales, mutantes

y variantes genomanipuladas.

En referencia a la figura 11, y en preparación para la síntesis, se proporciona un ácido nucleico diana (110) y los extremos se pulen en preparación para la ligación de extremos romos de adaptadores. La etapa I muestra la ligación de cebadores de horquilla (120) al ácido nucleico diana. El extremo 5' libre de los cebadores se bloquea con un grupo de bloqueo extraíble (119). Los cebadores cebarán ambas hebras del ácido nucleico diana. En general, los adaptadores se añaden en exceso. Los grupos de bloqueo en los extremos calientes de los cebadores se retiran en la etapa II y las dos hebras molde se separan por desnaturización. En la etapa III, el molde monocatenaria cebado (111) se pone en contacto con una colección de construcciones de sustrato (representada por la construcción (112) para propósito de ilustración) y con ligasa, L, en condiciones permisibles para hibridación del sustrato de sonda complementario (113) y ligación en el extremo reactivo del cebador, como se muestra en la etapa IV. En general, la hibridación y ligación se realiza a una temperatura mayor que la temperatura de fusión del sustrato para reducir las reacciones secundarias inespecíficas. Cada construcción de sustrato en este ejemplo contiene una atadura dispuesta con tres indicadores. Cada sustrato de sonda tiene un enlace escindible selectivamente (indicado con una "V") entre los dos sitios de unión de la atadura. En la etapa V, se añade una segunda construcción de sustrato (114) por hibridación dirigida por molde y ligación, etc. En la etapa VI, se demuestra la formación de un intermedio de Xpandómero completamente extendido (117). Este intermedio se puede desnaturizar de la hebra molde y escindir selectivamente en los sitios de escisión que se muestran, formando de este modo un producto Xpandómero adecuado para secuenciación. En algunos aspectos, no es necesaria la desnaturización y la hebra molde se puede digerir en su lugar.

La figura 12 es un esquema condensado de un segundo procedimiento, en el presente documento para elaborar otro Xpandómero de clase I. En 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 cebadores de horquilla doblemente bloqueados al ácido nucleico diana. Un extremo de los cebadores de horquilla híbridos se bloquea con grupos de bloqueo extraíbles (125a, 125b, 125c, 125d) que pretenden evitar la ligación y concatenación de las hebras molde o adaptadores. En general, los adaptadores se añaden en exceso. Los grupos de bloqueo se retiran en la etapa II y las dos hebras molde se separan por desnaturización. En la etapa III, los cebadores de horquilla autohibridan, formando sitios de cebado (126, 127) para la ligación posterior de construcciones de sustrato, que pueden proseguir bidireccionalmente, es decir, tanto en una dirección de '3 a 5' como de '5 a 3'. En la etapa IV, los moldes cebados se ponen en contacto con una colección de construcciones de sustrato (128) en condiciones permisibles para hibridación de sustratos de sonda complementarios y ligación. La ligación prosigue de forma incremental (es decir, extendiendo los extremos en crecimiento con aparente procesividad) por un procedimiento de hibridación de sustratos de sonda complementarios y ligación en los extremos de las hebras hijas emergentes. Cada construcción de sustrato en este ejemplo contiene un bucle de atadura dispuesto con grupos indicadores. En la etapa V, se representa la formación de un intermedio de Xpandómero completo (129). Opcionalmente, la hebra molde se puede retirar por digestión con nucleasas, liberando el Xpandómero. El intermedio se puede escindir selectivamente en los sitios de escisión que se muestran, formando de este modo un producto Xpandómero adecuado para secuenciación. Los productos Xpandómeros se forman en solución libre.

En la figura 13, se muestra un procedimiento que se basa en hebras molde inmovilizadas. En el presente documento, las hebras molde están ancladas a una microesfera (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 se adaptan de modo que se produce la hibridación. Se puede observar que se forman "islas" de construcciones de sustrato colindantes e hibridadas. En la etapa II, la adición de ligasa, L, da como resultado la ligación de las construcciones de sustrato colindante, formando de este modo múltiples secuencias contiguas de 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 emparejado erróneamente, 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 procedimiento sin cebador es denominado en el presente documento "ligación indiscriminada". La ligación se puede extender bidireccionalmente y las uniones melladas se pueden sellar con ligasa, llenando de este modo los huecos. En la etapa V, después de la optimización de las longitudes deseadas del producto, se lavan los híbridos inmovilizados para retirar el sustrato sin reaccionar y la ligasa. A continuación, en la etapa VI, las hebras hijas (que aquí se muestran como un intermedio de Xpandómero monocatenario) (138, 139) se disocian del molde. La escisión selectiva de enlaces escindibles selectivamente del intermedio da como resultado la formación del producto Xpandómero (no se muestra). En este aspecto, el molde inmovilizado se puede reusar. Una vez que se secuencian los productos Xpandómeros, los cóntigos se pueden ensamblar mediante algoritmos bien conocidos para superponer y alinear los datos para construir una secuencia consenso.

En referencia a la figura 14, se muestra un procedimiento para usar cebadores inmovilizados. Los moldes adaptados en el extremo (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 un colección de construcciones de sustrato, cuyos miembros se muestran como (144), y se ajustan las condiciones para hibridación dirigida por molde. En este ejemplo, el extremo 3' OH de la construcción de sustrato de los miembros de sonda (grupo R) se ha sustituido (146) para bloquear reversiblemente una extensión adicional. En la etapa III, los extremos colindantes de la construcción de sustrato contigua y el cebador, o el extremo libre de la hebra hija en crecimiento emergente, se ligan y el extremo 3' OH de la hebra hija emergente se activa por retirada del grupo R de bloqueo (146). Como se indica en las etapas IV y V, este procedimiento de adición gradual cíclica se puede repetir varias

veces. Típicamente, se usa una etapa de lavado para retirar los sustratos sin reaccionar entre cada etapa de extensión.

Por tanto, el procedimiento es análogo a lo que se denomina "extensión de base única cíclica", pero se denominaría más apropiadamente aquí "extensión de sonda única cíclica". Si bien se muestra la ligasa, L, el procedimiento se puede realizar con una ligasa, una polimerasa o mediante cualquier protocolo de acoplamiento químico adecuado para unir oligómeros en una síntesis dirigida por molde. El acoplamiento químico se puede producir espontáneamente en los extremos colindantes de las sondas hibridadas o se puede añadir un agente de condensación al comienzo de la etapa III y cada etapa V subsiguiente del ciclo. El grupo R de bloqueo terminal está configurado de modo que no se puede producir 5 polimerización continua libre en el molde o en solución. La etapa VI muestra la formación de un intermedio de Xpandómero 10 completo (149); no se puede añadir más sustrato. Este intermedio se puede disociar del molde, el producto monocatenario 15 se escinde a continuación para abrir la cadena principal como se describe anteriormente.

Este procedimiento se puede adaptar para secuenciación selectiva de dianas particulares en una mezcla de ácidos 15 nucleicos y para procedimientos de secuenciación analizada en matrices de secuenciación, por ejemplo, por selección no aleatoria de los cebadores inmovilizados. De forma alternativa, se pueden usar cebadores universales o aleatorios como se muestra.

La figura 15 describe un procedimiento para hibridación indiscriminada en un molde inmovilizado (150) (etapa I), donde 20 las construcciones de sustrato de la colección (152) se modifican con un grupo funcional químico que es selectivamente reactivo (156), representado como un triángulo abierto, con una sonda colindante. Un detalle del grupo funcional químico de las construcciones de sustrato se muestra en la porción expandida mostrada por el círculo sombreado (figura 15a). A determinada densidad de hibridación, se inicia el acoplamiento como se muestra en la etapa II, lo que da como resultado intermedios de Xpandómero de alto peso molecular unidos por el producto reticulado (157), representado como un 25 triángulo lleno, de la reacción de acoplamiento. Un detalle de las sondas reticuladas se muestra en la porción expandida mostrada en el círculo sombreado (figura 15b) en el producto de la etapa II. Este procedimiento puede ir acompañado de etapas para disociación selectiva y retirada de productos de bajo peso molecular y cualquier posible producto emparejado erróneamente. Las químicas de acoplamiento para este procedimiento de acoplamiento químico indiscriminado son 30 conocidas por los expertos en la técnica e incluyen, por ejemplo, las técnicas divulgadas en la patente de EE. UU. n.º 6.951.720 a Burgin *et al.*

En otro aspecto, se divultan procedimientos basados en polimerasa para ensamblar productos Xpandómeros. En general, 35 los trifosfatos de sustrato (Xmers) son el sustrato apropiado para reacciones que implican una polimerasa. La selección de una polimerasa adecuada es parte de un procedimiento de optimización del protocolo experimental. Como se muestra en la figura 16 a modo de ilustración, y si bien no pretende ser una limitación, una mezcla de reacción que contiene un molde (160) y un cebador (161) se pone en contacto con una colección de construcciones de sustrato (162) y una 40 polimerasa (P), en condiciones optimizadas para polimerización dirigida por molde. En la etapa I, la polimerasa comienza a añadir procesivamente Xmers de dinucleótidos (ataduras con dos indicadores) a la hebra molde. Este procedimiento continúa en las etapas II y III. Cada subunidad de sonda añadida es una especie particular seleccionada por unión específica al siguiente oligómero contiguo del molde para formar una copia complementaria contigua del molde. Si bien 45 no está vinculado por la teoría, se cree que la polimerasa ayuda a garantizar que las especies de sonda entrantes añadidas a la cadena emergente sean específicamente complementarias al siguiente segmento contiguo disponible del molde. Loeb y Patel describen ADN polimerasas mutantes con actividad incrementada 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 50 se pueden modificar para una velocidad incrementada de incorporación y reducción de la tasa de error, vinculando claramente la tasa de error no con la exactitud de hibridación sino con la procesividad de la polimerasa. La etapa III da como resultado un intermedio de Xpandómero completo (168). El intermedio de Xpandómero monocatenario se trata después por un procedimiento que puede implicar la desnaturización de la cadena molde (no se muestra). La cadena principal primaria de la hebra hija se escinde selectivamente para expandir las ataduras, formando de este modo un producto Xpandómero adecuado para su uso en un protocolo de secuenciación, como se explica previamente.

Como se muestra en la figura 17, la síntesis dirigida por molde accionada por polimerasas de un Xpandómero se puede lograr por técnicas alternativas. En el presente documento, un cebador inmovilizado (170) al que se hibrida una hebra molde procesada (171) en la etapa I. En la etapa II, la polimerasa, P, acopla procesivamente construcciones de sustrato 55 específicamente complementarias (175) de una colección de dichas construcciones (representada por 174) en la mezcla de reacción. Las condiciones y soluciones de reactivos se ajustan para favorecer la actividad polimerasa procesiva. Como se muestra aquí, la hibridación en la etapa II y la polimerización en la etapa III son actividades separadas, pero no es necesario aislar las actividades de la polimerasa de esa manera. En la etapa IV, la adición procesiva incremental de construcciones de sustrato complementarias continúa cíclicamente (continuamente sin interrupción), lo que da como resultado el intermedio de Xpandómero (177) completamente cargado, como se representa como resultado de la etapa 60 IV. El intermedio de Xpandómero se puede disociar y expandir en preparación para su uso en un protocolo de secuenciación como se describe previamente. Hay que tener en cuenta que este procedimiento también se presta a procedimientos de secuenciación analizados por selección de cebadores inmovilizados adecuados. Además, los procedimientos para estirar el molde para aliviar la estructura secundaria se adaptan fácilmente a este procedimiento y se analizan en secciones posteriores.

65 Las polimerasas incluyen, pero 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 Klenow, ADN polimerasa de *Thermophilus aquaticus*, ADN polimerasa Tth, ADN polimerasa VentR® (New England Biolabs), ADN polimerasa Deep VentR® (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 eucariota beta, telomerasa, polimerasa ThermoQ™ (New England Biolabs), ADN polimerasa KOD HiFi™ (Novagen), ADN polimerasa KOD1, Q-beta replicasa, transferasa terminal, retrotranscriptasa de AMV, retrotranscriptasa de M-MLV, retrotranscriptasa de Phi6, retrotranscriptasa de HIV-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 naturales, mutantes y variantes genomanipuladas.

Construcciones oligoméricas de clase II y III

En referencia a las figuras 18A a 18E, se describen construcciones oligoméricas de clase II con más detalle, que (junto con las construcciones oligoméricas de clase III isoméricas) pueden ser Xsondas o Xmeros.

Las figuras 18A a 18C se leen de izquierda a derecha, mostrándose primero la construcción de sustrato de sonda (precursor oligomérico de Xpandómero), a continuación la hebra hija híbrida intermedia en el centro y, a la derecha, el producto Xpandómero preparado para secuenciación.

Como se muestra en la figura 18A, una construcción de sustrato de clase III tiene un miembro de sonda oligomérico (-P¹-P²-) (180) y un miembro de atadura, T (181). La atadura se une mediante un enlace de un solo extremo (184) de un primer resto de extremo al resto de sonda P². En el extremo distal de la atadura (186), un segundo resto de extremo tiene un grupo conector δ y se sitúa próximo a R². El segundo resto de extremo también tiene una reticulación intraatadura escindible (187) para restringirlo a esta localización. La reticulación escindible (187) se indica mediante una línea punteada, que puede indicar, por ejemplo, un enlace disulfuro. Estas restricciones evitan que la atadura se alargue o se expanda y, por tanto, en una configuración restringida. Un segundo grupo conector ε se sitúa cerca del extremo distal (189) del miembro de sonda cerca de R¹. En el ensamblaje dirigido por molde, los sustratos forman un híbrido con el molde diana, de modo que los sustratos son colindantes. En condiciones controladas, los grupos conectores δ y ε de los sustratos colindantes se unen para formar un enlace x entre las construcciones de sustrato contiguas (que se muestran en las figuras 18B y 18C). Estos grupos de enlace se sitúan en la construcción de sustrato para limitar estas reacciones de enlace con construcciones de sustrato colindantes. La construcción de sustrato preferentemente no se une consigo misma. Las químicas de enlace y protección/desprotección adecuadas para δ, ε y x se detallan en la descripción general de la construcción oligomérica.

R¹ y R² son grupos de extremo configurados como es adecuado para el protocolo de síntesis en el que se usa la construcción de sustrato. Por ejemplo, R¹ = 5'-fósforo y R² = 3'-OH encontrarán uso en un protocolo de ligación, y R¹ = 5'-trifosfato y R² = 3'-OH para un protocolo de polimerasa. Opcionalmente, R² se puede configurar con un grupo de bloqueo reversible para la adición cíclica de un solo sustrato. De forma alternativa, R¹ y R² se pueden configurar con grupos de extremo conectores para acoplamiento químico o sin grupos conectores para un protocolo solo de hibridación. R¹ y R² pueden ser del tipo general XR, en el que X es un grupo de enlace y R es un grupo funcional.

Las construcciones de sustrato son reactivos que se usan para el ensamblaje dependiente de molde de una hebra hija, una composición intermedia para producir Xpandómeros. La figura 18B muestra la hebra hija híbrida, un heteropolímero con subunidades repetitivas (que se muestran entre paréntesis). Se muestran la cadena principal primaria de la hebra hija (~P¹-P²-) y la hebra molde diana (-P¹'-P²') como un híbrido (185). Cada subunidad de la hebra hija es un motivo repetitivo que comprende un miembro de sonda y un miembro de atadura. Los motivos tienen variabilidad específica de especie, indicada aquí por el superíndice α. Cada subunidad particular en la cadena hija se selecciona de una colección de motivos por un procedimiento dirigido por molde y su sonda se une a una secuencia correspondiente de nucleótidos complementarios en la hebra molde. De esta manera, la secuencia de residuos de nucleobase de las sondas forma una copia complementaria y contigua de la hebra molde diana.

Cada tilde (~) indica un enlace escindible selectivamente. El enlace interno entre los restos P¹ y P² de un miembro de sonda no son enlaces escindibles selectivamente, pero los enlaces entre sondas (entre subunidades) son necesariamente escindibles selectivamente como se requiera para expandir las ataduras y el Xpandómero. En un aspecto, no se forma un enlace directo entre las sondas de subunidades separadas, eliminando de este modo la necesidad de una escisión selectiva posterior.

La hebra hija se compone del Xpandómero precursor llamado "Xpandómero restringido" que además se compone de ataduras en la "configuración restringida". Cuando las ataduras se convierten a su "configuración expandida", el Xpandómero restringido se convierte en el producto Xpandómero. Las ataduras se restringen por los enlaces x formados por un puente con los miembros de sonda de subunidades contiguas y, opcionalmente, los enlaces intraataduras si todavía están presentes. El enlace x une el miembro de atadura de una primera subunidad al extremo colindante de una segunda subunidad contigua y se forma uniendo los grupos conectores colocados δ, de la primera subunidad, y ε, de la segunda subunidad.

- Se puede observar que la hebra hija tiene dos cadenas principales, un "cadena principal primario" y la cadena principal del "Xpandómero restringido". La cadena principal primaria está compuesta de los sustratos de sonda colindantes. La "cadena principal del Xpandómero restringido" evita el enlace escindible selectivamente entre los sustratos de subunidades y está formado por restos de cadena principal unidos por el enlace χ , siendo cada resto de cadena principal un enlace lineal de atadura, a P^2 , a P^1 , uniendo cada enlace χP^1 a la atadura del siguiente resto de cadena principal. Se puede observar que la cadena principal del Xpandómero restringido forma un puente o forma un bucle sobre los enlaces escindibles selectivamente de la cadena principal primaria, y permanecerá covalentemente intacto cuando estos enlaces escindibles selectivamente se escindan y la cadena principal primaria se fragmente.
- En la figura 18B, los grupos conectores δ y ϵ se han reticulado y ahora forman un enlace χ entre subunidades. Después de que se forma el enlace χ , el enlace intraatadura se puede romper, aunque aquí se muestra intacto (línea punteada en el sustrato). En general, la formación del enlace χ depende de la proximidad del grupo conector δ en la primera subunidad y la posición del grupo conector ϵ de una segunda subunidad colindante, de modo que se colocan y se ponen en contacto durante o después del ensamblaje dirigido por molde de construcciones de sustrato.
- En otros aspectos, la reticulación depende solo de la hibridación con el molde para unir los dos grupos conectores. Todavía en otros aspectos, el enlace del enlace χ está precedido por acoplamiento enzimático de los miembros de sonda P a lo largo de la cadena principal primaria, con formación de enlaces fosfodiéster entre sondas contiguas. En la estructura que se muestra en el presente documento, se ha formado la cadena principal primaria de la hebra hija, y los enlaces entre sustratos se representan con una tilde (~) para indicar que son escindibles selectivamente. Después de disociar o degradar la hebra molde diana, escindiendo los enlaces escindibles selectivamente (que incluyen los enlaces intraatadura), el Xpandómero restringido se libera y se convierte en el producto Xpandómero.
- La figura 18C es una representación del producto Xpandómero de clase II después de la disociación de la hebra molde y después de la escisión de los enlaces escindibles selectivamente (incluyendo los de la cadena principal primaria y, si no están ya escindidos, los enlaces intraatadura). Los procedimientos para la disociación de la hebra molde incluyen la desnaturalización por calor o digestión selectiva con una nucleasa o degradación química. La cadena del producto Xpandómero contiene una pluralidad de subunidades κ , donde κ denota la $\kappa^{\text{ésima}}$ subunidad en una cadena de m subunidades que componen la cadena hija, donde $\kappa = 1, 2, 3 \dots m$, donde $m > 3$, y en general $m > 20$, y preferentemente $m > 50$, y más preferentemente $m > 000$. Cada subunidad está formada por una atadura y restos de sonda P^1 y P^2 . Se observa atadura, T (181), en su configuración expandida y se estira hasta su longitud entre P^2 y P^1 de subunidades contiguas. Cada subunidad, un motivo de subunidad α , contiene información genética específica de especie establecida por ensamblaje dirigido por molde del intermedio de Xpandómero (hebra hija).
- La figura 18D muestra la construcción de sustrato de la figura 18A como un modelo molecular, donde el miembro de sonda (180), representado con cuatro residuos de nucleobase (81, 82, 83, 84), se une a la atadura (181) por un enlace del primer resto de extremo de la atadura (184). Un enlace intraatadura (85) de un segundo extremo está en el extremo distal de la atadura. Un grupo conector (δ) (86) también está dispuesto en el segundo resto de extremo y el segundo grupo conector correspondiente (ϵ) (87) está anclado al extremo de la sonda opuesto al grupo conector (δ). El bucle de atadura que se muestra aquí tiene tres indicadores (78, 79, 80), que también pueden ser específicos de la especie del motivo.
- La figura 18E muestra la construcción de sustrato después de la incorporación en el producto Xpandómero. Las subunidades se escinden y expanden y se unen mediante enlaces χ (88), formados uniendo los grupos conectores δ y ϵ a los que se hace referencia en la figura 18A. Una subunidad se indica mediante líneas punteadas que enmarcan verticalmente la subunidad repetitiva, como se representa por corchetes en la figura 18C adjunta. "E" indica de nuevo expansión.
- En el producto Xpandómero (figura 18E), la cadena principal primaria se ha fragmentado y no es covalentemente contiguo porque se ha escindido cualquier enlace directo entre las sondas de subunidades contiguas. A través del procedimiento de escisión, el Xpandómero restringido se libera para convertirse en el producto Xpandómero. Los miembros de atadura que estaban anteriormente en una configuración restringida, están ahora en una configuración expandida, funcionando de este modo para extender linealmente la información de secuencia de la diana del molde. La expansión de las ataduras reduce la densidad lineal de la información de secuencia a lo largo del Xpandómero y proporciona una plataforma para incrementar el tamaño y la abundancia de los indicadores, lo que a su vez mejora la relación señal-ruido para la detección y descodificación de la secuencia molde.
- Si bien la atadura se representa como una construcción de indicadores con tres grupos indicadores, se pueden organizar diversas configuraciones de indicadores en la atadura y pueden comprender indicadores individuales que identifican los constituyentes de sonda, indicadores individuales que identifican especies de sonda, códigos de barras moleculares que identifican especies de sonda o la atadura puede ser de polímero desnudo. En algunos casos, uno o más precursores indicadores se disponen en la atadura, y los indicadores se unen por afinidad o se unen covalentemente después del ensamblaje del producto Xpandómero.
- Las construcciones oligoméricas de clase III, ilustradas en las figuras 19A a 19E, son isómeros de las construcciones de clase II analizadas anteriormente. No se incluye más descripción porque la descripción de la clase II es adecuada para

comprender esta clase.

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

La figura 20 representa un esquema resumido de un procedimiento para elaborar un Xpandómero de clase II; el procedimiento ilustra la elaboración y uso de sustratos y productos que se muestran en las figuras 18D y 18E. El procedimiento se realiza con químicas de fase sólida. Los procedimientos para aliviar la estructura secundaria en el molde se analizan en una sección posterior. Las condiciones adecuadas adaptadas para hibridación y acoplamiento químico son bien conocidas en la técnica y las condiciones se pueden optimizar fácilmente por un experto en este campo.

La etapa I de la figura 20 muestra una mezcla de reacción que contiene un molde inmovilizada (200) y una colección de reactivos de sustrato (201). Se observa que las construcciones de sustrato se unen específicamente al molde en una hibridación dirigida por molde. Se ajustan las condiciones para optimizar la complementariedad y fidelidad de la unión. Como se muestra en el inserto de la figura (véase la figura 20a), cada construcción de sustrato colindanteacerca el grupo funcional δ (202) en el aspecto distal de la atadura, que se muestra aquí unida al tallo de la atadura por una reticulación intraatadura (203), representada por los triángulos contiguos, y el grupo funcional ϵ (204) del miembro de sonda colindante.

En la etapa II, se produce una reacción de reticulación entre los extremos colindantes de forma próxima hibridados de los miembros de sonda que implican los dos grupos funcionales δ y ϵ , formando de este modo un enlace de unión atadura a sonda entre subunidades χ (205), representado como un óvalo abierto, como se muestra en el inserto de la figura (véase la figura 20b). La hibridación se produce en paralelo en diversos sitios en el molde, de forma indiscriminada, y el acoplamiento químico se puede producir en un ciclo de hibridación (etapa III), fusión y/o lavado riguroso (etapa IV) y acoplamiento químico (etapa V). El ciclo se puede repetir para incrementar el número de subunidades contiguas ensambladas para formar el intermedio de Xpandómero. La etapa VI ilustra un intermedio de Xpandómero completo con dos cadenas de producto contiguas de longitud variable. Se puede emplear un procedimiento similar con los Xpandómeros de clase III.

La figura 21 ilustra un procedimiento de ligación procesiva de sustratos de clase II en un molde inmovilizado. La etapa I muestra un cebador (210) que hibrida con el molde (212), el cebador adaptado con un grupo funcional químicamente reactivo ϵ que se muestra en el inserto de la figura como (214) (véase la figura 21a). A continuación, se añade en la etapa II una mezcla de reacción que contiene sustratos de clase II (216). Como se muestra en el inserto de la figura (figura 21a), estas construcciones de sustrato tienen reactividad δ (217) y ϵ (214) en extremos opuestos del miembro de sonda-atadura. Se observa que una primera construcción de sustrato se une específicamente al molde en una hibridación dirigida por molde. Se ajustan las condiciones para optimizar la complementariedad y fidelidad de la unión. A continuación, se usa una ligasa para unir covalentemente la primera sonda al cebador (etapa II).

En las etapas III y IV, el procedimiento de hibridación procesiva y ligación de construcciones de sustrato continúa para desarrollar el intermedio de Xpandómero que se muestra formado en la etapa IV. Después de esto, en la etapa V, se realiza la reticulación entre los grupos δ (217) y ϵ (214) (véase la figura 21b), lo que da como resultado un enlace χ como se muestra en la figura 21c como (219). Como se muestra en los insertos de las figuras (figuras 21b y 21c), el grupo funcional δ (217) en la atadura está restringido por una reticulación intraatadura (211), representado por los triángulos contiguos, hasta que se forma el enlace χ . El intermedio de Xpandómero completo se disocia opcionalmente de la hebra molde y se escinde para formar un producto Xpandómero adecuado para secuenciación. Se puede emplear un procedimiento similar con los Xpandómeros de clase III. Este procedimiento también se puede adaptar para su uso con una polimerasa sustituyendo construcciones de sustrato de trifosfato.

Construcciones oligoméricas de clase IV y V

En referencia a las figuras 22A a 22E, se describen construcciones oligoméricas de clase IV con más detalle.

Las figuras 22A a 22C se leen de izquierda a derecha mostrándose primero la construcción de sustrato de la sonda (precursores Xsonda o Xmero de Xpandómero), a continuación la hebra hija híbrida intermedia en el centro y, a la derecha, el producto Xpandómero preparado para secuenciación.

La figura 22A muestra una construcción de sustrato de clase IV que tiene un miembro de sonda oligomérico (229) con restos de sonda P^1 y P^2 que se unen a la atadura, T (220). La atadura T está unida a P^1 y P^2 por un enlace apropiado con los restos de extremo primero y segundo de la atadura, respectivamente. Los grupos conectores ϵ del primer resto de extremo y δ del segundo resto de extremo se sitúan cerca de los extremos R^1 y R^2 de la sonda, respectivamente (las posiciones de los grupos funcionales se pueden invertir). En condiciones controladas, los grupos funcionales δ (222) y ϵ (221) reaccionarán para formar un enlace χ como se muestra en la figura 22B. Estos grupos de enlace se sitúan en la construcción de sustrato para limitar estas reacciones de enlace con construcciones de sustrato colindantes. La construcción de sustrato preferentemente no se une consigo misma. Las químicas de enlace y protección/desprotección adecuadas para δ , ϵ y χ se detallan en la descripción general de la construcción oligomérica.

R¹ y R² son grupos de extremo configurados como es adecuado para el protocolo de síntesis en el que se usa la construcción de sustrato. Por ejemplo, R¹ = 5'-fosfato y R² = 3'-OH encontrarían uso en un protocolo de ligación, y R¹ = 5'-trifosfato y R² = 3'-OH para un protocolo de polimerasa. Opcionalmente, R² se puede configurar con un grupo de bloqueo reversible para la adición cíclica de un solo sustrato. De forma alternativa, R¹ y R² se pueden configurar con grupos de extremo conectores para acoplamiento químico o sin grupos conectores para un protocolo solo de hibridación. R¹ y R² pueden ser del tipo general XR, en el que X es un grupo de enlace y R es un grupo funcional.

Las construcciones de sustrato son reactivos que se usan para el ensamblaje dirigido por molde de una hebra hija, una composición intermedia para producir Xpandómeros. La figura 22B muestra la hebra hija híbrida, un heteropolímero con subunidades repetitivas (que se muestran entre paréntesis). Se muestran la cadena principal primaria de la hebra hija (-P¹~P²-) y la hebra molde diana (-P¹·P²-) como un híbrido (228). Cada subunidad de la hebra hija es un motivo repetitivo que comprende un miembro de sonda y un miembro de atadura. Los motivos tienen variabilidad específica de especie, indicada aquí por el superíndice α. Cada subunidad particular en la cadena hija se selecciona de una colección de motivos por un procedimiento dirigido por molde y su sonda se une a una secuencia correspondiente de nucleótidos complementarios en la hebra molde. De esta manera, la secuencia de residuos de nucleobase de las sondas forma una copia complementaria y contigua de la hebra molde diana.

La tilde (~) indica un enlace escindible selectivamente. El enlace interno entre los restos P¹ y P² de un miembro de sonda es necesariamente escindible selectivamente como se requiere para expandir las ataduras y el Xpandómero. En un aspecto, no se forma enlace directo entre las sondas de subunidades separadas.

La hebra hija se compone de un precursor de Xpandómero llamado "Xpandómero restringido" que además se compone de ataduras en la "configuración restringida". Cuando las ataduras se convierten a su "configuración expandida", como se muestra en la figura 22C, el Xpandómero restringido se convierte en el producto Xpandómero. Las ataduras se restringen por los enlaces χ formados por un puente al miembro de atadura de subunidades contiguas y por los enlaces de sonda. El enlace χ une el miembro de atadura de una primera subunidad a la atadura de una segunda subunidad contigua y se forma uniendo los grupos conectores colocados, δ de la primera subunidad, y ε de la segunda subunidad.

Se puede observar que la hebra hija tiene dos cadenas principales, un "cadena principal primario" y la cadena principal de la "cadena principal del Xpandómero restringido". La cadena principal primaria está compuesta de los sustratos de sonda colindantes. La "cadena principal de Xpandómero restringido" es el enlace lineal de las ataduras en cada subunidad unidas entre sí por los enlaces χ que evitan los sustratos de sonda de la subunidad. El enlace χ resulta de una reacción del grupo funcional ε de una primera subunidad con el grupo funcional δ de una segunda subunidad colindante. Se puede observar que la cadena principal del Xpandómero restringido forma un puente o forma un bucle sobre los enlaces escindibles selectivamente de la cadena principal primaria, y permanecerá covalentemente intacto cuando estos enlaces escindibles selectivamente se escindan y la cadena principal primaria se fragmente.

En la figura 22B, los grupos conectores δ y ε se han reticulado y ahora forman un enlace χ entre subunidades. En general, la formación del enlace χ depende de la colocación del grupo conector δ, en la primera subunidad, y el grupo conector ε, de una segunda subunidad colindante, de modo que se pongan en contacto durante o después del ensamblaje dirigido por molde de construcciones de sustrato.

En otros aspectos, la reticulación del enlace χ depende solo de la hibridación con el molde para unir los dos grupos conectores. En otros aspectos más, la formación del enlace χ está precedida por acoplamiento enzimático de los miembros de sonda P a lo largo de la cadena principal primaria con enlaces fosfodiéster entre sondas contiguas. En la estructura que se muestra en la figura 22B, se ha formado la cadena principal primaria de la hebra hija, y el enlace entre restos de sonda se representa por una tilde (~) para indicar que es escindible selectivamente. Después de disociar o degradar la hebra molde diana, escindiendo los enlaces escindibles selectivamente, el Xpandómero restringido se libera y se convierte en el producto Xpandómero como se muestra en la figura 22C.

A este respecto, la figura 22C es una representación del producto Xpandómero de clase IV después de la disociación de la hebra molde y después de la escisión de los enlaces escindibles selectivamente de la cadena principal primaria. Los procedimientos para la disociación de la hebra molde incluyen la desnaturización por calor o digestión selectiva con una nucleasa o degradación química. La cadena del producto Xpandómero contiene una pluralidad de subunidades κ, donde κ denota la κ^{ésima} subunidad en una cadena de m subunidades que componen la cadena hija, donde m>3, y en general m>20, y preferentemente m>50, y más preferentemente m> 000. Cada subunidad está formada por una atadura (220) y restos de sonda colgantes P¹ y P². Se observa atadura, T, en su configuración expandida y se estira hasta su longitud entre de subunidades contiguas. Cada subunidad, un motivo de subunidad α, contiene información genética específica de especie establecida por ensamblaje dirigido por molde del intermedio de Xpandómero (hebra hija).

La figura 22D muestra la construcción de sustrato de la figura 22A como un modelo molecular, donde el miembro de sonda, representado con cuatro residuos de nucleobase (círculos abiertos), se une a la atadura por un enlace del primer resto de extremo de la atadura. El grupo conector (221), que se muestra como ε en la figura 22A, también es del primer resto de extremo de la atadura. Un grupo conector (222), que se muestra como δ en la figura 22A, está dispuesto en un segundo resto de extremo en el extremo distal de la atadura (220). El bucle de atadura que se muestra aquí tiene tres indicadores (800, 801, 802), que también pueden ser específicos de la especie del motivo. Un enlace escindible

selectivamente, que se muestra como una "V" (225), está localizado dentro del miembro de sonda (229).

La figura 22E muestra la construcción de sustrato después de la incorporación en el producto Xpandómero. Las subunidades se escinden, mostrado como líneas (225a, 225b) y expanden y se unen mediante enlaces χ (223, 224), formados uniendo los grupos conectores δ y ϵ a los que se hace referencia en la figura 22A. Una subunidad se indica mediante líneas punteadas que enmarcan verticalmente la subunidad repetitiva, como se representa por corchetes en la figura 22C adjunta.

En el producto Xpandómero de la figura 22E, la cadena principal primaria se ha fragmentado y no es covalentemente contiguo porque se ha escindido cualquier enlace directo entre las sondas de subunidades contiguas. A través del procedimiento de escisión, el Xpandómero restringido se libera para convertirse en el producto Xpandómero. Los miembros de atadura que estaban anteriormente en una configuración restringida, están ahora en una configuración expandida, funcionando de este modo para extender linealmente la información de secuencia de la diana del molde. La expansión de las ataduras reduce la densidad lineal de la información de secuencia a lo largo del Xpandómero y proporciona una plataforma para incrementar el tamaño y la abundancia de los indicadores, lo que a su vez mejora la relación señal-ruido para la detección y descodificación de la secuencia molde.

Si bien la atadura se representa como una construcción de indicadores con tres grupos indicadores, se pueden organizar diversas configuraciones de indicadores en la atadura y pueden comprender indicadores individuales que identifican los constituyentes de sonda, indicadores individuales que identifican especies de sonda, códigos de barras moleculares que identifican especies de sonda o la atadura puede ser de polímero desnudo. En algunos casos, uno o más precursores indicadores se disponen en la atadura, y los indicadores se unen por afinidad o se unen covalentemente después del ensamblaje del producto Xpandómero.

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

La figura 23A ilustra una construcción de sustrato de clase V que tiene restos de extremo primero y segundo de atadura T (239) unidos con dos enlaces de extremo escindibles selectivamente (234, 238) (representados como dos " \sim " verticales). Estos enlaces escindibles se unen a continuación a los restos de sonda primero y segundo, P¹ y P², de un miembro de sonda oligomérico (235). Los grupos conectores ϵ (230) y δ (231) de dichos restos de extremo primero y segundo se sitúan cerca de los extremos R¹ y R² de la sonda (de nuevo, las posiciones de estos grupos funcionales se pueden invertir). En condiciones controladas, los grupos funcionales δ y ϵ reaccionan para formar un enlace χ . Estos grupos de enlace se sitúan en la construcción de sustrato para limitar estas reacciones de enlace con construcciones de sustrato colindantes. La construcción de sustrato preferentemente no se une consigo misma. Las químicas de enlace y protección/desprotección adecuadas para δ , ϵ y χ se detallan en la descripción general de la construcción oligomérica.

R¹ y R² son grupos de extremo configurados como es adecuado para el protocolo de síntesis en el que se usa la construcción de sustrato. Por ejemplo, R¹ = 5'-fosfato y R² = 3'-OH encontrarían uso en un protocolo de ligación como se encuentra en Xsondas, y R¹ = 5'-trifosfato y R² = 3'-OH para un protocolo de polimerasa como se encuentra en Xmeros. Opcionalmente, R² se puede configurar con un grupo de bloqueo reversible para la adición cíclica de un solo sustrato. De forma alternativa, R¹ y R² se pueden configurar con grupos de extremo conectores para acoplamiento químico o sin grupos conectores para un protocolo solo de hibridación. R¹ y R² pueden ser del tipo general XR, en el que X es un grupo de enlace y R es un grupo funcional.

Las construcciones de sustrato son reactivos que se usan para el ensamblaje dirigido por molde de una hebra hija, una composición intermedia para producir Xpandómeros. La figura 23B muestra la hebra hija híbrida, un heteropolímero con subunidades repetitivas (que se muestran entre paréntesis). Se muestran la cadena principal primaria de la hebra hija (-P¹-P²-) y la hebra molde diana (-P¹'-P²') como un híbrido (236). Cada subunidad de la hebra hija es un motivo repetitivo que comprende un miembro de sonda y un miembro de atadura. Los motivos tienen variabilidad específica de especie, indicada aquí por el superíndice α . Cada subunidad particular en la cadena hija se selecciona de una colección de motivos por un procedimiento dirigido por molde y su sonda se une a una secuencia correspondiente de nucleótidos complementarios en la hebra molde. De esta manera, la secuencia de residuos de nucleobase de las sondas forma una copia complementaria y contigua de la hebra molde diana.

La tilde (~) indica un enlace escindible selectivamente. Los enlaces conectan los restos P¹ y P² de un miembro de sonda con la atadura y son necesariamente escindibles selectivamente como se requiere para expandir las ataduras y el Xpandómero. En un aspecto, no se forma enlace directo entre las sondas de subunidades separadas.

La hebra hija se compone de un precursor de Xpandómero llamado "Xpandómero restringido" que además se compone de ataduras en la "configuración restringida". Cuando las ataduras se convierten a su "configuración expandida", como se muestra en la figura 23C, el Xpandómero restringido se convierte en el producto Xpandómero. Las ataduras se restringen por los enlaces χ formados por un puente a los miembros de atadura de subunidades contiguas y por los enlaces escindibles selectivamente (234, 238). El enlace χ une el miembro de atadura de una primera subunidad a la atadura de

una segunda subunidad contigua y se forma uniendo los grupos conectores colocados, δ de la primera subunidad, y ϵ de la segunda subunidad.

5 Se puede observar que la hebra hija tiene dos cadenas principales, un "cadena principal primario" y la cadena principal de la "cadena principal del Xpandómero restringido". La cadena principal primaria está compuesta de los sustratos de sonda colindantes. La "cadena principal de Xpandómero restringido" es el enlace lineal de las ataduras en cada subunidad unidas entre sí por los enlaces χ que evitan los sustratos de sonda de la subunidad. El enlace χ resulta de una reacción del grupo funcional ϵ de una primera subunidad con el grupo funcional δ de una segunda subunidad colindante. Se puede observar que la cadena principal del Xpandómero restringido forma un puente o forma un bucle sobre los enlaces escindibles selectivamente que se conecta a la cadena principal primaria, y permanecerá covalentemente intacto cuando estos enlaces escindibles selectivamente se escindan y la cadena principal primaria se disocie o fragmente de otro modo.

10

15 En la figura 23B, los grupos conectores δ y ϵ se han reticulado y ahora forman un enlace χ entre subunidades. En general, la formación del enlace χ depende de la colocación del grupo conector δ , en la primera subunidad, y el grupo conector ϵ , de una segunda subunidad colindante, de modo que se pongan en contacto durante o después del ensamblaje dirigido por molde de construcciones de sustrato.

20 En algunos protocolos, la reacción de reticulación depende solo de la hibridación con el molde para unir los dos grupos reactivos. En otros protocolos, la unión está precedido por acoplamiento enzimático de los miembros de sonda, con formación de enlaces fosfodiéster entre sondas contiguas. En la estructura que se muestra en la figura 23B, se ha formado la cadena principal primaria de la hebra hija. La atadura, ahora unida a subunidades contiguas por enlaces χ , y comprende la cadena principal del Xpandómero restringido. Tras la escisión de los enlaces escindibles selectivamente (~), el Xpandómero restringido se separa de la cadena principal primaria para convertirse en el producto Xpandómero, y sus ataduras ahora sin restricciones se expanden linealmente en toda su longitud, como se muestra en la figura 23C.

25

A este respecto, la figura 23C es una representación del producto Xpandómero de clase V después de la escisión de los enlaces escindibles selectivamente que disocian la cadena principal primaria. La cadena del producto Xpandómero contiene una pluralidad de subunidades κ , donde κ denota la $\kappa^{\text{ésima}}$ subunidad en una cadena de m subunidades que componen la cadena hija, donde $m > 3$, y en general $m > 20$, y preferentemente $m > 50$, y más preferentemente $m > 000$.

30 Cada subunidad está formada por una atadura, T (239), como se observa en su configuración expandida y se estira en toda su longitud entre subunidades contiguas. Cada subunidad, un motivo de subunidad α , contiene información genética específica de especie establecida por ensamblaje dirigido por molde del intermedio de Xpandómero (hebra hija).

35 La figura 23D muestra la construcción de sustrato de la figura 23A como un modelo molecular, donde el miembro de sonda (235), representado con cuatro residuos de nucleobase (círculos abiertos), se une a los restos de extremo primero y segundo de la atadura mediante dos enlaces escindibles (234, 238). El grupo conector (232), que se muestra como ϵ en la figura 23A, es del resto de extremo primero de la atadura y el grupo conector (233), que se muestra como δ en la figura 23A, es del resto de extremo segundo de la atadura. El bucle de atadura que se muestra aquí tiene tres indicadores (237a, 237b, 237c), que también pueden ser específicos de la especie del motivo.

40 La figura 23E muestra la construcción de sustrato después de la incorporación en el producto Xpandómero. Las subunidades se escinden (234a, 234b, 238a, 238b) y expanden y se unen mediante enlaces χ (249, 248), formados uniendo los grupos conectores δ y ϵ a los que se hace referencia en la figura 23A. Una subunidad se indica mediante líneas punteadas que encierran verticalmente la subunidad repetitiva, como se representa por corchetes en la figura 23C adjunta.

45 En el producto Xpandómero de la figura 23E, la cadena principal primaria (235) se ha escindido (disociado). A través del procedimiento de escisión, el Xpandómero restringido se libera para convertirse en el producto Xpandómero. Los miembros de atadura que estaban anteriormente en una configuración restringida, están ahora en una configuración expandida, funcionando de este modo para extender linealmente la información de secuencia de la diana del molde. La expansión de las ataduras reduce la densidad lineal de la información de secuencia a lo largo del Xpandómero y proporciona una plataforma para incrementar el tamaño y la abundancia de los indicadores, lo que a su vez mejora la relación señal-ruido para la detección y descodificación de la secuencia molde.

55 Si bien la atadura se representa como una construcción de indicadores con tres grupos indicadores, se pueden organizar diversas configuraciones de indicadores en la atadura y pueden comprender indicadores individuales que identifican los constituyentes de sonda, indicadores individuales que identifican especies de sonda, códigos de barras moleculares que identifican especies de sonda o la atadura puede ser de polímero desnudo. En algunos casos, uno o más precursores indicadores se disponen en la atadura, y los indicadores se unen por afinidad o se unen covalentemente después del ensamblaje del producto Xpandómero.

60 En la figura 24 se ilustra la elaboración y uso de un Xpandómero de clase V. Al preparar la síntesis en la etapa I, se pone en contacto un molde monocatenario (245) y se hibrida con el cebador de secuenciación (246). El ensamblaje de cebador (247) se pone en contacto a continuación con una colección 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, se completa la polimerización de la cadena principal primaria de la hebra hija y los grupos funcionales reactivos

de los brazos laterales de la atadura colindantes se reticulan, formando los enlaces χ de atadura a atadura. Finalmente, en la etapa V, se cortan los enlaces escindibles en los tallos de los bucles de atadura, liberando la cadena principal sintética de atadura a atadura de la hebra hija oligomérica y el molde. Por tanto, este Xpandómero (249) esta construido enteramente de enlaces de atadura y se muestra que se expande espontáneamente a medida que se aleja del resto del intermedio sintético. En el presente documento, la información genética correspondiente a la secuencia de polinucleótidos diana se codifica en las subunidades contiguas de las ataduras.

Elaboración y uso de Xmeros

10 La clase I incluye Xsondas y Xmeros. Los Xsondas son monofosfatos, mientras que los Xmeros son trifosfatos. Los "Xmeros" son construcciones de sustrato de oligonucleótido trifosfato expandibles que se pueden polimerizar en una síntesis dirigida por molde y dependiente de enzimas de un Xpandómero. Al igual que las Xsondas, las construcciones de sustrato de Xmeros tienen una forma característica de "sonda-bucle" como se ilustra en las figuras 10A y 10C, donde R¹ es 5'-trifosfato y R² es 3'-OH. Hay que tener en cuenta que las construcciones de sustrato son trifosfatos de 15 oligonucleobase o trifosfatos de análogos de oligómeros, pero los miembros de sonda (es decir, el oligómero) se han modificado con una construcción de atadura y un enlace escindible selectivamente entre los enlaces de extremo de la atadura como se muestra en la figura 10D, cuya función se ilustra además en la figura 10E.

20 Las ADN y ARN polimerasas pueden incorporar oligonucleótidos de dinucleótido, trinucleótido y tetranucleótido trifosfato con un nivel de eficacia y fidelidad en un procedimiento procesivo dependiente de cebador como se divulga en la patente de EE. UU. n.º 7.060.440 concedida a Kless. Los trifosfatos de oligonucleótidos modificados con atadura de longitud n (n = 2, 3, 4 o más) se pueden usar como sustratos para su incorporación basada en polimerasa en Xpandómeros. Las 25 enzimas adecuadas para su uso en los procedimientos que se muestran en las figuras 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 Klenow, ADN polimerasa de *Thermophilus aquaticus*, ADN polimerasa Tth, ADN polimerasa VentR® (New England Biolabs), ADN polimerasa Deep VentR® (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 30 polimerasa eucariota beta, telomerasa, polimerasa Therminator™ (New England Biolabs), ADN polimerasa KOD HiFi™ (Novagen), ADN polimerasa KOD1, Q-beta replicasa, transferasa terminal, retrotranscriptasa de AMV, retrotranscriptasa de M-MLV, retrotranscriptasa de Phi6, retrotranscriptasa de HIV-1, polimerasas novedosas descubiertas por bioprospección y polimerasas citadas en los documentos US 2007/0048748, US 6329178, US 6602695 y US 6395524. Estas 35 polimerasas incluyen isoformas naturales, mutantes y variantes genomanipuladas.

35 La polimerización de Xmeros es un procedimiento para la síntesis de Xpandómero y se ilustra en la figura 16, por ejemplo, donde el sustrato dimérico se proporciona como un trifosfato. Debido a que los Xmeros se polimerizan de forma procesiva, las etapas de extensión, reticulación, activación de extremo y lavado de alta restricción típicamente asociadas con la secuenciación cíclica por procedimientos de síntesis se eliminan opcionalmente con este enfoque. Por tanto, la reacción 40 se puede realizar en solución. La síntesis de Xpandómeros con Xmeros también se puede realizar con moldes inmovilizados, como se ilustra en la figura 17, donde un trifosfato de Xmero tetramérico se polimeriza procesivamente en una síntesis dirigida por molde dependiente de un cebador.

45 Se puede emplear una variedad de procedimientos para la síntesis robusta de Xmeros de 5' trifosfato. Como describen Burgess y Cook ("Syntheses of Nucleoside Triphosphates", Chem. Rev. 100(6):2047-2060, 2000), estos procedimientos incluyen (pero sin limitaciones) reacciones que usan fosforamiditas de nucleósidos, síntesis por medio de ataque nucleofílico de pirofosfato sobre monofosfatos de nucleósidos activados, síntesis por medio de ataque nucleofílico de fosfato sobre pirofosfato de nucleósido activado, síntesis por medio ataque nucleofílico de difosfato sobre sintón de fosfato activado, síntesis que implica fosfitos activados o fosforamiditas derivadas de nucleósidos, síntesis que implica desplazamiento directo de grupos salientes 5'-O por nucleófilos de trifosfato y procedimientos biocatalíticos. Un 50 procedimiento representativo para producir sustratos de dinucleótidos compatibles con polimerasa usa N-metilimidazol para activar el grupo 5' monofosfato; la reacción posterior con pirofosfato (sal de tributílamonio) produce el trifosfato (Abramova *et al.*, "A facile and effective synthesis of dinucleotide 5'-triphosphates", Bioorganic and Med Chem 15, 6549-6555, 2007).

55 Como se analiza con más detalle a continuación, la construcción de atadura con Xmeros está relacionada en diseño, composición y enlace con la de atadura usadas para Xsondas. En muchos aspectos, se codifica información genética en la atadura y, por lo tanto, cada atadura de cada construcción de sustrato es una atadura específica de especie. La información codificada en la atadura se codifica con un código de indicador que digitaliza la información genética. Por ejemplo, la codificación binaria de cinco bits en las ataduras produciría 32 códigos de secuencia únicos (2⁵). Esta estrategia se puede usar para codificar de forma única las 16 combinaciones de dos residuos de nucleobase por miembro de sonda de una colección de dímeros, independientemente de la orientación de la atadura. De forma similar a la codificación de Xsondas, se puede considerar una variedad de estrategias de funcionalización y marcado para Xmeros, que incluyen (pero sin limitaciones): dendrímeros funcionalizados, polímeros, polímeros ramificados, nanopartículas y nanocróstales como parte de la estructura de la atadura, así como también químicas de indicador y señales de indicador, que se va a detectar con la tecnología de detección adecuada. Se pueden introducir marcadores específicos de bases

(por medio de la unión a la atadura) antes o bien después de la polimerización de Xmersos, por unión covalente o dirigida por afinidad.

Diseño y síntesis de Xsondas y Xmersos

5

A continuación se presenta una visión general de las estrategias sintéticas y de escisión, comenzando con los oligómeros de sonda con enlaces escindibles selectivamente, seguidos de las construcciones de atadura y atadura de indicador.

10 Un objetivo de un procedimiento de SBX basado en Xsondas o Xmersos es ensamblar una réplica del ácido nucleico diana tan completa y eficazmente como sea posible mediante una síntesis dirigida por molde, en general un procedimiento o una combinación de procedimientos seleccionados entre hibridación, ligación, polimerización, o reticulación química de composiciones de precursores adecuadas, denominadas aquí "sustratos". Los sustratos de Xsondas y Xmersos se suministran como colecciones de reactivos (por ejemplo, como partes de kits para secuenciación) con este fin. Las colecciones son, en general, de naturaleza combinatoria y contienen miembros de sonda seleccionados para que se unan específicamente a cualquiera o todas las secuencias complementarias tal como las que se encontrarían en un polinucleótido diana. El número de sondas requeridas en una colección para este propósito es una función del tamaño de la sonda. Se puede considerar que cada sonda es un fragmento de secuencia, y debe estar presente variedad suficiente de miembros de sonda para formar una copia contigua de la secuencia contigua de fragmentos de secuencia complementarios del polinucleótido diana. Para las sondas en las que cada oligómero es un dímero, existen 16 posibles combinaciones de especies de A, T, C y G. Para las sondas en las que cada oligómero es un trímero, existen entonces 64 posibles combinaciones de especies de A, T, C y G, etc. Cuando se secuencian fragmentos genómicos aleatorios, es probable que se requieran todas esas especies en una colección de reactivos.

15 25 Las Xsondas y los Xmersos son construcciones de sustrato oligomérico que se dividen en cinco clases funcionales diferentes. Las construcciones de sustrato oligomérico tienen dos componentes funcionales distintivos: una oligonucleobase modificada o miembro de "sonda" y un miembro de atadura ("T"). La sonda se une al miembro de atadura mediante una construcción de "sonda-bucle", donde el bucle de atadura es un precursor del miembro de atadura linealizado del producto final Xpandómero. Cada atadura T se puede codificar con indicadores (comúnmente denominados "marcas" o "marcadores"), o combinaciones de los mismos, que identifican de forma única la secuencia de sonda a la que 30 está atada. De esta manera, la información de secuencia del Xpandómero ensamblado se detecta más fácilmente.

35 40 El oligómero es la porción de la sonda de la Xsonda. La sonda es una oligonucleobase modificada que tiene una cadena de x desoxirribonucleótidos, ribonucleótidos o, más en general, residuos de nucleobase (donde x puede ser 2, 3, 4, 5, 6 o más). En estos análisis, una sonda con 2, 3, 4, 5 o 6 residuos de nucleobase de longitud se puede denominar dímero, trímero, tetrámero, pentámero o hexámero, respectivamente.

45 Los reactivos de la construcción de sustrato se pueden sintetizar con una cadena principal de oligonucleótido 5'-3' fosfodiéster, teniendo el oligómero los nucleótidos A, T, G y C (estructuras que se muestran en la tabla de la figura 25), u otros análogos de ácidos nucleicos hibridables tales como los que tienen una cadena principal de péptido, cadena principal de fosfonopéptido, cadena principal de serina, cadena principal de hidroxiprolina, cadena principal mixta de péptido-fosfono-péptido, cadena principal mixta de péptido-hidroxiprolina, cadena principal mixta de hidroxiprolina-fosfono-péptido, cadena principal mixta de serina-fosfono-péptido, cadena principal de treosa, cadena principal de glicol, cadena principal de morfolino y similares, como se conocen en la técnica. También se pueden usar como sondas oligómeros de ácido desoxirribonucleico y oligómeros de ácido ribonucleico, y oligómeros mixtos de los dos. También se pueden sustituir otras bases, tal como timina por uracilo e inosina como base degenerada. También se pueden usar residuos fragmentarios de nucleobases que tienen complementariedad.

50 Una lista más completa de bases degeneradas y tambaleantes conocidas en la técnica incluye, pero sin limitaciones, xantina, hipoxantina o un derivado heterocíclico, análogo o tautómero de xantina e hipoxantina, 8-azapurina, purinas sustituidas en la posición 8 con metil- 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-(C3-C6)-alquinilcitosina, 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 bases nitrogenadas descritas en las patentes de EE. UU. n.º 5.432.272 y 6.150.510, PCT WO 92/002258 publicadas, los documentos WO 93/10820, WO 94/22892 y WO 94/22144, y en Fasman, Practical Handbook of Biochemistry and Molecular Biology, págs. 385-394, CRC Press, Boca Raton, LA, 1989.

55 60 Como es conocido en la técnica, se pueden diseñar oligómeros para incluir modificadores de nucleótidos. En algunos aspectos, estos sirven como puntos de unión para el miembro o miembros de atadura. Los derivados de purina y pirimidina adecuados para la síntesis de oligómeros derivatizados son bien conocidos en la técnica. Dos de dichas 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.

65 Como se ilustra en las figuras 27A y 27B, tomando como ejemplo un oligómero tetramérico (aquí ilustrado como 5'-monofosfato), cualquiera de las cuatro posiciones de la base en el oligómero se puede modificar para crear puntos de unión mediante sustancias químicas conocidas. Los nucleótidos modificados en los residuos de sonda 2 y 3 (en lados

opuestos de un enlace escindible de selectivamente, representado como "V") se ilustran en la figura 27A. Esta figura ilustra un oligómero tetramérico con conectores amino unidos a la citosina y guanosina del oligómero. La figura 27B ilustra un oligómero tetramérico con grupos funcionales de benzaldehído para la citosina y guanosina del oligómero. Los detalles son ilustrativos de procedimientos bien conocidos en la técnica. Para simplificar, la mayoría de las ilustraciones proporcionadas en el presente documento asumirán tetrámeros a menos que se indique de otro modo, pero se entiende que se pueden emplear otras colecciones de construcciones de sustrato o combinaciones de colecciones en la práctica de la presente invención.

Escisión

En general, las construcciones de sustrato de Xsondas y Xmeros tienen enlaces escindibles selectivamente que permiten la expansión controlada de la atadura. Como se menciona previamente, dicha escisión selectiva se puede lograr por cualquier número de técnicas conocidas por un experto en la técnica, que incluyen, pero sin limitaciones, escisión de la cadena principal de fosforotiolato con cationes metálicos, escisión ácida de modificaciones de la cadena principal de fosforamidato, escisión de nucleasa selectiva de enlaces fosfodiéster estándar usando modificaciones de fosforotioato resistente a nucleasas para la protección de la cadena principal, fotoescisión de conectores de la cadena principal modificados con nitrobencilo y reducción de enlaces disulfuro.

En las figuras 28A a 28D y en las figuras 29A a 29D se ilustra la modificación de las sondas de sustrato para incluir enlaces escindibles selectivamente. La figura 28A muestra un ejemplo de un dímero de Xsonda con un grupo ribosil 2'-OH susceptible de escisión por ribonucleasa H en un intermedio de Xpandómero híbrido de ADN/ARN. Por tanto, el enlace es escindible selectivamente, siempre que el(los) otro(s) nucleótido(s) en la Xsonda sea(n) resistente(s) a la escisión de la RNasa (por ejemplo, nucleobases de 2'-o-metil pentosa, 2' desoxirribosa, nucleobases de LNA "bloqueadas" y nucleobases unidas 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.

La figura 28B muestra una Xsonda con un enlace fosfodiéster que acopla dos nucleótidos. En esta figura, así como en las figuras 28A, 28C y 28D, se indican en (282) y (284) ataduras para formar un punto con el enlace escindible selectivamente de la sonda. Este enlace es escindible selectivamente con Mung Bean Nuclease, nucleasa S1, DNase I u otras DNAsas, por ejemplo, si otros enlaces que unen las ataduras de las subunidades son resistentes a nucleasas. La síntesis de una colección de dímeros, por ejemplo, con un enlace de fosfato estándar entre los puntos de unión de la atadura y el(los) enlace(s) de fosforotioato en la(s) posición(es) de la cadena principal del nucleótido que van a permanecer intactos, proporciona el patrón de escisión deseado. La figura 28C es un dímero de Xsonda que se mantiene unido por un enlace 3'-fosforotiolato, y en la figura 28D se mantiene unido por un enlace 5'-fosforotiolato. Estos enlaces son escindibles selectivamente por ataque químico, por ejemplo, con yodoetanol como describen 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 describen Vyle *et al.* ("Sequence- and strand-specific cleavage in oligodeoxyribonucleotides and DNA containing 3'-thiethylimidine". Biochemistry 31(11): 3012-8, 1992). Otras opciones de escisión de la cadena principal incluyen, pero sin limitaciones, escisión fotorredox inducida por UV (como por adaptación de grupos de fotoescisión de nitrobencilo) como describen Vallone *et al.* ("Genotyping SNPs using a UV-photo cleavable oligonucleotide in MALDI-TOF MS", Methods Mol. Bio. 297:169-78, 2005), escisión ácida de enlaces de fosforamidato 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 de la cadena principal de 3'-O-B-D-ribofuranosil-2'-desoxi como describen en 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).

Al igual que con las Xsondas, la escisión de la cadena principal del poli-Xmero para producir un Xpandómero se logra de diversas maneras. Como se muestra en la figura 29A, por ejemplo, un Xmero que contiene una base de ribonucleótido digerible por RNasa se puede escindir selectivamente en esa 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 de 2' desoxirribosa, nucleobases de LNA "bloqueadas" y nucleobases unidas a glicol o péptido). Para el Xmero descrito en la figura 29A, la base en 5' es una 2' hidroxilo ribonucleótido citidina estándar y la base en 3' es una de de 2' desoxirribonucleótido guanina resistente a RNasa. El diseño de los Xmeros permite la escisión selectiva por RNasa de la cadena principal del Xmero para expandir un Xpandómero. De forma alternativa, como se muestra en la figura 29B, la DNase se puede usar para escindir todos los enlaces de la cadena principal no protegidos distintos de fosforotioato. En consecuencia, una colección de dímeros, por ejemplo, con un enlace de fosfato estándar entre los puntos de unión de la atadura y el(los) enlace(s) de fosforotioato en la(s) posición(es) de la cadena principal del nucleótido que van a permanecer intactos, proporciona el patrón de escisión deseado. La figura 29C es un Xmero dímerico que se mantiene unido por un enlace 3'-fosforotiolato, y la figura 29D se mantiene unida por un enlace 5'-fosforotiolato. Estos enlaces son escindibles selectivamente por ataque químico, por ejemplo, con yodoetanol o por escisión con cationes metálicos divalentes como se menciona previamente. Otras opciones de escisión de la cadena principal incluyen (pero sin limitaciones) escisión fotorredox inducida por UV (como por adaptación de grupos de fotoescisión de nitrobencilo) y escisión ácida de enlaces de fosforamidato, ambos de los cuales se citan anteriormente en la figura 28. En las figuras 29A a 29D, se indican en (292) y (294) ataduras para formar un punto con el enlace escindible selectivamente de la sonda.

Volviendo ahora a la figura 30, en un primer esquema general para la síntesis de construcciones de sustrato de "sonda-bucle" de clase I, se modifican dos residuos de nucleobase (círculos) en las posiciones segunda y tercera de la sonda para crear puntos de unión L1 y L2 para los dos extremos L1' y L2' de la atadura. La atadura se muestra aquí como preensamblada por separado y se une al miembro de sonda en un etapa sintética (flecha). Los enlaces disulfuro intraatadura (representados por los dos triángulos) se pueden usar en el ensamblaje y uso de estas construcciones de sustrato. La introducción de un agente reductor en el producto Xpandómero romperá selectivamente los puentes disulfuro que mantienen juntas la atadura, permitiendo, por tanto, la expansión de la cadena principal del Xpandómero. Los enlaces fotoescindibles también son útiles para plegar ataduras durante el ensamblaje con posterior liberación y despliegamiento tras la exposición a la luz.

En otros aspectos, la cadena principal de fosfodiéster del sustrato se puede modificar para crear puntos de unión para la atadura, como se divulga en 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 phosphodiesters in DNA", Curr Med Chem. 8(10):1147-55, 2001), Cook *et al.* (patente de EE. UU. n.º 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 componen el miembro de sonda se pueden sustituir con análogos de nucleobases para alterar la funcionalidad de la Xsonda. Por ejemplo, se pueden usar ácidos nucleicos bloqueados ("LNA") para aumentar la estabilidad del híbrido de la sonda. Si se pretende el acoplamiento químico de Xsondas (en lugar de la ligación enzimática), los extremos 5' y 3' de la sonda se pueden derivatizar aún más para permitir la reticulación química.

Diseño, composición y síntesis de construcciones de indicadores

En un modo de realización, las ataduras se codifican con "construcciones de indicadores" que identifican de forma única la secuencia de residuos de nucleobase (o "sonda" de Xsondas, Xmeros y otros sustratos de oligómeros de la figura 8) o nucleobase (como en XNTP, RT-NTP y sustratos monoméricos de la figura 9) a los que se ata. Los indicadores son indicadores o combinaciones de indicadores asociados en general con las ataduras que sirven para "analizar" o "codificar" la información de secuencia inherente a los sustratos e inherente al orden en que los sustratos se incorporan al Xpandómero. En algunos modos de realización, la atadura es solo un espaciador y los indicadores son, o se asocian con, el sustrato.

La figura 31 muestra un procedimiento para el ensamblaje de la atadura del sustrato similar al de la figura 30, pero la atadura preensamblada incluye grupos indicadores (que se muestran como las tres porciones rectangulares de la atadura) y se denomina "construcción de indicadores". Las construcciones de indicadores y las ataduras se pueden elaborar mediante diversas químicas de polímero, y su uso y síntesis se analizan con más detalle en el presente documento.

En la práctica de la presente invención, las ataduras pueden cumplir una variedad de funciones, por ejemplo: (1) como atadura para unir secuencialmente, directa o indirectamente, ataduras contiguas a lo largo de la cadena principal de la nucleobase, (2) como espaciador para estirar o expandir para formar una cadena alargada de subunidades atadas, denominadas un Xpandómero, tras escisión de la cadena principal, y/o (3) comprende opcionalmente construcciones de indicadores o precursores de indicadores que codifican la información de la secuencia oligomérica o de nucleobases de la construcción de sustrato individual a la que se asocia la atadura.

Las construcciones de indicadores son manifestaciones físicas de los códigos de indicadores, que son de naturaleza bioinformática y digital. Los códigos de indicadores analizan o codifican la información genética asociada con la sonda o el fragmento de secuencia de nucleobases al que se une la construcción de indicadores y atadura. Las construcciones de indicadores están diseñadas para optimizar la detectabilidad del código del indicador ajustando separaciones espaciales, abundancia e intensidad de señal de los indicadores constituyentes. Las construcciones de indicadores pueden incorporar una amplia gama de elementos estructurales y de señal que incluyen, pero sin limitaciones, polímeros, dendrímeros, microesferas, aptámeros, ligandos y oligómeros. Estas construcciones de indicadores se elaboran mediante una variedad de químicas de polímeros y se analizan en más detalle a continuación.

En un modo de realización, las construcciones de indicadores se unen a la sonda o nucleobase mediante una atadura polimérica. Las ataduras se pueden construir de uno o más polímeros duraderos, solubles en agua o en disolventes, incluidos, pero sin limitaciones, el siguiente segmento o segmentos: polietilenglicoles, poliglicoles, polipiridinas, poliisocianuros, poliisocianatos, poli(triarilmelil) metacrilatos, polialdehídos, polipirrolinonas, poliureas, fosfodiésteres de poliglicol, poliacrilatos, polimetacrilatos, poliacrilamidas, ésteres de polivinilo, poliestirenos, poliamidas, poliuretanos, policarbonatos, polibutiratos, polibutadienos, polibutiroacetonas, polipirrolidinonas, polivinilfosfonatos, poliacetamidas, polisacáridos, polihialuronatos, poliamidas, poliimidás, poliésteres, polietilenos, polipropilenos, poliestirenos, policarbonatos, politereftalatos, polisilanos, poliuretanos, poliéteres, poliaminoácidos, poliglicinas, poliprolinas, polilisina N-sustituida, polipéptidos, péptidos N-sustituidos de cadena lateral, glicina poli-N-sustituida, peptoides, péptidos

sustituidos con carboxilo de cadena lateral, homopeptídos, oligonucleótidos, oligonucleótidos de ácido ribonucleico, oligonucleótidos de ácido desoxinucleico, oligonucleótidos modificados para evitar el emparejamiento de bases de Watson-Crick, análogos de oligonucleótido, ácido policitidílico, ácido poliadenílico, ácido poliuridílico, politimidina, polifosfato, polinucleótidos, polirribonucleótidos, fosfodiésteres de polietilenglicol, análogos de polinucleótidos peptídicos, 5 análogos de treosil-polinucleótido, análogos de glicol-polinucleótido, análogos de morfolino-polinucleótido, análogos de oligómero de nucleótido bloqueado, análogos de polipéptido, polímeros ramificados, polímeros de peine, polímeros de estrella, polímeros dendríticos, copolímeros aleatorios, de gradiente y de bloque, polímeros aniónicos, polímeros catiónicos, polímeros que forman tallo-bucle, segmentos rígidos y segmentos flexibles. Dichos polímeros se pueden circularizar en los puntos de unión en una construcción de sustrato como se describe, por ejemplo, en las figuras 30 y la 10 figura 31.

La atadura es en general resistente al enredo o se pliega para que sea compacta. El polietilenglicol (PEG), el óxido de polietileno (PEO), el metoxipolietilenglicol (mPEG) y una amplia variedad de derivados de PEG construidos de forma similar (PEG) son polímeros ampliamente disponibles que se pueden usar en la práctica de la presente invención. Los 15 PEG modificados están disponibles con una variedad de agentes de reticulación de extremo bifuncionales y heterobifuncionales y se sintetizan en una amplia gama de longitudes. En general, los PEG son solubles en agua, metanol, benceno, diclorometano y muchos disolventes orgánicos comunes. En general, los PEG son polímeros flexibles que típicamente no interactúan inespecíficamente con productos químicos biológicos.

20 La figura 32A ilustra la estructura repetitiva de un polímero de PEG. La figura 32B muestra una Xsonda o Xmero con una atadura de PEG desnuda fijada a la cadena principal de la sonda, por ejemplo, por conectores finalizados en amina (no se muestran) usando químicas de conector estándar. La figura 32C muestra la misma construcción de sustrato después de la escisión de la cadena principal de la sonda en un enlace escindible selectivamente ("V"), acomodando el polímero PEG de forma flexible el alargamiento del Xpandómero. En algunos modos de realización, los segmentos de polímero 25 PEG se ensamblan por partes en la atadura para proporcionar una longitud de expansión o para minimizar problemas estéricos, tal como, por ejemplo, en los tallos de las ataduras próximos a los enlaces de los extremos terminales que conectan los brazos de la atadura al sustrato.

30 Otros polímeros que se pueden emplear como ataduras y proporcionan "estructura" para los indicadores incluyen, por ejemplo, poliglicina, poliprolina, polihidroxiprolina, policisteína, poliserina, ácido poliaspártico, ácido poliglutámico y similares. Se pueden usar funcionalidades de la cadena lateral para construir estructuras ricas en grupos funcionales para una capacidad de señal o complejidad añadida.

35 La figura 33A muestra la estructura de la polilisina. En las construcciones de atadura descritas en la figura 33B a 33D, los segmentos de atadura de polilisina crean una estructura para la unión del indicador. En la figura 33B, los grupos ε-amino de las cadenas laterales de lisina (indicadas por flechas) proporcionan funcionalidad para unir una pluralidad de elementos indicadores a una construcción de sustrato, amplificando el código del indicador. La figura 33C ilustra un dendrímero en forma de estrella unido a una construcción de sustrato con cadenas laterales de polilisina (flechas).

40 La figura 33D ilustra la carga de los oligómeros de dendrímero que se pueden detectar añadiendo oligómeros complementarios marcados en una etapa de marcado y "amplificación de la señal" posterior al ensamblaje. Esto proporciona un procedimiento útil para preparar una unión universal, uniendo un complejo de dendrímero no marcado con múltiples grupos indicadores oligoméricos a una sonda, y después tratando el dendrímero unido a la sonda con una selección de una o dos sondas marcadas complementarias, se obtiene un dendrímero "pintado" específico para la especie 45 de sustrato individual. Se pueden pintar diferentes construcciones de sonda/atadura con diferentes sondas marcadas complementarias.

En otro aspecto de este enfoque, la cadena principal del sistema indicador comprende ocho oligonucleótidos únicos que se codifican espacialmente de forma binaria usando dos indicadores fluorescentes discernibles, cada uno excitado por 50 medio del mismo donante de FRET. Antes o después del acoplamiento de las construcciones de atadura a su respectiva construcción de sustrato, las construcciones de atadura se codifican por secuencia hibridando la mezcla adecuada de elementos indicadores fluorescentes para crear el código binario específico de sonda apropiado. Se emplean variaciones de este enfoque usando dendrímeros, polímeros, polímeros ramificados o microesferas codificados y no marcados como cadena principal del sistema indicador. Se pueden usar de nuevo oligonucleótidos para la codificación binaria de la 55 construcción de indicadores. Una ventaja con este enfoque es que la intensidad de señal se amplifica significativamente y que la codificación no depende de un solo acontecimiento de hibridación, lo que disminuye la posibilidad de error de medición y/o codificación.

Todavía otro aspecto reemplaza la estrategia de codificación de oligonucleótidos descrita anteriormente con ligandos 60 heteroespecíficos unidos por afinidad para producir, por ejemplo, una construcción de indicadores codificada binariamente de forma similar. Usando una estrategia de codificación binaria de 9 bits, esta construcción de indicadores no universal, en su forma más simple, emplea solo una química de acoplamiento única para marcar simultáneamente todas las ataduras.

65 Dada la flexibilidad del enfoque SBX, se usa una amplia gama de indicadores para producir señales únicas y medibles. Cada atadura está codificada de forma única por uno o muchos segmentos de indicadores distintos. La estructura a la

que se unen los restos de indicadores se puede construir usando una amplia gama de rasgos característicos estructurales existentes que incluyen, pero sin limitaciones, dendrímeros, microesferas, polímeros y nanopartículas. Dependiendo del esquema de codificación, se pueden usar uno o muchas estructuras de indicadores claramente separadas para el código de indicador de cada atadura. Cualquier cantidad de opciones está disponible para la unión directa e indirecta de restos 5 de indicadores a la estructura de indicador, que incluyen (pero sin limitaciones): codificación de indicadores de polímero(s) químicamente reactivo(s) integrado(s) en las construcciones de atadura; codificación de indicadores de grupos superficiales químicamente reactivos en dendrímero(s) integrado(s) en la cadena principal de la atadura; y codificación de indicador de grupos superficiales químicamente reactivos en microesfera(s) integrada(s) en la atadura. En este contexto, una "microesfera" se toma ampliamente para que indique cualquier partícula o microesfera cristalina, polimérica, de látex 10 o compuesta. Para los tres ejemplos, la abundancia de indicadores se puede incrementar significativamente uniendo, a las estructuras de indicadores, polímeros que se cargan con múltiples de indicadores. Estos polímeros pueden ser tan simples como una polilisina de 100 residuos o más avanzados, tal como sondas oligoméricas marcadas.

También se pueden usar construcciones de atadura de tamaño reducido. Por ejemplo, las construcciones de atadura se 15 pueden alargar en una etapa de posprocesamiento usando procedimientos dirigidos para insertar unidades de espaciado, reduciendo, por tanto, el tamaño de la atadura del indicador.

La reducción del tamaño y la masa de la construcción de sustrato también se puede lograr mediante usando ataduras no 20 marcadas. Eliminando los indicadores voluminosos (y la estructura de indicadores tales como dendrímeros, que para algunos modos de realización de codificación comprenden más de un 90 % de la masa de la atadura), se puede potenciar la hibridación y/o la cinética de acoplamiento. A continuación, se puede emplear el marcado de ataduras posterior al ensamblaje. Los indicadores están unidos a una o más químicas de enlace que se distribuyen a lo largo de las construcciones de atadura usando estrategias espaciales o combinatorias para codificar la información de secuencia de la base. Un esquema de codificación binaria simple puede usar solo una sustancia química de enlace reactivo para el 25 marcado posterior al ensamblaje del Xpandómero. Los esquemas de marcado más complicados, que pueden requerir cientos de conectores únicos, usan una estrategia basada en oligonucleótidos para el marcado del Xpandómero. Otro aspecto posterior al marcado de las Xsondas o Xmers es usar las secuencias de nucleótidos resultantes derivadas de P¹ y P² (véase la figura 10) que quedan después de la escisión y expansión del Xpandómero para la unión del indicador 30 por hibridación de una colección de sondas marcadas. De forma similar, otras técnicas de marcado y/o detección pueden identificar directamente los nucleótidos más resueltos espacialmente.

Las construcciones de atadura e indicador se emplean para que se unan a las construcciones de sustrato con un nivel de exactitud deseado, ya que el acoplamiento inexacto da lugar a ineficiencias en la detección y también puede dar lugar a la finalización del polímero o la codificación aleatoria del código indicador (por ejemplo, si se usa un código de indicador 35 asimétrico). La fidelidad del procedimiento de SBX se puede unir a la pureza de síntesis de las construcciones de sustrato. Después de la purificación de la construcción de atadura/indicador para enriquecer el producto de longitud completa, la construcción se puede acoplar 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 procedimientos de síntesis de polímeros, la purificación (tamaño, afinidad, HPLC, electroforesis, etc.) se usa después de 40 que se complete la síntesis y el ensamblaje de la construcción de sustrato para garantizar una alta pureza de las construcciones de sonda expandible de longitud completa.

Síntesis de construcciones de sustrato de clase I que presentan indicadores

- 45 La síntesis de construcciones de sustrato de clase I con indicadores o precursores de indicadores que se presentan en las ataduras se puede lograr de diversas maneras. Se puede usar un procedimiento gradual para ensamblar un polímero de atadura en horquilla que se conecta cerca de los extremos de unión de la atadura de la construcción del sustrato por medio de un puente disulfuro. Esto orienta los extremos reactivos de modo que se favorece altamente el acoplamiento de la atadura a la sonda. Las fosforamiditas C6 modificadoras de amino están disponibles comercialmente para los cuatro 50 nucleótidos (Glen Research, EE. UU.) y se usan para unir la atadura para formar la construcción de sustrato completa, por ejemplo. De forma alternativa, se puede emplear una sustancia química conectora en forma de nucleótidos modificados con benzaldehído. La purificación por tamaño y/o afinidad es útil para enriquecer las construcciones de sustrato ensambladas correctamente.
- 55 La heterobifuncionalización de la sonda se puede realizar de forma ventajosa en una matriz de soporte sólida, como es habitual para la síntesis de oligonucleótidos y péptidos, o en solución con procedimientos de purificación apropiados. Se encuentra disponible una amplia gama de reactivos de reticulación heterobifuncionales y homobifuncionales listos para usar para modificar, por ejemplo, restos de amina, carboxilo, tiol e hidroxilo, y para producir diversas químicas de unión sólidas y selectivas. Dado que los modificadores de amino C6 están disponibles para los cuatro desoxirribonucleótidos y 60 pueden estar disponibles para los cuatro ribonucleótidos, las estrategias de funcionalización descritas aquí usan procedimientos de reticulación basados en aminas listos para usar junto con químicas de protección/desprotección de aminas bien establecidas. Sin embargo, dada la amplia gama de fosforamiditas y químicas de reticulación conocidas en la técnica, también se pueden considerar procedimientos no descritos aquí y dar como resultado productos equivalentes.
- 65 La necesidad de heterobifuncionalización de la sonda se puede eliminar si la estrategia de codificación del indicador produce una codificación simétrica digital o usa puntos de referencia direccionales (bits de paridad) para identificar la

orientación del código. En este caso, dos modificadores de sonda de amina internos son suficientes, ya que cualquiera de las orientaciones de acoplamiento de las construcciones de indicadores en la atadura produciría una identificación de secuencia específica de sonda única.

- 5 Uno o muchos segmentos de polímero se pueden ensamblar secuencialmente usando enlaces covalentes catalizados por químicas (por ejemplo, agentes de reticulación) o enzimas (por ejemplo, hibridación de sonda de ácido nucleico y ligación) para formar una atadura funcionalizada en el extremo circular. Dado el estado actual de la técnica para los procedimientos de síntesis de polímeros, el enfoque de síntesis de reticulación química constituye un representante. Como es habitual en muchos procedimientos de síntesis de polímeros, se puede usar una matriz de soporte sólido como estructura para la síntesis. Los segmentos de polímero se pueden ensamblar uno cada vez en base a la funcionalización de extremo, como segmentos de pares mixtos que tienen diferentes funcionalizaciones de extremo, o como pares unidos: dos segmentos poliméricos con diferentes restos de extremo homobifuncionales (por ejemplo, hidrazida y amina) emparejados por medio de puentes disulfuro.
- 10 15 Las químicas de marcado, que incluyen tanto el conector como restos de elemento indicador, se desarrollan y optimizan en base a un alto rendimiento y estabilidad de la señal, baja reactividad cruzada y entrelazamiento de polímeros, y la rigidez estructural (endurecimiento) que estas químicas imparten a la cadena principal del Xpandómero, lo que puede ser importante para la preparación y detección de muestras, como se analiza a continuación.
- 20 25 30 35 En la figura 31 analizada anteriormente, la construcción completa de atadura de horquilla/indicador se ensambla independientemente de la sonda de oligonucleótidos y, a continuación, se une mediante químicas de conector homobifuncional o heterobifuncional al miembro de sonda. En una alternativa, como se muestra en la figura 34, las ataduras se circularizan de forma gradual mediante la construcción sobre secuencias de sonda inmovilizadas. La construcción de indicador y la atadura se sintetizan con segmentos de atadura heterobifuncionales (direccionales) u homobifuncionales (simétricos u orientados) unidos directamente a una sonda de oligonucleótidos. La secuencia de sonda que se muestra en la figura 34A es un tetrámero e incluye restos de sonda P¹ y P² (los círculos segundo y tercero) separados por un enlace escindible selectivamente ("V"). Las técnicas de síntesis en estado sólido se usan en la síntesis de la construcción de indicadores. Esta síntesis se integra con el cierre del lazo de la atadura. En la etapa I de la figura 34A, se añade un primer segmento de atadura (341) con el primer grupo indicador (342) usando la química de grupos funcionales específicos indicada por L1 y L1' (el conector L1 en uno de los restos de sonda está bloqueado, como se muestra en el pequeño rectángulo). En la etapa II, se añade un segundo segmento de atadura (344) con un segundo grupo indicador (345) usando la química de grupos funcionales específicos indicada por L2' y L2'. En la etapa III, se añade un tercer segmento de atadura (346) con un grupo indicador (347) usando la química de grupos funcionales específicos indicada por L2 y L1'. En la etapa IV, y después de la retirada del grupo de bloqueo del sitio L1 en el resto P² de la sonda (de nuevo representado por el rectángulo pequeño), el bucle se cierra después del acoplamiento de L1' y L1.

- Aún en otra alternativa, como se ilustra en la figura 34B, la química del conector heterobifuncional se puede usar de nuevo para garantizar que la atadura esté situada direccionalmente en la sonda (aunque esto no es necesario para todas las estrategias de codificación). Se representa de nuevo un tetrámero de sonda con restos de sonda P¹ y P² (los círculos segundo y tercero) separados por un enlace escindible selectivamente ("V"). En la etapa I, dos ataduras (341, 344) con segmentos de indicador (342, 345) se ponen en contacto con los grupos funcionales L1 y L2 en P¹ y P²; las químicas son específicas para cada atadura. Las ataduras en esta etapa se pueden estabilizar con enlaces intraatadura (representados por los triángulos contiguos). En la etapa II, se usa una tercera atadura (346) con un segmento de indicador (347) para "proteger" las ataduras después de la retirada de grupos de bloqueo (que se muestran como pequeños rectángulos) en los segmentos de predecesores. El segmento de protección también se puede estabilizar con un enlace intrasegmento (de nuevo representado como triángulos contiguos).

- 40 45 Volviendo ahora a la figura 34C, se muestra una alternativa basada en la adición de restos de sonda P¹ y P² para separar los enlaces de extremo de una atadura preformada. En la etapa I, la atadura preformada se hace reaccionar primero con P¹ poniendo en contacto L1 con L1'. En la etapa II, la atadura se hace reaccionar después con P² poniendo en contacto L2 con L2'. La atadura se puede estabilizar mediante un enlace intraatadura (representado por triángulos contiguos), que aproxima P¹ con P². A continuación, los dos restos de sonda se ligan en la etapa III para formar el enlace selectivamente escindible ("V") entre P¹ y P² (los círculos segundo y tercero). La ligación de P¹ y P² se puede facilitar opcionalmente duplicando dichos restos de sonda a un molde complementario.
- 55 60 65 En la figura 35A, se divulga otro aspecto para la síntesis de un segmento de construcción de indicadores. Usando procedimientos químicos de estado sólido, un conector escindible (351) se ancla primero en un sustrato sólido (350). Un primer segmento de atadura con un enlace reversible (352) se hace reaccionar con el conector en la etapa I, y a continuación en la etapa II se hace reaccionar con una colección combinatoria de monómeros M1 y M2, que se muestra en este ejemplo como una mezcla estequiométrica 4:1 de los respectivos monómeros. La síntesis aleatoria de copolímeros se realiza de esta manera para producir composiciones únicas de segmentos de atadura. Si se usan monómeros de péptidos o aminoácido, esto se puede hacer con química de anhídridos mixtos, por ejemplo, lo que da como resultado ataduras peptídicas de copolímero aleatorias 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, se añade a continuación al segmento peptídico un elemento conector de extremo (356) bloqueado en L1 (representado como un

pequeño rectángulo). Después de la escisión del segmento del soporte sólido (no se muestra), el segmento de atadura se puede incorporar en una construcción de indicadores usando diversos procedimientos, algunos de los cuales se describen en la figura 34.

- 5 En la figura 35B, se demuestra una alternativa. Una construcción de indicadores con fragmentos peptídicos incorporados aleatoriamente se sintetiza como antes en las etapas I, II y III, pero en la etapa III el grupo de conector de final (358) se proporciona con un conector heterofuncional L2. A diferencia del ejemplo anterior, los monómeros se proporcionan en proporciones iguales. Después de la escisión del soporte sólido (no se muestra), la construcción de indicadores está disponible para su posterior incorporación en las ataduras de los ejemplos anteriores y es un conector heterobifuncional.
- 10 Las construcciones de indicadores sirven para codificar la información genética del fragmento de la secuencia de la sonda, como se describe con más detalle a continuación.

Los polipéptidos son construcciones de indicadores útiles y también sirven como ataduras. Se pueden usar copolímeros aleatorios, periódicos, alternos y de bloque junto con el homopolímero, por ejemplo, para el segmento de atadura y la composición de la construcción de atadura. Los segmentos de polipéptido se pueden funcionalizar en los extremos usando agentes de reticulación heterobifuncionales que contienen succinimidilo para la conversión de amina en una hidrazida o 4-formilbenzoato (4FB), por ejemplo. Los polipéptidos se pueden producir por síntesis química o por clonación y sobreexpresión en sistemas biológicos (bacterias, levaduras, baculovirus/insectos, mamíferos). Dependiendo de la longitud deseada del segmento, los segmentos de atadura pueden variar de $N > 2$ (segmentos cortos) a $N > 1000$ (segmentos largos). Se pueden usar químicas de protección de grupo lateral amina cuando sea apropiado. Como alternativa al uso de agentes de reticulación estándar, se pueden sintetizar químicamente polipéptidos con restos de hidrazida, 4FB y NHS unidos directamente.

- 15
 - 20
 - 25
 - 30
 - 35
- Los segmentos de polipéptido funcionalizados en los extremos usando agentes de reticulación heterobifuncionales que contienen maleimido para la conversión de tiol en una hidrazida o 4-formilbenzoato (4FB), también se pueden usar en la síntesis de ataduras y construcciones de indicadores. De forma similar, se pueden emplear segmentos de polipéptido funcionalizados en los extremos usando agentes de reticulación de EDC para la conversión de carboxilo en una hidrazida (HZ) o 4-formilbenzoato.
- La síntesis de las cinco clases de construcciones de sustratos oligoméricos que tienen ataduras modificadas con construcciones de indicadores se logra mediante el uso de los procedimientos sintéticos anteriores. De forma similar, la síntesis de sustratos monoméricos que tienen ataduras modificadas con construcciones de indicadores también se puede lograr mediante los mecanismos sintéticos anteriores. Las químicas para estas variantes son en general aplicables a los géneros de especies de Xpandómero que se muestran en las figuras 8 y 9. A continuación se pasa a codificar estrategias y reglas para transmitir información genética en Xpandómeros con construcciones de indicadores.

Construcciones de indicadores y estrategias de código de indicador

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

La codificación digital requiere que los códigos de indicador sigan ciertas reglas. Por ejemplo, se requieren al menos 256 códigos de indicador para identificar todas las combinaciones posibles de una colección de Xsonda tetraméricas. Tener más códigos de indicador que combinaciones posibles de construcción de indicador es ventajoso porque los estados adicionales se pueden usar para otros fines, tal como marcado de espacios, proporcionar información de posición o identificar errores de paridad o errores de alto orden.

Se pueden considerar una serie de estrategias para representar físicamente un código de indicador. La atadura se puede dividir en uno o varios segmentos codificables, cada uno de los cuales se puede marcar antes o después del ensamblaje del Xpandómero. Se pueden usar niveles de señal variables (cantidad de marcador), longitudes (duración de la señal) y formas de segmentos de construcciones de atadura marcados para incrementar las opciones de codificación. La codificación también se puede expandir usando marcadores multihibridables. Por ejemplo, usando un enfoque de marcadores con marca masiva, se puede usar una amplia colección de marcas espectralmente distintas para codificar de forma única un solo segmento de indicador; 14 marcas de masa distintas usadas en combinaciones de tres marcas agrupadas en un solo segmento de construcciones de atadura crearían 364 espectros de 3 masas únicos. Para ataduras multisegmentadas, el marcado posterior al ensamblaje del Xpandómero de la mayor parte o la totalidad de la cadena principal de la atadura puede tener el beneficio adicional de incrementar la rigidez del Xpandómero, facilitando potencialmente su manipulación para la detección y mejorando la estabilidad.

60 La figura 36A (y también la figura 2A) ilustra una atadura con un solo segmento de indicador de una Xsonda o un Xmero. Este enfoque se beneficia del marcado de indicadores altamente multihibridables, tales como marcas del espectrómetro de masas, para producir una amplia colección de resultados espectralmente distintos. Una marca de masa escindible es una molécula o un complejo molecular de indicadores escindibles que se pueden ionizar fácilmente 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 unos pocos cientos de indicadores de dichas marcas de masa. Este

ejemplo de código de indicador no necesita información de posición del indicador en la cadena principal de la atadura para determinar el estado del código (aunque se requiere información de posición para distinguir un código de indicador del siguiente). Esta característica simplifica la construcción de atadura y acorta potencialmente los requisitos de longitud de la atadura.

5

En un modo de realización, usando marcadores con marcas de masa escindibles, el Xpandómero alargado se puede presentar para su detección por medio de una fuente de iones de nanoporos (ionización por electropulverización, ionización química a presión atmosférica, fotoionización) o por deposición superficial (nanocomb, nanocanal, flujo laminar, electroforético y similar) seguida de ionización por desorción con láser, haz de iones o fuentes de haz de electrones (ionización por desorción láser asistida por matriz "MALDI", ionización por electropulverización por desorción "DESI", ionización por desorción en sílice "DIOS", espectrometría de masas de iones secundarios "SIMS").

10

En la figura 36B se describe un ejemplo de construcciones de atadura de 9 bits con 2 estados de detección ("1" y "0"), que produce 512 identidades de código. En un modo de realización, las construcciones de atadura consisten en segmentos "1" y "0" que producen dos niveles de impedancia eléctrica medida en un detector de nanoporos de tipo Coulter. En un segundo modo de realización, las construcciones de atadura consisten en segmentos conductores eléctricos "1" y segmentos no conductores "0". En un tercer modo de realización, las construcciones de atadura consisten en segmentos fluorescentes "1" y segmentos no fluorescentes "0". Para este tipo de codificación se puede considerar una pluralidad de elementos indicadores diferentes. Para cualquiera de estos enfoques, la química del ensamblaje de atadura y la unión de la sonda puede ser idéntica; solo sería necesario cambiar la composición del segmento de indicador. Con este formato simple, el marcado se puede hacer antes o después del ensamblaje del Xpandómero. El marcado posterior es deseable ya que el segmento de indicador no marcado es significativamente menos masivo y, de este modo, se puede usar a una concentración mucho mayor que una construcción de indicadores totalmente marcada. Dependiendo de la estrategia, los segmentos de polímero pueden ser: (1) codificados por medio de químicas superficiales conjugables o reactivas (por ejemplo, polilisina, ácido poliglutámico), (2) no reactivos (por ejemplo, PEG, polímeros de baja reactividad) o (3) una mezcla de polímeros reactivos y no reactivos. Los grupos reactivos incluyen, entre otros, aminas primarias (-NH₂), carboxilos (-COOH), tioles (-SH), hidroxilos (-OH), aldehídos (-RCOH) y restos de hidrazida (-RNN). El marcado de segmentos reactivos, que puede incluir desprotección de grupos reactivos, se puede hacer directamente en las construcciones de sustrato, después de la formación del intermedio del Xpandómero, después de la escisión de la cadena principal para producir el Xpandómero, o en cualquier otro momento en el procedimiento de SBX, como sea apropiado, para producir los mejores resultados.

20

25

30

35

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

40

Volviendo a la figura 36D, se muestra que una construcción de indicadores de 7 bits, cada indicador con tres estados de detección, da como resultado 2187 identidades de código detectables. El uso de espaciadores poliméricos flexibles se puede usar por razones estéticas.

En la figura 36E, se muestra un equivalente rígido de la construcción de indicadores de 7 bits previa.

45

La figura 36F describe un ejemplo de construcciones rígidas de atadura de 4 segmentos con un total de ocho estados distintos o combinaciones de marcadores por segmento. El uso de siete de las combinaciones de marcadores para marcar tres de los segmentos de construcciones de atadura producirá 343 identidades de código únicas y dejará 1 segmento disponible para la identificación de límites de subunidades, paridad u otro fin funcional. Esto puede usar un enfoque de marcado mixto en el que se pueden incorporar de 1 a 3 marcadores diferentes en cada segmento para producir al menos ocho combinaciones únicas por segmento, como se muestra en la figura 37, donde se contemplan combinaciones de diferentes tipos de indicadores y química de construcción de indicadores. Las opciones de marca masiva y marcado fluorescente, entre otras, se pueden usar como se ha descrito. El marcado de construcciones de atadura posteriores al ensamblaje del Xpandómero está dirigido por la abundancia e identidad de tres restos de agente de reticulación, como se describe en la figura 37. Las longitudes de los segmentos de atadura pueden ser del orden de 100-1000 nm para mediciones limitadas por la difracción o < 100 nm para mediciones de campo cercano, si se desea usar estas tecnologías de detección. Se pueden usar ataduras más cortas para otros procedimientos de detección.

50

55

60

En la figura 36G se describe una construcción rígida de atadura de 3 segmentos con 22 combinaciones de marcadores distintas en total por segmento. El uso de 21 de las combinaciones de marcadores descritas en la figura 37 para marcar dos de los segmentos de construcciones de atadura producirá 441 identidades de construcciones de atadura únicas y dejará un segmento disponible para la identificación de delimitadores de subunidades de Xsonda. Esto usa un enfoque de marcado mixto donde se pueden incorporar de uno a tres marcadores diferentes en cada segmento para producir hasta 22 combinaciones únicas por segmento (figura 37). Se pueden usar opciones de marca masiva y marcado fluorescente, entre otras. Como se describe en la figura 37, el marcado de la construcción de atadura posterior al ensamblaje del Xpandómero está dirigido por la abundancia e identidad de tres restos químicos.

Diseñando segmentos de indicador/atadura en los que la abundancia de grupos reactivos y la dimensión espacial (distancia radial desde la cadena principal del polímero) se pueden variar, se pueden lograr niveles de codificación en los que son posibles al menos tres estados de código totales: Alto "2", Medio "1" y Bajo "0" (figura 36H). Por otra parte, una estrategia de codificación de tres marcadores y dos niveles (es decir, 21 estados por de indicador) requeriría solo dos segmentos de codificación de indicador para producir 441 códigos y podría usar un segmento adicional para la orientación del código (figura 36I).

La codificación del indicador se puede diseñar para reducir errores inherentes a la construcción de indicador y las tecnologías de detección asociadas usando varios enfoques. En el caso de la detección tipo Coulter de nanoporos, la velocidad a la que pasa un Xpandómero a través del poro y la modulación de corriente que provoca pueden depender de muchos factores, que incluyen, pero sin limitaciones, el estado de carga de la porción del Xpandómero dentro del poro, las concentraciones de electrolitos, los estados de carga de la superficie del nanoporo, el potencial aplicado, los efectos de fricción que limitan el movimiento del Xpandómero y las dimensiones relativas tanto del Xpandómero como del nanoporo. Si la velocidad no es predecible, la descodificación de la modulación de corriente no puede usar el tiempo (y la velocidad constante) para resolver las asignaciones de medición del indicador. Un modo de realización de codificación resuelve este problema usando codificación de 3 estados. La señal del indicador es la impedancia que un marcador provoca a la conductividad del electrolito a través del nanoporo. Proporcionando tres posibles niveles de impedancia para un indicador, la transición al siguiente marcador codifica un bit de información. Por diseño, esta transición es siempre un cambio a uno de los otros dos estados. Si los tres estados están marcados como A, B y C, entonces una secuencia de indicadores nunca tiene 2 Aes, 2 Bes o 2 Ces juntas. De esta manera, la información se codifica en transiciones de nivel y, por lo tanto, es independiente de la velocidad a través del nanoporo. Un esquema de codificación 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 les asigna el valor "1". Por ejemplo, la secuencia detectada ABACBCA se descodifica en 0,1,1,1,0,0.

Aunque la programación no puede resolver de forma fiable marcadores secuenciales, puede ser suficiente para diferenciar la separación de la secuencia de marcadores en construcciones de una sola atadura de la de la siguiente atadura secuencial. Las longitudes adicionales de la atadura del espaciador en cada extremo de la secuencia de marcador indicadora (en los puntos de unión de la construcción de sustrato) pueden proporcionar una gran hueco de programación que delinea los códigos de las construcciones de la atadura a tiempo.

En los casos en que dicha programación sea insuficiente, se puede producir un error de desplazamiento del marco de lectura. Los errores de desplazamiento del marco de lectura se producen cuando el detector lee una serie de marcadores de múltiples códigos (marcos de lectura) de construcciones de atadura, pero no delinea correctamente el inicio del código (inicio del marco de lectura). Esto da como resultado códigos incorrectos. Un modo de realización para resolver esto es añadir más bits en el código de los necesarios para identificar las secuencias de bases correspondientes (que típicamente tienen de 1 a 4 bases de longitud). Por ejemplo, se requieren ocho bases para identificar de forma única una secuencia de 4 bases. Cada par de 2 bits describe una sola base. Añadiendo un bit de paridad para cada base, el código de construcción de atadura se incrementa a 12 bits. Una alta tasa de error de paridad (cerca de un 50 %) indicaría un error de desplazamiento del marco de lectura que resultaría de un cambio de estado perdido. Además de un cambio de estado perdido, otro tipo de error que se puede producir en esta tecnología de detección de nanoporos es el de un estado leído erróneamente. Los errores de un solo indicador se pueden aislar a la base particular usando el bit de paridad y, a continuación, se le puede asignar el valor "base desconocida".

Los Xpandómeros marcados usando segmentos que impiden la electricidad, que conducen la electricidad o fluorescentes se pueden medir en solución usando una variedad de formatos de nanoporos, nanopeines o nanocanales. De forma alternativa, los Xpandómeros se pueden depositar en la superficie como polímeros lineales espacialmente distintos usando procedimientos de depósito electroforético, de flujo laminar o de nanopeine, entre otros. Al igual que con el enfoque de solución, la detección directa del Xpandómero alargado en la superficie se puede hacer midiendo las características de la señal de los segmentos de atadura marcados. Dependiendo de la composición del material de depósito (conductor, aislante) y el sustrato subyacente (conductor, aislante, fluorescente), se puede considerar una variedad de tecnologías de detección para la detección de marcadores.

Otros procedimientos de SBX

En las figuras anteriores (figuras 11-17, 20, 21, 24) se ilustraron una serie procedimientos de SBX que usan construcciones de sustratos oligoméricos de clase IV. A continuación, los inventores consideran procedimientos opcionales que complementan estos protocolos.

Funcionalización de extremo

La figura 38 ilustra la preparación y uso de adaptadores de extremo diana. La figura 38A ilustra oligonucleótidos complementarios funcionalizados en los extremos hibridados para formar un adaptador con un extremo conjugable bifuncional ("L1" y "L2") y un extremo ligable enzimáticamente (5'-fosfato y 3'-OH). Estos adaptadores también se pueden diseñar con funcionalidades adicionales. Por ejemplo, como se muestra en la figura 38B, se pueden sintetizar adaptadores funcionalizados en los extremos con agentes de reticulación anidados ("L3") y un enlace escindible ("V"). Para simplificar, el enlace escindible V puede ser la misma química de enlace escindible (o enzimología) usada para liberar o expandir el

Xpandómero, aunque se pueden usar 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, se puede usar una microesfera magnética con un oligonucleótido atado a la superficie complementario a cada hebra de adaptador (dos mezclas de microesferas diferentes) para ensamblar segmentos de oligonucleótidos modificados diferencialmente. Una vez ensamblados, los segmentos se pueden ligar enzimáticamente para unir covalentemente los segmentos hibridados. Con este enfoque, cada segmento se puede modificar individualmente de una manera que no está disponible para la síntesis de oligonucleótidos estándar.

La manipulación del Xpandómero puede ser útil para una presentación y detección eficientes de la muestra. Por ejemplo, se pueden usar marcadores de afinidad terminal para modificar selectivamente un extremo del Xpandómero para permitir el alargamiento electroforético. La unión de un modificador voluminoso, de carga neutra, al extremo 3' o bien 5' (no a ambos) produce un arrastre electroforético en el Xpandómero que provoca que el extremo no modificado se alargue a medida que viaja hacia el detector. Se pueden usar modificadores de extremo, que incluyen, pero sin limitaciones, microesferas, nanopartículas, nanocristales, polímeros, polímeros ramificados, proteínas y dendrímeros, para influir en la estructura (alargamiento), posición y velocidad a la que el Xpandómero se presenta al detector impartiendo propiedades únicas y diferenciadoras a sus extremos, tal como carga (+/-neutra), flotabilidad (+/-neutra), hidrofobia y paramagnetismo, por nombrar algunas. En los ejemplos proporcionados, las modificaciones de extremo producen una fuerza de arrastre que permite que el Xpandómero se alargue; sin embargo, también se puede emplear la estrategia opuesta en la que se usa la modificación del extremo para tirar del Xpandómero hacia y a través del detector. Con este enfoque, tirar del modificador de extremo facilita el alargamiento del Xpandómero. El adaptador de extremo en la hebra molde también puede contener opcionalmente uno o más ácidos nucleicos que se usarán para sintetizar señales de validación y registro de armazón en el Xpandómero terminado (véase la figura 54).

La incorporación de un modificador de afinidad se puede hacer antes, durante o después de la síntesis del Xpandómero (hibridación, ligación, lavado, escisión). Por ejemplo, los cebadores marcados por afinidad terminal que son complementarios a la secuencia del adaptador se pueden precargar en el ADN monocatenario diana en condiciones altamente específicas antes de la síntesis del Xpandómero. El cebador y su marca de afinidad se pueden incorporar en el Xpandómero de longitud completa y se pueden usar para modificar selectivamente su extremo. Un enfoque potencialmente más elegante es incorporar modificadores de extremo enzimáticamente. La transferasa terminal (TdT), por ejemplo, es una polimerasa independiente del molde que cataliza la adición de desoxinucleótidos al extremo 3' hidroxilo de moléculas de ADN mono o bicatenario. Se ha demostrado que la 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). Una amplia gama de enzimas es adecuada para este propósito, incluyendo (pero sin limitaciones) ARN ligasas, ADN ligasas y ADN polimerasas.

En las figuras 38E y 38F, se ilustran adaptadores de horquilla. Los adaptadores de horquilla encuentran uso en la ligación de extremos romos para proporcionar hebras molde autocebantes. Una ventaja de este enfoque es que la hebra hija permanece acoplada covalentemente a la hebra molde y se reasocia más rápidamente después de una fusión para retirar material no ligado y fragmentos de bajo peso molecular. Como se muestra en la figura 38F, estos adaptadores pueden contener funcionalidades de conector preformadas incrustadas para la purificación o manipulación posterior, y también pueden contener sitios de escisión para una cosecha más eficiente de las hebras hijas del Xpandómero.

Preparación y análisis de moldes diana

Para realizar la secuenciación del genoma completo de ADN largo y continuo usando los procedimientos de SBX basados en Xsondas, se presupone que el ADN se prepara en una forma manejable para hibridación, ligación, llenado de huecos si se requiere, expansión y medición. El procedimiento de ensamblaje del Xpandómero se puede mejorar por inmovilización superficial del ADN diana para (1) reducir la complejidad y los efectos de hibridación cruzada, (2) mejorar el lavado, (3) permitir la manipulación de la diana (alargamiento) para facilitar una hibridación mejorada y/o (4) si se usa un sensor de nanoporos, proporcionar una interfase perfecta con el procedimiento de detección. Como se describe con detalle a continuación, se prevé que los procedimientos para la preparación de moldes diana, el análisis y la unión a superficie mejoren la calidad de los datos y el ensamblaje de secuencias.

La mayoría de los procedimientos de secuenciación del genoma completo requieren la fragmentación del genoma diana en piezas más manejables. El cromosoma más grande del genoma humano (cromosoma 1) tiene ~227 Mb y el cromosoma más pequeño (cromosoma 22) tiene ~36 Mb. Para la mayoría de los modos de realización altamente procesivos y continuos descritas en el presente documento, 36 Mb de secuenciación continua es demasiado largo para que se secuenciado con alta eficacia. Sin embargo, para aprovechar la capacidad de longitud de lectura larga inherente de SBX, se seleccionan longitudes de fragmentos de ADN de >1kb. En consecuencia, se pueden emplear una serie de estrategias para lograr la fragmentación del genoma y preparar un conjunto de ADN diana compatibles con el procedimiento de SBX como se divulga en el presente documento.

Un modo de realización implica fragmentar el genoma total en piezas de 1-10 Kb (promedio de 5 Kb). Esto se puede hacer mediante enzimas de restricción o bien mediante cizallamiento hidrodinámico/mecánico. A continuación, se vuelven romos los extremos de los fragmentos y se reparan en preparación para la ligación de extremos romos a un adaptador de secuencia ("SA") o una construcción de adaptador de secuencia-dendrímero ("SAD"). El extremo no romo de la

construcción SA o SAD está diseñado para que no se pueda ligar para evitar el ensamblaje de multímeros de dichos adaptadores. Cualquier diana ligada a SA o SAD que esté presente se puede purificar por afinidad lejos de adaptadores libres. En este punto, si el dendrímero de captura no se introduce con la construcción de SA, entonces se puede hacer esto (usando un exceso eficiente de dendrímero) seguido de purificación para aislar el complejo diana-SAD. Una vez purificado, el complejo está listo para su unión a la superficie de ensayo. Con este fin, se desea que solo se una un único complejo diana por localización de reacción. De forma alternativa, el dendrímero de captura se puede asociar con la superficie de ensayo, en cuyo caso el complejo diana-SA purificado se puede unir directamente al dendrímero que ya está localizado y unido covalentemente a la superficie. Un procedimiento similar para el ensamblaje ADN diana en superficies se ha divulgado por Hong et al. ("DNA microarrays on nanoscale-controlled surface", Nucleic Acids Research, 33(12): e106, 2005).

Otro enfoque, llamado procedimiento de "análisis", implica la fragmentación gruesa del genoma en piezas de 0,5-5 Mb (usando enzimas de restricción de corte raras o por cizallamiento hidrodinámico/mecánico), seguido de la captura de dichos fragmentos en una micromatriz modificada (o superficie dividida). compuesta por sondas de captura de oligonucleótidos específicas de genes/locus. Las micromatrices personalizadas y los conjuntos de sondas de captura de oligonucleótidos están ampliamente disponibles en una serie fuentes comerciales (ArrayIt, Euorfins-Operon, Affymetrix Inc.). Este análisis adicional puede proporcionar una ventaja para el ensamblaje de secuencias de *backend*. La captura de estos fragmentos grandes implica la desnaturalización parcial o completa de la diana para permitir que las sondas de captura se unan o hibriden con ADN diana específico. Para reducir la hibridación inespecífica, puede ser necesario cargar la matriz de captura en condiciones diluidas para evitar la hibridación cruzada entre moldes.

Cada conjunto de sondas de captura divididas, que se pueden disponer en una gran superficie, está diseñado para proporcionar una resolución del genoma lineal de ~3 Mb. Para proporcionar un análisis eficiente del genoma, cada conjunto de sonda de captura individual puede estar compuesto de 3 a 5 oligonucleótidos específicos de gen/locus, separados linealmente en el genoma por 0,5-1,5 Mb cada uno. Las sondas de captura se seleccionan para que sean completamente exclusivas del fragmento diana, proporcionando, por tanto, especificidad y redundancia al procedimiento. Dada una resolución de 3 Mb de un genoma de 3 Gb, este enfoque requiere una matriz de captura compuesta de aproximadamente 1000 sondas específicas de gen/locus. Los conjuntos de sondas de captura de oligonucleótidos listos para micromatrices específicos para dianas de genes humanos están ampliamente disponibles en el mercado (Operon Biotechnologies, Huntsville AL, EE. UU.).

Una vez que se han eliminado los acontecimientos de unión inespecífica, cada matriz de captura puede entonces proseguir como reacciones independientes durante el resto del procedimiento de preparación del genoma de la misma manera que se analiza anteriormente para el procedimiento sin análisis. La principal diferencia es que la muestra diana ahora se puede analizar en su posición en una superficie de ensayo de SBX o como reacciones basadas en soluciones individuales, lo que reduce la complejidad del ensamblaje de la secuencia posterior a la adquisición de datos.

Otro modo de realización es tener hibridación Xsonda-diana no atada y producción de Xpandómero. En este caso, el ensayo de SBX se realiza en solución libre. Este enfoque puede usar una o una combinación de manipulaciones físicas, tal como usar marcas electroforéticas, magnéticas, de arrastre, o funcionalidades de extremo de flotabilidad positiva/negativa en condiciones de flujo estático o laminar, como un medio para alargar el ADN diana antes y durante la hibridación de la sonda y ligación, por ejemplo, y, para alargar o expandir el Xpandómero escindido antes de la detección. La síntesis en solución libre de Xpandómeros, sin inmovilización, se puede hacer usando polimerasas y ligasas (con y sin cebadores) y también se puede hacer usando procedimientos de ligación química. Se pueden usar tanto construcciones de sustrato de trifosfato como construcciones de sustrato de monofosfato. Se concibe la síntesis simultánea de Xpandómeros a partir de dianas de ácidos nucleicos múltiples y mixtas. En general, las construcciones de trifosfato de sustrato son capaces de polimerización continua y procesiva en solución y se pueden adaptar a protocolos de un solo tubo para secuenciación masivamente paralela de una sola molécula en solución libre, por ejemplo.

50 Ensamblaje superficial de dianas de ácido nucleico

Un procedimiento de preparación de ácidos nucleicos diana para la secuenciación usa una diana de ADN bicatenario funcionalizado en el extremo como se muestra en la figura 39. Para este ejemplo, cada adaptador típicamente se proporciona con un grupo ANH (hidrazida reactiva) útil para procesamiento adicional, por ejemplo, usando la química de enlace de acetona hidrazona de succinimidil 4-hidrazinonicotinato (SANH) y un híbrido de oligonucleótido-adaptador de extremo modificado con amina. De forma similar, se puede usar un SANH reactivo con amina para crear restos de hidrazida reactivos. SANH se conjuga fácilmente con aldehídos como 4-formilbenzoato (4-FB) para formar enlaces de hidrazona covalentes estables. Se puede usar C6-succinimidil 4-formilbenzoato (C6 SFB) reactivo con amina para crear un resto de benzaldehído reactivo. Los oligonucleótidos complementarios funcionalizados en los extremos se hibridan para formar un extremo conjugable bifuncional (SANH y amina) y un extremo ligable enzimáticamente (3' OH y 5' fosfato). La ligación del adaptador crea una diana de ADN bicatenario funcionalizado en el extremo que se puede reticular con grupos aldehído atados a la superficie. La figura 39 ilustra el extremo diana de ADN bicatenario funcionalizado usando SANH e híbrido de oligonucleótido-adaptador de extremo modificado con amina.

65 Para muchos de los procedimientos de SBX descritos, las dianas de ácido nucleico se pueden unir covalentemente a un soporte sólido recubierto plano (acero inoxidable, silicio, vidrio, oro, polímero). Como se describe en la figura 40, los puntos

de unión de la diana se pueden producir por derivatización en 4FB de una monocapa de SAM. El procedimiento usa el adaptador ANH de la figura 39, y en la etapa I de la figura 40, el ANH se hace reaccionar con las cabezas de 4FB de la monocapa y el molde se desnaturaliza (etapa II). La otra hebra del molde, que también está marcada en el extremo, se puede capturar en una reacción separada. En la etapa III, la amina libre en el extremo 3' del molde monocatenario se hace reaccionar con una microesfera, por ejemplo, que aquí se muestra como una microesfera de flotabilidad, de modo que el molde diana se pueda estirar. Estos complejos de captura se pueden ensamblar mediante autoensamblaje aleatorio de una mezcla estequiométricamente equilibrada de polímeros funcionalizados en los extremos (por ejemplo, tiol-PEG-hidrazida para la unión de la diana; tiol-PEG-metoxi para protección de "SAM" de la monocapa autoensamblada) o mediante formación de patrones de puntos de unión reactivos resueltos espacialmente usando técnicas litográficas. El procedimiento litográfico con patrón puede producir puntos de unión de dianas espaciados de forma constante, aunque esto es difícil de hacer para la unión de una sola molécula, mientras que el procedimiento de autoensamblaje aleatorio probablemente produciría un espaciado de unión de dianas más variable, pero tiene un alto porcentaje de uniones de una sola molécula. Una amplia gama de agentes de reticulación monofuncionales, bifuncionales y heterobifuncionales están disponibles comercialmente de una variedad de fuentes. Los polímeros monofuncionales, bifuncionales y heterobifuncionales compatibles con agentes de reticulación (polietilenglicol, poli-L-lisina) también están disponibles en una amplia gama de fuentes comerciales.

La densidad de dianas de ADN de mil millones de dianas en una superficie de 100 cm² requeriría un promedio por área de diana de 10 um². El espaciado de diana en este intervalo proporciona suficiente separación de las dianas para evitar una reacción cruzada significativa de ácidos nucleicos diana de 5000 bases de longitud atados a microesferas (100-1000 nm de diámetro de microesfera; 5 Kb de ADN bicatenario = 1700 um). El área de las dianas se puede expandir fácilmente si se determina que la reactividad cruzada de la diana y/o de la microesfera (>1 microesfera/diana) es inaceptablemente alta.

25 Alargamiento de la diana usando microesferas o nanopartículas

Como se muestra en la figura 40, una microesfera o nanopartícula (400) atada en el extremo libre del ADN diana se puede usar para alargar y mantener la diana en su conformación monocatenaria durante, por ejemplo, la hibridación de la colección de Xsondas. Retener la diana en una conformación alargada reduce significativamente la frecuencia y la estabilidad de las estructuras secundarias intramoleculares de las dianas que se forman a temperaturas menores. La reducción o incluso la eliminación de las influencias de la estructura secundaria promueve un ensamblaje de construcción de sustrato eficiente y de alta fidelidad.

35 Se puede emplear una variedad de enfoques para aplicar una fuerza de alargamiento a la diana monocatenaria. Por ejemplo, microesferas/nanopartículas/polímeros paramagnéticos permiten el uso de un campo magnético para aplicar una fuerza direccional controlada a la diana atada a la superficie. Además, controlando la dirección de las líneas del campo magnético, se pueden usar microesferas/partículas paramagnéticas para guiar y mantener la diana cargada con la Xsonda en la superficie del sustrato durante las etapas de lavado. Secuestrando las dianas a lo largo de la superficie, se puede minimizar la pérdida de dianas debida a fuerzas de cizallamiento. Como alternativa al alargamiento del campo magnético, 40 se pueden usar microesferas/partículas de flotabilidad positiva y negativa que son más o menos densas que el agua, respectivamente, para proporcionar una fuerza de alargamiento. Todas las microesferas/partículas se recubren en la superficie según sea necesario para minimizar las interacciones inespecíficas (agregación de microesferas, unión a sonda) y se functionalizan para permitir la reticulación covalente con dianas modificadas por adaptador.

45 Las figuras 41A a 41D ilustran estrategias de alargamiento de dianas basadas en microesferas representativas. En la figura 41A se ilustra el alargamiento de dianas usando microesferas/partículas paramagnéticas atadas en el extremo atraídas por un campo magnético externo (B). El secuestro de dianas en la superficie del sustrato (para reducir el cizallamiento de la diana) usando microesferas/partículas paramagnéticas atadas en el extremo atraídas por un campo magnético externo se ilustra en la figura 41B. Las figuras 41C y 41D ilustran el alargamiento de dianas usando 50 microesferas/partículas de flotabilidad negativa atadas en el extremo (densidad mayor que el agua) y microesferas/partículas de flotabilidad positiva (densidad menor que el agua), respectivamente. Los procedimientos de solución libre para alargamiento de dianas usando restos atados en el extremo, por ejemplo, usando una perla de flotabilidad tanto positiva como negativa para functionalizar extremos opuestos de una diana, proporcionan una alternativa elegante para reducir la estructura secundaria de la diana.

55 55 El uso de estas estrategias de alargamiento en la preparación de Xpandómeros se ilustra en la figura 41E. En el presente documento, en la etapa I, el molde inmovilizado (416) se estira usando una microesfera flotante hibridada a la diana por medio del cebador adaptador (417). A continuación, el molde se pone en contacto en la etapa II con construcciones de sustrato y estas se ligan a continuación a la diana monocatenaria, produciendo un intermedio de Xpandómero bicatenario (411) con una construcción de sonda-bucle característica y dos cadenas principales, una a través de la cadena principal primaria del polinucleótido y la otro a través de la cadena principal del Xpandómero restringido. En la etapa III, la hebra molde se desnaturaliza, y en la etapa IV, el intermedio de Xpandómero monocatenario se escinde en enlaces selectivamente escindibles en la cadena principal primaria, lo que da como resultado un despliegamiento de los bucles y alargamiento del producto Xpandómero con una cadena principal sustituta totalmente extendida (419).

65 Alargamiento de la diana de ADN usando modificaciones de extremo de polímero

- Una alternativa a los procedimientos descritos en la figura 40 usa polímeros funcionalizados largos (en lugar de microesferas) unidos covalentemente a los extremos libres de las dianas de ADN atadas a la superficie para alargar la diana. El electroestiramiento y enhebrado de las modificaciones de los extremos de polímero a través de un sustrato 5 poroso seguido de la captura (secuestro) del polímero dentro del sustrato produce ADN diana monocatenario completamente alargado significativamente libre de estructura secundaria. La figura 42A ilustra este procedimiento, que muestra el enhebrado de las hebras diana a través de los poros en un sustrato. Los sustratos porosos incluyen, pero sin limitaciones, matriz de gel, óxido de aluminio poroso y membranas porosas. La captura del polímero completamente 10 alargado se puede lograr por funcionalización química controlada en la que la reticulación o unión del polímero al sustrato poroso se pueden iniciar selectivamente después del alargamiento completo de la diana. Existen disponibles diversas estrategias de reticulación o unión. Por ejemplo, la reticulación del polímero alargado funcionalizado con carboxilo a un sustrato poroso funcionalizado con amina se puede lograr con la introducción del agente de reticulación clorhidrato de 1- 15 etil-3-[3-dimetilamino propil]carbodiimida (EDC).
- En la figura 42B se ilustra otro modo de realización que consiste en electroestirar el polímero hacia una superficie 20 funcionalizada que se reticula o une al polímero. En este caso, el polímero no necesariamente enhebra el sustrato, sino que se reticula o une a grupos funcionales en la superficie. Para este enfoque, los sustratos incluyen, pero sin limitaciones, matriz de gel, óxido de aluminio poroso, membranas porosas y superficies conductoras no porosas como el oro. Tanto los sustratos porosos como los no porosos se pueden funcionalizar usando diversas químicas de reticulación (hidrazida-aldehído; oro-tiol) o de unión (biotina-estreptavidina), listas para usar.

Como alternativa, se puede emplear un sistema de activación eléctrica cerrado que usa regiones de campo eléctrico variable (alta, baja, tierra) para capturar segmentos de polímero cargados. El procedimiento, que se ilustra en la figura 25 42C, produce un tipo de recinto de jaula de Faraday que se puede usar para mantener el polímero y el ADN diana en su posición alargada. Este enfoque requiere una división o confinamiento eficaz de los campos eléctricos y tiene la ventaja de permitir un alargamiento ajustable de la diana (a lo largo de la síntesis de SBX) modificando los parámetros de activación. Fuera de la jaula de Faraday, las influencias del campo eléctrico en las construcciones de sustrato son mínimas.

- En un procedimiento relacionado, se usa un polímero elástico para proporcionar una fuerza de estiramiento constante al 30 ácido nucleico diana menos elástico. Este procedimiento se puede usar como complemento de cualquiera de los procedimientos de alargamiento descritos en el presente documento, incluyendo microesferas magnéticas, deformación por flotabilidad, deformación por densidad (gravedad) o electroestiramiento. La fuerza elástica almacenada dentro del polímero proporciona una fuerza de alargamiento amortiguada más consecuente la hebra diana.

35 Sustrato diana basado en matriz de gel con electroenderezamiento

En la figura 43A, se muestra un procedimiento alternativo para la formación de matrices de ADN diana monocatenario (frente a la unión a un soporte sólido) en el que la diana se une covalentemente a una matriz de gel de tamizado (430). En la figura 43b (inserto), se pueden observar construcciones de sustrato asociadas con los moldes anclados, que se han 40 estirado en línea recta. Se pueden usar campos eléctricos para alargamiento de la diana y presentación de la construcción de sustrato. Este enfoque tiene ventajas sobre otros procedimientos para la matriz de dianas de ADN en términos de densidad de la diana y alargamiento de la diana (electroenderezamiento). Por ejemplo, la reticulación de ADN con acrilamida o medio de adaptadores de extremo modificados con acrilato (unidos a oligonucleótidos) se hace de forma rutinaria. Acrydite™ es una fosforamidita disponible (Matrix Technologies, Inc., Hudson, NH, EE. UU.) que se ha usado ampliamente para incorporar modificaciones en el extremo 5' de metacrilato a los oligonucleótidos: el doble enlace en el 45 grupo Acrydite reacciona con los 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 comercialmente (G Biosciences, St. Louis, MO, EE. UU.) y se puede usar de forma similar para que se reticle con ADN usando químicas de reticulación reactiva con amina disponibles comercialmente (Spagna *et al.*, "Stabilization of a β-glucosidase from *Aspergillus niger* by binding to an amine agarose gel", *J. of Mol. Catalysis B: Enzymatic* 11(2,3): 63-69, 2000).

Al igual que con los procedimientos de unión en soporte sólido descritos previamente, los fragmentos de ADN se 55 funcionalizan de forma útil con adaptadores de extremo para producir, por ejemplo, una modificación con Acrydite en el extremo 5'. Para garantizar la uniformidad de las dianas, las dianas de ADN bicatenario se desnaturalizan (y se mantienen en un estado desnaturalizado) mientras se reticulan con la matriz de tamizado. La diana se puede desnaturalizar usando una variedad de técnicas existentes que incluyen, pero sin limitaciones, urea, pH alcalino y fusión térmica. Una vez 60 reticulado a una matriz de gel, el ADN diana desnaturalizado se puede enderezar aplicando un campo eléctrico apropiado (similar a la electroforesis). La difusión, la electroforesis y la neutralización de los desnaturalizantes dentro de la matriz del gel junto con un estricto control de la temperatura, con la adición de soluciones compatibles con hibridación, producen un entorno propicio para la hibridación de construcciones de sustrato. Por ejemplo, se puede presentar Xsondas continuamente a dianas atadas por electroforesis. Este formato permite el reciclaje de sondas y proporciona un buen control sobre el caudal de las Xsondas.

65 La densidad diana de una disposición de matriz de gel de 5 mm x 50 mm x 20 mm produce un volumen de reactivo de 5,0 x 10¹² μm³, que es sustancialmente mayor que el que se puede disponer en una superficie plana. La figura 43C muestra

las estructuras químicas para acilamida, bis-acrilamida y la modificación del extremo del oligonucleótido con Acrydite. La figura 43D ilustra una matriz de gel de poliacrilamida con una diana de ADN unida covalentemente.

El procesamiento enzimático de construcciones de sustrato se puede hacer dentro de la matriz de gel. Los ensayos enzimáticos que usan polimerasas (PCR) y ligasas se usan de forma rutinaria para la modificación de oligonucleótidos acoplados covalentemente a almohadillas de gel 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 administra a Xsonda-dianas ensambladas dentro de la matriz del gel por difusión pasiva, electroforesis o bien una combinación de ambas. Se considera una amplia gama de densidades de matriz y tamaños de poro; sin embargo, la difusión pasiva se beneficia de una matriz de densidad relativamente baja y poros grandes. El suministro activo de ligasa por electroforesis se puede lograr mediante la unión de modificadores de carga (polímeros y dendrímeros) a la enzima para potenciar su migración a través de la matriz.

Una variación del procedimiento de matriz de gel es la electroforesis de polinucleótidos diana monocatenarios modificados en el extremo a través de una matriz de gel. Se pueden usar modificaciones de extremo para crear arrastre en un extremo de la diana. Se puede realizar una modificación de extremo catiónica en el otro extremo de la diana para arrastrarlo a través del gel. También se puede inducir una zona o gradiente de temperatura en el gel para aumentar o modular la restricción de la hibridación a medida que las dianas avanzan a través de la matriz del gel. El procesamiento enzimático se puede hacer dentro de la matriz de gel o después de que los complejos diana-sustrato hibridados salgan de la matriz de gel. El procedimiento se puede repetir como sea necesario para producir la longitud promedio deseada del Xpandómero.

La figura 44 ilustra el uso de marcas de arrastre para la manipulación del Xpandómero posterior a la síntesis. La marca de arrastre (representada como un rombo) sirve como solución equivalente a 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 se unen mediante química de conector a los adaptadores funcionalizados en los extremos 5' del molde monocatenario (L1' de la marca de arrastre se combina con L1 del molde). En la etapa que se muestra, la adición procesiva de construcciones de sustrato ya se muestra en progreso. En la figura 44B, la marca de arrastre se añade asociándola a un adaptador complementario. En la figura 44C, el Xpandómero se trata primero para unir un conector a un adaptador en 3' (que se muestra como el cuadrado pequeño con L1'), y, a continuación, el conector (L1) se usa en la etapa II para unir la marca de arrastre al mismo. La transferasa terminal, que extiende el extremo 3'-OH del ADN monocatenario o bicatenario, polimeriza trifosfatos de nucleótido modificados con conectores, tal como nucleótidos biotinilados. También se pueden usar otras enzimas para añadir bases u oligómeros modificados. En la figura 44C, se añade una marca de arrastre al extremo 3' del intermedio de Xpandómero. Las marcas de arrastre pueden incluir, pero sin limitaciones, nanopartículas, microesferas, polímeros, polímeros ramificados y dendrímeros.

Huecos y relleno de huecos con hibridación de construcción de sustrato y ligación

Se pueden emplear múltiples variantes en los procedimientos de ligación para abordar el manejo y llenado de huecos. En un modo de realización, se usa ligación gradual cíclica, donde las sondas se ensamblan secuencialmente a partir de un cebador atado a la superficie hibridado al polinucleótido diana. En otro modo de realización, la hibridación "indiscriminada" y la ligación de construcciones de sustrato se producen simultáneamente en toda la secuencia de ADN diana, en general sin cebador.

Para los enfoques de ligación tanto cíclica como indiscriminada, es posible que, por ejemplo, una parte de las construcciones de sustrato triméricas (por ejemplo, un 25 %) y alguna porción de las construcciones de sustrato pentaméricas (por ejemplo, un 20 %) hibriden de forma contigua para permitir la ligación. De estos híbridos contiguos, un porcentaje puede contener secuencias con emparejamiento erróneo que darán como resultado un error de medición (si se ligan) o imposibilidad de ligación. La mayoría de los casos de imposibilidad de ligación (debido a un emparejamiento erróneo) no tienen consecuencias, ya que las sondas no ligadas se retiran antes de la siguiente ronda de hibridación. El ciclo de hibridación/lavado/ligación/lavado se puede repetir 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 procedimiento cíclico (ilustrado con Xsondas en la figura 14), las Xsondas se ensamblan secuencialmente desde un extremo de un cebador atado a la superficie hibridado al ADN diana; una Xsonda por ciclo. Siendo este el caso, la longitud de lectura de la diana se puede calcular usando los siguientes supuesto: un 25 % de hibridación contigua (de tetrameros) con un 20 % de fidelidad de híbrido-sonda perfecta produciría solo 20 ligaciones secuenciales de tetrameros ($400 \times 25\% \times 20\%$ después de 400 ciclos. Usando estos supuestos, un ensayo de 400 ciclos (1,5 días a 6'/ciclo) produciría una longitud de producto promedio de 80 nucleótidos.

En el procedimiento indiscriminado (ilustrado con Xsondas de clase I en la figura 45), si se permite que las reacciones de hibridación y ligación de Xsondas se produzcan de manera espontánea y simultánea a lo largo del molde diana de ADN, se puede lograr la replicación de un molde de ADN mucho más largo usando muchos menos ciclos. Después de cada ciclo de hibridación y lavado, se puede usar ligación para conectar los híbridos restantes que son contiguos y 100 % correctos. Se puede usar otro lavado más riguroso para eliminar productos de ligación más pequeños (todos de octámeros

y menos; algunos de dodecámeros), así como todos los tetrámeros no ligados. Con las reacciones de ligación que se producen a lo largo de la secuencia diana de 1000-10 000 nucleótidos, la procesividad del ensayo aumenta drásticamente. Dado que el enfoque indiscriminado para hibridación y ligación de Xsondas es independiente de la longitud del molde diana de ADN, la mayoría del molde diana se puede replicar en una fracción de los ciclos requeridos para el procedimiento en serie. Se pueden usar tiempos de ciclo más largos para compensar las limitaciones cinéticas si dicho ajuste incrementa la fidelidad y/o cantidad de reacciones de hibridación/ligación.

La figura 45 ilustra una progresión secuencial del procedimiento promiscuo. La etapa I ilustra la hibridación de Xsondas tetraméricas en múltiples locus a lo largo de la diana de ADN. La etapa II ilustra la hibridación y ligación de Xsondas contiguas; las Xsondas ligadas se estabilizan y, por tanto, permanecen hibridadas mientras que las Xsondas no ligadas se desprenden por fusión. La etapa III y la etapa IV ilustran otro ciclo térmico de hibridación seguido de ligación y fusión térmica de Xsondas no ligadas. Cada ciclo extiende preferentemente las cadenas de Xsonda 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 la longitud de una sonda, donde no se ha producido hibridación. Para completar el Xpandómero, las Xsondas se unen a través de los huecos por medios enzimáticos o químicos, como se ilustra en la etapa V.

Llenado de huecos estándar

Como se ilustra con Xsondas en la etapa V de la figura 45, después de completar el procedimiento cíclico de hibridación y ligación, los huecos de la secuencia se pueden llenar para producir un Xpandómero continuo. Los huecos a lo largo de la cadena principal de ADN diana se pueden llenar usando procedimientos de llenado de huecos basado 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 contiguas de Xsondas se encuentran entre sí y tienen una longitud de hueco de 1, 2 o 3 nucleótidos entre ellas (suponiendo una Xsonda de clase 1 tetramérica, como se ilustra en la figura 45). El llenado de huecos también se puede hacer por medio de reticulación química (Burgin *et al.*). Una vez que se completa el llenado de huecos, el complemento de ADN diana, que se compone principalmente de Xsondas ligadas con rellenos periódicos de 1, 2 o 3 nucleótidos, se puede procesar (purificación, escisión, modificación de extremo, marcado de indicadores) como sea apropiado para producir un Xpandómero medible. Dado que los productos del ensayo de SBX se preparan y purifican en un procesamiento por lotes antes de la etapa de detección, la detección es eficiente y no está limitada en velocidad por ningún procesamiento bioquímico concurrente.

Para diferenciar los huecos de la señal indicadora específica de nucleótido, se pueden usar desoxinucleótidos trifosfatos modificados (que ya están marcados o bien que se pueden marcar después del ensayo) para identificar la secuencia del hueco. Los desoxinucleótidos trifosfatos adecuados se ilustran en las figuras 46A (bases modificadas con conector) y 46B (bases modificadas con biotina).

La longitud y frecuencia de los huecos dependen de una serie de variables de síntesis, que incluyen número de ciclos, restricción de la hibridación, estrategia de la colección (indiscriminada frente a subcolección con ajuste estequiométrico), longitud del molde diana (100b - 1 Mb) y densidad de reacción (0,1-10 B). La frecuencia de los huecos y longitud de los huecos se pueden reducir significativamente usando condiciones de máxima restricción, siempre que las condiciones sean compatibles con el intervalo de tiempo del ensayo especificado. La longitud de la diana y la densidad de la reacción también son factores importantes con respecto tanto a la fidelidad de la alineación como a la probabilidad de llenar huecos.

El procedimiento indiscriminado de hibridación puede usar procedimientos de ciclos térmicos para mejorar la restricción de la hibridación e incrementar la frecuencia de alineaciones de construcciones de sustratos contiguos. La hibridación, ligación y fusión térmica continúan con una rutina de ciclo térmico precisa hasta que la mayoría de la secuencia diana se hibrida con la sonda. Los híbridos de sonda-diana inespecíficos y débilmente unidos se pueden retirar mediante un simple etapa de lavado, de nuevo con un control térmico preciso. La ligación enzimática o química se puede realizar para unir cualquier construcción oligomérica hibridada de forma contigua a lo largo del ADN diana, generando secuencias más largas y más estabilizadas. La ligación enzimática tiene el beneficio añadido de proporcionar una verificación cruzada de la fidelidad del híbrido de adición, ya que las sondas-dianas emparejadas erróneamente no se ligan eficazmente. Con un segundo lavado en condiciones precisas de ciclo térmico, las construcciones de sustrato no unidas se pueden desprender por fusión de la cadena principal de ADN; permanecerán intermedios más largos. Las secuencias ligadas más largas crecen en múltiples locus a lo largo del ADN diana hasta que la replicación se completa en su mayor parte. El procedimiento de hibridación/lavado/ligación/lavado se repite de cero a muchas veces hasta que se ha replicado la mayoría del molde diana.

El ciclo de temperatura junto con la estabilización de la construcción de sustrato debido al apilamiento de bases se puede usar para enriquecer los acontecimientos de hibridación contiguos. Las condiciones de los ciclos térmicos, similares en concepto a la "PCR de toma de contacto", se pueden usar para eliminar híbridos no contiguos menos estables. Por ejemplo, mediante ciclos repetitivos entre la temperatura de hibridación y la temperatura de fusión superior ("Tm") estimada para la colección de oligonucleótidos (como se determinar por fusión de una sola sonda, es decir, no estabilizada con apilamiento de bases), se pueden seleccionar positivamente acontecimientos de hibridación de sonda contiguos. Con suficiente eficacia de enriquecimiento, el número de ciclos se puede reducir significativamente. Esto se puede realizar en condiciones de un solo baño o de múltiples baños.

Con un tamaño de colección pequeño (256, construcciones de sustrato triméricas, por ejemplo), la colección se puede dividir en subcolecciones en base a un intervalo incluso más estrecho de temperatura de fusión. Esta estrategia puede beneficiar significativamente el procedimiento si el sesgo de T_m no se puede controlar a través de modificación de bases, adyuvantes de la solución de hibridación y similares. Además, la estequiométría de la sonda se puede ajustar (hacia arriba o hacia abajo) para compensar cualquier sesgo residual que todavía pueda existir en la colección/subcolecciones.

5

Se realizó un modelado informático para evaluar estadísticos de aparición de huecos y las longitudes de construcciones de sustrato conectadas consecutivamente que se hibridan en un molde de ADN diana. El modelo simula un ciclo de fusión térmica de hibridación/ligación con una colección completa de 256, Xsondas tetraméricas. Los resultados del modelo que se presentan en el presente documento se basan en el siguiente procedimiento del modelo:

10

i) La etapa de hibridación/ligación simula Xsondas tetraméricas aleatorias que se encuentran aleatoriamente con un molde de ADN secuenciado aleatoriamente en posiciones aleatorias y que hibridan si hay emparejamiento. Ninguna Xsonda se puede superponer a ningún nucleótido. La hibridación continúa hasta que se hibridan todas las localizaciones en la diana mayor de 3 nucleótidos de longitud.

15

ii) La etapa de fusión térmica se simula retirando todas las Xsondas que están en cadenas más cortas que las M Xsondas tetraméricas de longitud (donde $M = 2,3,4,5,\dots$). Las cadenas más largas permanecen en el molde de ADN. Una "cadena" se define en el presente documento como múltiples Xsondas consecutivas sin huecos entre ellas.

20

iii) Repetir el ciclo definido por i) y ii) para que las Xsondas se construyan aleatoriamente a partir de los locus existentes de cadenas de múltiples Xsondas. El ciclo se detiene cuando no se produce ningún cambio entre 2 ciclos consecutivos.

25

La figura 47 y la figura 48 tienen dos gráficos que ilustran los estadísticos de ejecutar el modelo 100 veces en un molde de ADN aleatorio. Para fines de procesamiento, se seleccionó que la diana de ADN tuviera una longitud de 300 nucleótidos (pero los estadísticos analizados aquí no cambian si la longitud del molde es de 5000 nucleótidos). La figura 47 muestra los resultados para $M = 3$, donde las cadenas más cortas que $M=3$ se "desprenden por fusión" del molde de ADN. La figura 47A muestra los estadísticos de las longitudes de las cadenas en el último ciclo. La distribución parece ser casi exponencial en un intervalo que va de dodecámero ($M = 3$) hasta algunos raros heptamirahexacontámero grandes con una media de ~28 nucleótidos de longitud. La figura 47B muestra que es necesario un intervalo de 5 a 20 ciclos para completar el molde de ADN usando este procedimiento de "crecimiento" de cadena y que un 90 % de las ejecuciones se completan después de 12 ciclos. Las figuras 48A y 48B muestran el mismo tipo de información para el caso donde $M=4$, una fusión térmica más rigurosa. En este caso, la longitud promedio de la cadena aumenta a ~40 nucleótidos, pero requiere ~18 ciclos para completar un 90 % de las ejecuciones. Las longitudes de cadena más largas reducen la probabilidad de tener un hueco como se esperaba. 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).

30

35

Relleno de huecos con construcciones de sustrato triméricas y diméricas

40

Una extensión del modo de realización de hibridación básica de tetrámeros es usar construcciones de sustrato triméricas y diméricas para llenar los huecos de secuencia de 3 bases y 2 bases, respectivamente. La figura 49 muestra un ejemplo de llenado de huecos mediante adición secuencial o simultánea de Xsondas más cortas. Si la mayoría del molde diana se hibrida con éxito con Xsondas después de la hibridación de tetrámeros, la mayoría de los huecos restantes tienen una longitud de 1, 2 o bien 3 bases. El tamaño promedio del Xpandómero se puede incrementar significativamente si un porcentaje de moderado a alto de los huecos de 2 y 3 bases se puede llenar y ligar con Xsondas más pequeñas. Dado que la hibridación de estas sondas más pequeñas se hace después de hibridar la mayor parte de la diana, la restricción de la temperatura se puede reducir para permitir la hibridación de trímeros y, a continuación, de dímeros, respectivamente. La estabilidad de híbridos de las Xsondas más pequeñas se incrementa debido a la estabilización por apilamiento de bases. Después de la ligación, los huecos restantes se llenan usando una polimerasa para insertar nucleótidos, seguido de nuevo por ligación para unir todos los 5' fosfatos y 3' hidroxilos contiguos restantes. Los nucleótidos modificados con conector incorporado a la polimerasa permiten el marcado de huecos, ya sea antes o después de la incorporación. Estos marcadores pueden o no identificar la base. En base a modelos estadísticos, resumidos en los gráficos de la figura 50, los polímeros de Xpandómero de longitud completa tienen una cobertura de secuencia diana mayor de un 95 % (mayor de 16 bases). De acuerdo con el modelo, el uso llenado de huecos en Xsondas triméricas/diméricas reduce el porcentaje de bases no marcadas en un factor de 6 (suponiendo una incorporación de alta eficacia), lo que, para una secuencia diana de 5000 bases, daría como resultado 5000 bases de secuencia en su mayoría contigua con menos de 50 huecos de base única no discriminada a lo largo del replicón de la diana. De este modo, a 16X se identifica una cobertura mayor de un 99 % de las secuencias de huecos no identificadas.

45

50

55

60

65

65

El llenado de huecos con estas Xsondas triméricas y diméricas más pequeñas, por ejemplo, extiende la longitud promedio del Xpandómero multiplicándola por tres. Los gráficos que resumen los resultados del modelado estadístico como se muestra en las figuras 50A y 50B indican que la longitud promedio de los Xpandómeros con relleno de huecos triméricos/diméricos, en las condiciones descritas, está en el intervalo de 130 bases de secuencia. Además, con 5000 gigabases de secuencia de Xpandómero sintetizadas de esta manera, lo que es equivalente a convertir solo 2,5 ng de secuencia diana en Xpandómero, la selección del tamaño de un 10 % más largo de fragmentos de Xpandómero de esa población produciría 500 gigabases de secuencia con longitudes promedio en el intervalo de 381 bases, mientras que la

selección de tamaño de un 2 % más largo de fragmentos de Xpandómero produciría 100 gigabases de secuencia con longitudes promedio en el intervalo de 554 bases. Estas longitudes de fragmentos se pueden lograr sin necesidad de llenar huecos de una sola base con polimerasa. Si la hibridación con Xsondas triméricas se hace sin llenado de espacios, los modelos estadísticos indican que las longitudes promedio para un 10 % y un 2 % más largo de los fragmentos serían 5 106 y 148 bases, respectivamente.

Adición de adyuvantes para reducir la estructura secundaria diana

El ADN diana monocatenario atado a la superficie y largo puede suponer una diana de hibridación desafianta debido a 10 estructuras secundarias intramoleculares. La adición de una microesfera atada en el extremo, como se describe en la figura 41E, reduce la aparición de estas formaciones intramoleculares y las desestabiliza cuando se producen. Sin embargo, puede resultar necesario disminuir aún más la formación de estructuras secundarias, que pueden bloquear eficazmente la hibridación de construcciones de sustrato, con la adición de adyuvantes.

La adición de sondas de oligonucleótidos diméricos y/o triméricos no marcadas a, por ejemplo, una mezcla de hibridación 15 de Xsondas puede servir para este fin. Para eliminar la posibilidad de incorporación en el Xpandómero, estas sondas se sintetizan para que no se puedan ligar (sin 5' fosfato o 3' hidroxilo). Debido a su pequeño tamaño, es poco probable que los adyuvantes diméricos y/o triméricos formen híbridos estables a las temperaturas de hibridación de Xsondas tetraméricas; sin embargo, su presencia en la mezcla de reacción en concentraciones de moderadas a altas reduce la 20 frecuencia y estabilidad de la estructura secundaria diana mediante el bloqueo débil y transitorio del acceso de secuencias de nucleótidos intramoleculares que, de otro modo, podrían hibridar. La figura 51 es una ilustración de cómo la adición de adyuvantes diméricos o triméricos inhibe la formación de estructuras secundarias (solo por simplicidad, las Xsondas triméricas no se mostraron en la figura). Junto con la fuerza de alargamiento proporcionada por la atadura de microesferas, 25 los adyuvantes disminuyen significativamente la frecuencia de formación de estructuras secundarias.

Estos adyuvantes de dinucleótidos y/o trinucleótidos 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 secuencias necesarias. La figura 52 muestra sustituciones de bases universales comunes que se usan para este propósito.

Otro enfoque para reducir la estructura secundaria diana es replicar el ADN diana para producir una diana de ADNc sintético que tenga una estabilidad reducida de la estructura secundaria en comparación con el ADN natural. 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 se pueden incorporar en la diana de ADNc sintética descrita. Se ha demostrado que estos análogos, que incluyen (pero sin limitaciones) N4-etildesoxicitidina, 2-aminoadenosina-5'-monofosfato, 2-tiouridina-monofosfato, inosina-monofosfato, pirrolopirimidina-monofosfato y 2-tiocitidina-monofosfato, 30 reducen la estructura secundaria del ADNc. Este enfoque se puede usar junto con adyuvantes de alargamiento e hibridación de la diana para reducir la estructura secundaria.

40 Detección y medición

Como se ha mencionado previamente, el Xpandómero se puede marcar y medir mediante una serie de técnicas. El potencial de salida masiva de datos del procedimiento SBX se adapta bien a las matrices de sensores basadas en nanoporos o tecnologías equivalentes. En un modo de realización, la matriz de nanoporos puede funcionar como fuente 45 de ionización en el extremo frontal de un espectrómetro de masas, en el que los códigos de indicadores en el Xpandómero son marcas de espectroscopia de masas escindibles. Otros modos de realización implican el uso de un sensor de nanoporos tal como impedancia/conductancia eléctrica o FRET.

Un modo de realización de detección usa una matriz de nanoporos como se muestra en la figura 54. En este modo de 50 realización, el Xpandómero se ensambla usando los procedimientos descritos previamente, excepto que la diana de ADN no está anclado a un soporte sólido inmovilizado sino que está anclada a una microesfera magnética que tiene su sonda de anclaje enhebrada a través de un sustrato de nanoporo. Además, una multitud de fluoróforos donantes de FRET, para dos longitudes de onda de excitación, se atan a la entrada del nanoporo (que se muestra como pequeños cuadrados). Los fluoróforos aceptores de FRET constituyen los indicadores incorporados en el Xpandómero. Después de que las 55 Xsondas unidas del Xpandómero se escinden, el Xpandómero se estira por una fuerza que puede ser electrostática, magnética, gravitacional y/o mecánica y se puede facilitar por la microesfera usada para extender la diana de ADN original (que se muestra unida a la parte superior del Xpandómero). En la etapa de medición final, el Xpandómero se extrae a través del nanoporo aplicando una fuerza magnética sobre la microesfera magnética que se muestra debajo de la estructura del nanoporo. Los fluoróforos donantes en la entrada del nanoporo se excitan con una fuente de luz (λ_1) y, 60 cuando los indicadores pasan proximales a los fluoróforos donantes, se excitan y emiten su fluorescencia característica (λ_2) que descodifica la secuencia de nucleótidos asociada.

La figura 54 también ilustra un modo de realización de codificación de la secuencia en los códigos de indicadores. En 65 cada atadura, hay 4 sitios de indicadores, cada uno cargado con una combinación de dos tipos de fluoróforos aceptores de FRET. Esto proporciona cuatro estados medibles en cada sitio usando el nivel de intensidad relativa. Si la fluorescencia de los fluoróforos aceptores fuera roja y verde, la codificación de estos estados a los 4 nucleótidos puede ser: A:

verde>rojo, C: solo rojo, G: solo verde y T: rojo > verde.

En otro modo de realización, el Xpandómero se marca con indicadores de marcas de masa que se miden usando un espectrómetro de masas. El Xpandómero se dosifica en un capilar estrecho que alimenta secuencialmente a los indicadores en un ionizador de electropulverización. Para permitir la medición por el espectrómetro de masas de los indicadores discretos de marcas de masa, las marcas de masa se pueden fotoescindir de la estructura del indicador justo antes, durante o después de la ionización de marcas de masa. Se puede usar un 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 espaciadas por masa. Esto se puede usar para mejorar la sensibilidad y el rendimiento del instrumento. Además, emplear un instrumento que tiene múltiples canales para realizar la ionización y detección en paralelo incrementa el rendimiento en órdenes de magnitud.

Un modo de realización de detección con un enfoque de espectrómetro de masas es un detector de masas TOF multicanal que puede leer >100 canales simultáneamente. Un instrumento adecuado usaría una fuente de ionización multicanal que alimenta el Xpandómero en múltiples canales a una concentración y velocidad que maximiza el uso del canal, maximizando, por tanto, la velocidad de salida de datos de buena calidad. Una fuente de ionización de este tipo requiere una separación adecuada de los segmentos indicadores de marcas de masa. Las marcas de masa se pueden fotoescindir a medida que emergen de un nanoporo y se ionizan. Los requisitos de dispersión no son altos, de modo que todo lo que se requiere es un tubo de vuelo corto. Es posible obtener resultados de medición extremadamente altos con un enfoque de detección por espectrómetro de masas multicanal. Por ejemplo, una matriz de 100 canales iónicos de nanoporos que leen a una velocidad de 10 000 códigos de indicadores por segundo (con 4 mediciones de nucleótidos por código de indicador) lograría un rendimiento del instrumento de > 4 M bases/segundo.

En otro modo de realización, el nanoporo se usa de manera similar a un contador Coulter (figura 55). En esta implementación, la densidad de carga del Xpandómero está diseñada para que sea similar a la del ADN natural. Los indicadores están diseñados para producir 3 niveles de impedancia medidos en el detector de nanoporos, por ejemplo. Los Xpandómeros se presentan en una solución libre que tiene altas concentraciones de electrolito, tal como KCl 1 M. El nanoporo tiene un diámetro de 2 a 15 nm y una longitud de 4 a 30 nm. Para lograr una buena resolución de los marcadores de construcción de atadura, se seleccionan para que tengan casi la misma longitud o más. El diámetro de los marcadores puede ser diferente para los 3 niveles. El segmento del Xpandómero cerca de los enlaces de la construcción de atadura no tiene indicadores (por ejemplo, PEG), este segmento tendrá un nivel de impedancia particular. Uno de los tres niveles de indicador puede ser equivalente. Se pueden producir diferentes firmas de impedancia, que implican tanto nivel de impedancia como respuesta temporal, variando las longitudes de los segmentos, la densidad de carga y la densidad molecular. Por ejemplo, para lograr 3 niveles de impedancia diferentes, los segmentos de construcciones de atadura marcados se pueden codificar químicamente para que se acoplen a uno de los tres tipos de polímeros diferentes, cada uno con una longitud y densidad de carga diferentes.

Cuando el Xpandómero pasa a través del detector de nanoporos, la corriente se modula de acuerdo con el tipo de marcador presente. La cantidad de polímero cargado que reside en el nanoporo afecta tanto a la corriente de las especies de electrolitos como a la velocidad de translocación.

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

En otro aspecto de un aparato de detección basado en nanoporos (figura 56), se usan electrodos laterales adheridos al nanoporo para medir la impedancia o la conductividad de lado a lado en el nanoporo mientras se aplica un voltaje por toda la película de soporte sólido. Esto tiene la ventaja de separar la función de translocación de la función de impedancia. A medida que el Xpandómero se transporta a través del nanoporo, se mide de nuevo la modulación de corriente. Se emplean técnicas de microfluidos y micropipetas, junto con marcas de arrastre, microesferas magnéticas, técnicas de estiramiento electroforético, etc., para transportar el Xpandómero a través del nanoporo. Por ejemplo, la electroforesis en solución libre marcada en el extremo, también denominada ELFSE, es un procedimiento para romper el equilibrio de carga frente a fricción del ADN de drenaje libre que se puede usar para la electroforesis del Xpandómero en solución libre (Slater *et al.*, "End-labeled free-solution electrophoresis of DNA", Electrophoresis 26: 331-350, 2005).

Los procedimientos para atar, estirar, marcar y medir 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 una sola nucleobase para la secuenciación del genoma completo de ácidos nucleicos naturales está más allá de las capacidades de estas técnicas. En la figura 57 a 59, se representan varios procedimientos de detección de Xpandómeros de "molécula única". En la figura 57, se representa un microscopio y se puede concebir cualquiera de una variedad de técnicas de formación de imágenes directas que aprovechan la mayor resolución espacial de la estructura del Xpandómero. Una molécula de Xpandómero se coloca plana sobre una superficie en general plana y se hace un barrido a lo largo de su longitud. Los ejemplos incluyen microscopía óptica y microscopía de fluorescencia. El uso de microscopía de alta resolución está limitado en la mayoría de los casos a una resolución de >100 nm, que todavía es posible con ataduras largas (100 nm por indicador, por ejemplo). Se puede lograr una resolución más alta (que requiere ataduras <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 MJ *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 interacciones de excitación de FRET localizadas, se pueden localizar a < 10 nm, dando lugar a longitudes de atadura de ~10 nm/indicador.

5

La detección y análisis de moléculas grandes de ADN mediante microscopía electrónica están bien establecidos (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 exacta y de alto rendimiento de polinucleótidos usando estos procedimientos ha resultado difícil. En la figura 58, la microscopía electrónica de transmisión (TEM) y de barrido (SEM) se indican conceptualmente para la detección de un Xpandómero. Aquí, se usa un haz de electrones enfocado para hacer un barrido de un Xpandómero, que de nuevo es en general plano sobre una superficie. Aspectos de la estructura del Xpandómero sirven para descodificar la información genética en la cadena principal. Las técnicas de fijación de muestras y recubrimiento por pulverización catódica, que permiten la formación de imágenes de rasgos característicos individuales y del tamaño de un átomo de las moléculas, se pueden usar para potenciar la ampliación.

También son de interés la espectroscopia de conductancia de tunelización de electrones activada por nanoelectrodos, en la que un haz de electrones de tunelización entre dos puntas de nanoelectrodos se modula mediante el transporte del Xpandómero entre las puntas (Lee *et al.*, "Nanoelectrode-Gated Detection of Individual Molecules with Potential for Rapid Secuenciación de ADN", Fenómenos de estado sólido 121-123:1379-1386, 2007). El Xpandómero perturba la corriente de tunelización por su efecto de cribado-conducción, que se puede amplificar sobre el ADN natural mediante el uso de indicadores adecuados. Esta técnica tiene la ventaja de que se evita la fijación de la muestra y el requisito de vacío y, en teoría, se pueden emplear matrices masivamente paralelas de compuertas de electrodos para leer muchos Xpandómeros en paralelo.

25

En la figura 59, se ilustra conceptualmente la microscopía de fuerza atómica. En un aspecto simple, un nanotubo montado sobre un voladizo sensible barrió una superficie y las fuerzas de atracción y repulsión entre la sonda y la superficie de la muestra se traducen en una imagen topológica de la superficie de la que se hace el barrido. Esta técnica puede lograr una resolución muy alta pero tiene velocidades de barrido relativamente lentes (M. Miles, *Science* 277, 1845-1847 (1997)) La microscopía electrónica de tunelización de barrido (STM) es una tecnología relacionada para la formación de imágenes de superficies; sin embargo, la sonda no toca la superficie, sino que se mide una corriente de tunelización entre la superficie y la sonda. En este caso, el Xpandómero se puede colocar plano sobre una superficie y barrer físicamente con la punta de la sonda, como una aguja de fonógrafo en un disco.

35 Ensamblaje de secuencias

La secuencia de referencia del genoma humano publicada (u otra secuencia de referencia) se puede usar como una herramienta de alineación para ayudar a ensamblar las cantidades masivas de datos de secuencia producidos con SBX. A pesar de la probable inclusión de pequeños huecos de secuencia identificados en posiciones, las capacidades de longitud larga de lectura descritas para SBX basado en Xpandómero simplifican y mejoran la fidelidad del ensamblaje de secuencias genómicas completas. Como se ha analizado anteriormente, el procedimiento se puede simplificar aún más dividiendo fragmentos contiguos en localizaciones dimensionalmente limitadas en la superficie de reacción del ensayo. En este aspecto, el procedimiento de análisis puede reducir drásticamente el tiempo y error de ensamblaje.

45 Construcciones monoméricas

La figura 9 proporciona una visión general de las 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 analizará individualmente a continuación.

50

Las construcciones monoméricas de clases VI a X se distinguen de las construcciones oligoméricas de clase I a V en que usan un residuo de una sola nucleobase como sustrato. En la siguiente descripción, N1 se refiere a cualquier residuo de nucleobase, pero típicamente es un nucleótido trifosfato o un análogo en el presente documento. Tiene puntos de unión en una atadura (también descritos en el presente documento) por ejemplo, a los anillos heterocíclicos de la base, al grupo ribosa o al α-fosfato del residuo de nucleobase. Como se describe, el procedimiento principal para la síntesis dirigida por molde usa polimerasa, pero cualquier procedimiento que pueda realizar la síntesis dirigida por molde es apropiado, incluyendo procedimientos de ligación química y enzimática.

Para las construcciones de sustrato donde se usan grupos conectores ε y δ para crear enlaces entre subunidades, se puede adaptar una amplia gama de químicas disponibles comercialmente (Pierce, Thermo Fisher Scientific, EE. UU.) con este fin. Las químicas de conector comunes incluyen, por ejemplo, ésteres de NHS con aminas, maleimidas con sulfhidrilos, imidoésteres con aminas, EDC con carboxilos para reacciones con aminas, disulfuros de piridilo con sulfhidrilos y similares. Otros implican el uso de grupos funcionales como hidrazida (HZ) y 4-formilbenzoato (4FB) que, a continuación, pueden reaccionar para formar enlaces. Más específicamente, una amplia gama de agentes de reticulación (hetero y homobifuncionales) están ampliamente disponibles (Pierce) que incluyen, entre otros, Sulfo-SMCC (sulfosuccinimidil 4-[N-maleimidometil] ciclohexano-1-carboxilato), SIA (yodoacetato de N-succinimidilo), Sulfo-EMCS

(éster de sulfosuccinimida[Ne-Maleimidocaproiloxi]), Sulfo-GMBS (éster de N-[g-maleimidobutiriloxi]sulfosuccinimida), AMAS N-(a-maleimidoacetoxi) éster de succinimida, BMPS (N EMCA (ácido ne-maleimidocaproico)-[β -maleimidopropiloxi] succinimida éster), 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]isoci, BMPH (N-[ácido β -maleimidopropiónico]hidrazida, sal de ácido trifluoroacético)anato), EMCH ([ácido ne-maleimidocaproico]hidrazida, sal de ácido trifluoroacético), SANH (succinimidil 4-hidrazinonicotinato acetona hidrazone), SHTH (succinimidil 4-hidrazidotereftalato clorhidrato) y C6-SFB (C6-succinimidil 4 -formilbenzoato). Además, el procedimiento descrito por Letsinger et al. ("Oligonucleótidos de fosforotioato que tienen enlaces internucleosídicos modificados", Patente de EE.UU. N.º 6.242.589) pueden adaptarse para formar enlaces de fosforotiolato.

Además, están ampliamente disponibles químicas de protección/desprotección bien establecidas para restos conectores comunes (Benoitou, "Chemistry of Peptide Synthesis", CRC Press, 2005). La protección de amino incluye, pero sin limitaciones, carbamato de 9-fluorenilmetilo (Fmoc-NRR), carbamato de t-butilo (Boc-NRR), carbamato de bencílico (Z-NRR, Cbz-NRR), trifluoroacetamida de acetamida, ftalimida, bencilamina (Bn-NRR), trifenilmetilamina (Tr-NRR) y bencildenamina p-toluenosulfonamida (Ts-NRR). La protección de carboxilos incluye, pero sin limitaciones, éster metílico, éster t-butílico, éster bencílico, éster de S-t-butilo y 2-alquil-1,3-oxazolina. Los carbonilos incluyen, pero sin limitaciones, dimetilacetal 1,3-dioxano y 1,3-ditiano N,N-dimetylhidrazone. La protección de hidroxilos incluye, pero sin limitaciones, éter metoximetílico (MOM-OR), éter tetrahidropiranílico (THP-OR), éter t-butílico, éter alílico, éter bencílico (Bn-OR), éter t-butildimetsilsilílico (TBDMS-OR), éter t-butildifenilsilílico (TBDPS-OR), éster de ácido acético, éster de ácido piválico y éster de ácido benzoico.

En el presente documento, la atadura a menudo se representa como una construcción de indicadores con tres grupos de indicadores, se pueden organizar diversas configuraciones de indicadores en la atadura y pueden comprender un solo indicador para identificar el sustrato, múltiples indicadores para identificar el sustrato, o la atadura puede estar desnuda de polímero (sin indicadores). Hay que tener en cuenta que los indicadores se pueden usar para sincronización de detección, corrección de errores, redundancia u otras funciones. En el caso del polímero desnudo, los indicadores pueden ser el propio sustrato o pueden estar en una segunda atadura unida al sustrato. En algunos casos, uno o más precursores indicadores se disponen en la atadura, y los indicadores se unen por afinidad o se unen covalentemente después del ensamblaje del producto Xpandómero.

Las estrategias de codificación de indicadores se divultan anteriormente y se analizan con más detalle a continuación. Por ejemplo, la codificación binaria de dos bits de cada monómero produciría cuatro secuencias de código únicas (11, 10, 01, 00) que se pueden usar para identificar cada base de la secuencia (adenina "A", citosina "C", guanina "G", timidina "T"), suponiendo que el acoplamiento del sustrato es direccional. Si no es de dirección, entonces un tercer bit proporciona una codificación inequívoca. De forma alternativa, una única construcción de indicadores multihibridada de 4 estados proporciona un código de indicador único para cada base de la secuencia. Se pueden considerar diversas estrategias de funcionalización y marcado para construcciones de atadura, incluyendo, por ejemplo, dendrímeros, polímeros, polímeros ramificados, nanopartículas y nanocrystales funcionalizados como parte de una estructura de indicador, así como químicas de indicadores con una característica de detección que se va a detectar con la tecnología de detección apropiada que incluye, por ejemplo, fluorescencia, emisores o excitadores de FRET, densidad de carga, tamaño o longitud. Se pueden incorporar marcadores específicos de la base en la atadura como sustratos marcados antes o bien después del ensamblaje del Xpandómero. Una vez que el Xpandómero se libera por completo y se alarga, los códigos de indicadores se pueden detectar y analizar usando una gama de procedimientos de detección.

Las colecciones de sustratos adecuados como sustratos monoméricos incluyen (pero sin limitaciones) ATP, GTP, CTP, TTP y UTP modificados.

Construcciones monoméricas de clase VI

La figura 60 describe construcciones de sustrato monoméricas de clase VI (un tipo de RT-NTP) con más detalle. Las figuras 60A a 60C se leen de izquierda a derecha, mostrándose primero la construcción de sustrato monoméricas (precursor de Xpandómero que tiene un residuo de nucleobase), a continuación la hebra hija híbrida intermedia en el centro y, a la derecha, el producto Xpandómero preparado para secuenciación.

Como se muestra en la figura 60A, las construcciones de sustrato monoméricas de clase VI tienen una atadura, T (600), unida por un enlace (601) de un primer resto de extremo a un residuo de nucleobase de sustrato, N. Se dispone un grupo conector, ϵ en el primer resto de extremo (603) de la atadura próximo a R¹. En el extremo distal (604) de la atadura, un segundo resto de extremo con un segundo grupo conector, δ , se sitúa próximo a R². El segundo resto de extremo de la atadura se fija al primer resto de extremo próximo a la nucleobase por un reticulación escindible selectivamente intraatadura (o por otra restricción). La reticulación escindible intraatadura (605) se indica aquí mediante una línea punteada, que puede indicar, por ejemplo, un enlace disulfuro o un conector fotoescindible. Esta restricción evita que la atadura se alargue o expanda y se dice que está en su "configuración restringida". En el ensamblaje dirigido por molde, los sustratos forman un híbrido con el molde diana, de modo que los sustratos son colindantes. En condiciones controladas, los grupos conectores colocados δ y ϵ de los sustratos colindantes se unen para formar un enlace X entre las construcciones de sustrato contiguas. Los grupos conectores δ y ϵ de una construcción de sustrato monomérica no forman un enlace intrasustrato debido a restricciones de posicionamiento. Las químicas de enlace y protección/desprotección

adequadas para δ , ϵ y χ se detallan en la descripción general de la construcción monomérica.

- R¹ y R² son grupos de extremo configurados como es adecuado para el protocolo de síntesis en el que se usa la construcción de sustrato. Por ejemplo, R¹ = 5'-fosfato y R² = 3'-OH encontrarán uso en un protocolo de ligación, y R¹ = 5'-trifosfato y R² = 3'-OH para un protocolo de polimerasa. Opcionalmente, R² se puede configurar con un grupo de bloqueo reversible para la adición ciclica de un solo sustrato. De forma alternativa, R¹ y R² se pueden configurar con grupos de extremo conectores para acoplamiento químico o sin grupos conectores para un protocolo solo de hibridación. R¹ y R² pueden ser del tipo general XR, en el que X es un grupo de enlace y R es un grupo funcional.
- 10 Durante el ensamblaje, la construcción de sustrato monomérica se polimeriza primero en el extremo extensible de la hebra hija emergente por un procedimiento de polimerización dirigida por molde que usa un molde monocatenario como guía. En general, este procedimiento se inicia a partir de un cebador y prosigue en la dirección 5' a 3'. En general, se usa una ADN polimerasa u otra polimerasa para formar la cadena hija, y se seleccionan las condiciones de modo que se obtenga una copia complementaria de la cadena molde. Posteriormente, el grupo conector δ , que ahora está colocado con el grupo conector ϵ de la atadura de la subunidad contigua, se reticula selectivamente para formar un enlace χ , que es un enlace entre ataduras y entre subunidades. Los enlaces χ unen las ataduras en una cadena continua, formando un intermedio denominado "hebra hija híbrida", como se muestra en la figura 60B. Después de que se forma el enlace χ , el enlace intraatadura se puede romper.
- 15 20 La hebra hija híbrida (figura 60B) es un heterocopolímero con las subunidades que se muestran entre paréntesis. La cadena principal primaria ($\sim N \sim$)_k, la hebra molde (-N'-)_k y la atadura (T) se muestran como una hebra hija híbrida, donde k indica una pluralidad de subunidades repetitivas. Cada subunidad de la hebra hija es un "motivo" repetitivo y los motivos tienen una variabilidad específica de especie, indicada en el presente documento por el superíndice α . La hebra hija se forma a partir de especies de construcción de sustrato monomérica seleccionadas mediante un procedimiento dirigido por molde de una colección de especies de motivo, uniéndose el sustrato monomérico de cada especie de construcción de sustrato a un nucleótido complementario correspondiente en la hebra molde diana. De esta manera, la secuencia de residuos de nucleobase (es decir, la cadena principal primaria) de la hebra hija es una copia complementaria contigua de la hebra molde diana.
- 25 30 35 Cada tilde (~) indica un enlace escindible selectivamente que se muestra aquí como enlaces entre sustratos. Estos se pueden escindir selectivamente para liberar y expandir las ataduras (y el Xpandómero) sin degradar el propio Xpandómero.
- 40 La hebra hija se compone de un precursor de Xpandómero llamado "Xpandómero restringido" que además se compone de ataduras en la "configuración restringida". Cuando las ataduras se convierten a su "configuración expandida", el Xpandómero restringido se convierte en el producto Xpandómero. Las ataduras se restringen por los enlaces χ entre subunidades, la unión al sustrato y, opcionalmente, los enlaces intraatadura si todavía están presentes. El enlace χ une el primer resto de extremo de la atadura de una primera subunidad al segundo resto de extremo de la atadura en el extremo colindante de una segunda subunidad y se forma uniendo los grupos conectores colocados, ϵ de la primera subunidad y δ de la segunda subunidad.
- 45 50 Se puede observar que la hebra hija tiene dos cadenas principales, una "cadena principal primaria" y la cadena principal del "Xpandómero restringido". La cadena principal primaria está compuesta por sustratos monoméricos polimerizados y colindantes. La "cadena principal del Xpandómero restringido" evita el enlace escindible selectivamente entre los sustratos monoméricos y se forma por restos de cadena principal unidos por enlaces χ , siendo cada resto de cadena principal una atadura. Se puede observar que la cadena principal del Xpandómero restringido forma un puente sobre los enlaces escindibles selectivamente de la cadena principal primaria, y permanecerá covalentemente intacto cuando estos enlaces escindibles selectivamente se escindan y la cadena principal primaria se fragmente.
- 55 60 65 La figura 60C es una representación del producto Xpandómero de clase VI después de la disociación de la hebra molde y después de la escisión de los enlaces escindibles selectivamente (incluyendo los de la cadena principal primaria y, si no están ya escindidos, los enlaces intraatadura). La hebra del producto Xpandómero contiene una pluralidad de subunidades κ , donde κ denota la késima subunidad en una cadena de m subunidades que componen la cadena hija, donde $\kappa = 1, 2, 3 \dots m$, donde $m > 10$, en general $m > 50$, y típicamente $m > 500$ o > 5000 . Cada subunidad está formada por una atadura en su configuración expandida y se estira en toda su longitud entre los enlaces χ de las subunidades contiguas. El sustrato

colgante está unido a la atadura en cada subunidad. Cada subunidad, un motivo de subunidad α, contiene información genética específica de especie establecida por ensamblaje dirigido por molde del intermedio de Xpandómero (hebra hija).

5 La figura 60D muestra la construcción de sustrato de la figura 60A como un modelo molecular, donde el miembro de sustrato monomérico, representado con un residuo de nucleobase (606), se une a la atadura mediante un enlace (607) del primer resto de extremo de la atadura. También dispuesto en el primer resto de extremo está un grupo conector (608), que se muestra como ε en la figura 60A. Un segundo grupo segundo conector (609), que se muestra como δ en la figura 60A, está dispuesto en un segundo resto de extremo en el extremo distal de la atadura. Se muestra un enlace intraatadura escindible selectivamente (602), representado por los triángulos contiguos, que restringe la atadura uniendo el primer y el segundo resto de extremo. Los grupos conectores ε y δ se sitúan para que no interactúen y para que se alineen preferentemente cerca de los lados R¹ y R² del sustrato, respectivamente. El bucle de atadura (600) que se muestra aquí tiene tres indicadores (590, 591, 592), que también pueden ser específicos de la especie del motivo.

10 15 La figura 60E muestra la construcción de sustrato después de la incorporación en el producto Xpandómero. Las subunidades se escinden y expanden y se unen mediante enlaces χ (580.581), representados como un óvalo abierto, formados uniendo los grupos conectores δ y ε a los que se hace referencia en la figura 60A. Una subunidad se indica mediante líneas punteadas que enmarcan verticalmente la subunidad repetitiva, como se representa por corchetes en la figura 60C adjunta.

20 25 En el producto Xpandómero de la figura 60E, la cadena principal primaria se ha fragmentado y no es covalentemente contiguo porque se ha escindido cualquier enlace directo entre los sustratos de subunidades contiguas. A través del procedimiento de escisión, el Xpandómero restringido se libera para convertirse en el producto Xpandómero. Los miembros de atadura que estaban anteriormente en una configuración restringida, están ahora en una configuración expandida, funcionando de este modo para extender linealmente la información de secuencia de la diana del molde. La expansión de las ataduras reduce la densidad lineal de la información de secuencia a lo largo del Xpandómero y proporciona una plataforma para incrementar el tamaño y la abundancia de los indicadores, lo que a su vez mejora la relación señal-ruido para la detección y descodificación de la secuencia molde.

30 Si bien la atadura se representa como una construcción de indicadores con tres grupos indicadores, se pueden organizar diversas configuraciones de indicadores en la atadura y pueden comprender indicadores individuales que identifican el monómero o la atadura puede ser un polímero desnudo. En algunos casos, uno o más precursores indicadores se disponen en la atadura, y los indicadores se unen por afinidad o se unen covalentemente después del ensamblaje del producto Xpandómero.

35 Volviendo ahora a la figura 61, se muestra un procedimiento para la extensión de una sola base (SBE) con construcciones de sustrato monoméricas. Los moldes diana adaptados al extremo (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 colección de construcciones de sustrato monoméricas, un miembro de la cual se muestra a modo de ilustración (612), y polimerasa (P). Las condiciones se ajustan para 40 polimerización dirigida por molde. En este ejemplo, los extremos 5' del primer sustrato monomérico (R² de las figuras 60A y 60D) se polimerizan para iniciar la hebra hija emergente. El sustrato se ha sustituido en los extremos 3' (R¹ de las figuras 60A y 60D), para bloquear reversiblemente una mayor extensión. Esto se muestra con más detalle en la figura 61a ampliada (líneas punteadas), donde se muestra el sustrato monomérico colindando con el cebador. Este sustrato es un 5'- trifosfato y se forma un enlace fosfodiéster con el cebador por acción de la polimerasa. Cabe señalar que el grupo funcional reactivo (δ de la figura 60A) que se muestra en la figura 61c típicamente se desactiva (por reacción u otros medios) en el primer sustrato añadido (como se muestra en la figura 61a por el rectángulo).

45 En la etapa I de la figura 61, se retira el grupo de bloqueo (rectángulo libre) para permitir la adición de otro monómero (extensión). Los procedimientos para bloquear reversiblemente el extremo 3' incluyen el uso de Pd para catalizar la eliminación de un grupo alilo para regenerar un extremo hidroxilo 3' viable, o el uso de nucleótidos finalizados en 3'-O-(2-nitrobencilo), donde los hidroxilos activos se pueden regenerar por exposición a una fuente UV para fotoescindir el resto de finalización 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 con más detalle en la vista ampliada de la figura 61b. En esta vista, los extremos colindantes de la construcción de sustrato adyacente y el cebador se polimerizan y el extremo 3'OH de la hebra hija emergente se activa retirando el grupo de bloqueo. En la etapa II, el molde se pone en contacto con una colección de construcciones de sustrato monoméricas, donde los grupos funcionales δ son reactivos en condiciones controladas, y un sustrato monomérico se polimeriza en la hebra hija emergente. Esta polimerización coloca los grupos ε (608) y δ (609) de las construcciones de sustrato colindantes como se muestra en la figura 61c. En la etapa III, en condiciones controladas, estos grupos reaccionan para formar un enlace χ (580), como se muestra con más detalle en la figura 61d.

50 Como se indica, el ciclo a través de las etapas II, III y IV extiende la hija emergente cada vez por sustrato adicional (construcción). Tipicamente, se puede usar un etapa de lavado para retirar los reactivos sin reaccionar entre las etapas. 55 El procedimiento es, por tanto, análogo a lo que se denomina en la literatura "extensión de base única" cíclica. El procedimiento se muestra con polimerasa, P, pero se puede adaptar para una ligasa o un protocolo de ligación química

adecuado para unir construcciones de sustrato en una síntesis dirigida por molde. La etapa V muestra el intermedio de hebra hija para el Xpandómero (605). Este intermedio se puede disociar del molde y el cebador, por ejemplo, con una nucleasa que ataca la cadena principal primaria de la hebra hija, aliviando de este modo las ataduras restringidas y liberando el producto Xpandómero.

5

Al igual que con todos los procedimientos SBE, el lavado eficiente entre ciclos es útil para reducir las reacciones secundarias indeseables. Para facilitar aún más la incorporación de una sola base a través de regiones del molde con estructura secundaria alta, la temperatura de extensión se puede variar a lo largo de cada ciclo de extensión y/o también se pueden añadir aditivos y/o adyuvantes, tales como betaina, TMACL o PEG, para neutralizar los efectos de la estructura secundaria (como se hace en los protocolos de extensión con polimerasa convencionales). Y si es necesario, la estequiométría de las especies de las construcciones de sustrato se puede variar para compensar el sesgo de reacción que favorece determinadas bases, como C o G.

10

Un procedimiento alternativo para producir Xpandómeros de clase VI es hacer polimerización procesiva basada en polimerasa. Se pueden considerar para este enfoque ADN y ARN polimerasas, así como cualquier enzima de funcionamiento similar que se demuestre que cataliza una polimerización exacta de RT-NTP, en ausencia de grupos R de bloqueo terminal reversible.

15

Una amplia gama de químicas de reticulación son conocidas en este campo, y son útiles para la formación de enlaces χ . Estos incluyen el uso de ésteres de NHS con aminas, maleimidas con sulfhidrilos, imidoésteres con aminas, EDC con sulfhidrilos y carboxilos para reacciones con aminas, disulfuros de piridilo con sulfhidrilos, etc. Otros implican el uso de grupos funcionales como hidrazida (HZ) y 4-formilbenzoato (4FB) que, a continuación, pueden reaccionar para formar subunidades de enlace. En una opción, dos químicas de enlace diferentes, $\varepsilon 1/\delta 1$ y $\varepsilon 2/\delta 2$ (también denominadas L1/L1' y L2/L2' en otras partes del presente documento) que reaccionan para formar enlaces $\chi 1$ y $\chi 2$, respectivamente, y se pueden usar para funcionalizar diferencialmente dos conjuntos de construcciones de sustrato de RT-NTP. Por ejemplo, si un ciclo de SBE se realiza con un conjunto de RT-NTP funcionalizado con $\varepsilon 1/\delta 2$, el siguiente ciclo de SBE usaría el conjunto $\varepsilon 2/\delta 1$, lo que da como resultado una reacción del par $\delta 2/\varepsilon 2$ colocado para formar $\chi 1$. La activación ordenada de los pares de reticulación es útil para minimizar los errores de reticulación y las reacciones secundarias no deseadas.

20

30 Construcciones monoméricas de clase VII

Las moléculas de clase VII son análogas a las de clase VI descritas previamente. La principal diferencia es que el enlace escindible es entre la atadura y el sustrato en lugar de entre los sustratos. En la figura 62, las construcciones de sustrato monoméricas de clase VII (un tipo de RT-NTP) se divultan con más detalle. Las figuras 62A a 62C se leen de izquierda a derecha, mostrándose primero la construcción de sustrato monoméricas (precursor de Xpandómero que tiene un residuo de nucleobase), a continuación la hebra hija híbrida intermedia en el centro y, a la derecha, el producto Xpandómero preparado para secuenciación.

35

Como se muestra en la figura 62A, las construcciones de sustrato monomérica de Clase VII tienen una atadura, T (620), unida por un conector escindible selectivamente (621) a un primer resto de extremo a un residuo de nucleobase de sustrato, N. Otro grupo conector, ε (622), se dispone en el primer resto de extremo de la atadura próximo a R¹. En el extremo distal de la atadura T, un segundo resto de extremo con un segundo grupo conector, δ (623), se sitúa próximo a R². El segundo resto de extremo de la atadura se fija al primer resto de extremo próximo a la nucleobase por un reticulación escindible selectivamente intraatadura (624) (o por otra restricción). La reticulación escindible intraatadura se indica aquí mediante una línea punteada, que puede indicar, por ejemplo, un enlace disulfuro o un conector fotoescindible. Esta restricción evita que la atadura se alargue o se expanda y está en una "configuración restringida". En el ensamblaje dirigido por molde, los sustratos forman un híbrido con el molde diana, de modo que los sustratos son colindantes. En condiciones controladas, los grupos conectores colocados δ y ε de los sustratos colindantes se unen para formar un enlace χ entre las construcciones de sustrato contiguas. Los grupos conectores δ y ε de una construcción de sustrato monomérica no forman un enlace intrasustrato debido a restricciones de posicionamiento. Las químicas de enlace y protección/desprotección adecuadas para δ , ε y χ se detallan en la descripción general de la construcción monomérica.

40

R¹ y R² son grupos de extremo configurados como es adecuado para el protocolo de síntesis en el que se usa la construcción de sustrato. Por ejemplo, R¹ = 5'-fosfato y R² = 3'-OH encontrarían uso en un protocolo de ligación, y R¹ = 5'-trifosfato y R² = 3'-OH para un protocolo de polimerasa. Opcionalmente, R² se puede configurar con un grupo de bloqueo reversible para la adición cíclica de un solo sustrato. De forma alternativa, R¹ y R² se pueden configurar con grupos de extremo conectores para acoplamiento químico o sin grupos conectores para un protocolo solo de hibridación. R¹ y R² pueden ser del tipo general XR, en el que X es un grupo de enlace y R es un grupo funcional.

45

R¹ y R² son grupos de extremo configurados como es adecuado para el protocolo de síntesis en el que se usa la construcción de sustrato. Por ejemplo, R¹ = 5'-fosfato y R² = 3'-OH encontrarían uso en un protocolo de ligación, y R¹ = 5'-trifosfato y R² = 3'-OH para un protocolo de polimerasa. Opcionalmente, R² se puede configurar con un grupo de bloqueo reversible para la adición cíclica de un solo sustrato. De forma alternativa, R¹ y R² se pueden configurar con grupos de extremo conectores para acoplamiento químico o sin grupos conectores para un protocolo solo de hibridación. R¹ y R² pueden ser del tipo general XR, en el que X es un grupo de enlace y R es un grupo funcional.

50

Durante el ensamblaje, la construcción de sustrato monomérica se polimeriza primero en el extremo extensible de la hebra hija emergente por un procedimiento de polimerización dirigida por molde que usa un molde monocatenario como guía. En general, este procedimiento se inicia a partir de un cebador y prosigue en la dirección 5' a 3'. En general, se usa una ADN polimerasa u otra polimerasa para formar la cadena hija, y se seleccionan las condiciones de modo que se obtenga una copia complementaria de la cadena molde. Posteriormente, el grupo conector δ , que ahora está colocado con el grupo conector ε de la atadura de la subunidad contigua, se reticula y forma un enlace χ , que es un enlace entre ataduras y entre subunidades. Los enlaces χ unen las ataduras en una cadena continua, formando un intermedio denominado "hebra

"hebra hija híbrida", como se muestra en la figura 62B. Después de que se forma el enlace χ , el enlace intraatadura se puede romper.

5 La hebra hija híbrida (figura 62B) es un heteropolímero con las subunidades que se muestran entre paréntesis. La cadena principal primaria ($\sim N \sim$) k , la hebra molde ($-N'$) k y la atadura (T) se muestran como una hebra hija hibridada, donde k indica una pluralidad de subunidades repetitivas. Cada subunidad de la hebra hija es un "motivo" repetitivo y los motivos tienen una variabilidad específica de especie, indicada en el presente documento por el superíndice α . La hebra hija se forma a partir de especies de construcción de sustrato monomérica seleccionadas mediante un procedimiento dirigido por molde de una colección de especies de motivo, uniéndose el sustrato monomérico de cada especie de construcción de sustrato a un nucleótido complementario correspondiente en la hebra molde diana. De esta manera, la secuencia de residuos de nucleobase (es decir, la cadena principal primaria) de la hebra hija es una copia complementaria contigua de la hebra molde diana.

10 15 La tilde (\sim) indica un enlace escindible selectivamente que se muestra aquí como el conector de atadura a sustrato. Estos se pueden escindir selectivamente para liberar y expandir las ataduras (y el Xpandómero) sin degradar el propio Xpandómero.

20 25 La hebra hija se compone de un precursor de Xpandómero llamado "Xpandómero restringido" que además se compone de ataduras en la "configuración restringida". Cuando las ataduras se convierten a su "configuración expandida", el Xpandómero restringido se convierte en el producto Xpandómero. Las ataduras se restringen por los enlaces χ entre subunidades, el enlace escindible al sustrato γ , opcionalmente, los enlaces intraatadura si todavía están presentes. El enlace χ une el primer resto de extremo de la atadura de una primera subunidad al segundo resto de extremo de la atadura en el extremo colindante de una segunda subunidad y se forma uniendo los grupos conectores colocados, ε de la primera subunidad y δ de la segunda subunidad.

30 35 40 Se puede observar que la hebra hija tiene dos cadenas principales, un "cadena principal primario" y una "cadena principal del Xpandómero restringido". La cadena principal primaria está compuesta por sustratos monoméricos polymerizados y colindantes. La cadena principal del Xpandómero restringido evita el enlace escindible selectivamente que se conecta al sustrato y está formado por restos de cadena principal unidos por enlaces χ , siendo cada resto de cadena principal una atadura. Se puede observar que la cadena principal del Xpandómero restringido forma un puente sobre los enlaces escindibles selectivamente conectado a la cadena principal primaria, y permanecerá covalentemente intacto cuando estos enlaces escindibles selectivamente se escindan y la cadena principal primaria se disocie o fragmente.

45 50 55 El enlace χ de la atadura (reticulación de los grupos conectores δ y ε) está precedido, en general, por acoplamiento enzimático de los sustratos monoméricos para formar la cadena principal primaria, por ejemplo, con enlaces fosfodiéster entre bases contiguas. En la estructura que se muestra aquí, se ha formado la cadena principal primaria de la hebra hija, y los entre sustratos se representan con una tilde (\sim) para indicar que son escindibles selectivamente. Después de disociar o degradar la hebra molde diana, escindiendo los enlaces escindibles selectivamente (que incluyen los enlaces intraatadura), el Xpandómero restringido se libera y se convierte en el producto Xpandómero. Los procedimientos para la disociación de la hebra molde incluyen la desnaturalización por calor o digestión selectiva con una nucleasa o degradación química. Un procedimiento para la escisión seleccionable usa digestión con nucleasas donde, por ejemplo, los enlaces fosfodiéster de la cadena principal primaria se digieren por una nucleasa y los enlaces de atadura a atadura son resistentes a nucleasas.

60 65 70 La figura 62C es una representación del producto Xpandómero de clase VII despues de la disociación de la hebra molde y despues de la escisión de los enlaces escindibles selectivamente (incluyendo los unidos a la cadena principal primaria γ , si no están ya escindidos, los enlaces intraatadura). La hebra del producto Xpandómero contiene una pluralidad de subunidades k , donde k denota la késima subunidad en una cadena de m subunidades que componen la cadena hija, donde $k = 1, 2, 3$ a m , donde $m > 10$, en general $m > 50$, y típicamente $m > 500$ o > 5000 . Cada subunidad está formada por una atadura (620) en su configuración expandida y se estira en toda su longitud entre los enlaces χ de las subunidades contiguas. La cadena principal primaria se ha retirado por completo. Cada subunidad, un motivo de subunidad α , contiene información genética específica de especie establecida por ensamblaje dirigido por molde del intermedio de Xpandómero (hebra hija).

75 80 85 La figura 62D muestra la construcción de sustrato de la figura 62A como un modelo molecular, donde el miembro de sustrato monomérico, representado con un residuo de nucleobase (626), unido por un enlace escindible selectivamente (625) al primer resto de extremo de la atadura. También dispuesto en el primer resto de extremo está un grupo conector (629), que se muestra como ε en la figura 62A. Un segundo grupo segundo conector (628), que se muestra como δ en la figura 62A, está dispuesto en un segundo resto de extremo en el extremo distal de la atadura. Se muestra un enlace intraatadura escindible selectivamente (627), representado por los triángulos contiguos, que restringe la atadura uniendo el primer y el segundo resto de extremo. Los grupos conectores ε (629) y δ (628) se sitúan para que no interactúen y para que se alineen preferentemente cerca de los lados R^1 y R^2 del sustrato, respectivamente. El bucle de atadura que se muestra aquí tiene tres indicadores (900, 901, 902), que también pueden ser específicos de la especie del motivo.

90 95 La figura 62E muestra la construcción de sustrato despues de la incorporación en el producto Xpandómero. Las subunidades se escinden y expanden y se unen mediante enlaces χ (910), representados como un óvalo abierto, formados

uniendo los grupos conectores δ y ϵ a los que se hace referencia en la figura 62A. Una subunidad se indica mediante líneas punteadas que enmarcan verticalmente la subunidad repetitiva, como se representa por corchetes en la figura 62C adjunta.

- 5 En el producto Xpandómero de la figura 62E, la cadena principal primaria se ha disociado o fragmentado y se separa del Xpandómero. A través del procedimiento de escisión, el Xpandómero restringido se libera para convertirse en el producto Xpandómero. Los miembros de atadura que estaban anteriormente en una configuración restringida, están ahora en una configuración expandida, funcionando de este modo para extender linealmente la información de secuencia de la diana del molde. La expansión de las ataduras reduce la densidad lineal de la información de secuencia a lo largo del Xpandómero y proporciona una plataforma para incrementar el tamaño y la abundancia de los indicadores, lo que a su vez mejora la relación señal-ruido para la detección y descodificación de la secuencia molde.
- 10

Si bien la atadura se representa como una construcción de indicadores con tres grupos indicadores, se pueden organizar diversas configuraciones de indicadores en la atadura y pueden comprender indicadores individuales que identifican el monómero o la atadura puede ser un polímero desnudo. En algunos casos, uno o más precursores indicadores se disponen en la atadura, y los indicadores se unen por afinidad o se unen covalentemente después del ensamblaje del producto Xpandómero.

Construcciones monoméricas de clase VIII

20 Las moléculas de clase VIII son análogas a las de clase VI descritas previamente. La principal diferencia es que el grupo conectador, ϵ , está conectado directamente al sustrato en lugar de a la atadura. En la figura 63, los autores de la invención describen construcciones de sustrato monoméricas de clase VIII (un tipo de RT-NTP) con más detalle. Las figuras 63A a 25 63C se leen de izquierda a derecha, mostrándose primero la construcción de sustrato monoméricas (precursor de Xpandómero que tiene un residuo de nucleobase), a continuación la hebra hija híbrida intermedia en el centro y, a la derecha, el producto Xpandómero preparado para secuenciación.

30 Como se muestra en la figura 63A, las construcciones de sustrato monoméricas de clase VIII tienen una atadura, T (630), unida por un enlace (631) de un primer resto de extremo a un residuo de nucleobase de sustrato, N. En el extremo distal de la atadura (632), un segundo resto de extremo con un segundo grupo conectador, δ , se coloca preferentemente próximo a R². El segundo resto de extremo de la atadura se fija al primer resto de extremo próximo a la nucleobase por un reticulación escindible selectivamente intraatadura (o por otra restricción). La reticulación escindible intraatadura (633) se indica aquí mediante una línea punteada, que puede indicar, por ejemplo, un enlace disulfuro o un conector fotoescindible. Esta restricción evita que la atadura se alargue o expanda y se dice que está en su "configuración restringida". Un grupo conectador, ϵ (635), se une al sustrato monomérico preferentemente próximo a R¹. En el ensamblaje dirigido por molde, los sustratos forman un híbrido con el molde diana, de modo que los sustratos son colindantes. En condiciones controladas, se colocan los grupos conectores δ y ϵ de los sustratos colindantes y se unen para formar un enlace X entre las construcciones de sustrato contiguas. Los grupos conectores δ y ϵ de una construcción de sustrato monomérica no forman un enlace intrasustrato debido a restricciones de posicionamiento. Las químicas de enlace y protección/desprotección adecuadas para δ , ϵ y X se detallan en la descripción general de la construcción monomérica.

40 R¹ y R² son grupos de extremo configurados como es adecuado para el protocolo de síntesis en el que se usa la construcción de sustrato. Por ejemplo, R¹ = 5'-fosfato y R² = 3'-OH encontrarían uso en un protocolo de ligación, y R¹ = 5'-trifosfato y R² = 3'-OH para un protocolo de polimerasa. Opcionalmente, R² se puede configurar con un grupo de bloqueo reversible para la adición cíclica de un solo sustrato. De forma alternativa, R¹ y R² se pueden configurar con grupos de extremo conectores para acoplamiento químico o sin grupos conectores para un protocolo solo de hibridación. R¹ y R² pueden ser del tipo general XR, en el que X es un grupo de enlace y R es un grupo funcional.

50 Durante el ensamblaje, la construcción de sustrato monomérica se polimeriza primero en el extremo extensible de la hebra hija emergente por un procedimiento de polimerización dirigida por molde que usa un molde monocatenario como guía. En general, este procedimiento se inicia a partir de un cebador y prosigue en la dirección 5' a 3'. En general, se usa una ADN polimerasa u otra polimerasa para formar la cadena hija, y se seleccionan las condiciones de modo que se obtenga una copia complementaria de la cadena molde. Posteriormente, el grupo conectador δ , que ahora está colocado con el grupo conectador ϵ de la atadura de la subunidad contigua, se reticula y forma un enlace X, que es un enlace entre subunidades. 55 Los enlaces X proporcionan un segundo enlace entre las subunidades (los enlaces entre sustratos polimerizados son los primeros) y forman un intermedio denominado "cadena hija híbrida", como se muestra en la figura 63B.

60 La hebra hija híbrida (figura 63B) es un heteropolímero con las subunidades que se muestran entre paréntesis. La cadena principal primaria (~N~)_k, la hebra molde (-N'-)_k y la atadura (T) se muestran como una hebra hija híbrida, donde k indica una pluralidad de subunidades repetitivas. Cada subunidad de la hebra hija es un "motivo" repetitivo y los motivos tienen una variabilidad específica de especie, indicada en el presente documento por el superíndice α . La hebra hija se forma a partir de especies de construcción de sustrato monomérica seleccionadas mediante un procedimiento dirigido por molde de una colección de especies de motivo, uniéndose el sustrato monomérico de cada especie de construcción de sustrato a un nucleótido complementario correspondiente en la hebra molde diana. De esta manera, la secuencia de residuos de nucleobase (es decir, la cadena principal primaria) de la hebra hija es una copia complementaria contigua de la hebra molde diana.

Cada tilde (~) indica un enlace escindible selectivamente que se muestra aquí como enlaces entre sustratos. Estos son necesariamente escindibles selectivamente para liberar y expandir las ataduras (y el Xpandómero) sin degradar el propio Xpandómero.

5 La hebra hija se compone de un precursor de Xpandómero llamado "Xpandómero restringido" que además se compone de ataduras en la "configuración restringida". Cuando las ataduras se convierten a su "configuración expandida", el Xpandómero restringido se convierte en el producto Xpandómero. Las ataduras se restringen por los enlaces χ entre subunidades, las uniones al sustrato y, opcionalmente, los enlaces intraatadura si todavía están presentes. El enlace χ une el sustrato de una primera subunidad a la atadura del segundo resto de extremo en el extremo colindante de una segunda subunidad y se forma uniendo los grupos conectores colocados, ϵ de la primera subunidad y δ de la segunda subunidad.

10 15 Se puede observar que la hebra hija tiene dos cadenas principales, una "cadena principal primaria" y la cadena principal del "Xpandómero restringido". La cadena principal primaria está compuesta por sustratos monoméricos polimerizados y colindantes. La "cadena principal del Xpandómero restringido" evita el enlace escindible selectivamente entre los sustratos monoméricos y se forma por restos de cadena principal unidos por enlaces χ , siendo cada resto de cadena principal una atadura unida a un sustrato que a continuación se une al siguiente resto de cadena principal con un enlace χ . Se puede observar que la cadena principal del Xpandómero restringido forma un puente sobre los enlaces escindibles 20 selectivamente de la cadena principal primaria, y permanecerá covalentemente intacto cuando estos enlaces escindibles selectivamente se escindan y la cadena principal primaria se fragmente.

25 30 El enlace χ de la atadura (reticulación de los grupos conectores δ y ϵ) está precedido, en general, por acoplamiento enzimático de los sustratos monoméricos para formar la cadena principal primaria, por ejemplo, con enlaces fosfodiéster entre bases contiguas. En la estructura que se muestra aquí, se ha formado la cadena principal primaria de la hebra hija, y los enlaces entre sustratos se representan con una tilde (~) para indicar que son escindibles selectivamente. Después de disociar o degradar la hebra molde diana, escindiendo los enlaces escindibles selectivamente (que incluyen los enlaces intraatadura), el Xpandómero restringido se libera y se convierte en el producto Xpandómero. Los procedimientos para la 35 40 disociación de la hebra molde incluyen la desnaturalización por calor o digestión selectiva con una nucleasa o degradación química. Un procedimiento para la escisión seleccionable usa digestión con nucleasas donde, por ejemplo, los enlaces fosfodiéster de la cadena principal primaria se digieren por una nucleasa y los enlaces de atadura a atadura son resistentes a nucleasas.

45 50 La figura 63C es una representación del producto Xpandómero de clase VIII después de la disociación de la hebra molde y después de la escisión de los enlaces escindibles selectivamente (incluyendo los de la cadena principal primaria y, si no están ya escindidos, los enlaces intraatadura). La hebra del producto Xpandómero contiene una pluralidad de subunidades κ , donde κ denota la $\kappa^{\text{ésima}}$ subunidad en una cadena de m subunidades que componen la cadena hija, donde $\kappa = 1, 2, 3$ a m , donde $m > 10$, en general $m > 50$, y típicamente $m > 500$ o > 5000 . Cada subunidad está formada por una atadura en su configuración expandida y se estira en toda su longitud entre los enlaces χ de las subunidades contiguas. El sustrato colgante está unido a la atadura en cada subunidad. Cada subunidad, un motivo de subunidad α , contiene información genética específica de especie establecida por ensamblaje dirigido por molde del intermedio de Xpandómero (hebra hija).

55 La figura 63D muestra la construcción de sustrato de la figura 63A como un modelo molecular, donde el miembro de sustrato monomérico, representado con un residuo de nucleobase (634), se une a la atadura mediante un enlace (631) del primer resto de extremo de la atadura. Un segundo grupo segundo conector (639), que se muestra como δ en la figura 63A, está dispuesto en un segundo resto de extremo en el extremo distal de la atadura. Se muestra un enlace intraatadura escindible selectivamente (633), representado por los triángulos contiguos, que restringe la atadura uniendo el primer y el segundo resto de extremo. También unido al sustrato hay un grupo conector (638), que se muestra como ϵ en la figura 63A. Los grupos conectores ϵ (638) y δ (639) se sitúan para que no interactúen y para que se alineen preferentemente cerca de los lados R^1 y R^2 del sustrato respectivamente. El bucle de atadura que se muestra aquí tiene tres indicadores (900, 901, 902), que también pueden ser específicos de la especie del motivo. En la figura 63D y la figura 63E se muestra la reticulación intraatadura escindible (633) que se va a situar más cerca del sustrato que el δ (639). En las figuras 60D y 60E, las posiciones están intercambiadas. Este posicionamiento puede estar en cualquier lugar y ambas funciones de conector pueden estar en un solo grupo multifuncional.

60 65 La figura 63E muestra la construcción de sustrato después de la incorporación en el producto Xpandómero. Las subunidades se escinden y expanden y se unen mediante enlaces χ (970.971), representados como un óvalo abierto, formados uniendo los grupos conectores δ y ϵ a los que se hace referencia en la figura 63A. Una subunidad se indica mediante líneas punteadas que encierran verticalmente la subunidad repetitiva, como se representa por corchetes en la figura 63C adjunta.

En el producto Xpandómero de la figura 63E, la cadena principal primaria se ha fragmentado y no es covalentemente contiguo porque se ha escindido cualquier enlace directo entre los sustratos de subunidades contiguas. A través del procedimiento de escisión, el Xpandómero restringido se libera para convertirse en el producto Xpandómero. Los miembros de atadura que estaban anteriormente en una configuración restringida, están ahora en una configuración expandida, funcionando de este modo para extender linealmente la información de secuencia de la diana del molde. La

expansión de las ataduras reduce la densidad lineal de la información de secuencia a lo largo del Xpandómero y proporciona una plataforma para incrementar el tamaño y la abundancia de los indicadores, lo que a su vez mejora la relación señal-ruido para la detección y descodificación de la secuencia molde.

- 5 Si bien la atadura (630) se representa como una construcción de indicadores con tres grupos indicadores, se pueden organizar diversas configuraciones de indicadores en la atadura y pueden comprender indicadores individuales que identifican el monómero o la atadura puede ser un polímero desnudo. En algunos casos, uno o más precursores indicadores se disponen en la atadura, y los indicadores se unen por afinidad o se unen covalentemente después del ensamblaje del producto Xpandómero.

10

Construcciones monoméricas de clase IX

Una construcción de sustrato de clase IX se distingue de los otros RT-NTP porque tiene dos puntos de unión de ataduras a los que se une una atadura libre después de ensamblar la cadena principal primaria. La figura 64 describe construcciones de sustrato monoméricas de clase IX (un tipo de RT-NTP) con más detalle. Las figuras 64A a 64C se leen de izquierda a derecha, mostrándose primero la construcción de sustrato monoméricas (precursor de Xpandómero que tiene un residuo de nucleobase), a continuación la hebra hija híbrida intermedia en el centro y, a la derecha, el producto Xpandómero preparado para secuenciación.

- 20 Como se muestra en la figura 64A, la construcción de sustrato monomérica de clase IX tiene un residuo de nucleobase de sustrato, N, con dos sitios de unión, los grupos conectores δ_1 y δ_2 . También se muestra una atadura libre, T (640), con grupos conectores ϵ_1 y ϵ_2 de un primer y segundo resto de extremo de la atadura. Los restos de extremo primero y segundo de la atadura restringen la atadura libre por una reticulación selectivamente escindible intraatadura (647) y sirven para colocar los grupos conectores ϵ_1 y ϵ_2 . La reticulación escindible intraatadura se indica aquí mediante una línea punteada,

25 puede indicar, por ejemplo, un enlace disulfuro o un conector fotoescindible. Esta restricción evita que la atadura se alargue o expanda y se dice que está en su "configuración restringida". Los grupos conectores, δ_1 y δ_2 , se unen al sustrato monomérico orientados próximos a R¹ y R², respectivamente. En la síntesis dirigida por molde, los sustratos forman un híbrido con el molde diana y el grupo conector de una construcción de sustrato, por ejemplo δ_1 , y se colocan el grupo conector de la construcción de sustrato colindante, por ejemplo δ_2 . En condiciones controladas, estos conectores

30

colocados ponen en contacto los conectores ϵ colocados en el extremo de la atadura libre. Se producen dos reacciones de enlaces selectivas, ϵ_1 y δ , para formar X¹ y ϵ_2 y δ_2 para formar X²; los sustratos contiguos ahora se une formando un puente por la atadura restringida. Las químicas de enlace y protección/desprotección adecuadas para δ , ϵ y X se detallan en la descripción general de la construcción monomérica.

- 35 R¹ y R² son grupos de extremo configurados como es adecuado para el protocolo de síntesis en el que se usa la construcción de sustrato. Por ejemplo, R¹ = 5'-fosfato y R² = 3'-OH encontrarán uso en un protocolo de ligación, y R¹ = 5'-trifosfato y R² = 3'-OH para un protocolo de polimerasa. Opcionalmente, R² se puede configurar con un grupo de bloqueo reversible para la adición ciclica de un solo sustrato. De forma alternativa, R¹ y R² se pueden configurar con grupos de extremo conectores para acoplamiento químico o sin grupos conectores para un protocolo solo de hibridación. R¹ y R²

40

pueden ser del tipo general XR, en el que X es un grupo de enlace y R es un grupo funcional.

Durante el ensamblaje, la construcción de sustrato monomérica (sin atadura) se polimeriza en el extremo extensible de la hebra hija emergente por un procedimiento de polimerización dirigida por molde que usa un molde monocatenario como guía. En general, este procedimiento se inicia a partir de un cebador y prosigue en la dirección 5' a 3'. En general, se usa una ADN polimerasa u otra polimerasa para formar la cadena hija, y se seleccionan las condiciones de modo que se obtenga una copia complementaria de la cadena molde. El grupo conector δ_1 se coloca ahora con el grupo conector δ_2 del sustrato monomérico contiguo. Después de la polimerización de la cadena principal primaria, las ataduras libres se reticulan para formar enlaces X¹ y enlaces X² entre los dos extremos de la atadura y dos sustratos contiguos.

50

En un aspecto, las ataduras libres no tienen información de secuencia y se denominan "desnudas". En este caso, se puede usar una química de un solo conector para que los grupos conectores δ_1 y δ_2 sean los mismos y los grupos conectores ϵ_1 y ϵ_2 sean los mismos con un tipo de enlace X.

55

En el aspecto en el que la atadura libre comprende información del tipo de base (como en los indicadores) hay ataduras libres específicas de la especie. En general, hay cuatro tipos de base que requerirían cuatro tipos de ataduras libres con la información de la base correspondiente. Se pueden usar varios procedimientos para unir la especie de atadura libre correcta a su base correcta. En un procedimiento, se usan cuatro químicas de enlace heteroespecíficas para diferenciar aún más el par conector δ_1 y ϵ_1 , ahora expresado como δ_{1a} y ϵ_{1a} , donde a es uno de los cuatro tipos y forma tipos de enlace X^{1a}. δ_{1a} y ϵ_{1a} de un tipo a se unirán solo entre sí. En este procedimiento, se provoca que el par conector δ_2 y ϵ_2 se

60

unian solo después de que se hayan formado δ_{1a} y ϵ_{1a} . En un segundo procedimiento, se usan diferentes grupos de protección desprotegibles selectivamente, donde cada grupo de protección se asocia con un tipo de conector, para bloquear selectivamente grupos conectores. En un primer ciclo, no hay protección en un tipo de base y su tipo de atadura libre asociado está unido ϵ_1 a δ_1 para formar un enlace X¹. En un segundo ciclo, se retira un tipo de grupo de protección de un tipo de base y su tipo de atadura libre asociada se une ϵ_1 a δ_1 para formar un enlace X¹. Este último ciclo se repite para los siguientes dos tipos de base y, una vez completado, se provoca que se una el par conector δ_2 y ϵ_2 . Entre cada etapa se incluyen etapas de lavado para reducir los errores de unión. Sin pérdida de generalidad, la descripción restante

se hará en términos de enlaces χ^1 y enlaces χ^2 . Hay que tener en cuenta que, después de que se une la atadura y se forman los enlaces χ , es posible que se rompa el enlace intraatadura.

Como se muestra en la figura 64B, los enlaces χ^1 proporcionan un enlace entre las subunidades (además de los enlaces entre sustratos polimerizados) y forman un intermedio denominado "cadena hija híbrida". La cadena principal primaria ($\sim N \sim \kappa$), la hebra molde ($-N' \sim \kappa$) y la atadura (T) se muestran como una hebra hija hibridada, donde κ indica una pluralidad de subunidades repetitivas. Cada subunidad de la hebra hija es un "motivo" repetitivo y los motivos tienen una variabilidad específica de especie, indicada en el presente documento por el superíndice α . La hebra hija se forma a partir de especies de construcción de sustrato monomérica seleccionadas mediante un procedimiento dirigido por molde de una colección de especies de motivo, uniéndose el sustrato monomérico de cada especie de construcción de sustrato a un nucleótido complementario correspondiente en la hebra molde diana. De esta manera, la secuencia de residuos de nucleobase (es decir, la cadena principal primaria) de la hebra hija es una copia complementaria contigua de la hebra molde diana.

Cada tilde (\sim) indica un enlace escindible selectivamente que se muestra aquí como enlaces entre sustratos. Estos son necesariamente escindibles selectivamente para liberar y expandir las ataduras (y el Xpandómero) sin degradar el propio Xpandómero.

La hebra hija se compone de un precursor de Xpandómero llamado "Xpandómero restringido" que además se compone de ataduras en la "configuración restringida". Cuando las ataduras se convierten a su "configuración expandida", el Xpandómero restringido se convierte en el producto Xpandómero. Las ataduras se restringen por los enlaces χ a los sustratos contiguos y, opcionalmente, los enlaces intraatadura (si todavía están presentes).

Se puede observar que la hebra hija tiene dos cadenas principales, una "cadena principal primaria" y la cadena principal del "Xpandómero restringido". La cadena principal primaria está compuesta por sustratos monoméricos polimerizados y colindantes. La "cadena principal del Xpandómero restringido" evita el enlace escindible selectivamente entre los sustratos monoméricos y se forma por restos de cadena principal unidos por enlaces χ^1 , siendo cada resto de cadena principal una atadura unida a un sustrato (por un enlace χ^2 que a continuación se une al siguiente resto de cadena principal con un enlace χ^1). Se puede observar que la cadena principal del Xpandómero restringido forma un puente sobre los enlaces escindibles selectivamente de la cadena principal primaria, y permanecerá covalentemente intacto cuando estos enlaces escindibles selectivamente se escinden y la cadena principal primaria se fragmente.

El enlace χ de la atadura (reticulación de los grupos conectores δ y ϵ) está precedido, en general, por acoplamiento enzimático de los sustratos monoméricos para formar la cadena principal primaria, por ejemplo, con enlaces fosfodiéster entre bases contiguas. En la estructura que se muestra aquí, se ha formado la cadena principal primaria de la hebra hija, y los enlaces entre sustratos se representan con una tilde (\sim) para indicar que son escindibles selectivamente. Después de disociar o degradar la hebra molde diana, escindiendo los enlaces escindibles selectivamente (que incluyen los enlaces intraatadura), el Xpandómero restringido se libera y se convierte en el producto Xpandómero. Los procedimientos para la disociación de la hebra molde incluyen la desnaturalización por calor o digestión selectiva con una nucleasa o degradación química. Un procedimiento para la escisión seleccionable usa digestión con nucleasas donde, por ejemplo, los enlaces fosfodiéster de la cadena principal primaria se digieren por una nucleasa y los enlaces de atadura a atadura son resistentes a nucleasas.

La figura 64C es una representación del producto Xpandómero de clase IX después de la disociación de la hebra molde y después de la escisión de los enlaces escindibles selectivamente (incluyendo los de la cadena principal primaria y, si no están ya escindidos, los enlaces intraatadura). La hebra del producto Xpandómero contiene una pluralidad de subunidades κ , donde κ denota la $\kappa^{\text{ésima}}$ subunidad en una cadena de m subunidades que componen la cadena hija, donde $\kappa = 1, 2, 3$ a m , donde $m > 10$, en general $m > 50$, y típicamente $m > 500$ o > 5000 . Cada subunidad está formada por una atadura (640) en su configuración expandida unida a un sustrato monomérico y al enlace χ^1 con la siguiente subunidad contigua. Cada subunidad, un motivo de subunidad α , contiene información genética específica de especie establecida por ensamblaje dirigido por molde del intermedio de Xpandómero (hebra hija).

La figura 64D muestra la construcción de sustrato de la figura 64A como un modelo molecular, donde el miembro del sustrato monomérico, representado con un residuo de nucleobase (641), tiene dos enlaces (642, 643), que se muestran en la figura 64a como δ_1 y δ_2 , que serán puntos de unión para la atadura libre. La atadura libre se muestra con dos grupos conectores (644, 645), que se muestran como ϵ_1 y ϵ_2 en la figura 64A, en los restos de extremo primero y segundo de la atadura. Se muestra un enlace intraatadura escindible selectivamente (647) que restringe la atadura uniendo los restos de extremo primero y segundo. Los grupos conectores se sitúan para favorecer la reticulación de los extremos de unión entre los sustratos monomérico y evitar la reticulación por todo un sustrato monomérico.

El bucle de atadura que se muestra aquí tiene tres indicadores (900.901.902), que también pueden ser específicos de la especie del motivo, pero requiere un procedimiento para que se una correctamente a las bases correctas en la cadena principal primaria.

La figura 64E muestra la construcción de sustrato después de la incorporación en el producto Xpandómero. Las subunidades se escinden y expanden y están unidas por enlaces χ^1 entre ataduras (930, 931) y enlaces χ^2 entre ataduras (932, 933). Cada subunidad es una atadura unida a un sustrato monomérico y conectada al siguiente enlace χ^1 . Una

subunidad se indica mediante líneas punteadas que enmarcan verticalmente la subunidad repetitiva, como se representa por corchetes en la figura 64C adjunta.

En el producto Xpandómero de la figura 64E, la cadena principal primaria se ha fragmentado y no es covalentemente contiguo porque se ha escindido cualquier enlace directo entre los sustratos de subunidades contiguas. A través del procedimiento de escisión, el Xpandómero restringido se libera para convertirse en el producto Xpandómero. Los miembros de atadura que estaban anteriormente en una configuración restringida, están ahora en una configuración expandida, funcionando de este modo para extender linealmente la información de secuencia de la diana del molde. La expansión de las ataduras reduce la densidad lineal de la información de secuencia a lo largo del Xpandómero y proporciona una plataforma para incrementar el tamaño y la abundancia de los indicadores, lo que a su vez mejora la relación señal-ruido para la detección y descodificación de la secuencia molde.

Si bien la atadura se representa como una construcción de indicadores con tres grupos indicadores, se pueden organizar diversas configuraciones de indicadores en la atadura y pueden comprender indicadores individuales que identifican el monómero o la atadura puede ser un polímero desnudo. En algunos casos, uno o más precursores indicadores se disponen en la atadura, y los indicadores se unen por afinidad o se unen covalentemente después del ensamblaje del producto Xpandómero.

La figura 65 demuestra un procedimiento de síntesis de Xpandómeros usando construcciones de sustrato de 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 mediante la síntesis dirigida por molde de una hebra hija. Este procedimiento continúa en la etapa II, y en la figura 65a se muestra una vista ampliada (flecha punteada) de la hebra hija. Se ilustran la hebra molde, el cebador y las construcciones de sustrato de nucleobase de clase IX polimerizadas (sin ataduras), cada construcción de sustrato con funcionalidades químicas, representadas como símbolos de candado y llave, para la adición química de reactivos de atadura. Las funcionalidades se seleccionan de modo que cada monómero tenga un sitio de unión de atadura específico de base y un sitio de unión universal. En la etapa III, se introducen ataduras de horquilla con cuatro conectores específicos de especie. Las ataduras se unen a la cadena principal primaria de una manera específica de base de acuerdo con los conectores específicos de base como se muestra en la figura 65b, una vista ampliada (flecha punteada). En este caso, los círculos blancos y negros indican la química de unión química universal y las formas de rombo y tenedor indican las químicas de unión específicas de base. En la etapa IV y la figura 65c, una vista ampliada (flecha punteada), una vez completada toda la unión específica de base, se provoca que los conectores universales formen enlaces. Hay que tener en cuenta que los conectores universales tanto en las ataduras como en la cadena principal primaria están en estrecha proximidad con su homólogo de conector específico de base para evitar errores de enlace. La unión química de los reactivos de atadura se muestra como que se ha completado y se ha formado un Xpandómero restringido en el molde. A continuación, el Xpandómero se libera (no se muestra) disociándose del molde y escindiendo los enlaces escindibles selectivamente (de la cadena principal primaria y los enlaces intraatadura).

Construcciones monoméricas de clase X

Las construcciones de sustrato de clase X, también llamadas XNTP, se diferencian de los RT-NTP en que la atadura está contenida dentro de cada sustrato monomérico para formar una atadura intrasustrato, cada sustrato XNTP tiene un enlace escindible selectivamente dentro del sustrato que, una vez escindido, permite la expansión de la atadura restringida. En la figura 66, se describen las construcciones de sustrato monoméricas de clase X con más detalle. Las figuras 66A a 66C se leen de izquierda a derecha, mostrándose primero la construcción de sustrato monoméricas (preursor de Xpandómero que tiene un residuo de nucleobase), a continuación la hebra hija híbrida intermedia en el centro y, a la derecha, el producto Xpandómero preparado para secuenciación.

Como se muestra en la figura 66A, la construcción de sustrato monomérica de clase X tiene un residuo de nucleobase de sustrato, N, que tiene dos restos (662, 663) separados por un enlace escindible selectivamente (665), uniéndose cada resto a un extremo de una atadura (660). Los extremos de la atadura se pueden unir a las modificaciones del grupo conector en el heterociclo, el grupo ribosa o la cadena principal de fosfato. El sustrato monomérico también tiene un sitio de escisión intrasustrato situado dentro de la cadena principal de fosfororibosilo de modo que la escisión proporcionará la expansión de la atadura restringida. Por ejemplo, para sintetizar un monómero de ATP de clase X, el conector amino en 8-[(6-Amino)hexil]-amino-ATP o N6-(6-Amino)hexil-ATP se puede usar como primer punto de unión de la atadura, y un conector de estructura mixta, tal como la modificación sin puente (N-1-aminoalquil) fosforamidato o (2-aminoethyl) fosfonato, se puede usar como un segundo punto de unión de la atadura. Además, se puede usar una modificación de la cadena principal puente tal como un fosforamidato (3' OPN5') o un fosforotiolato (3' OPS5'), por ejemplo, para la escisión química selectiva de la cadena principal primaria.

R¹ y R² son grupos de extremo configurados como es adecuado para el protocolo de síntesis en el que se usa la construcción de sustrato. Por ejemplo, R¹ = 5'- trifosfato y R² = 3'-OH para un protocolo de polimerasa. El R¹ 5' trifosfato puede incluir modificaciones mixtas de la cadena principal, como un de aminoethyl fosfonato o un de 3'-OPS-5' fosforotiolato, para permitir el enlace de la atadura y la escisión de la cadena principal, respectivamente. Opcionalmente, R² se puede configurar con un grupo de bloqueo reversible para la adición cíclica de un solo sustrato. De forma alternativa, R¹ y R² se pueden configurar con grupos de extremo conectores para acoplamiento químico. R¹ y R² pueden ser del tipo general XR, en el que X es un grupo de enlace y R es un grupo funcional.

5 Durante el ensamblaje, la construcción de sustrato monomérica se polimeriza en el extremo extensible de la hebra hija emergente por un procedimiento de polimerización dirigida por molde que usa un molde monocatenario como guía. En general, este procedimiento se inicia a partir de un cebador y prosigue en la dirección 5' a 3'. En general, se usa una ADN polimerasa u otra polimerasa para formar la cadena hija, y se seleccionan las condiciones de modo que se obtenga una copia complementaria de la cadena molde.

10 Como se muestra en la figura 66B, los residuos de nucleobase se polimerizan de una subunidad a la siguiente y forman un intermedio denominado "cadena hija híbrida". La cadena principal primaria ($\sim N\sim k$), la hebra molde ($\sim N'\sim k$) y la atadura (T) se muestran como una hebra hija hibridada, donde k indica una pluralidad de subunidades repetitivas. Cada subunidad de la hebra hija es un "motivo" repetitivo y los motivos tienen una variabilidad específica de especie, indicada en el 15 presente documento por el superíndice α . La hebra hija se forma a partir de especies de construcción de sustrato monomérica seleccionadas mediante un procedimiento dirigido por molde de una colección de especies de motivo, uniéndose el sustrato monomérico de cada especie de construcción de sustrato a un nucleótido complementario correspondiente en la hebra molde diana. De esta manera, la secuencia de residuos de nucleobase (es decir, la cadena principal primaria) de la hebra hija es una copia complementaria contigua de la hebra molde diana.

20 La ("V") que se muestra en la figura 66B por encima del residuo de la nucleobase indica un enlace escindible selectivamente que divide el sustrato en los restos primero y segundo. Tras la escisión, el primer resto (669) de una subunidad permanecerá unido al segundo resto (668) de la subunidad contigua y, dentro de una subunidad, cada resto formará un puente, el uno con el otro mediante la atadura. Estos son necesariamente escindibles selectivamente para liberar y expandir las ataduras (y el Xpandómero) sin degradar el propio Xpandómero.

25 25 La hebra hija tiene dos cadena principales, una "cadena principal primaria" y la cadena principal del "Xpandómero restringido". La cadena principal primaria está compuesta por sustratos monoméricos polimerizados y colindantes. La "cadena principal del Xpandómero restringido" evita el enlace escindible selectivamente dentro del sustrato monomérico y está formada por los enlaces entre sustratos que unen los restos de cadena principal, siendo cada resto de cadena principal una atadura unida a dos restos del residuo de nucleobase todavía intacto. La cadena principal del Xpandómero restringido forma un puente sobre el enlace escindible selectivamente dentro de cada monómero, y permanecerá 30 covalentemente intacta cuando estos enlaces escindibles selectivamente se escinden y los monómeros se fragmenten en las porciones n^1 y n^2 que se muestran en la figura 66C.

35 35 La escisión está precedida por acoplamiento enzimático de los sustratos monoméricos para formar la cadena principal primaria, por ejemplo, con enlaces fosfodiéster o la cadena principal mixta entre bases contiguas. En la estructura que se muestra aquí, se ha formado la cadena principal primaria de la hebra hija. Después de disociar o degradar la hebra molde diana y escindir los enlaces escindibles selectivamente, el Xpandómero restringido se libera y se convierte en el producto Xpandómero. Los procedimientos para la disociación de la hebra molde incluyen, por ejemplo, desnaturización por calor.

40 40 La figura 66C es una representación del producto Xpandómero de clase X después de la disociación de la hebra molde y después de la escisión de los enlaces escindibles selectivamente. La hebra del producto Xpandómero contiene una pluralidad de subunidades k , donde k denota la $k^{\text{ésima}}$ subunidad en una cadena de m subunidades que componen la cadena hija, donde $k = 1, 2, 3 \dots m$, donde $m > 10$, en general $m > 50$, y típicamente $m > 500$ o > 5000 . Cada subunidad está formada por una atadura en su configuración expandida unida a las porciones n^1 y n^2 de un sustrato monomérico, y cada subunidad está unida a la siguiente por enlaces de polimerización de monómero. Cada subunidad, un motivo de subunidad α , contiene información genética específica de especie establecida por ensamblaje dirigido por molde del intermedio de Xpandómero (hebra hija).

50 50 La Figura 66D muestra la construcción de sustrato como un modelo molecular, donde el miembro de nucleobase (664) se une a un resto de nucleobase primero y segundo, cada resto con un sitio de unión a la atadura (662, 663). La atadura (660) comprende grupos indicadores (900, 901, 902). Un enlace escindible selectivamente que separa los dos restos de nucleobase se indica con una "V" (665).

55 55 La figura 66E muestra el producto Xpandómero. Las subunidades comprenden la atadura expandida (660) unida a las porciones de nucleobase (669, 668), que se muestran como n^1 y n^2 en la figura 66C, cada subunidad unida por enlaces entre nucleobases. A través del procedimiento de escisión, el Xpandómero restringido se libera para convertirse en el producto Xpandómero. Los miembros de atadura que estaban anteriormente en una configuración restringida, están ahora en una configuración expandida, funcionando de este modo para extender linealmente la información de secuencia de la diana del molde. La expansión de las ataduras reduce la densidad lineal de la información de secuencia a lo largo del Xpandómero y proporciona una plataforma para incrementar el tamaño y la abundancia de los indicadores, lo que a su vez mejora la relación señal-ruido para la detección y descodificación de la secuencia molde.

60 60 Si bien la atadura se representa como una construcción de indicadores con tres grupos indicadores, se pueden organizar diversas configuraciones de indicadores en la atadura y pueden comprender indicadores individuales que identifican el monómero o la atadura puede ser un polímero desnudo. En algunos casos, uno o más precursores indicadores se disponen en la atadura, y los indicadores se unen por afinidad o se unen covalentemente después del ensamblaje del producto Xpandómero.

La figura 67 muestra un procedimiento para ensamblar un Xpandómero con las construcciones de sustrato de XNTP de clase X. En la primera vista, se usa un cebador de horquilla para cebar una molde y el molde se pone en contacto con una polimerasa y sustratos monoméricos de clase X. La polimerización se muestra en la etapa I para extender la hebra hija emergente de manera procesiva mediante la adición de sustratos de monómero dirigida por molde. La vista ampliada (figura 67a) ilustra esto con más detalle. Se muestra que la hebra hija complementaria a la hebra molde está compuesta por sustratos de nucleobase modificados con un sitio de escisión interno ("V"). En la etapa II, se completa el procedimiento de formación del intermedio de Xpandómero, y en la etapa III, se muestra que está en marcha un procedimiento de escisión, disociación del intermedio y expansión de la hebra hija. En la vista ampliada (figura 67b), se muestra la escisión interna de los sustratos de nucleobase para aliviar la restricción en las ataduras, que se expanden, alargando la cadena principal del Xpandómero.

Ejemplo 1

15 Síntesis de una Xsonda dimérica químérica "ca" con enlace internucleotídico ribosil-5'-3' escindible selectivamente

Las construcciones de sustrato oligoméricas están compuestas por miembros de sonda y miembros de atadura y tienen una construcción general de "sonda-bucle". La síntesis de miembros de sonda se logra usando procedimientos bien establecidos de síntesis de oligómeros en fase sólida. En estos procedimientos, la adición de nucleobases a una cadena de sonda emergente en una resina se logra con química de fosforamidita (patentes de EE. UU. n.º 4.415.732 y 4.458.066), por ejemplo, y se pueden sintetizar económicamente miligramos o gramos de oligómero sintético usando sintetizadores automatizados fácilmente disponibles. La síntesis típica de oligonucleótidos en fase sólida implica la realización reiterativa de cuatro etapas: desprotección, acoplamiento, protección y oxidación. Sin embargo, al menos un enlace en la sonda de una construcción de sustrato de Xsonda de clase I es un enlace escindible selectivamente, y al menos dos restos de sonda se modifican para la aceptación de un miembro de atadura. El enlace escindible selectivamente se sitúa entre los restos de sonda seleccionados para la unión de la atadura (es decir, "entre" no se debe limitar a significar "entre miembros de nucleobase contiguos" porque los puntos de unión primer y segundo de la atadura solo se tienen que situar en cualquier lugar en un resto primero y segundo de la sonda, respectivamente, estando unidos los restos por el enlace escindible selectivamente). En este ejemplo, un enlace internucleotídico ribosilo 5'-3', que es escindible selectivamente por la ribonucleasa H, es el enlace escindible selectivamente y los dos puntos de unión son los residuos de nucleobase primero y segundo de una sonda dimérica.

35 La síntesis de Xsondas modificadas con conector se logra usando fosforamiditas disponibles comercialmente, 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 (Reino Unido) y otros, o se pueden sintetizar de forma personalizada. Se pueden usar procedimientos de síntesis bien establecidos para preparar la sonda descrita en la que la nucleobase 3', que para este ejemplo es una C6 desoxiadenosina modificadora de amino, se une primero a un soporte universal usando 5'-dimetoxitritil-N6-benzoyl-N8-[6-(trifluoroacetilamino)-hex-1-il]-8-amino-2'-desoxiadenosina-3'-(2-cianoethyl)-(N,N-diisopropil)]-fosforamidita seguido de la adición de un modificador amino C6 citidina usando 5'-dimetoxitritil-N-dimetilformamidina-5-[N(trifluoroacetilaminohexil)-3-acrilimido]-citidina,3'-(2-cianoethyl)-(N,N-diisopropil)]-fosforamidita, en la que la 5' citidina es un ribonucleótido. La síntesis se completa con la adición de un reactivo de fosforilación química seguido de procedimientos estándar de escisión, desprotección y purificación. El producto dinucleotídico es una 5' fosfato (aminoC6-citosina) (aminoC6-desoxiadenosina) 3' con un enlace ribosil-5',3' centralmente escindible y conectores amino en cada base.

45 La atadura de la Xsonda para este ejemplo se construye con éster de bis-N-succinimidil-[pentaetilenglicol] (Pierce, Rockford IL; Producto n.º 21581). Las aminas conectoras del oligómero pCA modificado se reticulan con bis(NHS)PEG5 de acuerdo con las instrucciones del fabricante. Se obtiene un producto del peso molecular previsto para la construcción 50 de la sonda de PEG circularizada.

Ejemplo 2

55 Síntesis de una xsonda tetramérica "tata" con enlace de fosforotiolato escindible selectivamente

55 Los tetrámeros de la Xsonda se pueden sintetizar con un enlace fosforotiolato como el enlace escindible selectivamente. Para el siguiente ejemplo, se describe la síntesis de un 5' fosfato (dT) (aminoC6-dA) (dT) (aminoC6-dA) 3' tetranucleótido. Primero se prepara una 5' mercapto-desoxitimidina como describe por Mag *et al.* ("Synthesis and selective cleavage of an 60 oligodeoxynucleotide containing a bridged internucleotide 5'-phosphorothioate linkage", Nucl Acids Res 19:1437-41, 1991). Se hace reaccionar la timidina con dos equivalentes de cloruro de p-toluenosulfonilo en piridina a temperatura ambiente, y el 5'-tosilato resultante se aísla mediante cristalización en etanol. El tosilato se convierte en 5'-(S-tritil)-mercaptop-5'-desoxi-timidina con cinco equivalentes de tritiotiolato de sodio (*preparado in situ*). El 5'-(S-tritil)-mercaptop-timidina nucleótido se purifica y se hace reaccionar con 2-cianoetoxi-bis-(N,N-diisopropilamino-fosfano) en presencia de tetrazol para formar el componente básico 3'-O-fosforamidita.

Para comenzar la síntesis automatizada, el amino modificador C6 desoxiadenosina se une primero a un soporte universal usando 5'-dimetoxitiril-N6-benzoil-N8-[6-(trifluoroacetilamino)-hex-1-il]-8-amino-2'-desoxiadenosina -3'-(2-cianoethyl)-(N,N-diisopropil)]-fosforamidita seguido de la adición de la mercaptotimidina fosforamidita preparada anteriormente. Antes de añadir la siguiente amino C6 dA fosforamidita, primero se desprotege el grupo S-trítilo con nitrato de plata acuoso 5 50 mM y se lava la resina con agua. A continuación, la resina se trata típicamente con un agente reductor como DTT para eliminar los disulfuros incidentales formados durante la escisión. A continuación, la columna se lava de nuevo con agua y con acetonitrilo, y el tiol libre se hace reaccionar en condiciones estándar con la amino C6 desoxiadenosina fosforamidita en presencia de tetrazol, formando de este modo "ATA" con un enlace de puente S3'→P5' fosforotiolato entre la desoxiadenosina terminal y la 3-mercaptop-timidina. En el siguiente ciclo, se añade desoxitimidina fosforamidita estándar. 10 Finalmente, la síntesis se completa con la adición de un reactivo de fosforilación química seguido de procedimientos estándar de escisión, desprotección y purificación. El producto tetranucleotídico es un 5' fosfato (dT) (aminoC6-dA)(dT) (aminoC6-dA) 3'.

El enlace fosforotiolato se puede escindir selectivamente, 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 escindible selectivamente se encuentra entre las nucleobases segunda y tercera, la atadura está diseñada para unir este enlace y se puede unir a dos nucleobases cualesquiera (o dos puntos de unión de la cadena principal primaria) a cualquier lado del enlace escindible selectivamente. Los procedimientos para la química del conector y del conector cero incluyen, por ejemplo, la disposición de conectores 20 con aminas primarias usadas en la síntesis del oligómero, como se describe en el ejemplo 1. Los conectores modificados con amina en el oligómero "TATA" típicamente se protegen durante la síntesis de oligómeros y se desprotegen en el curso normal de completar la síntesis de oligonucleótidos.

El atadura de la Xsonda para este ejemplo se construye a continuación a partir de poli(etilenglicol)diglicidiléter activado 25 con bis-epóxido (SigmaAldrich, St. Louis MO, Producto n.º 475696). La reacción de epóxido de las aminas con los grupos de extremo de PEG activados se lleva a cabo en una solución diluida para minimizar cualquier reacción de concatenación que compita. Se obtiene una reactividad similar con anhídridos mixtos o incluso con cloruros de ácido, y se pueden emplear grupos de enlace heterobifuncionales para orientar la unión de la atadura. Los productos de reacción atados se separan mediante HPLC preparativa y se caracterizan mediante espectroscopía de masas. Se obtiene un producto del 30 peso molecular aproximado para la construcción de sonda de PEG tetramérica circularizada (de aproximadamente 2,5 Kd). Con este procedimiento se obtiene una distribución de ataduras de PEG con aproximadamente $M_n=500$. Esto corresponde a una atadura de aproximadamente 40 Angstroms (a aproximadamente 3,36 Å/unidad de PEG).

Ejemplo 3

Síntesis de una xsonda trimérica "cta" con enlace fosfodiéster 5'-3' escindible selectivamente

Las Xsondas también se pueden sintetizar con un enlace fosfodiéster como enlace escindible selectivamente. Para el siguiente ejemplo, se describe la síntesis de un trinucleótido 3' 5'fosfato (aminoC6-dC) (aminoC6-dT)(dT) con una 40 modificación de fosforotioato sin formación de puentes. Los enlaces fosfodiéster son atacados por una diversidad de nucleasas. En este ejemplo se usa un enlace de fosforotioato, con azufre que no forma puentes, como enlace resistente a las nucleasas.

Para la síntesis automatizada en la dirección 3' a 5', se usa un soporte sólido de desoxiadenosina inmovilizada con CPG (5'-dimetoxitiril-N-benzoil-2'-desoxiadenosina, 3'-succinoyl-alquilamino de cadena larga-CPG500). En el primer ciclo, se acopla el modificador amino C6 desoxitimidina fosforamidita (5'-dimetoxitiril-5-[N-(trifluoroacetilamino)hexil]-3-acrilimido]-2'-desoxiuridina, 3'-(2-cianoethyl)-(N,N-diisopropil)]-fosforamidita). Antes de la protección, el dA inmovilizado se hace reaccionar con reactivo sulfurizante (Glen Research, Sterling VA; n.º de cat. 40-4036), también conocido como reactivo de Beaucage, siguiendo el protocolo del fabricante. El reactivo en general se añade a través de un puerto separado en el 50 sintetizador. Después de la tiolación, se acopla la desoxicitidina C6 modificada con amino (5'-dimetoxitirilo-N-dimethylformamidina-5-[N-(trifluoroacetilamino)hexil]-3-acrilimido]-2'-desoxicitidina, 3'-(2-cianoethyl)-(N,N-diisopropil)]-fosforamidita). Antes de la protección, el dC inmovilizado se hace reaccionar con reactivo sulfurizante (Glen Research, Sterling VA; n.º de cat. 40-4036), también conocido como reactivo de Beaucage, siguiendo el protocolo del fabricante. El trímero "CTA" resultante tiene un enlace fosforotioato entre la T y la A, y un enlace fosfodiéster entre la C y la T. La adición 55 de un reactivo de fosforilación química seguido de procedimientos estándar de escisión, desprotección y purificación completa la síntesis.

La resistencia de los enlaces de fosforotioato al ataque de nucleasas está bien caracterizada, 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.º 5770713. Tanto C como T están modificadas con un conector o con un conector cero, sirviendo la derivatización para la unión de un miembro de atadura.

Ejemplo 4

Síntesis de una xsonda trimérica "ata" con enlace 5' n-p-o 3' fosforamidato escindible selectivamente

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

Este enlace fosforamidato es escindible selectivamente en condiciones en las que los enlaces fosfodiéster permanecen intactos tratando el oligómero con ácido acético al 80 % como describen Mag *et al.* ("Synthesis and selective cleavage of oligodeoxynucleotides containing non-chiral internucleotide phosphoramidate linkages", Nucl. Acids Res. 17: 5973-5988, 1989). Uniendo una atadura para formar un puente con el enlace 5'N-P-O fosforamidato, un Xpandómero que contiene dímeros de este tipo se puede expandir por la escisión selectiva de los enlaces fosforamidato de la cadena principal primaria.

Ejemplo 5**Síntesis de una xsonda hexamérica "caccac" con un enlace fotoescindible interno**

Después de la síntesis estándar de 3' a 5' con desoxicitosina y desoxiadenosina fosforamiditas no modificadas, C6 dC (5'-dimetoxitritil-N-dimetilformamidina-5-[N-(trifluoroacetilamino)hexil]-3-acrilimido]-2'-desoxicitidina modificada con amino, se acopla 3'-(2-cianoetil)-(N,N-diisopropil)]-fosforamidita; Glen Research; n.º cat. 10-1019), formando de este modo 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-diisopropilo)]-fosforamidita; Glen Research; n.º de cat. 10-4920). En el siguiente ciclo, se añade una segunda dC modificada con amino seguido de dos rondas finales de adición estándar de fosforamiditas dA y dC, respectivamente. El producto resultante, "CAC-pc-CAC", contiene conectores de amino en las posiciones de la base tercera y cuarta, y se puede modificar mediante la adición de una atadura que forma un puente con el enlace escindible selectivamente formado por la construcción de nitrobenceno fotoescindible entre las dos bases modificadas con amino. La escisión selectiva de una cadena principal de fosfodiéster modificada por un conector fotoescindible se divulga 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-photo cleavable 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), por ejemplo.

Ejemplo 6**Síntesis de una construcción de sustrato xmérica**

Las construcciones de sustrato Xméricas están estrechamente relacionadas en diseño y composición con las Xsondas. Una colección de Xmers se sintetiza, por ejemplo, mediante 5'-pirofosforilación de Xsondas. Los 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 procedimiento, 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 trifenilfosfina (Ph_3P) y disulfuro de 2,2'-dipiridilo (PyS_2) en DMF/DMSO, usando DMAP (4-dimetilaminopiridina) o 1-Melm (1-metilimidazol) como catalizador nucleófilo. El producto se precipita con LiClO_4 en acetona y se purifica por cromatografía de intercambio aniónico.

Se pueden considerar diversos otros procedimientos para la síntesis robusta de xmers de 5' trifosfato. Como describen Burgess y Cook (Chem Rev 100(6):2047-2060), estos procedimientos incluyen, pero sin limitaciones, reacciones que usan fosforamiditas de nucleósidos, síntesis por medio de ataque nucleofílico de pirofosfato sobre monofosfatos de nucleósidos activados, síntesis por medio de ataque nucleofílico de fosfato sobre pirofosfato de nucleósido activado, síntesis por medio de ataque nucleofílico de difosfato sobre sintonía de fosfato activado, síntesis que implica fosfitos activados o fosforamiditas derivados de nucleósidos, síntesis que implica desplazamiento directo de grupos salientes 5'-O por nucleófilos de trifosfato y procedimientos biocatalíticos. Un procedimiento específico que produjo sustratos dinucleotídicos compatibles con polimerasa usa N-metilimidazol para activar el grupo 5' monofosfato; la reacción posterior con pirofosfato (sal de tributilamonio) produce el trifosfato (Bogachev, 1996). En otro procedimiento, Kayushin ha sintetizado de trinucleótidos fosforamidatos (Kayushin AL *et al.* 1996. A convenient approach to the synthesis of trinucleotide phosphoramidites. Nucl Acids Res 24:3748-55).

Ejemplo 7**Síntesis de Xpandómero con escisión de fosforamidato con 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.º 7.060.440 concedida a Kless describe el uso de polimerasas para polimerizar oligómeros de trifosfato y el procedimiento se adapta en el presente documento para la síntesis de Xpandómeros. La construcción de sustrato consiste en un miembro de sonda trimérico "pppCA" con un enlace internucleotídico de 5' N-P-O fosforamidato escindible selectivamente entre nucleótidos y una construcción de atadura-bucle de PEG. Se sintetiza una hebra molde y un cebador acompañante y se purifica antes de su uso. La secuencia: "TGTGTGTGTGTGTGTGATCTACCGTCCGTCCC" se usa como molde. La secuencia "5'GGGACGGACGGTAGAT" se usa como cebador. Se usa como marcador un HEX 5' terminal (5'hexaclorofluoresceína) en el cebador. La hibridación del cebador y el molde forma un molde con cebador híbrido y el nucleótido protuberante monocatenario 3'-OH y 5' terminal libre de veinte bases de longitud. A continuación, se añaden las construcciones de sustrato y la ADN polimerasa T7 recombinante de la marca Sequenase™ (US Biochemicals Corp., Cleveland, OH) y se continúa la polimerización durante 30 min en condiciones ajustadas para una polimerización óptima. Una muestra de la reacción de polimerización se mezcla con tampón de carga de gel y se somete a electroforesis en un gel de acrilamida TBE al 20 % (Invitrogen, EE. UU.) junto con un control negativo sin polimerasa y un marcador de MW para confirmar la polimerización del Xmero.

El intermedio de Xpandómero se trata con ácido acético al 80 % durante 5 horas a temperatura ambiente de acuerdo con 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 de fosforamidato, lo que también se confirma por electroforesis.

Ejemplo 8

25 Síntesis de Xpandómero con escisión de fosforotioato con 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.º 7.060.440 concedida a Kless describe el uso de polimerasas para polimerizar oligómeros de trifosfato. La construcción de sustrato es un "pppCA". La construcción de sustrato está diseñada con un enlace de cadena principal de fosforotiolato entre nucleótidos escindible selectivamente y una construcción de atadura-bucle de PEG 2000. Se sintetiza una hebra molde y un cebador acompañante y se purifica antes de su uso. La secuencia: "TGTGTGTGTGTGTGTGATCTACCGTCCGTCCC" se usa como molde diana. La secuencia "5'GGGACGGACGGTAGAT" se usa como cebador. Se usa como marcador un HEX 5' terminal (5'hexaclorofluoresceína) en el cebador. La hibridación del cebador y el molde forma un molde con cebador híbrido y el nucleótido protuberante monocatenario 3'-OH y 5' terminal libre de veinte bases de longitud. A continuación, se añaden las construcciones de sustrato y la ADN polimerasa Therminator™ (New England Biolabs, EE. UU.) con tampón y sales optimizadas para la polimerización. La polimerización continuó durante 60 min en condiciones ajustadas para una polimerización óptima. Una muestra de la reacción de polimerización se mezcla con tampón de carga de gel y se somete a electroforesis en un gel de acrilamida TBE al 20 % (Invitrogen, EE. UU.) junto con un control negativo sin polimerasa y un marcador de MW para confirmar la polimerización del Xmero.

Los enlaces fosforotiolato del intermedio de Xpandómero son escindibles selectivamente, 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 mediante electroforesis en gel.

Ejemplo 9

50 Síntesis de Xpandómero químérico con ligasa

La síntesis de un Xpandómero se realiza usando una construcción de sustrato preparada con extremos 5'-monofosfato y 3'-OH. La construcción de sustrato es un químérico "5' p dC rA~dC dA3'" tetramérico donde la penúltima 5' adenosina es un ribonucleótido y el resto del sustrato es un desoxirribonucleótido. La construcción del sustrato está diseñada con un enlace internucleotídico ribosil 5'-3'fosfodiéster escindible selectivamente (como se muestra por la "˜") y una construcción de atadura-bucle PEG 2000, donde la atadura se une a los "C" y "A" terminales del tetramero. Se sintetiza una hebra molde y un cebador acompañante y se purifica antes de su uso. La secuencia: "TGTGTGTGTGTGTGTGATCTACCGTCCGTCCC" se usa como molde diana. La secuencia "5'GGGACGGACGGTAGAT" se usa como cebador. Se usa como marcador un HEX 5' terminal (5'hexaclorofluoresceína) en el cebador. La hibridación del cebador y el molde forma un molde con cebador híbrido y el nucleótido protuberante monocatenario 3'-OH y 5' terminal libre de veinte bases de longitud. A continuación, se añaden las construcciones de sustrato y la ADN ligasa T4 (Promega Corp, Madison, WI, EE. UU.; n.º de catálogo M1801) con temperatura, tampón y sales optimizados para la hibridación y ligación transitoria de la sonda. La ligación se continúa durante 6 horas. Una muestra de la reacción de ligación se mezcla con tampón de carga de gel y se somete a electroforesis en un gel de acrilamida TBE al 20 % (Invitrogen, Carlsbad, CA, EE. UU.) junto con un control negativo sin ligasa y un marcador MW para confirmar la ligación de la Xsonda.

En una segunda etapa, el intermedio de Xpandómero se trata con ribonucleasa H para escindir el enlace fosfodiéster 5'-

3' RNasa lábil para producir un producto Xpandómero que también se confirma mediante electroforesis en gel.

Ejemplo 10

5 Preparación de una construcción de conector alfa-fosfato

Los conectores de atadura también se pueden unir al diéster de fosforotioato S o a una N-amida de fosforamidato, como se analiza por Agrawal ("Site specific functionalization of oligonucleotides for attaching two different reporter groups", Nuc. Acids Res. 18:5419-23, 1990). Se divulga un procedimiento para la funcionalización de dos enlaces entre nucleótidos de 10 cadenas principales diferentes: conector de fosforamidato de aminoxylo (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 atadura entre nucleobases 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).

15

Ejemplo 11

Vías heterobifuncionales para atar construcciones

20 La síntesis de oligómeros modificados que contienen un enlace escindible selectivamente se describe en los ejemplos 1-5. En este caso, se prepara un tetrámero con una base modificada con C6 amino en la segunda posición y una base modificada con conector de 4-formilbenzoato en la tercera posición del tetrámero mediante química sintética estándar de oligómeros. Las bases segunda y tercera se separan por un enlace escindible selectivamente seleccionado de enlace ribosil 5'-3' fosfodiéster, enlace desoxirribosil 5'-3' fosfodiéster, enlace fosforotiolato (5' O-P-S 3' o 5' SPO3'), 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 éster NHS activo. La síntesis del bucle de atadura heterobifuncional se lleva a cabo a continuación haciendo reaccionar el miembro de sonda modificado con una atadura funcionalizada con grupos terminales de extremo de amina e hidrazida. Se obtiene un producto final circularizado con atadura.

30 Ejemplo 12

Construcción de atadura marcada con tinte

35 Se prepara una construcción de sustrato de Xsonda con un solo indicador. Primero se sintetiza un dímero con aminas C6 en las bases primera y segunda usando procedimientos estándar. Las bases primera y segunda se separan por un enlace escindible selectivamente seleccionado de enlace ribosil 5'-3' fosfodiéster, enlace desoxirribosil 5'-3' fosfodiéster, enlace fosforotiolato (5' O-P-S 3' o 5' SPO3'), enlace fosforamidato (5' O-P-N 3' o 5' N-P-O 3') o un enlace fotoescindible. En este ejemplo, la atadura es una molécula de PEG 2000 funcionalizada en el extremo, específica de especie, con un solo grupo 40 conector interno, que en este caso es un conector funcionalizado con maleimido situado en la mitad de la atadura. El indicador es un dendrímero marcado con tinte unido al grupo conector maleimido en la atadura por medio de un resto sulfhidrilo en el dendrímero. La construcción de indicador resultante contiene, de este modo, un indicador en la atadura. Se usa como indicador un dendrímero de cistamina (Dendritic Nanotechnologies, Mt Pleasant, MI, EE. UU.; n.º de cat. DNT-294 G3) con un diámetro de 5,4 nm y 16 aminas superficiales por medio dendrímero. Uniendo tintes específicos o 45 combinaciones de tintes a los grupos amina en el dendrímero, la especie de la construcción de sustrato se marca de manera única para su identificación.

Ejemplo 13

Construcción de atadura marcada con péptido

50 Se prepara una construcción de sustrato de Xsonda con un solo indicador. Primero se sintetiza un dímero con aminas C6 en las bases primera y segunda usando procedimientos estándar. Las bases primera y segunda se separan por un enlace escindible selectivamente seleccionado de enlace ribosil 5'-3' fosfodiéster, enlace desoxirribosil 5'-3' fosfodiéster, enlace fosforotiolato (5' O-P-S 3' o 5' SPO3'), enlace fosforamidato (5' O-P-N 3' o 5' N-P-O 3') o un enlace fotoescindible. En este 55 ejemplo, la atadura es una molécula de PEG 2000 funcionalizada en el extremo, específica de especie, con un solo grupo conector interno, que en este caso es un conector funcionalizado con maleimido situado en la mitad de la atadura. El indicador es un dendrímero marcado con tinte unido al grupo conector maleimido en la atadura por medio de un resto sulfhidrilo en el dendrímero. La construcción de indicador resultante contiene, de este modo, un indicador en la atadura. Se usa como indicador un dendrímero de cistamina (Dendritic Nanotechnologies, Mt Pleasant, MI, EE. UU.; n.º de cat. DNT-294 G3) con un diámetro de 5,4 nm y 16 aminas superficiales por medio dendrímero. Uniendo tintes específicos o 60 combinaciones de tintes a los grupos amina en el dendrímero, la especie de la construcción de sustrato se marca de manera única para su identificación. Uniendo péptidos específicos al dendrímero, que está funcionalizado con amina, la especie de la construcción de sustrato se marca para su posterior identificación. Las dimensiones y la carga de los péptidos unidos se usan como características de detección en un aparato de detección.

65

Ejemplo 14

Vía heterobifuncional para atar construcciones de indicador

- 5 Se preparan construcciones de sustrato con múltiples indicadores. Se prepara una colección de dímeros con base modificada con amino C6 en la primera posición y base modificada con 4'-formilbenzoato (4FB) en la segunda posición del dímero mediante química orgánica estándar. Las bases primera y segunda se separan por un enlace escindible selectivamente seleccionado de enlace ribosil 5'-3' fosfodiéster, enlace desoxirribosil 5'-3' fosfodiéster, enlace fosforotiolato (5' OP-S 3' o 5' SPO3'), enlace fosforamidato (5' O-P-N 3' o 5' N-P-O 3') o un enlace fotoescindible. La amina C6 se hace reaccionar con Sulfo EGS para formar un grupo éster NHS 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 del segmento 1 se hace reaccionar con Sulfo EGS para formar un éster NHS activo. Una protección heterobifuncional específica de especie que consiste en un par de segmentos indicadores (segmentos 2 y 3) con grupos de extremo amina y 4FB se hace reaccionar a continuación con la construcción, cerrando el bucle de atadura.
- 10
- 15 La construcción de indicador resultante contiene, de este modo, cuatro indicadores en la construcción de la atadura. En este ejemplo, los indicadores son dendrones de poliamina con un conector de cistamina para que se una covalentemente a cada segmento de polímero mediante enlaces tioéter (Dendritic Nanotechnologies, Mt Pleasant MI, EE. UU.; n.º de cat. DNT-294: G4, 4,5 nm de diámetro, 32 aminas superficiales por medio dendrímero) y los miembros poliméricos son PEG 2000 funcionalizados en los extremos, cada uno con un grupo conector interno. Cada segmento de atadura se compone de un solo indicador. Con cuatro segmentos acoplados direccionalmente, están disponibles 2^4 posibles combinaciones de códigos de indicadores.
- 20

Ejemplo 15**Vía heterobifuncional a construcciones de indicadores con marcado posterior a la síntesis**

- 25 Se preparan construcciones de sustrato con múltiples indicadores. Se prepara una tetrámero con base modificada con amino C6 en la primera posición y base modificada con 4'-formilbenzoato (4FB) en la segunda posición del dímero mediante química orgánica estándar. Las bases segunda y tercera se separan por un enlace escindible selectivamente seleccionado de enlace ribosil 5'-3' fosfodiéster, enlace desoxirribosil 5'-3' fosfodiéster, enlace fosforotiolato (5' OP-S 3' o 5' SPO3'), enlace fosforamidato (5' O-P-N 3' o 5' N-P-O 3'), o un enlace fotoescindible. La amina C6 se hace reaccionar con Sulfo EGS para formar un grupo éster NHS 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 del segmento 1 se hace reaccionar con Sulfo EGS para formar un éster NHS activo. Una protección heterobifuncional específica de especie que consiste en un par de segmentos indicadores (segmentos 2 y 3) con grupos de extremo amina y 4FB se hace reaccionar a continuación con la construcción, cerrando el bucle de atadura.
- 30
- 35
- 40 La construcción de indicador resultante contiene, de este modo, cuatro indicadores en la construcción de la atadura. En este ejemplo, se describe el marcado posterior de la construcción de atadura. Unida covalentemente a cada segmento de atadura hay un oligómero hexadecamérico que se usa para la unión del indicador. Cada segmento de atadura está compuesto en parte de una molécula de PEG funcionalizada. El indicador es un dendrímero de poliamina marcado con tinte con un conector de cistamina (Dendritic Nanotechnologies, Mt Pleasant MI, EE. UU.; n.º de cat. DNT-294: G5, 4,5 nm de diámetro, 64 aminas superficiales por medio dendrímero) para acoplar a la sonda de oligómero hexadecamérica. Despues del ensamblaje del Xpandómero, los dendrímeros marcados con tinte se hibridan con los segmentos de atadura oligoméricos. Este enfoque de marcado es análogo al procedimiento descrito por DeMattei *et al.* ("Designed Dendrimer Syntheses by Self-Assembly of Single-Site, ssDNA Functionalized Dendrons", Nano Letters, 4:771-77, 2004).
- 45

Ejemplo 16**Elementos de indicadores marcados con tinte**

- 50 En referencia a los ejemplos 14 y 15, las aminas superficiales de los elementos indicadores dendriméricos se marcan mediante química de ésteres activos. Alexa Fluor 488 (verde) y Alexa Fluor 680 (rojo) están disponibles para unión en una sola etapa como ésteres activos de sNHS de Molecular Probes (Eugene OR). La densidad y la proporción de los colorantes se varían para producir una marca molecular distintiva en cada elemento indicador.

Ejemplo 17**Elementos de indicadores multiestado**

- 55 Diversas paletas de tintes se seleccionan mediante técnicas similares a las usadas en M-FISH y SKY, como es conocido para los expertos en la técnica, y se conjugan con un indicador dendrimérico. Por tanto, el elemento indicador de cada atadura constituye una "direcciónpectral", con lo que una sola construcción dendrimérica con una multiplicidad de sitios de unión de tintes puede crear una pluralidad de códigos de indicadores. En referencia a la construcción de atadura

descrita en los ejemplos 14 y 15, una dirección espectral de 5 estados da como resultado 625 combinaciones de códigos de indicadores.

Ejemplo 18

5

Atadura espaciadora Peg-5000

Los segmentos de atadura se construyen con un polímero duradero soluble en disolvente acuoso/orgánico que posee poca o ninguna afinidad de unión por los reactivos SBX. El PEG 5000 modificado se usa para los espaciadores de atadura flexibles que flanquean un indicador de polilisina 5000. Los extremos de PEG libres se funcionalizan para que se unan al miembro de sonda. El polímero se circulariza mediante reticulación a los puntos de unión de la sonda usando la química del conector heterobifuncional.

Ejemplo 19

15

Preparación de una composición de indicador de marcas de masa

Una atadura se sintetiza de la siguiente manera: Las marcas de masa escindible se acoplan covalentemente a dendrímeros G5 funcionalizados con péptido de poli-l-lisina (Dendritic Nanotechnologies, Mt Pleasant MI) usando química de acoplamiento de aminas estándar. Los dendrímeros G5 tienen un diámetro de ~5,7 nm y proporcionan 128 grupos superficiales reactivos. Una cadena de diez dendrímeros G5 funcionalizados con péptidos de polilisina de peso molecular 10 000 proporciona alrededor de 100 000 sitios de unión de indicadores conjuntamente en un segmento dendrimérico de ~ 57 nm. Un total de aproximadamente 10.000 marcas de masa están disponibles para detección en el segmento completamente ensamblado, suponiendo solo un 10 % de ocupación de los sitios de unión disponibles. Usando el procedimiento de codificación de 3 indicadores de marcas de masa descrito previamente (figura 37), se dispone de aproximadamente 3300 copias de cada marca de masa para la medición. Como alternativa, un único dendrímero G9 (con 2048 grupos reactivos) funcionalizado con polilisina 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 indicadores.

30 Para detectar una secuencia con marcas de masa, se usa un procedimiento para la liberación controlada de los indicadores de marcas de masa en el punto de medición mediante el uso de conectores fotoescindibles. La fragmentación secuencial de la atadura no es necesaria. Los indicadores de marcas de masa asociados con cada subunidad del polímero de Xpandómero se miden en una sola etapa. Por ejemplo, un conjunto de 13 indicadores de marcas masivas que van de 350 Dalton a 710 Dalton (es decir, una escalera de 30 Dalton de indicadores de marcas masivas) tiene 286 combinaciones de tres marcas masivas cada una. De esta manera, cualquiera de los 256 tetrámeros diferentes se asocia con solo una combinación particular de 3 indicadores de marcas de masa. La información de secuencia codificada del Xpandómero se detecta fácilmente por espectroscopía de masa 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.

40

Ejemplo 20

Análisis directo de construcciones de sustrato no marcadas

45 Se sintetiza una colección de construcciones de sustrato; las ataduras no contienen indicadores. Después de la preparación de un producto Xpandómero, las bases individuales del producto Xpandómero se analizan mediante espectroscopía de tunelización de electrones.

Ejemplo 21

50

Análisis asistido por hibridación de construcciones de sustrato no marcadas

Se sintetiza una colección de construcciones de sustratos oligoméricos; las ataduras no contienen indicadores. Después de la preparación de un producto Xpandómero, un conjunto completo de sondas marcadas se hibridan con el producto Xpandómero y las sondas hibridadas se analizan secuencialmente.

Ejemplo 22

60

Síntesis de una desoxiadenosina trifosfato con atadura de lys-cys-peg-poliglutamato-peg-cys-cooh

La lisina con una cadena lateral amino protegida por BOC se inmoviliza en una resina y se hace reaccionar con un residuo de cisteína usando procedimientos estándar de síntesis de péptidos. La cadena lateral amino de la lisina será el grupo funcional reactivo épsilon de la unión de RT-NTP (véase clase VI, VII). La cisteína formará la primera mitad de un enlace disulfuro intraatadura. La amina desprotegida en la cisteína se modifica con SANH (Pierce-Thermo Fisher, EE. UU.; n.º de cat. 22400: Bioconjugate Toolkit) para formar una hidrazida.

Por separado, se prepara un segmento espaciador a partir de bis-amino PEG 2000 (Creative PEGWorks, Winston Salem NC; n.º de cat. PSB330) por funcionalización de las aminas libres con C6-SFB (Pierce-Thermo Fisher, EE. UU.; n.º de cat. 22400: Bioconjugate Toolkit), formando un segmento espaciador bis-4FB PEG; el producto se purifica.

- 5 A continuación, el segmento espaciador bis-4FB PEG se acopla al conector de hidrazida en la cisteína, dejando un grupo 4FB como grupo reactivo terminal, y se lava la resina.

Por separado, se prepara un segmento de poliglutamato (cada glutamato derivatizado en el carboxilo gamma con 5 unidades de PEG de PEG protegido con metilo). El extremo c se convierte en una amina con agente de acoplamiento EDC y diamino hexano. SANH se usa para formar un segmento de poliglutamato finalizado en dihidrazida y el producto se purifica. El segmento de poliglutamato finalizado en dihidrazida se hace reaccionar con el grupo 4FB terminal de la resina, formando una hidrazida terminal y la resina se lava.

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

15 Un residuo de cisteína se acopla al carboxilo libre usando reactivos de síntesis de péptidos estándar. El carboxilo terminal de la cisteína se protege con O-bencilo. El producto resultante se lava de nuevo y a continuación se escinde de la resina. 20 El carboxilo libre generado por la escisión se modifica a continuación con EDC, aminohexilo y SANH para formar una hidrazida reactiva.

25 Por separado, una desoxiadenosina trifosfato modificada con amina C6 (N6-(6-amino)hexil-dATP, Jena Bioscience, Jena DE; n.º de cat. NU-835) se trata con SFB (Pierce Bioconjugate Toolkit, n.º de cat. 22419) para formar un grupo funcional 4FB. Combinando la base modificada con la hidrazida reactiva de las etapas anteriores, se ensambla una construcción de sustrato de atadura-sonda. El BOC de la cadena lateral de lisina se retira antes de su uso. En condiciones en general oxidantes, las cisteínas se asocian para formar un enlace disulfuro intraatadura, estabilizando la atadura en una forma compacta y restringida.

30 Ejemplo 23

Síntesis de una colección de trifosfato rt-ntp

35 Se preparan bases de nucleótido trifosfato atadas A, T, C y G con enlace intraatadura -S-S- como se describe en el ejemplo 22, pero los parámetros físicos y de carga de los segmentos de glutamato PEGilado usados para cada base se seleccionan para proporcionar una característica de indicador diferenciada.

Ejemplo 24

40 Síntesis de Xpandómeros por sbe usando rt-ntps atados con poliglutamato

45 Se preparan nucleótidos de adenosina trifosfato y guanosina RT-NTP modificados con enlaces intraatadura disulfuro. Las bases se modifican además para que se bloquen de forma reversible 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). Las ataduras de las bases modificadas se construyen con un grupo funcional delta y un grupo funcional épsilon en general como se muestra en la figura 61. El grupo funcional delta es un carboxilo de una cisteína colgante de la atadura y el grupo funcional épsilon es una amina de cadena lateral de una lisina cerca de la unión de la atadura a las purinas. Las ataduras se modifican además para que contengan segmentos de poliglutamato modificados específicos de nucleobase.

50 50 La secuencia "TCTCTCTCTCTCATCTACCGTCCGTCCC" se usa como molde. La secuencia "5'GGGACGGACGGTAGAT" se usa como cebador. Se usa como marcador un HEX 5' terminal (5'hexaclorofluoresceína) en el cebador. El procedimiento de síntesis del Xpandómero es esencialmente como se describe en 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 sola base a la hebra hija emergente en el extremo 3'-OH del cebador. Debido a que la construcción del sustrato ese bloquea en la posición 3', no se produce más polimerización.

60 60 El grupo conector de la cadena lateral amino (épsilon) en el primer RT-NTP añadido está protegido y permanecerá así a lo largo de la reacción SBE. El grupo carboxilo terminal de la atadura se desprotege y el 3' OH del sustrato se desbloquea; el complejo se lava antes de una siguiente ronda de SBE.

65 65 En un segundo ciclo de SBE, otra nucleobase se polimeriza en el intermedio de Xpandómero emergente. El enlace x se forma entre la amina libre del grupo conector épsilon en la primera nucleobase y el grupo conector carboxilo en la atadura de la segunda nucleobase usando EDC y Sulfo-NHS como agente de reticulación (Pierce, n.º de cat. 22980 y 24510). Se desprotege el carboxilo del grupo conector delta de la atadura y se desbloquea el 3' OH del sustrato; el complejo se lava antes de una próxima ronda de SBE.

El ciclo de SBE se puede repetir varias veces, formando de este modo un intermediario de Xpandómero en la configuración restringida. Cada atadura en la cadena creciente de nucleobases unidas por χ está en la configuración de Xpandómero restringida.

5

Ejemplo 25

Escisión por nucleasa y tcep para formar un Xpandómero de clase X

- 10 El intermedio de Xpandómero del ejemplo 24 se escinde con nucleasa, formando un producto Xpandómero compuesto de nucleobases individuales unidas por segmentos de unión y enlaces χ . La nucleasa también degrada el molde y cualquier cebador asociado, liberando el producto. Los enlaces disulfuro intraatadura se escinden mediante la adición de un agente reductor (TCEP, Pierce, n.^o de cat. 20490).
- 15 El producto Xpandómero se filtra y purifica para retirar los sintones truncados y los productos secundarios de la digestión con nucleasas. La detección y el análisis del Xpandómero linealizado se pueden hacer usando una amplia variedad de procedimientos existentes y de próxima generación.

Ejemplo 26

Síntesis de un trifosfato de desoxiadenosina con un conector fotoescindible intraatadura

La glicina se inmoviliza en una resina y se hace 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 finaliza en un carboxilo protegido con O-Bencilo, tal como un conector de 2-nitroveratrilmamina adaptado del que se describe 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" de la atadura RT-NTP. La amina desprotegida en el glutamato se modifica con SANH (Pierce-Thermo Fisher, EE. UU.; n.^o de cat. 22400: Bioconjugate Toolkit) para formar una hidrazida. La cadena lateral del glutamato formará un conector intraatadura fotoescindible después de la síntesis de la atadura.

Por separado, se prepara un segmento espaciador a partir de bis-amino PEG 2000 (Creative PEGWorks, Winston Salem, NC, EE. UU.; n.^o de cat. PSB330) por funcionalización de las aminas libres con C6-SFB (Pierce Bioconjugate Toolkit, n.^o de cat. 22423), formando un segmento espaciador bis-4FB PEG, y el producto se purifica. A continuación, el segmento espaciador bis-4FB PEG se acopla al conector de hidrazida en el glutamato, dejando un grupo 4FB como grupo reactivo terminal y se lava la resina.

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

Un segmento espaciador PEG-2000 (con extremos amino libres y amino protegidos por FMOC; n.^o de cat. PHB-0982, Creative PEGWorks) se modifica con C6 SFB para formar un segmento espaciador de PEG modificado con 4FB y FMOC-amino. El extremo 4FB se hace reaccionar con el grupo hidrazida en la resina, formando una cadena finalizada en amino FMOC. La resina se lava de nuevo.

Un residuo de lisina se acopla a la amina libre del espaciador mediante un enlace peptídico. El residuo de lisina se protege en la cadena lateral por BOC y la alfa-amina de la lisina se protege por FMOC. El carboxilo terminal O Bencilo del conector fotoescindible y la cadena lateral de la lisina protegida por BOC se desprotegen y se reticulan con EDC/Sulfo-NHS para circularizar la unión.

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 a continuación con EDC, aminohexilo y SANH para formar una hidrazida reactiva.

55 Por separado, una desoxiadenosina trifosfato modificada con amina C6 (N6-(6-amino)hexil-dATP, Jena Bioscience, Jena DE; n.^o de cat. NU-835) se trata con SFB (Pierce Bioconjugate Toolkit, n.^o de cat. 22419) para formar un grupo funcional 4FB. Combinando la base modificada con la hidrazida reactiva de las etapas anteriores, se ensambla una construcción de sustrato de atadura-sonda. El conector intraatadura fotoescindible estabiliza la atadura en una forma compacta y restringida. A continuación se retira el FMOC y la amina terminal libre se hace reaccionar con Sulfo-EMCS (Pierce; n.^o de cat. 22307) para introducir un grupo funcional maleimido terminal.

Ejemplo 27

Síntesis de Xpandómeros por sbe usando rt-htps fotoescindibles

- Se preparan nucleótidos de adenosina trifosfato y guanosina RT-NTP modificados con enlaces intraatadura fotoescindibles. Las bases se modifican además para que se bloqueeen de forma reversible 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). Las ataduras de las bases modificadas se construyen con un grupo funcional delta y un grupo funcional épsilon en general como se muestra en la figura 61. El grupo conector delta es una amina de una lisina colgante terminal de la atadura y el grupo conector épsilon es un sulfhidrilo de una cisteína cerca del punto de unión de la atadura. Las ataduras se modifican además para que contengan segmentos de poliglutamato modificados específicos de especie.
- 5 La secuencia "TCTCTCTCTCTCATCTACCGTCGTCCC" se usa como molde. La secuencia "5'GGGACGGACGGTAGAT" se usa como cebador. Se usa como marcador un HEX 5' terminal (5'hexaclorofluoresceína) en el cebador. El procedimiento de síntesis del Xpandómero es esencialmente como se describe en la figura 61. En un primer ciclo de cebado de SBE, la nucleobase modificada (A) se añade con polimerasa Klenow en condiciones adaptadas para la polimerización y se añade una sola base a la hebra hija emergente en el extremo 3'-OH del cebador. Debido a que la construcción del sustrato ese bloquea en la posición 3', no se produce más polimerización. A continuación, se lava el complejo cebador inmovilizado-molde para retirar el sustrato que no ha reaccionado.
- 10 15 20

El grupo conector de la cadena lateral sulfhidrilo (épsilon) en el primer RT-NTP añadido está protegido y permanecerá así a lo largo de la reacción SBE. El grupo amino terminal de la atadura se desprotege y el 3' OH del sustrato se desbloquea; el complejo se lava antes de la siguiente ronda de SBE.

En un segundo ciclo de SBE, otra nucleobase (G) se polimeriza en el intermedio de Xpandómero emergente. El enlace χ se forma entre el amino (grupo conector delta) en la primera nucleobase y el sulfhidrilo (grupo conector épsilon) en la atadura de la segunda base usando el reactivo de reticulación GMBS (Pierce; n.º de cat. 22309). El grupo conector amino delta en 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.

30 El ciclo de SBE se puede repetir varias veces, formando de este modo un intermediario de Xpandómero en la configuración restringida. Cada atadura en la cadena creciente de nucleobases unidas por χ está en la configuración de Xpandómero restringida.

Ejemplo 28

Escisión por nucleasas y fotoescisión para formar un Xpandómero de clase x

- 35 El intermedio de Xpandómero del ejemplo 27 se escinde con nucleasa, formando un producto Xpandómero compuesto de nucleobases individuales unidas por segmentos de unión y enlaces χ . La nucleasa también degrada el molde y cualquier cebador asociado, liberando el producto. Los enlaces fotoescindibles intraatadura se escinden por exposición a una luz ultravioleta.
- 40 El producto Xpandómero se filtra y purifica para retirar los sintones truncados y los productos secundarios de la digestión con nucleasas. La detección y el análisis del Xpandómero linealizado se pueden hacer usando una amplia variedad de procedimientos existentes y de próxima generación.

Ejemplo 29

Síntesis de una atadura rt-ntp *in situ*

- 50 Usando procedimientos de síntesis de péptidos estándar sobre un soporte sólido, se prepara un péptido que tiene la estructura (Resina-C')-Glu-Cys-(Gly-Ala)₁₀-Pro-Ser-Gly-Ser-Pro-(Ala-Gly)₁₀-Cys-Lys. La amina terminal se hace reaccionar con SANH (Pierce, n.º de cat. 22400) para crear un conector de hidrazida.

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

- 60 A continuación, la construcción se escinde de la resina. Después de la desprotección y en condiciones en general oxidantes, las cisteínas se asocian para formar un enlace disulfuro intraatadura, estabilizando la horquilla beta, que contiene un carboxilo libre terminal (un grupo conector delta colgante en la atadura) y una lisina cerca del punto de unión de la atadura (la amina de cadena lateral y un grupo conector épsilon).

- 65 El disulfuro es representativo de la estabilización intraatadura representada en las construcciones de sustrato de las clases II, III, VI, VII y VIII (véanse las figuras 8 y 9), aunque se ilustra aquí con una referencia más específica a las construcciones de sustrato monoméricas de las clases VI, VII y VIII. La longitud de la atadura desplegada, suponiendo una longitud del enlace peptídico CC residual de 3,8 Å, es de aproximadamente 10 nm, pero asume una forma compacta debido a los

enlaces de hidrógeno en la horquilla beta.

Como se describe en Gellman ("Foldamers, a manifesto", Acc Chem Res 31:173-80, 1998), una amplia variedad de polímeros, no simplemente péptidos, se pueden plegar en formas compactas. Dichos polímeros incluyen oligopiridinas, poliisocianuros, poliisocianatos, poli(triarilmethyl) metacrilatos, polialdehídos, poliprolina, ARN, oligopirrolinonas y oligoureas, todos los cuales han presentado la capacidad de plegarse en estructuras secundarias compactas y expandirse en condiciones adecuadas. Por lo tanto, los ejemplos de péptidos presentados aquí son representativos de una clase mucho más grande de químicas de ataduras, donde las restricciones en la atadura no expandida pueden incluir enlaces de hidrógeno e interacciones hidrófobas, por ejemplo, así como ligación entre ataduras.

10

Ejemplo 30

Síntesis de Xpandómeros usando Xsondas

En un ejemplo de SBX, se presenta una colección Xsondas de 256 Xsondas tetraméricas a una diana de ADN monocatenario alargada y atada a la superficie para hibridación. La etapa de hibridación continúa con una rutina de ciclo térmico precisa para favorecer cadenas largas de Xsonda. Los híbridos de sonda-diana inespecíficos y débilmente unidos se retiran mediante un simple etapa de lavado, de nuevo con un control térmico preciso. Se realiza ligación enzimática para unir cualquier cadena de Xsonda a lo largo del ADN diana, seguido de un segundo lavado. Repitiendo el ciclo de hibridación/lavado/ligación/lavado, las secuencias ligadas más largas crecen en múltiples locus a lo largo del ADN diana hasta que la replicación del molde diana está casi completa.

15

Los huecos sin llenar a lo largo del ADN diana se llenan usando un procedimiento bien establecido de llenado de huecos basado en ADN polimerasas y ligasas (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 intermedio de Xpandómero completo, que está compuesto de la diana de ADN original con Xsondas acopladas e hibridadas de forma complementaria con rellenos de huecos ocasionales de 1, 2 o 3 nucleótidos, se escinde para producir un Xpandómero. El conector escindible para este ejemplo es una modificación de la cadena principal del sustrato 3' O-P-N 5'. La escisión selectiva se cataliza mediante la adición de ácido acético a temperatura ambiente.

20

El Xpandómero se filtra y purifica para eliminar los productos truncados y, posteriormente, se alarga para formar una estructura lineal de códigos indicadores de enlazados. La detección y el análisis del producto Xpandómero se pueden hacer usando una amplia variedad de procedimientos existentes.

25

Ejemplo 31

Síntesis de Xpandómeros de Xntp usando polimerasa

30

La síntesis de un Xpandómero de clase X se realiza usando una construcción de sustrato de 8-[(6-amino)hexil]-amino-desoxiadenosina trifosfato modificado que tiene una estructura mixta que consta de un fosfonato de 2-aminoetilo que no forma puente y un fosforotiolato que forma puente (3' O-P-S 5') en el fosfato alfa. Un atadura intranucleotídica se une al conector de fosfonato de 2-aminoetilo y a un conector de amino C6 en la 8-[(6-amino)hexil]-amino-desoxiATP. La secuencia "TTTTTTTTTTTTTTTTATTATCTACCGTCCGTCCC" se usa como molde. La secuencia "5'GGGACGGACGGTAGAT" se usa como cebador. Se usa como marcador un HEX 5' terminal (5'hexaclorofluoresceína) en el cebador. La hibridación del cebador y el molde forma un molde con cebador híbrido y el nucleótido protuberante monocatenario 3'-OH y 5' terminal libre de veinte bases de longitud. A continuación, se añaden las construcciones de sustrato y la polimerasa y se continúa la polimerización durante 60 min en condiciones ajustadas para una polimerización óptima. Una muestra de la reacción de polimerización se mezcla con tampón de carga de gel y se somete a electroforesis en un gel de acrilamida TBE al 20 % (Invitrogen, EE. UU.) junto con un control negativo sin polimerasa y un marcador de PM para confirmar la polimerización de Xntp.

35

El 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 enlaces de fosforotiolato entre la unión de atadura y la desoxirribosa, que se confirma por electroforesis

40

Ejemplo 32

Síntesis de Xpandómeros de Xntp usando polimerasa

45

La síntesis de un Xpandómero de clase X se realiza usando una construcción de sustrato de N⁶-(6-amino)hexil-desoxiadenosina trifosfato modificada que tiene una estructura mixta que consta de un fosforamidato que no forma puente (N-1-aminoalquilo) y un fosforotiolato que forma puente (3' OPS5') en el fosfato alfa. Un atadura intranucleotídica se une al grupo N-1-aminoalquilo y a un conector amino C6 en la N⁶-(6-amino)hexil-desoxiATP. La secuencia "TTTTTTTTATCTACCGTCCGTCCC" se usa como molde. La secuencia "5'GGGACGGACGGTAGAT" se usa como cebador. Se usa como marcador un HEX 5' terminal (5'hexaclorofluoresceína) en el cebador. La hibridación del cebador

y el molde forma un molde con cebador híbrido y el nucleótido protuberante monocatenario 3'-OH y 5' terminal libre de veinte bases de longitud. A continuación, se añaden las construcciones de sustrato y la polimerasa y se continúa la polimerización durante 60 min en condiciones ajustadas para una polimerización óptima. Una muestra de la reacción de polimerización se mezcla con tampón de carga de gel y se somete a electroforesis en un gel de acrilamida TBE al 20 % (Invitrogen, EE. UU.) junto con un control negativo sin polimerasa y un marcador de PM para confirmar la polimerización de XNTP.

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

Ejemplo 33

Síntesis de Xpandómeros de Xntp usando ligasa

La síntesis de un Xpandómero de clase X se realiza usando una construcción de sustrato de de 8-[(6-amino)hexil]-amino-desoxiadenosina monofosfato modificado que tiene una estructura mixta que consta de un fosfonato de 2-aminoetilo que no forma puente y un fosforamidato que forma puente (3'-O-P-N-5') en el fosfato alfa. Una atadura intranucleotídica se une al conector de 2-aminoetilo y a un conector amino C6 en la 8-[(6-amino)hexil]-amino-desoxiAMP. La secuencia TTTTTTTTTT TTTTTTTTATCTACCGTCCGTCCC" se usa como molde. La secuencia "5'GGGACGGACGGTAGAT" se usa como cebador. Se usa como marcador un HEX 5' terminal (5'hexaclorofluoresceína) en el cebador. La hibridación del cebador y el molde forma un molde con cebador híbrido y el nucleótido protuberante monocatenario 3'-OH y 5' terminal libre de veinte bases de longitud. A continuación, se añaden las construcciones de sustrato y la ligasa y se continúa la ligación durante 5 horas en condiciones ajustadas para la ligación. Una muestra de la reacción de ligación se mezcla con tampón de carga de gel y se somete a electroforesis en un gel de acrilamida TBE al 20 % (Invitrogen, EE. UU.) junto con un control negativo sin polimerasa y un marcador de PM para confirmar la ligación de XNTP.

El intermedio de Xpandómero se trata con ácido acético al 80 % durante 5 horas a temperatura ambiente de acuerdo con 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 de fosforamidato entre el punto de unión de atadura y la desoxirribosa, que se confirma por electroforesis.

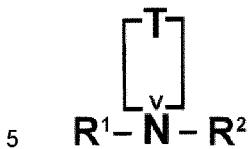
Los diversos modos de realización descritos anteriormente se pueden combinar para proporcionar otros modos de realización.

LISTADO DE SECUENCIAS

	<210> 4	
	<211> 32	
	<212> ADN	
5	<213> Secuencia artificial	
	<220>	
	<223> Secuencia de molde sintetizada	
10	<400> 4	
	tcttcgttcc tttttttttt atctaccgtc cgtccc	32
15	<210> 5	
	<211> 49	
	<212> PRT	
	<213> Secuencia artificial	
	<220>	
20	<223> Péptido sintetizado sobre soporte sólido	
	<400> 5	
	Glu Cys Gly Ala Gly Ala Gly Ala Gly Ala Gly Ala 1 5 10 15	
	Gly Ala Gly Ala Gly Ala Pro Ser Gly Ser Pro Ala Gly Ala 20 25 30	
	Gly Ala Gly Ala Gly Ala Gly Ala Gly Ala Gly Ala Gly Cys 35 40 45	
	Lys	
25	<210> 6	
	<211> 36	
	<212> ADN	
	<213> Secuencia artificial	
30	<220>	
	<223> Secuencia de molde sintetizada	
	<400> 6	
35	ttttttttttt tttttttttt atctaccgtc cgtccc	36

REIVINDICACIONES

1. Una construcción de sustrato monomérica, que tiene la siguiente estructura:



en la que

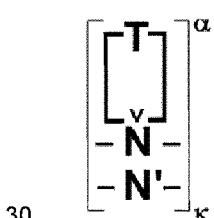
T representa la atadura, en la que la atadura es un polímero con un resto de extremo en cada uno de los dos extremos opuestos, y en la que el polímero se selecciona del siguiente segmento o segmentos: polietilenglicoles, poliglicoles, polipiridinas, polisocianuros, polisocianatos, poli(triarylmetil)metacrilatos, polialdehídos, polipirrolinonas, poliureas, fosfodiésteres de poliglicol, poliacrilatos, polimetacrilatos, poliacrilamidas, ésteres de polivinilo, poliestirenos, poliamidas, poliuretanos, policarbonatos, polibutiratos, polibutadienos, polibutirolactonas, polipirrolidinonas, polivinilfosfonatos, poliacetamidas, polisacáridos, polihialuranatos, poliamidas, poliimidas, poliésteres, polietilenos, polipropileno, poliestirenos, policarbonatos, politereftalatos, polisilanios, poliuretanos, poliéteres, poliaminoácidos, poliglicinas, poliprolinas, polilisina N-sustituida, polipéptidos, péptidos N-sustituidos de cadena lateral, glicina poli-N-sustituida, peptoides, péptidos sustituidos con carboxilo de cadena, homopéptidos, ácido policitidílico, ácido poliadénilico, ácido poliuridílico, politimidina, polifosfato, fosfodiésteres de polietilenglicol, análogos de polinucleótidos peptídicos, análogos de polinucleótidos treosilo, análogos de polinucleótidos glicol, análogos de polinucleótidos morfolinos, análogos de oligómeros de nucleótidos bloqueados, análogos de polipéptidos, polímeros aniónicos y polímeros catiónicos;

N representa un residuo de nucleósido;

V representa un sitio de escisión interno del resto de nucleósido que es un fosforamidato o un fosforotiolato; y

25 R¹ y R² representan grupos de extremo iguales o diferentes para la síntesis dirigida por molde de la hebra hija.

2. Una hebra hija híbrida que tiene la siguiente estructura:



en la que

T representa la atadura, en la que la atadura es un polímero con un resto de extremo en cada uno de los dos extremos opuestos, y en la que el polímero se selecciona del siguiente segmento o segmentos: polietilenglicoles, poliglicoles, polipiridinas, polisocianuros, polisocianatos, poli(triarylmetil)metacrilatos, polialdehídos, polipirrolinonas, poliureas, fosfodiésteres de poliglicol, poliacrilatos, polimetacrilatos, poliacrilamidas, ésteres de polivinilo, poliestirenos, poliamidas, poliuretanos, policarbonatos, polibutiratos, polibutadienos, polibutirolactonas, polipirrolidinonas, polivinilfosfonatos, poliacetamidas, polisacáridos, polihialuranatos, poliamidas, poliimidas, poliésteres, polietilenos, polipropileno, poliestirenos, policarbonatos, politereftalatos, polisilanios, poliuretanos, poliéteres, poliaminoácidos, poliglicinas, poliprolinas, polilisina N-sustituida, polipéptidos, péptidos N-sustituidos de cadena lateral, glicina poli-N-sustituida, peptoides, péptidos sustituidos con carboxilo de cadena, homopéptidos, ácido policitidílico, ácido poliadénilico, ácido poliuridílico, politimidina, polifosfato, fosfodiésteres de polietilenglicol, análogos de polinucleótidos peptídicos, análogos de polinucleótidos treosilo, análogos de polinucleótidos glicol, análogos de polinucleótidos morfolinos, análogos de oligómeros de nucleótidos bloqueados, análogos de polipéptidos, polímeros aniónicos y polímeros catiónicos;

N representa un residuo de nucleósido;

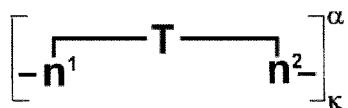
50 N' representa un resto de nucleótido de la hebra molde a la que N es complementario;

V representa un sitio de escisión interno dentro de la cadena principal de fosforribosilo, cuyo sitio de escisión es una modificación de la cadena principal puente que es un fosforamidato o un fosforotiolato;

κ representa la κ ^{ésima} subunidad en una cadena de m subunidades, donde m es un número entero mayor que diez; y

α representa una especie de un motivo de subunidad seleccionado de una colección de motivos de subunidad, en el que cada una de las especies es complementaria a la secuencia de nucleótidos contigua de una porción del ácido nucleico diana.

- 5 3. Un Xpandómero que tiene la siguiente estructura:



10 en el que un Xpandómero es un polímero sustituto de longitud extendida formado por síntesis dirigida por molde que conserva la información genética original del ácido nucleico diana, al tiempo que incrementa la separación lineal de los elementos individuales de la secuencia,

T representa la atadura, en la que la atadura es un polímero con un resto de extremo en cada uno de los dos extremos opuestos, y en la que el polímero se selecciona del siguiente segmento o segmentos: polietilenglicoles, poliglicoles, 15 polipiridinas, polisocianuros, polisocianatos, poli(triarilmetyl)metacrilatos, polialdehídos, polipirrolinonas, poliureas, fosfodiésteres de poliglicol, poliacrilatos, polimetacrilatos, poliacrilamidas, ésteres de polivinilo, poliestirenos, poliamidas, poliuretanos, policarbonatos, polibutiratos, polibutadienos, polibutirolactonas, polipirrolidinonas, polivinilfosfonatos, poliacetamidas, polisacáridos, polihialuranatos, poliamidas, poliimidas, poliésteres, polietilenos, polipropilenos, 20 poliestirenos, policarbonatos, politereftalatos, polisilanos, poliuretanos, poliéteres, poliaminoácidos, poliglicinas, poliprolinas, polilisina N-sustituida, glicina poli-N-sustituida, peptoides, ácido policitidílico, ácido poliadénílico, ácido poliuridílico, politimidina, polifosfato, fosfodiésteres de polietilenglicol, polímeros aniónicos y polímeros catiónicos;

n¹ y n² representan una primera porción y una segunda porción, respectivamente, de un residuo de nucleósido;

25 κ representa la κésima subunidad en una cadena de m subunidades, donde m es un número entero mayor que diez; y

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

30 4. La construcción monomérica, la hebra hija híbrida o el Xpandómero de una cualquiera de las reivindicaciones anteriores, en la que la atadura comprende una o más construcciones de indicadores que comprenden uno o más indicadores que pueden producir señales detectables, en la que la(s) señal(es) detectable(s) codifica(n) información de secuencia, en la que el indicador incluye al menos uno de los siguientes: donante o acceptor resonante de FRET, tinte, punto cuántico, microesfera, dendrímero, fluoróforo de conversión ascendente, partícula magnética, dispersor de electrones, boro, masa, microesfera de oro, resonancia magnética, grupo ionizable, grupo polar, grupo hidrófobo, 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, aloficocianina, 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, eritrosina, ácido cumárico, umbelifera, Texas red rodaina, tetrametil rodamina, Rox, 7-nitrobenzo-1-oxa-1-diazol (NBD), oxazol, tiazol, pireno, fluoresceína o lantamidas, radioisótopos, ³³P, ³H, ¹⁴C, ³⁵S, ¹²⁵I, ³²P or ¹³¹I, etidio, europio, rutenio, samario, marcadores quimioluminiscentes, 35 marcadores de espín, enzimas, peroxidásas, fosfatásas alcalinas, beta-galactosidas, oxidásas, fragmentos de anticuerpos, ligandos de afinidad, oligómeros, haptenos, aptámeros y pirimidinas modificados en la posición C5 o purinas 40 modificadas en la posición N7, en los que los grupos modificadores de masa son halógeno, éter, poliéter, alquilo, éster, poliéster o del general tipo XR, en el que X es un grupo de enlace y R es un grupo modificador de masa.

5. La construcción monomérica, la hebra hija híbrida o el Xpandómero de la reivindicación 4, en la que el indicador produce una señal detectable que es una señal de sincronización.

50 6. La construcción monomérica, la hebra hija híbrida o el Xpandómero de una cualquiera de las reivindicaciones 4 o 5, en la que el indicador produce una señal detectable que es una velocidad de translocación a través de un detector de nanoporos.

55 7. La construcción monomérica, la hebra hija híbrida o el Xpandómero de una cualquiera de las reivindicaciones 4 a 6, en la que la unión está construida con un polímero de polietilenglicol, un polímero de óxido de polietileno o un polímero de metoxipolietilenglicol.

8. La construcción monomérica o la hebra hija híbrida de cualquiera de las reivindicaciones anteriores, en la que el sitio 60 de escisión interno V es una modificación de la cadena principal puente que es un fosforamidato.

Fig. 1A

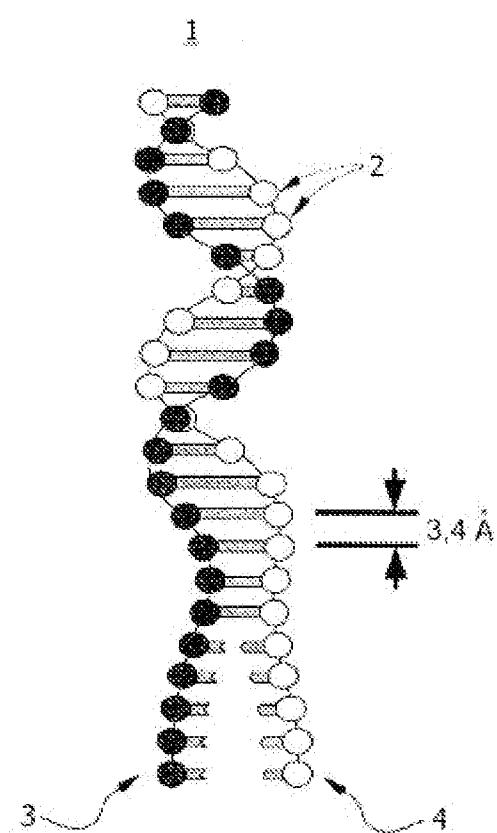


Fig. 1B

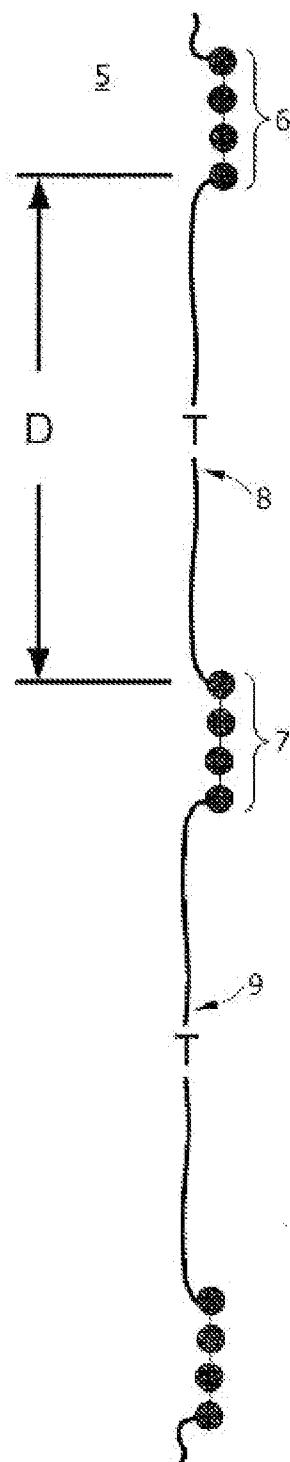


Fig. 2A

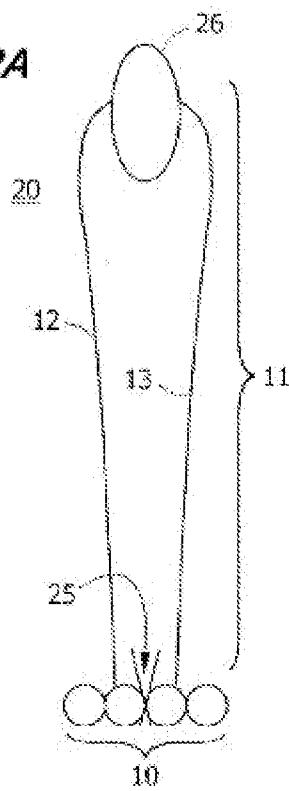


Fig. 2B

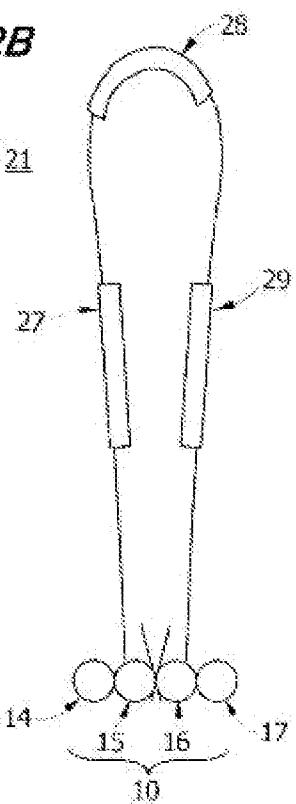


Fig. 2C

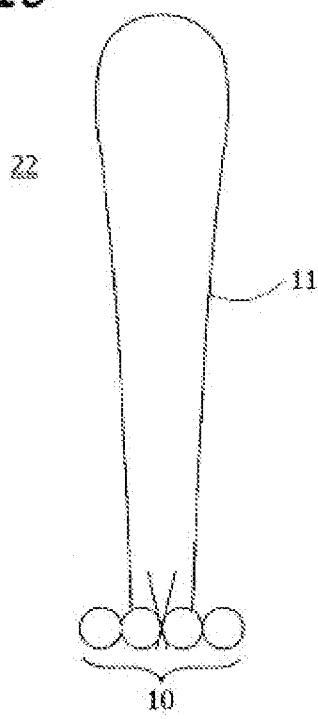


Fig. 2D

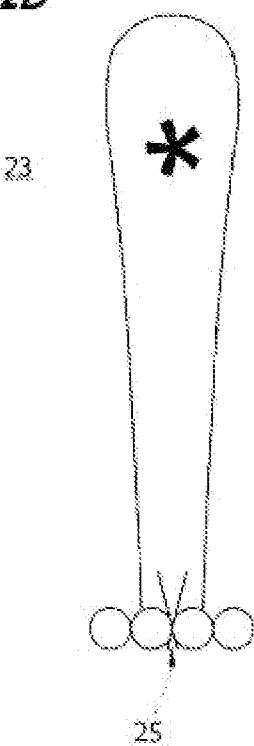


Fig. 3A

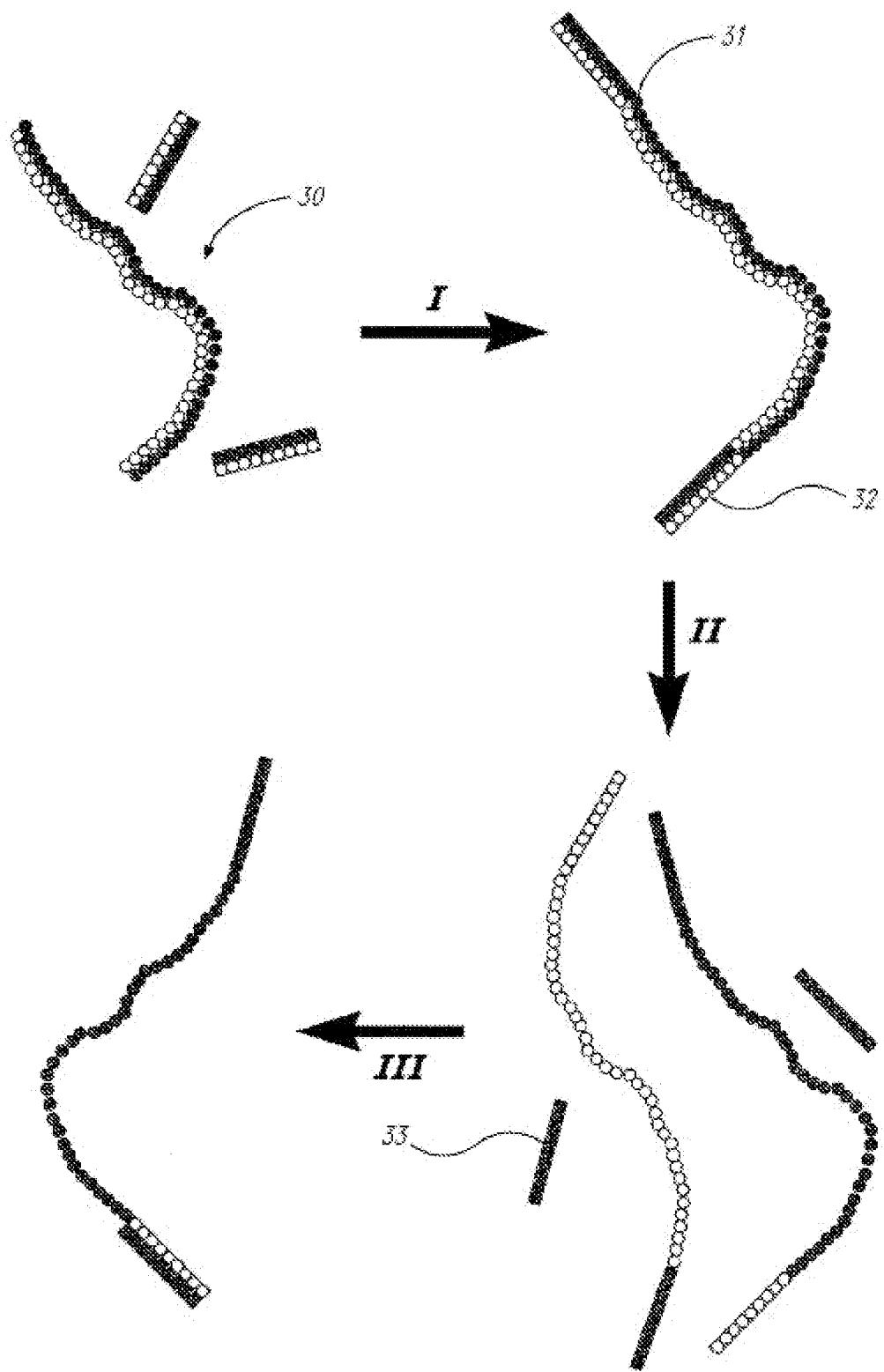


Fig. 3B

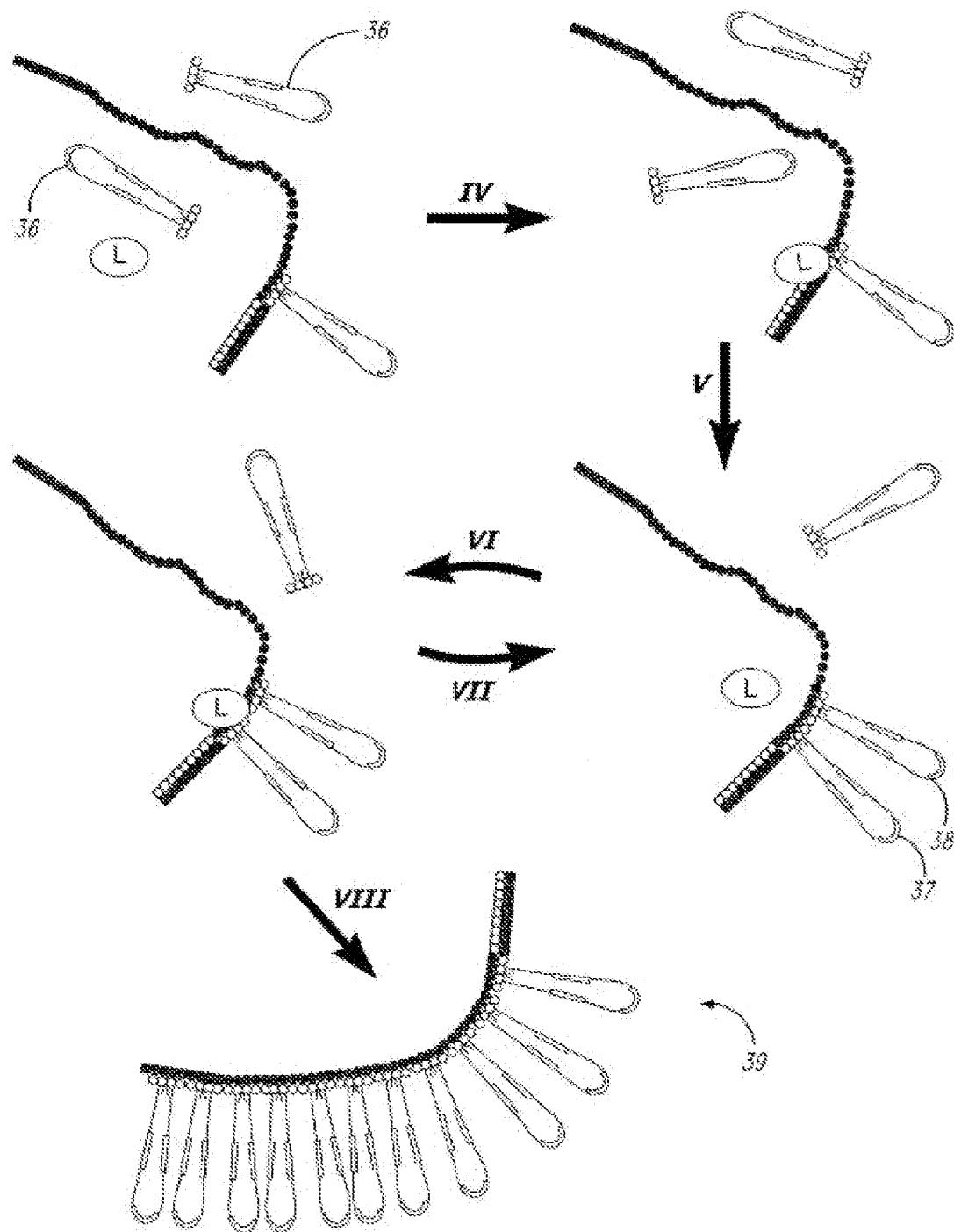


Fig. 3C

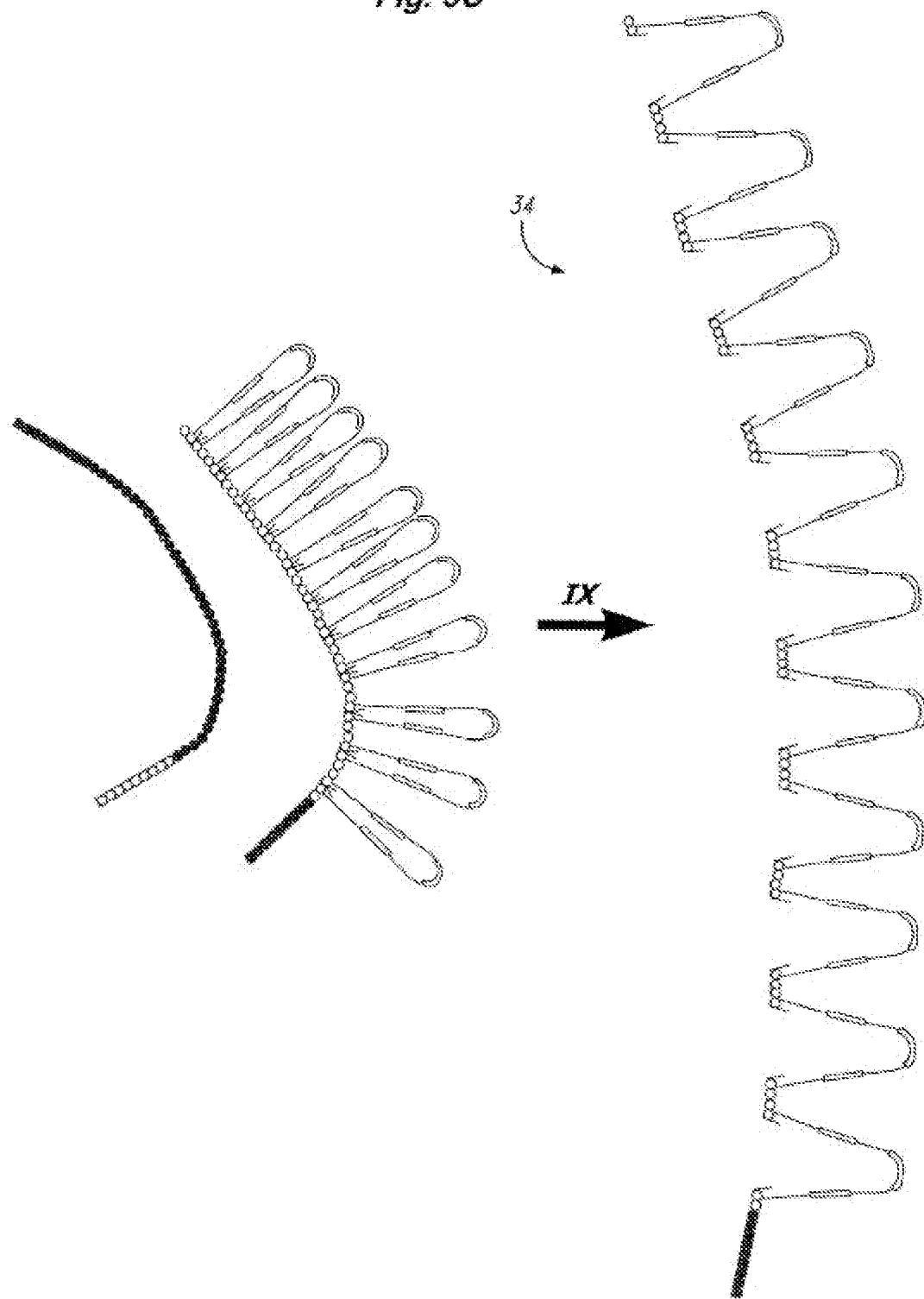
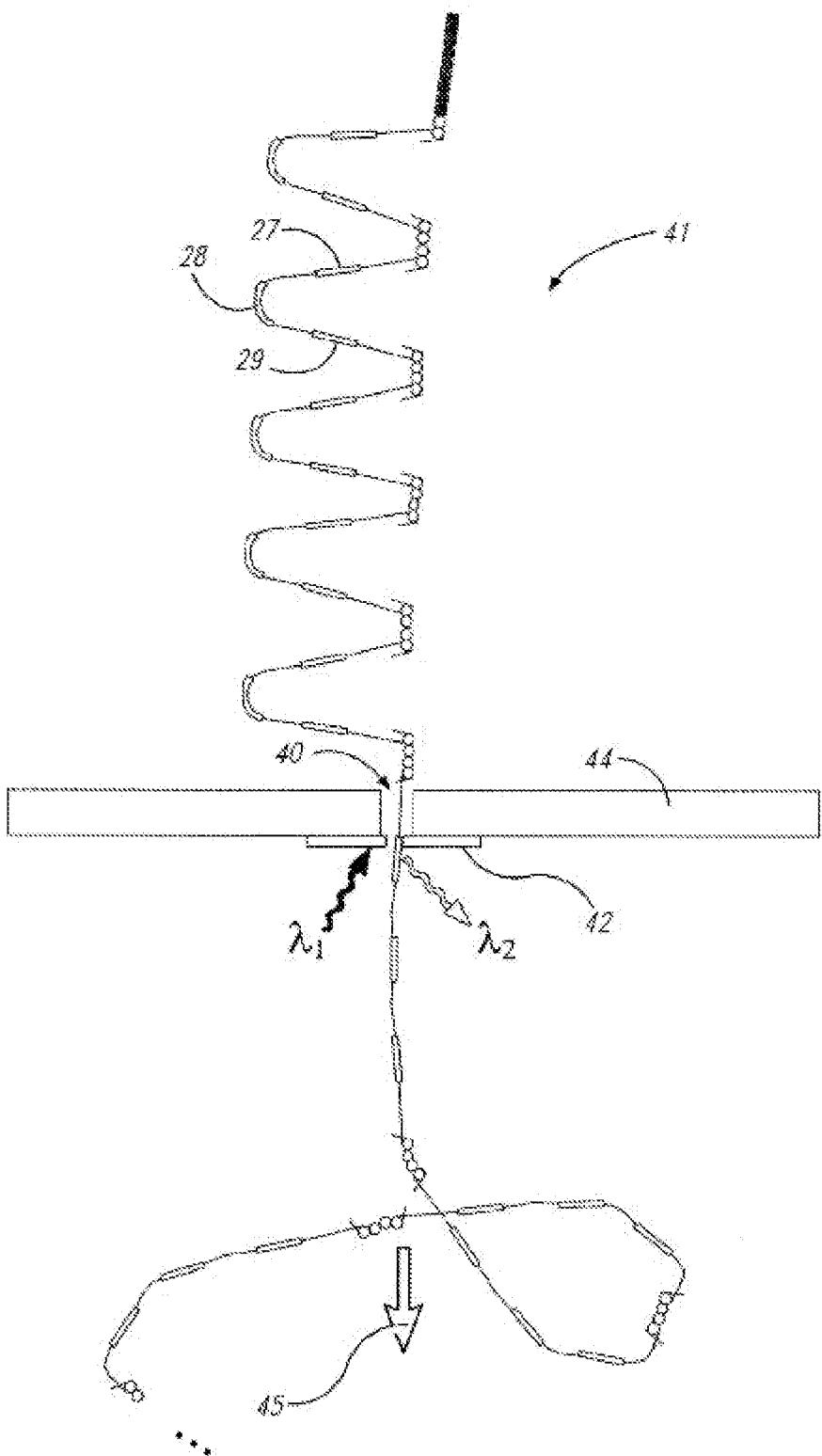


Fig. 4



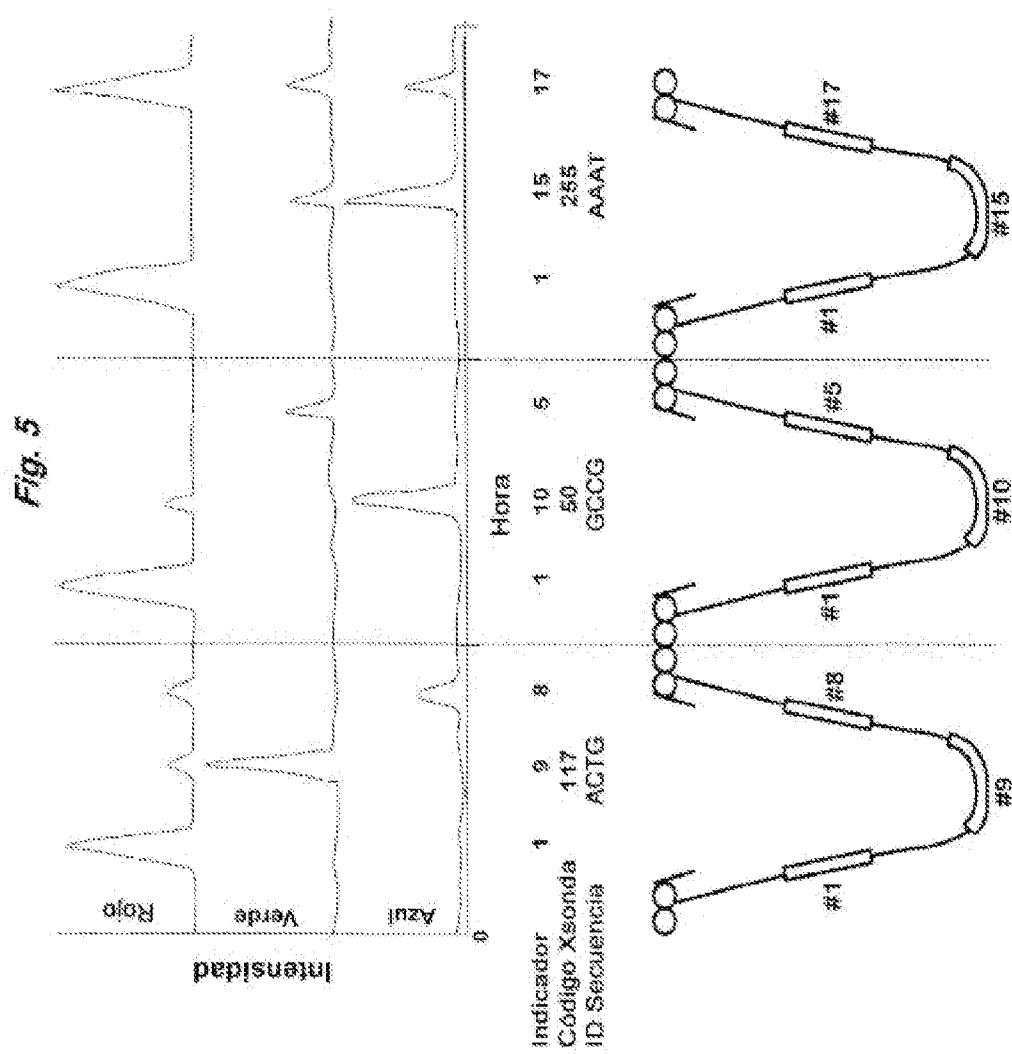


FIG. 6

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

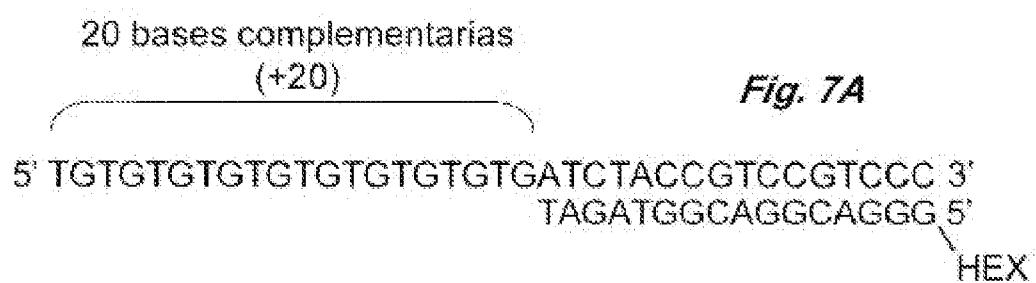


Fig. 7A

Fig. 7B

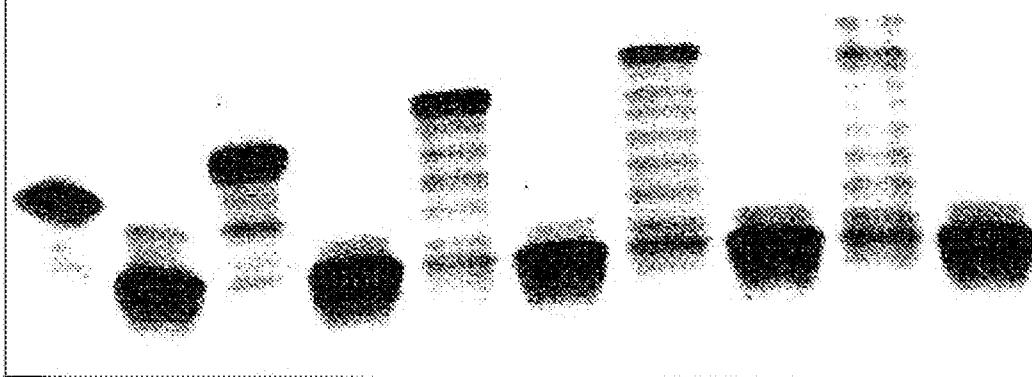


Fig. 7C

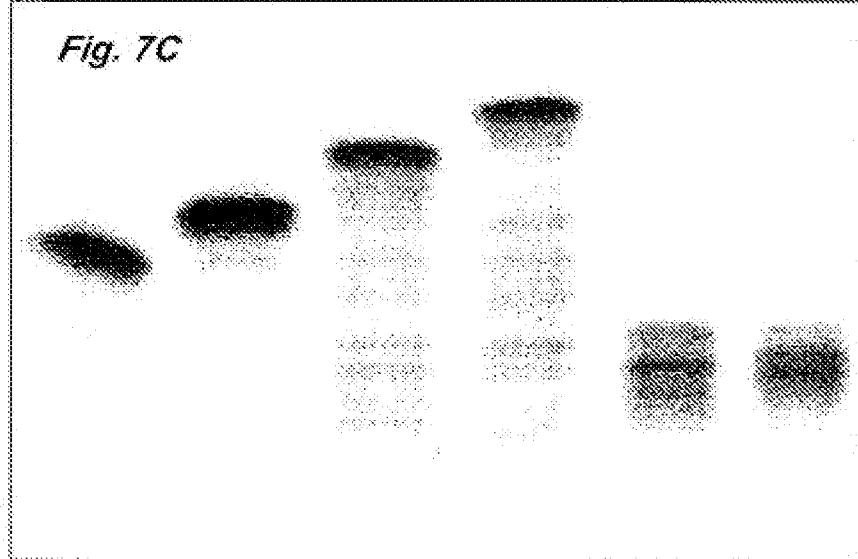


Fig. 7D

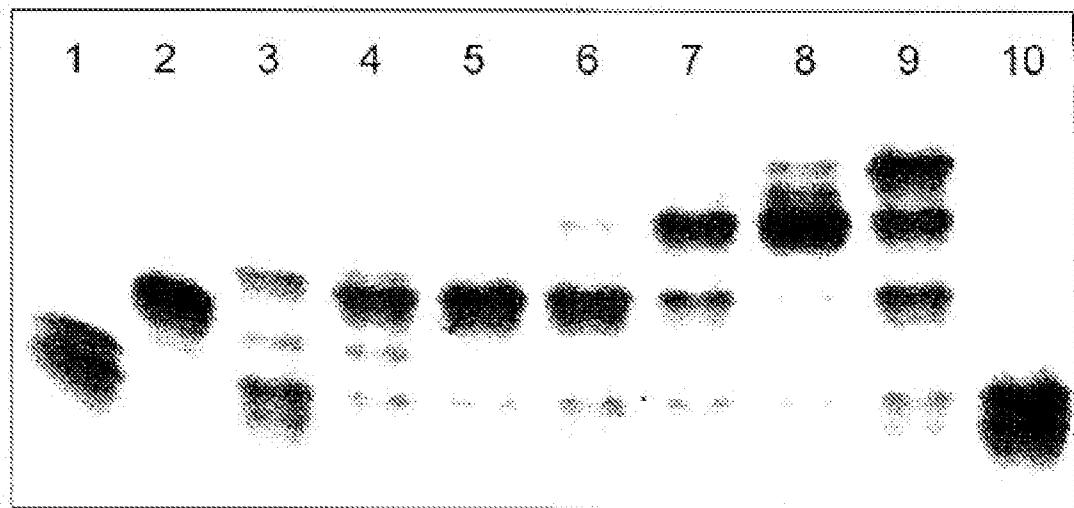


Fig. 7E

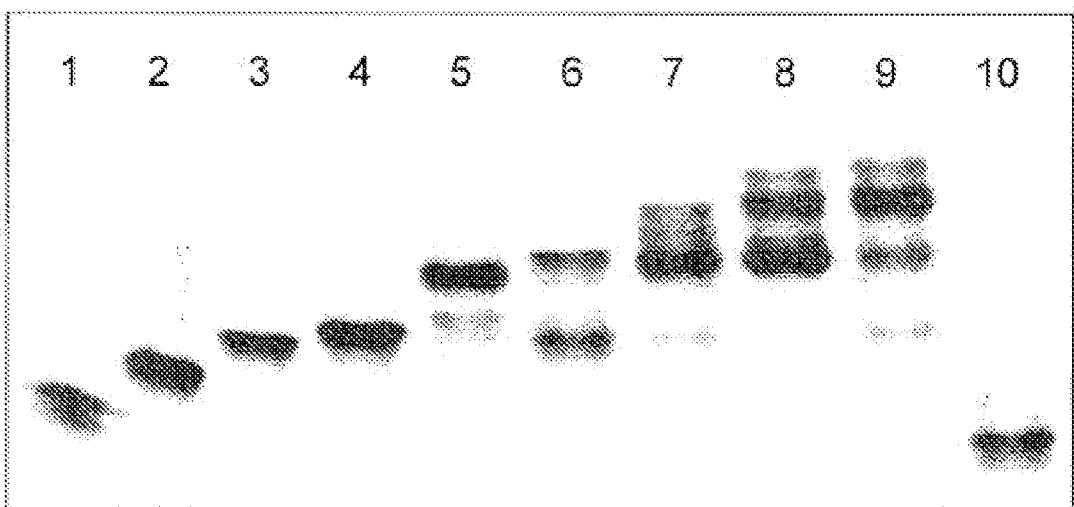


Fig. 8

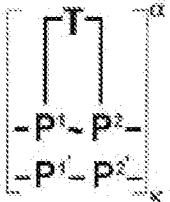
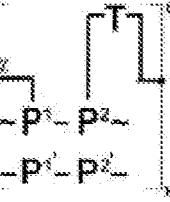
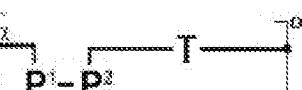
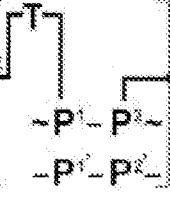
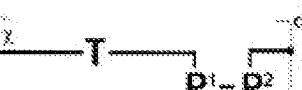
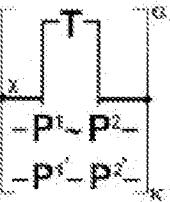
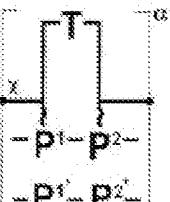
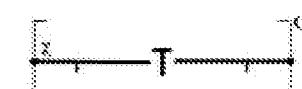
CLASE	CONSTRUCCIÓN DE SUSTRATO OLIGOMERICO	INTERMEDIO DE XPANDÓMERO	XPANDÓMERO
I.	$R^1-P^1-P^2-R^2$		
II.	$R^1-P^1-P^2-R^2$		
III.	$R^1-P^1-P^2-R^2$		
IV.	$R^1-P^1-P^2-R^2$		
V.	$R^1-P^1-P^2-R^2$		

Fig. 9

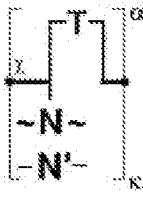
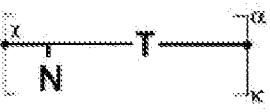
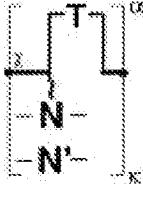
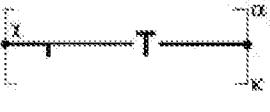
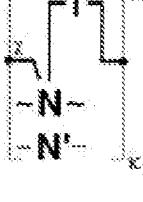
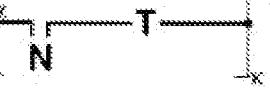
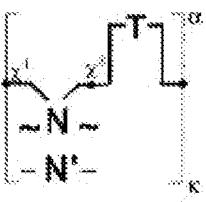
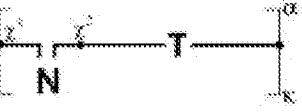
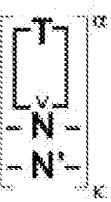
CLASE	CONSTRUCCIÓN DE SUSTRATO MONOMÉRICA	INTERMEDIO DE XPANDÓMERO	XPANDÓMERO
VI.	R^1-N-R^2		
VII.	R^1-N-R^2		
VIII.	R^1-N-R^2		
IX.	R^1-N-R^2 $\delta_1 \curvearrowright \delta_2$		
X.	R^1-N-R^2		

Fig. 10A

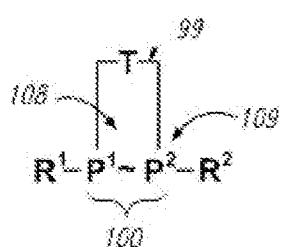


Fig. 10B

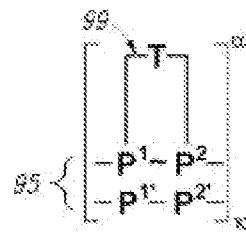


Fig. 10C

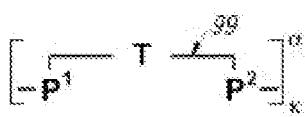


Fig. 10D

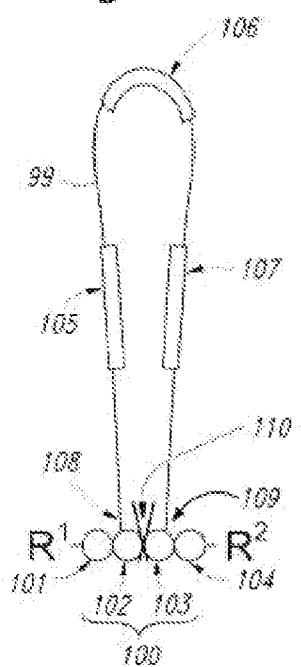


Fig. 10E

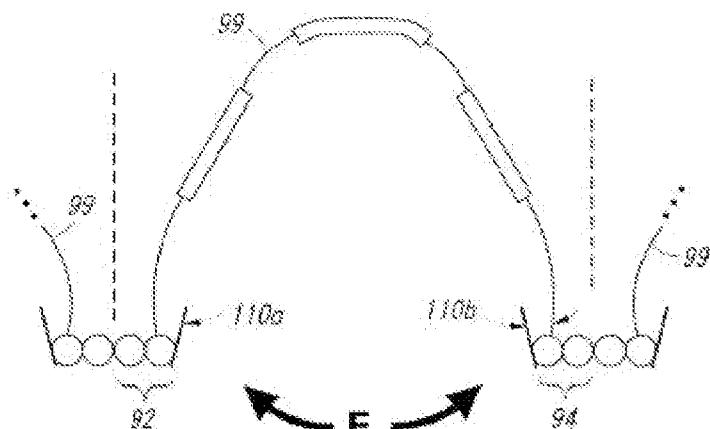


Fig. 11

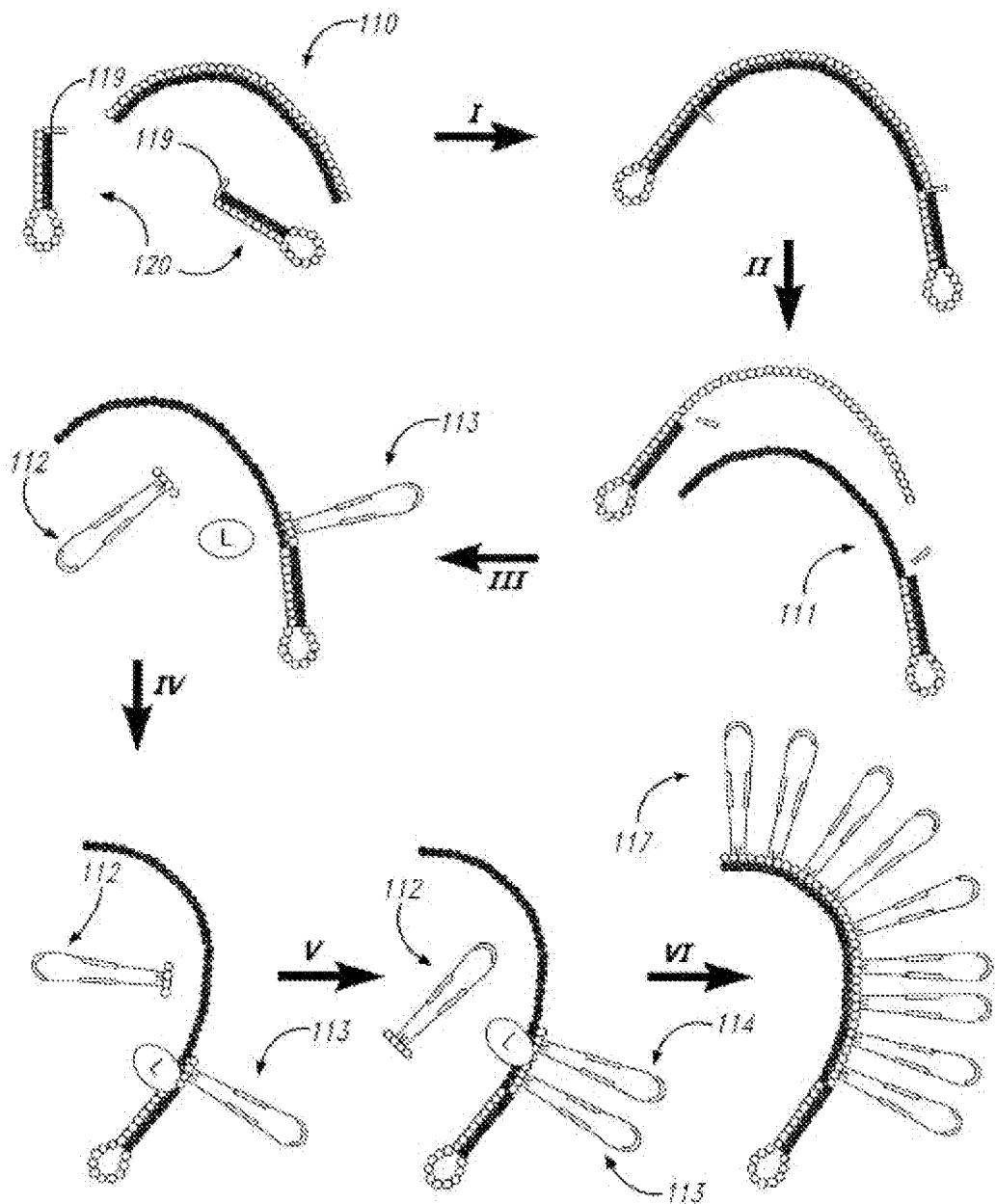


Fig. 12

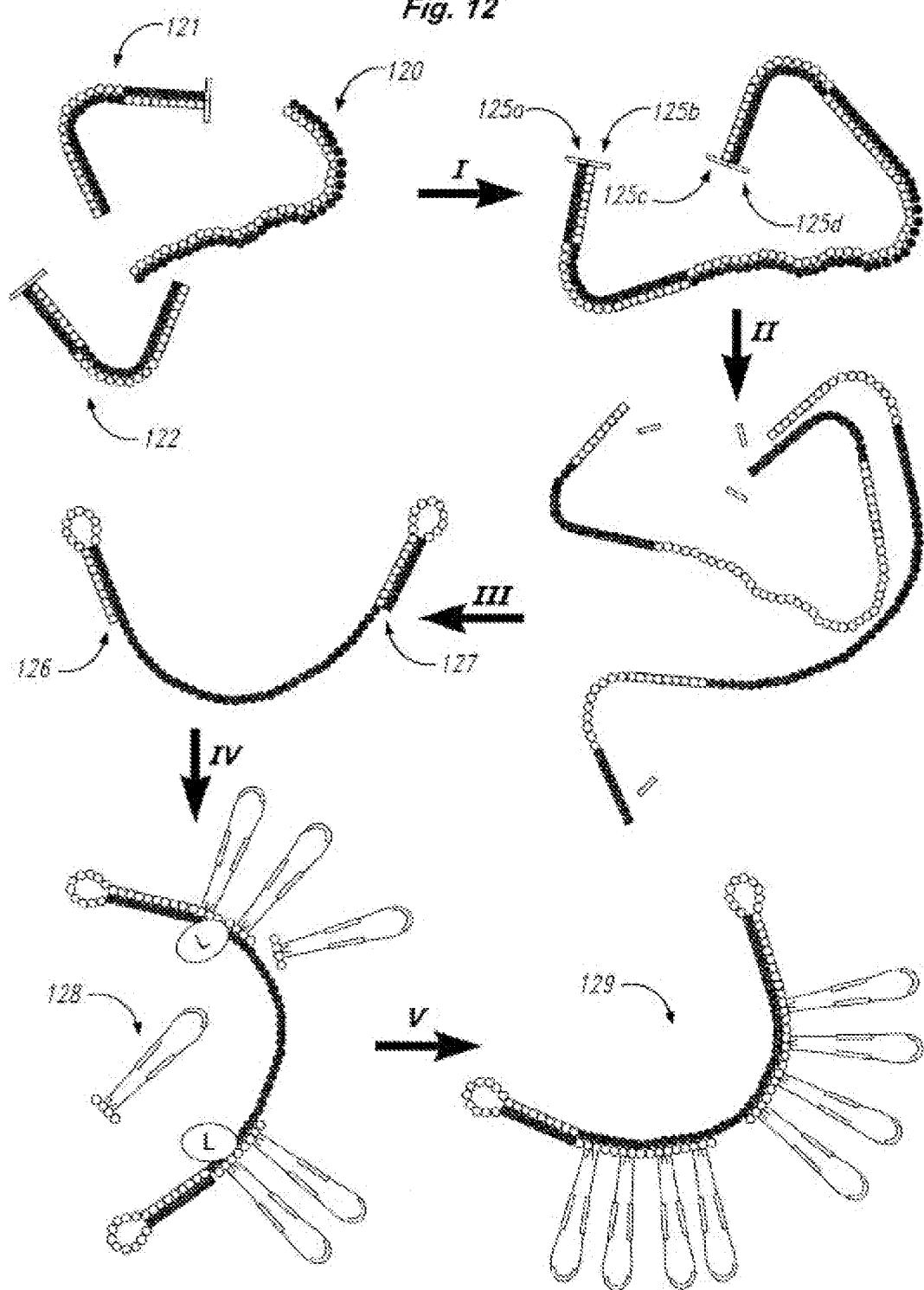


Fig. 13

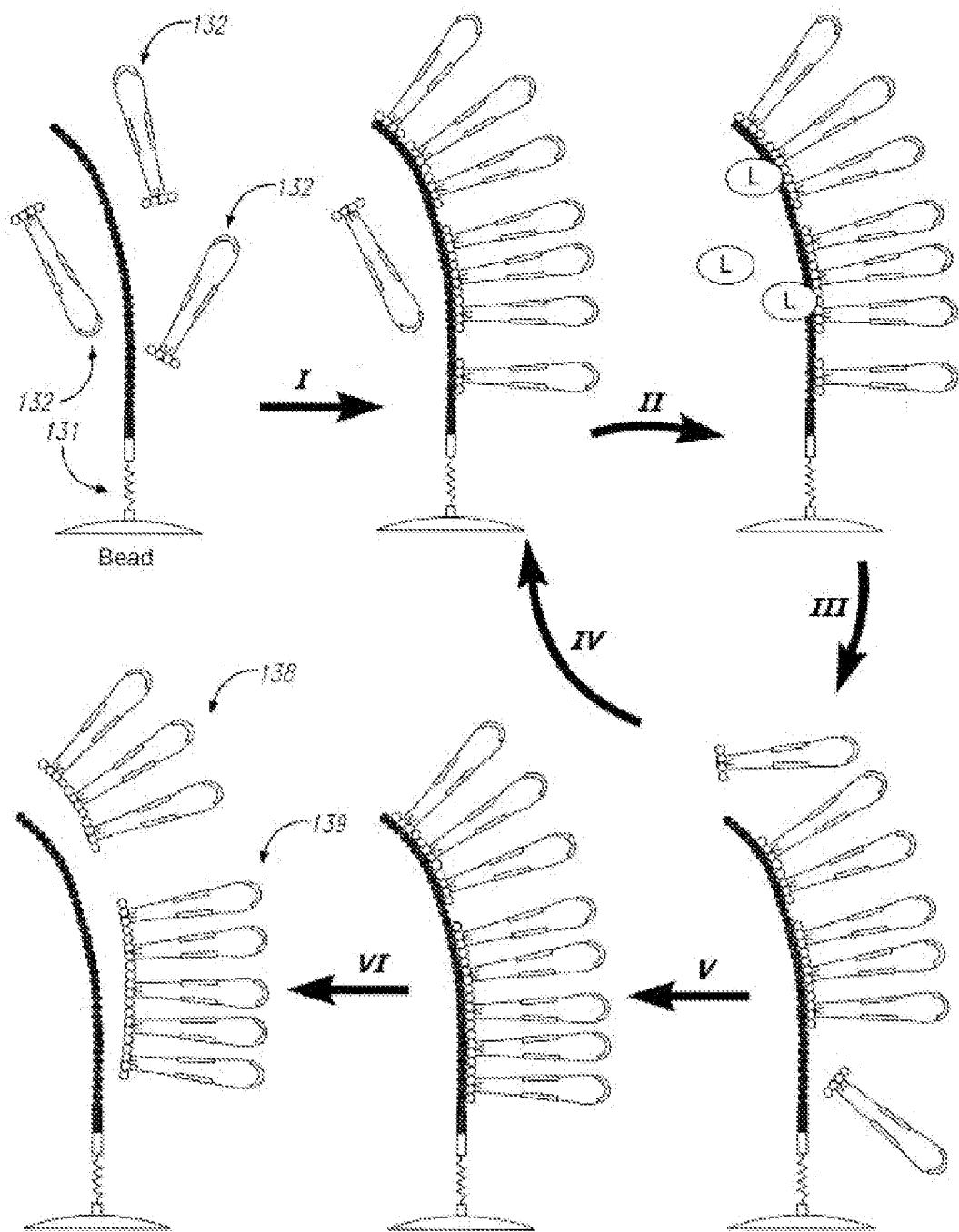


Fig. 14

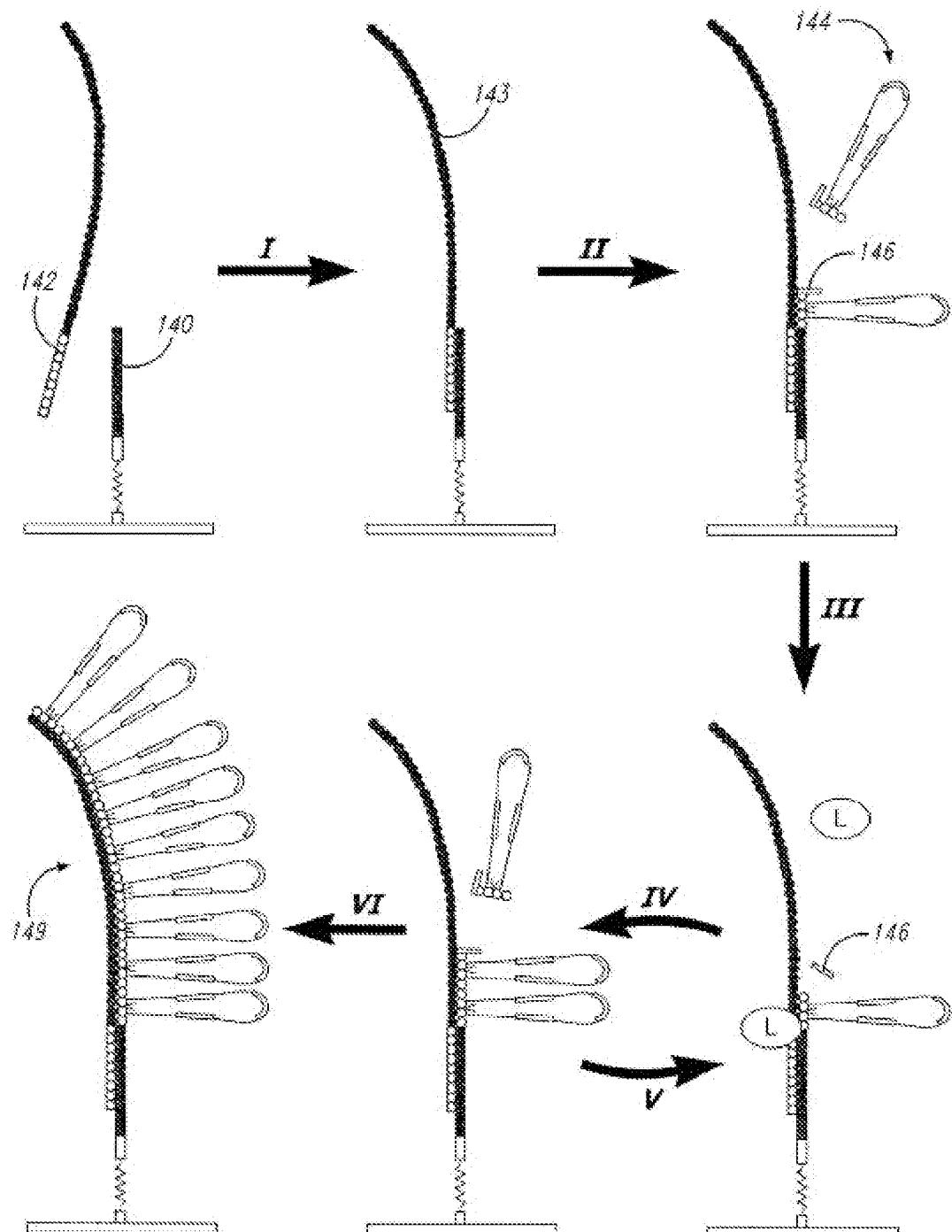


Fig. 15

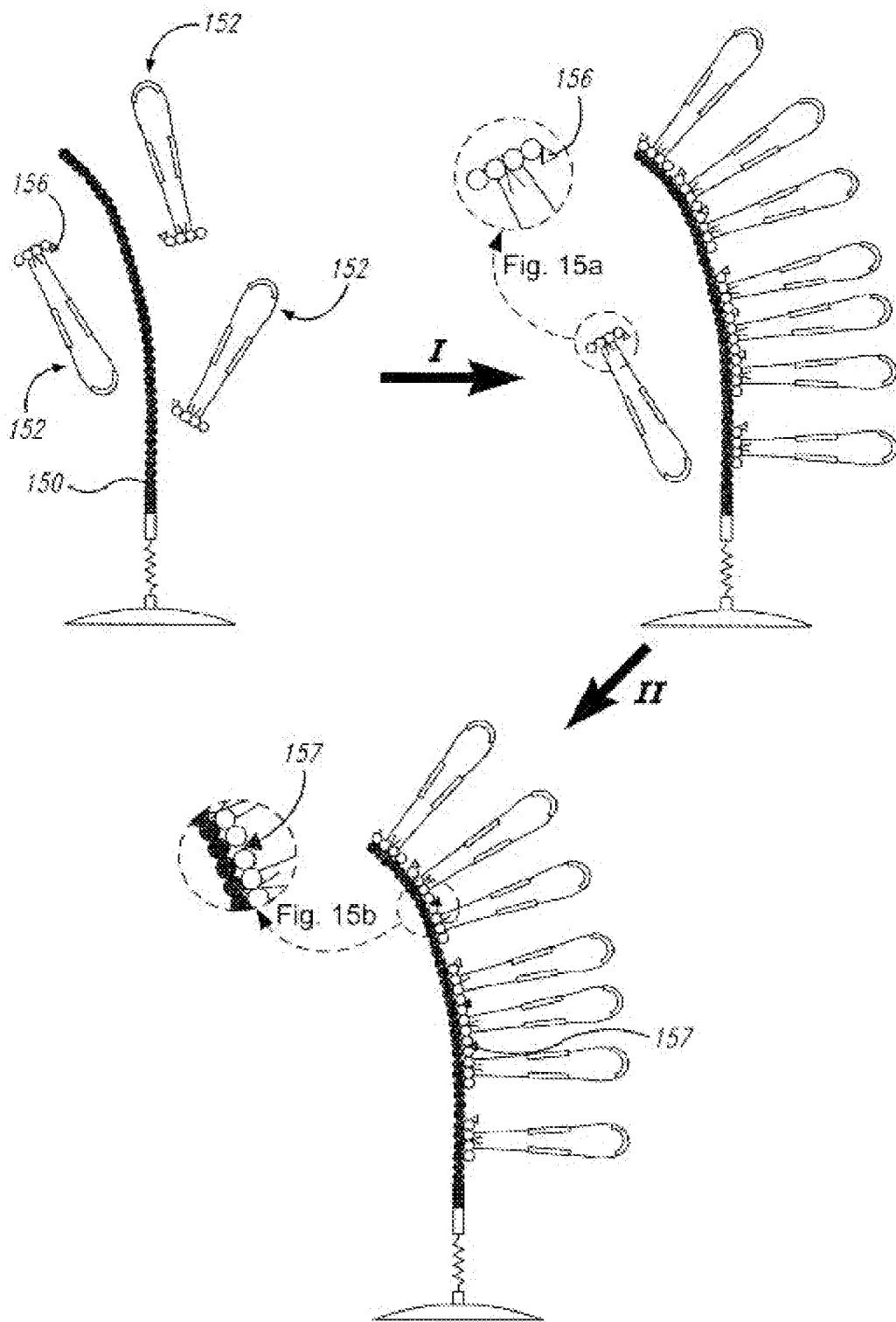


Fig. 16

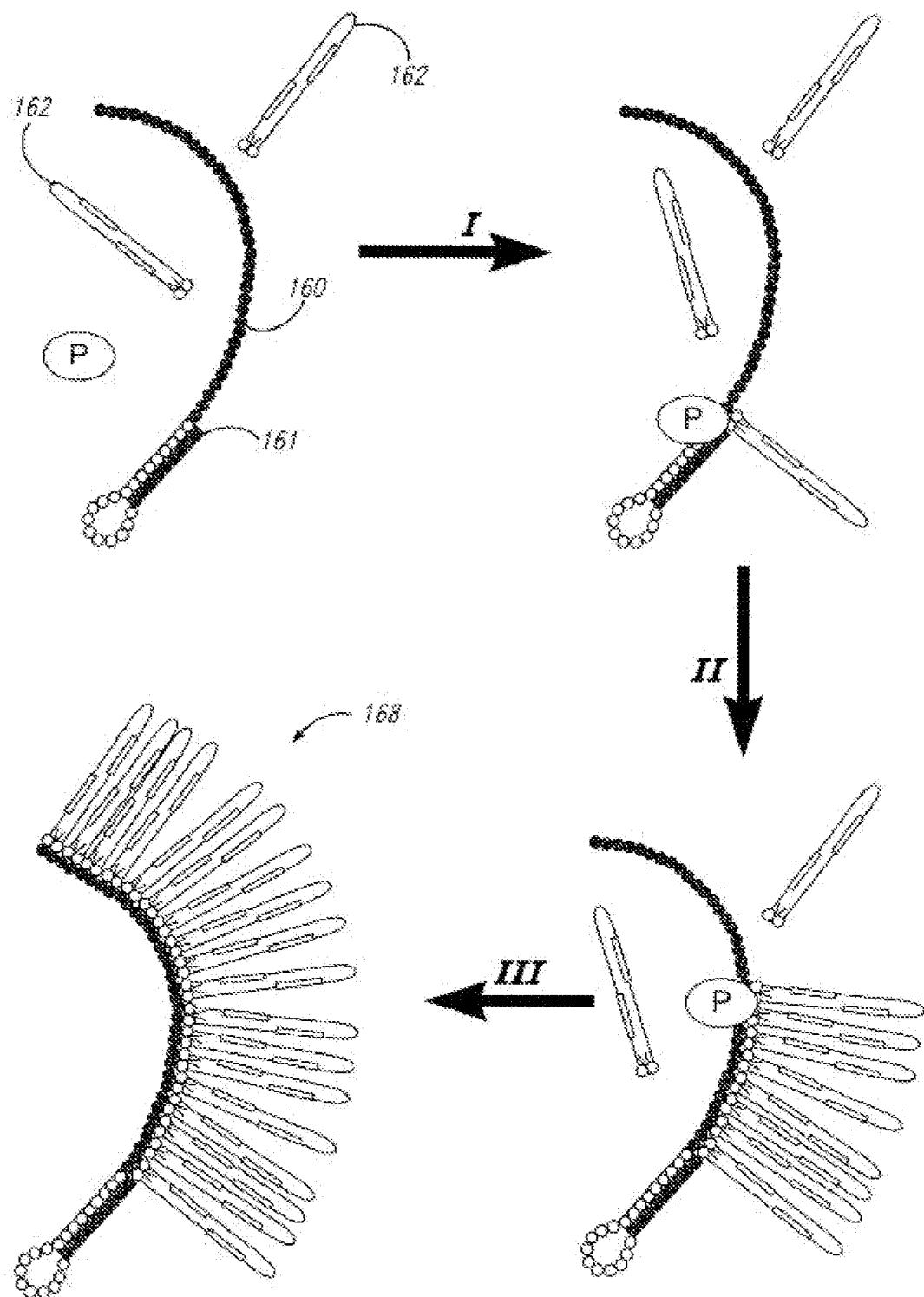


Fig. 17

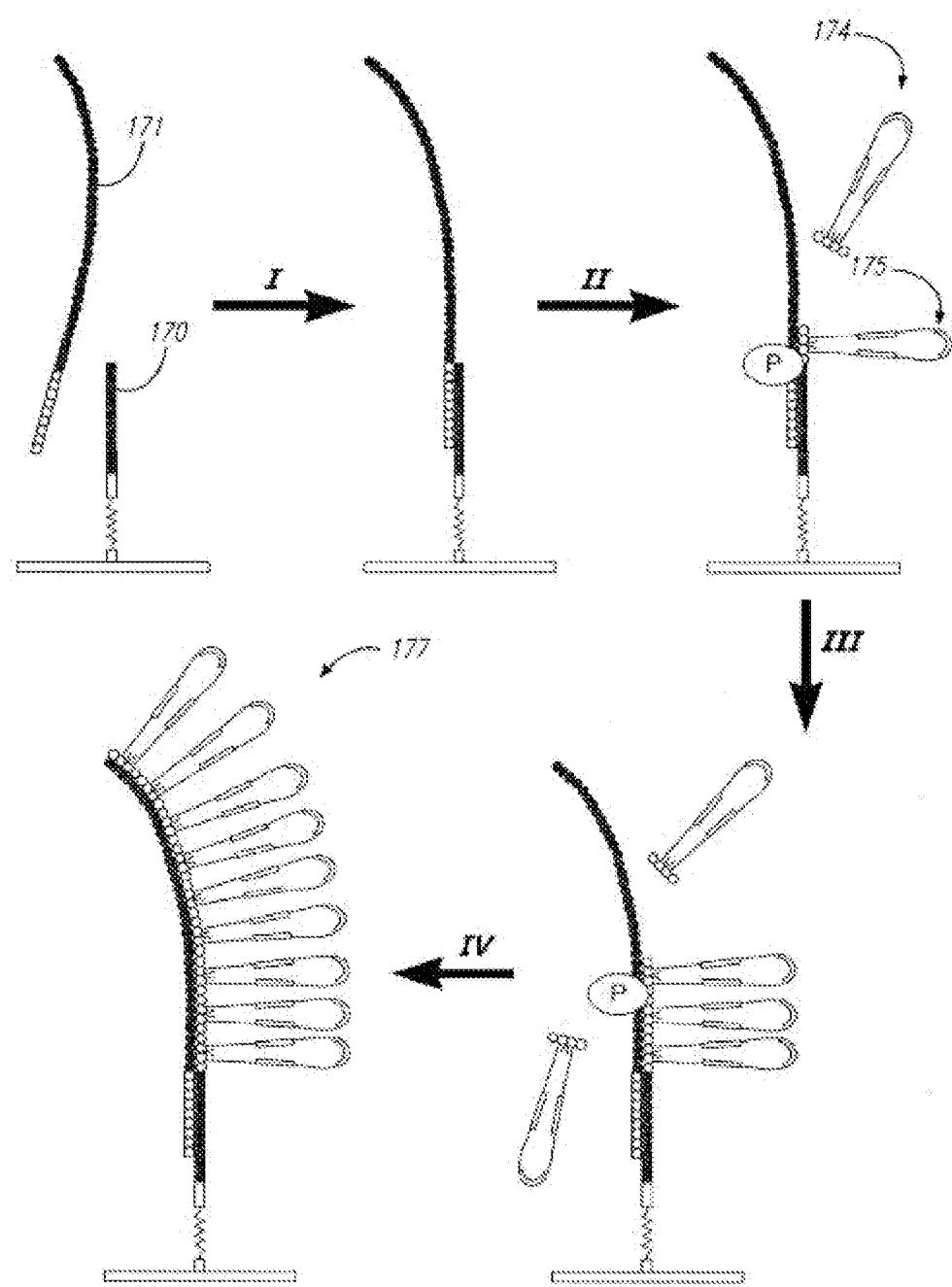


Fig. 18A

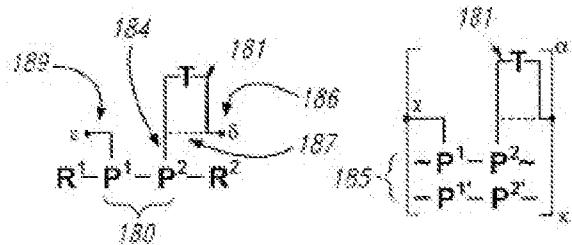


Fig. 18B

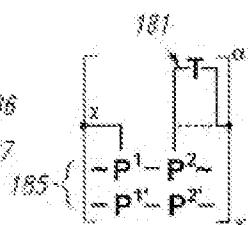


Fig. 18C

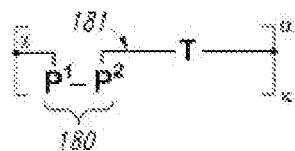


Fig. 18D

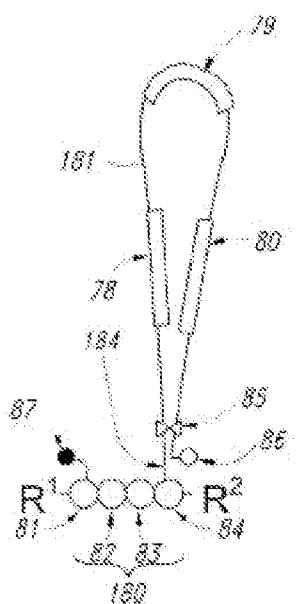


Fig. 18E

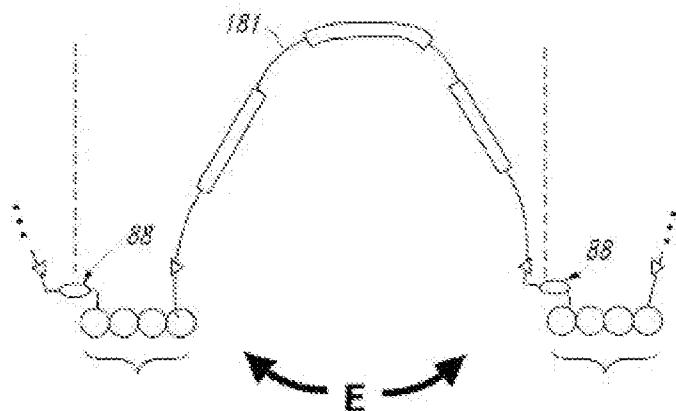


FIG. 19A

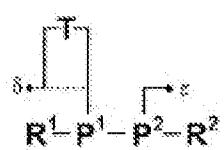


FIG. 19B

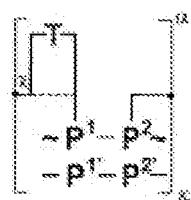


FIG. 19C

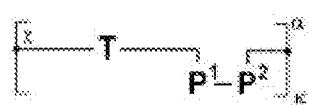


FIG. 19D

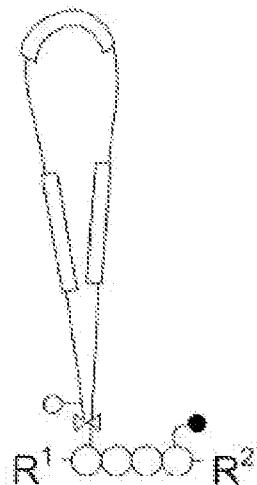
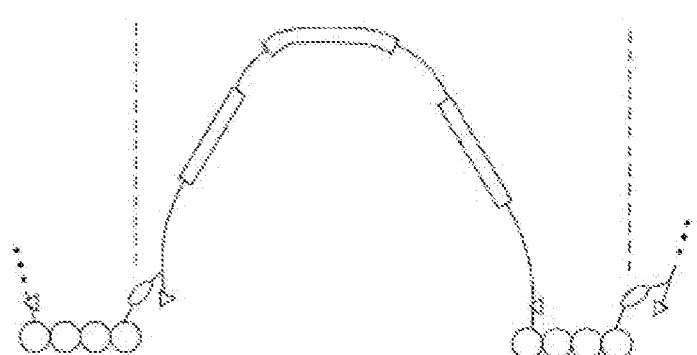


FIG. 19E



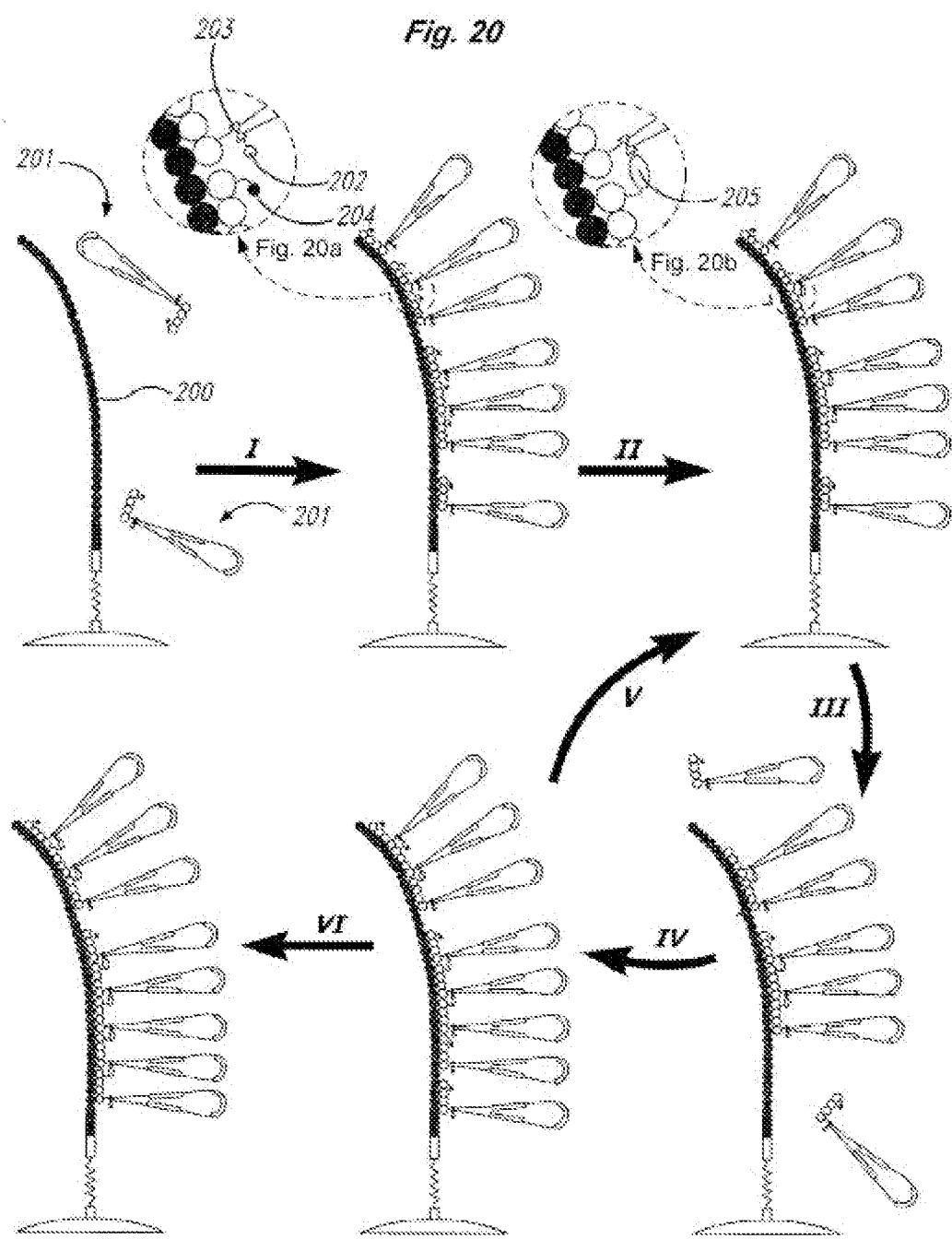


Fig. 21

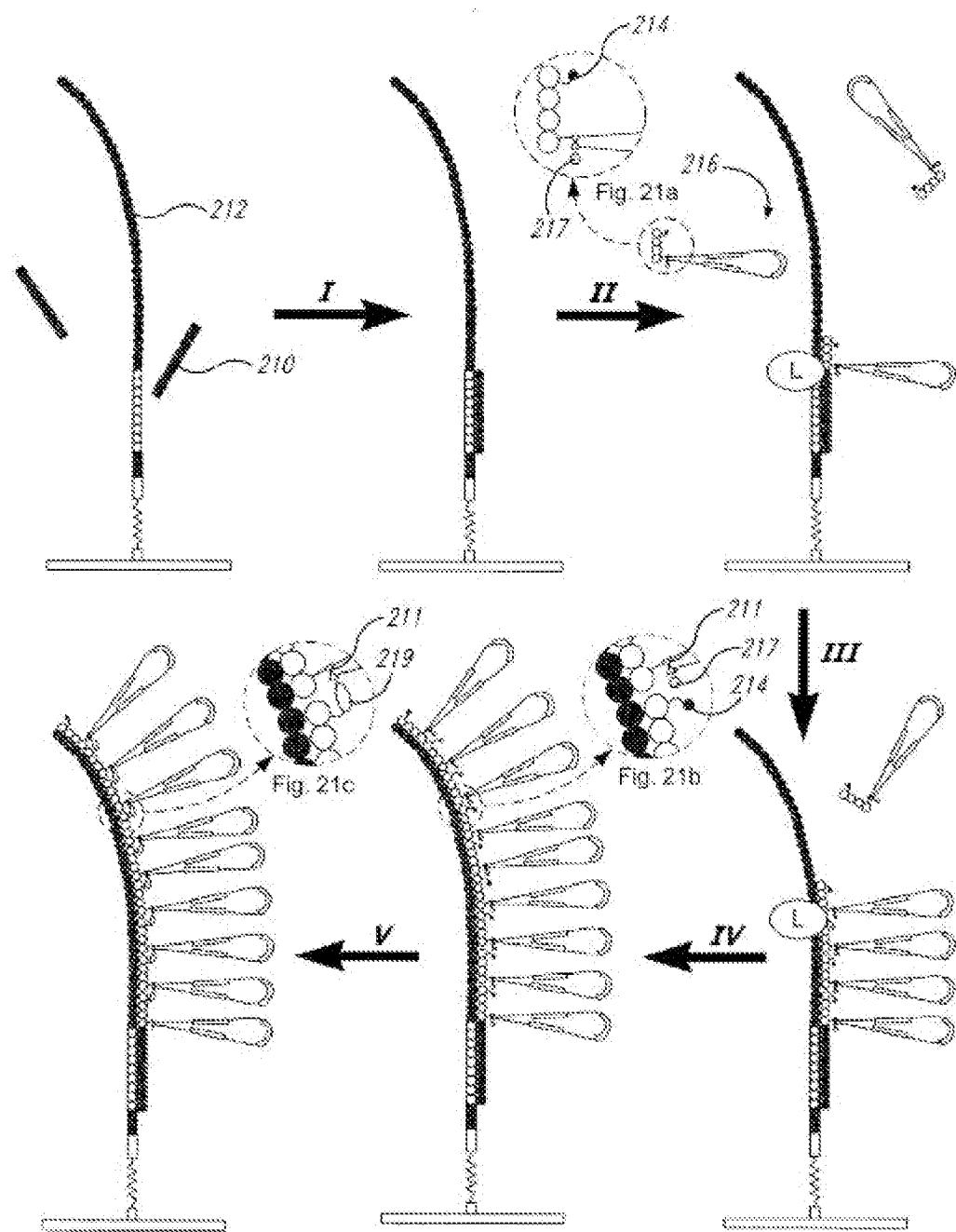


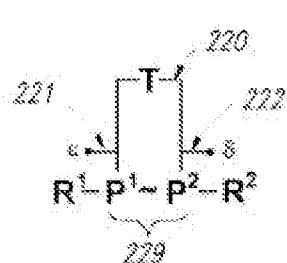
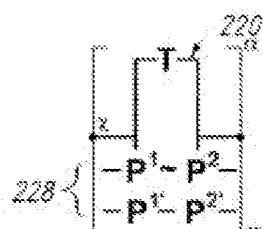
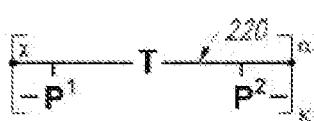
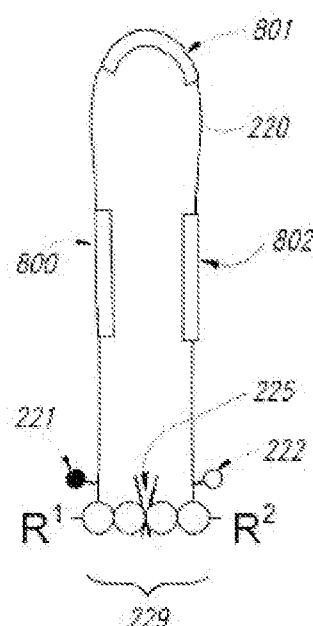
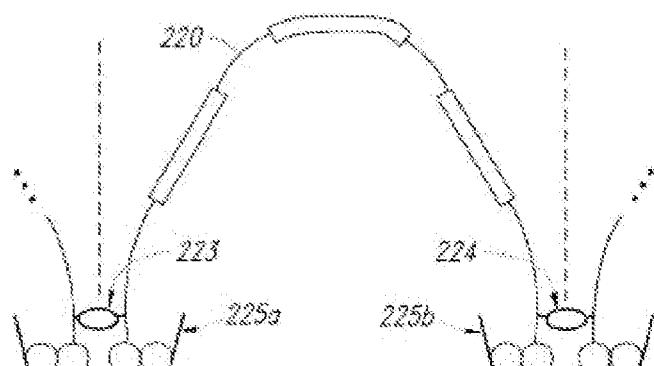
Fig. 22A*Fig. 22B**Fig. 22C**Fig. 22D**Fig. 22E*

Fig. 23A

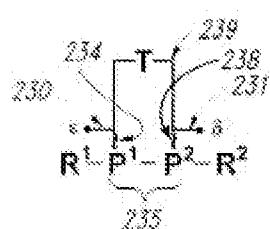


Fig. 23B

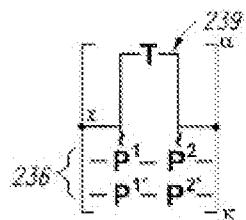


Fig. 23C

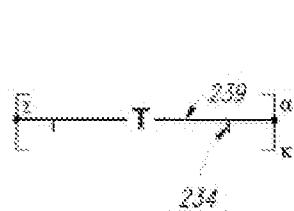


Fig. 23D

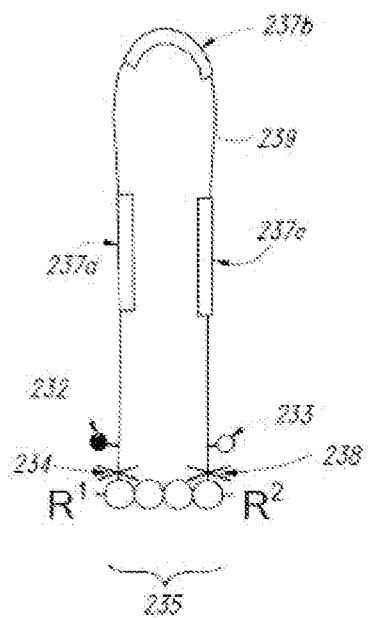


Fig. 23E

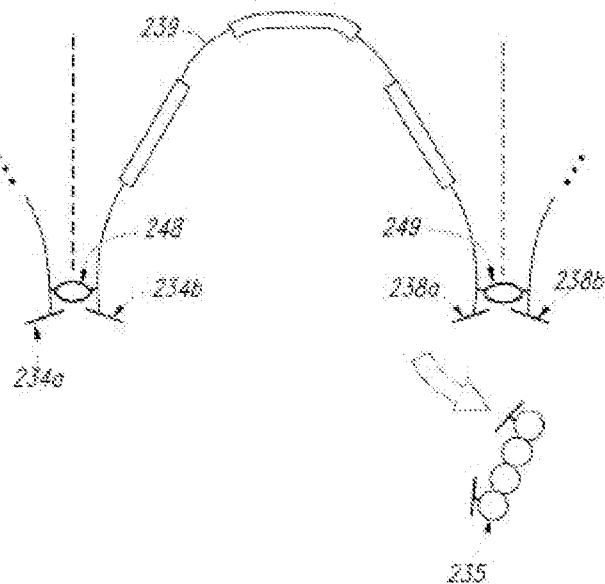


Fig. 24

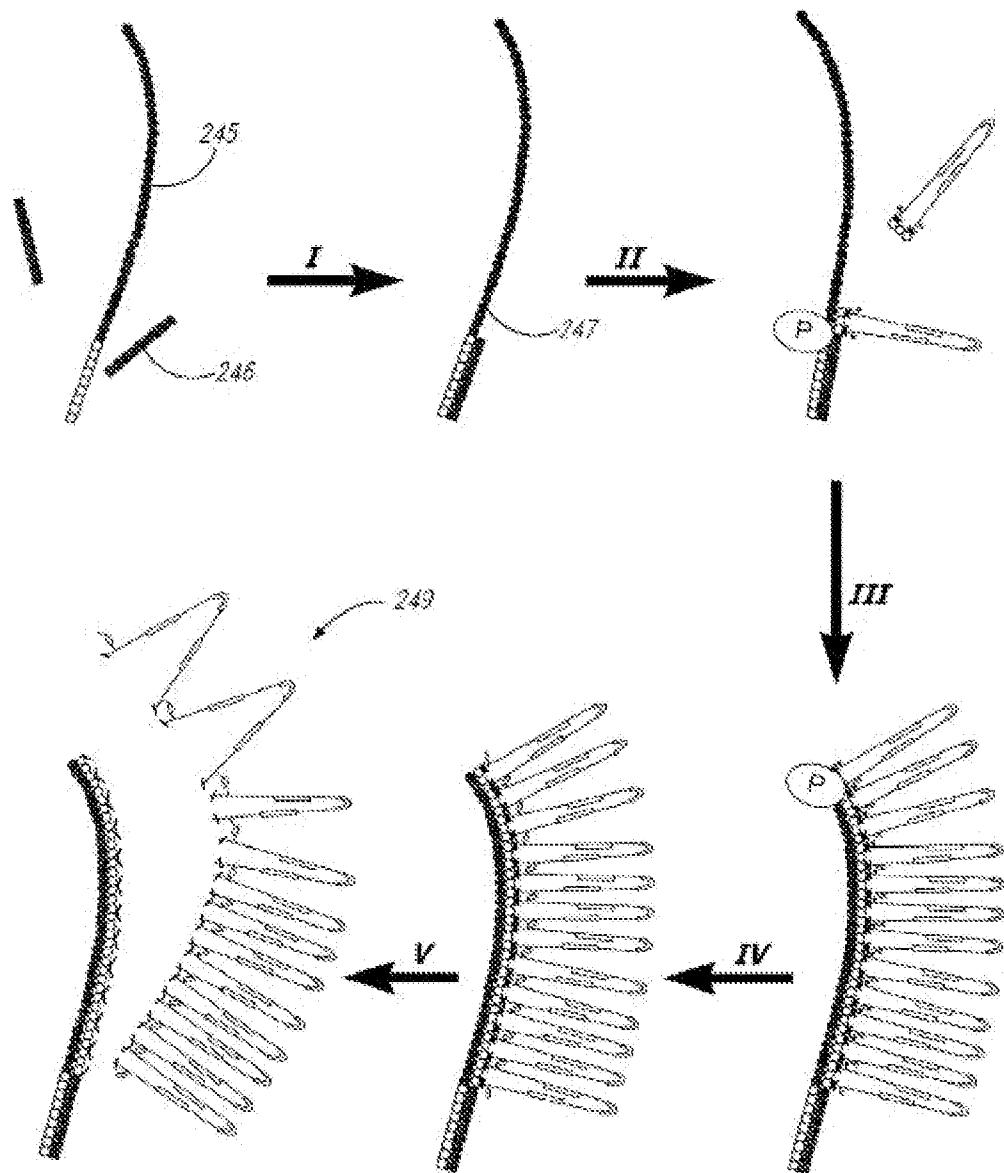


Fig. 25

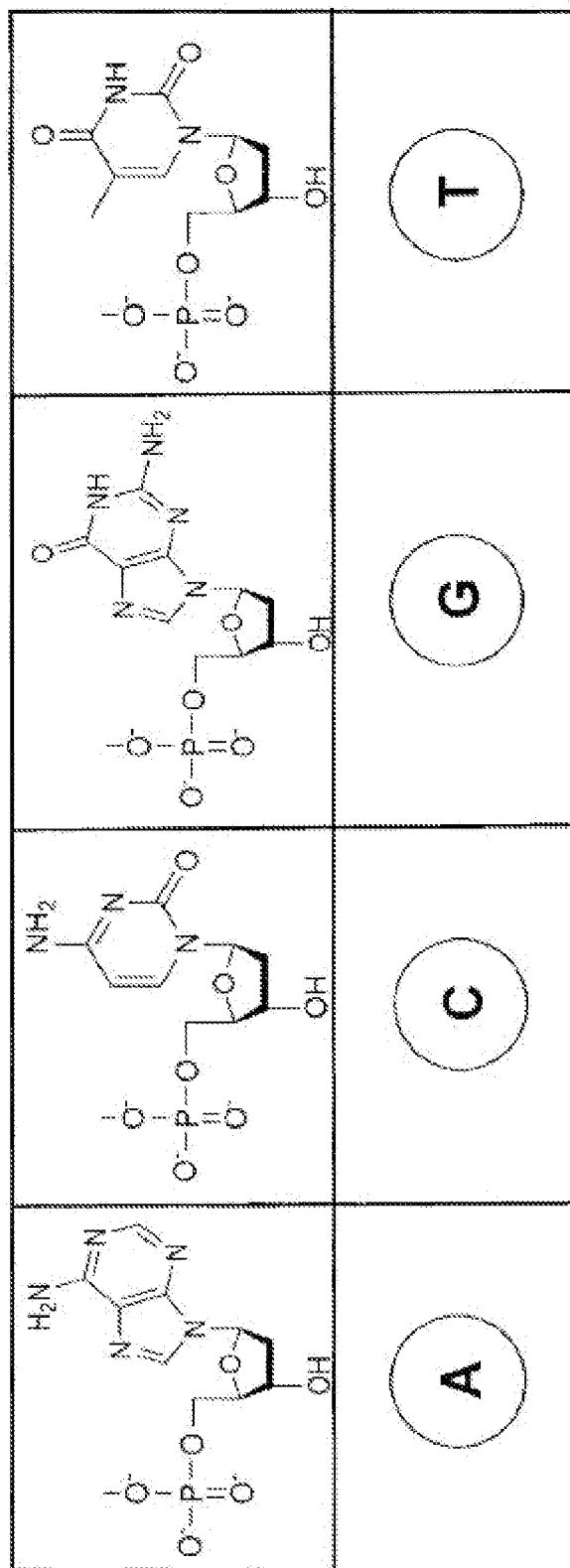


Fig. 26A

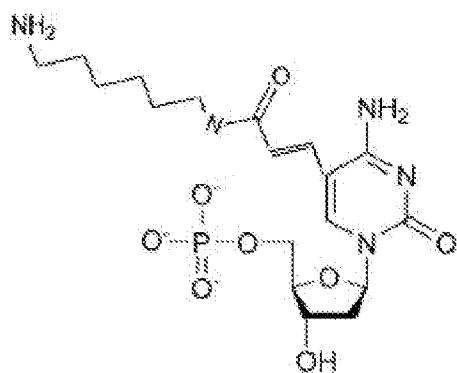


Fig. 26B

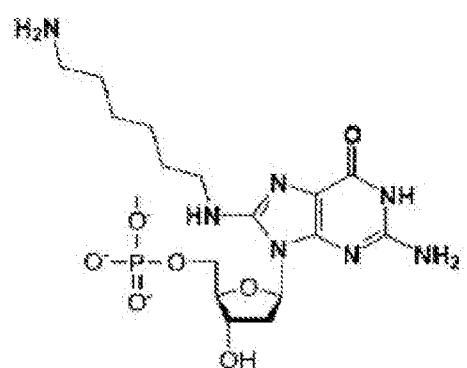


Fig. 27A

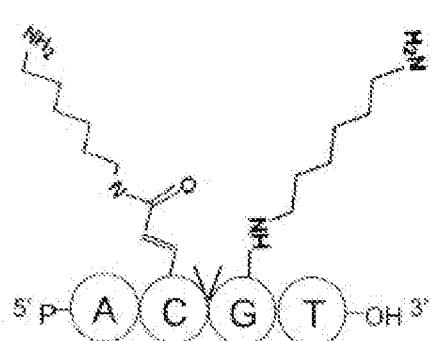


Fig. 27B

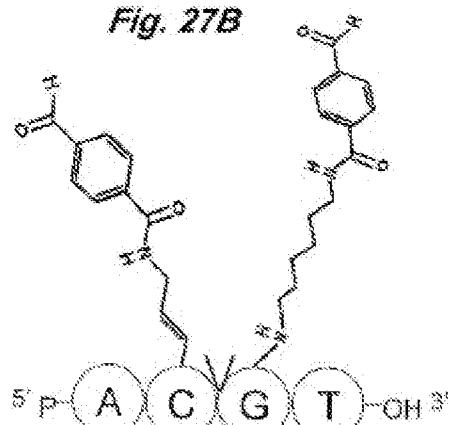


Fig. 28B

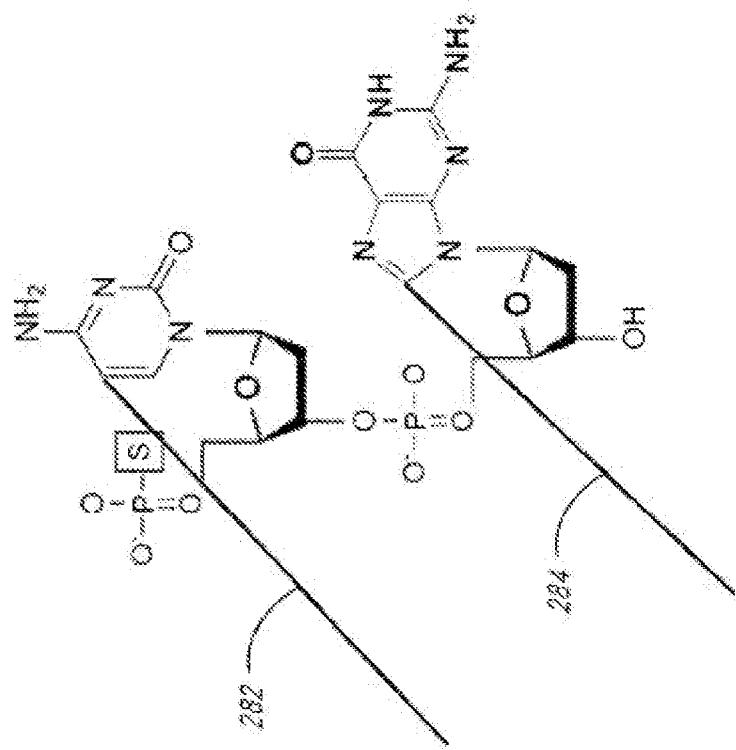


Fig. 28A

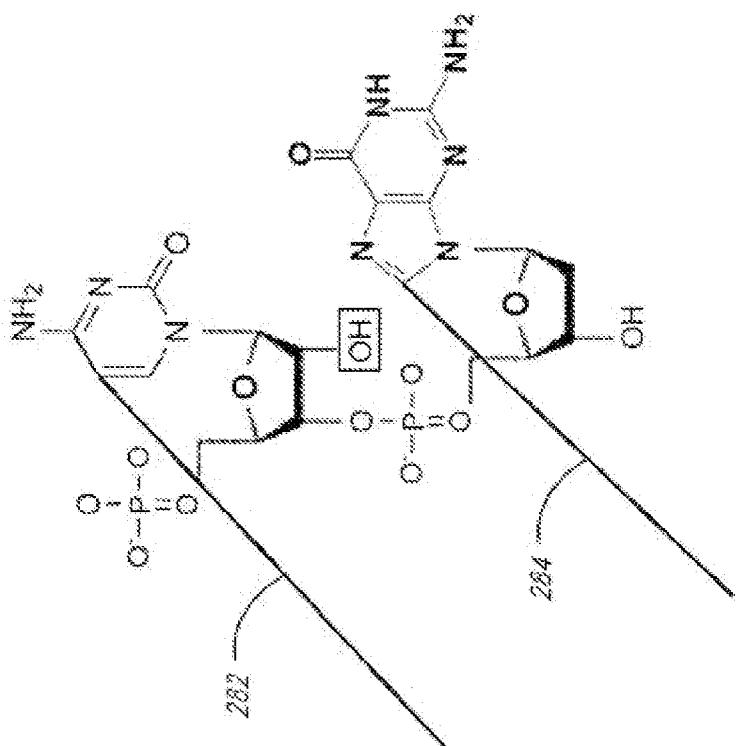


Fig. 28C

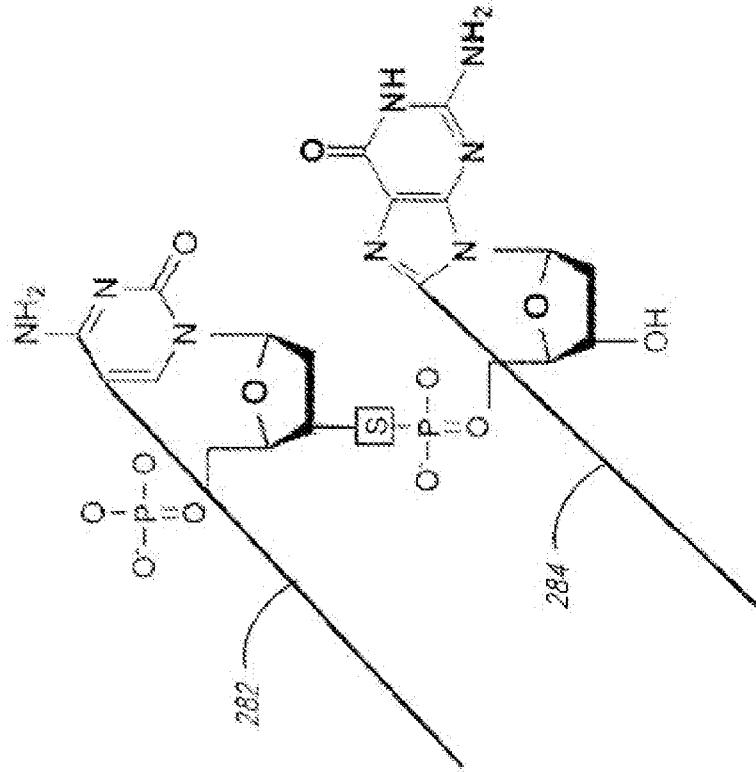


Fig. 28D

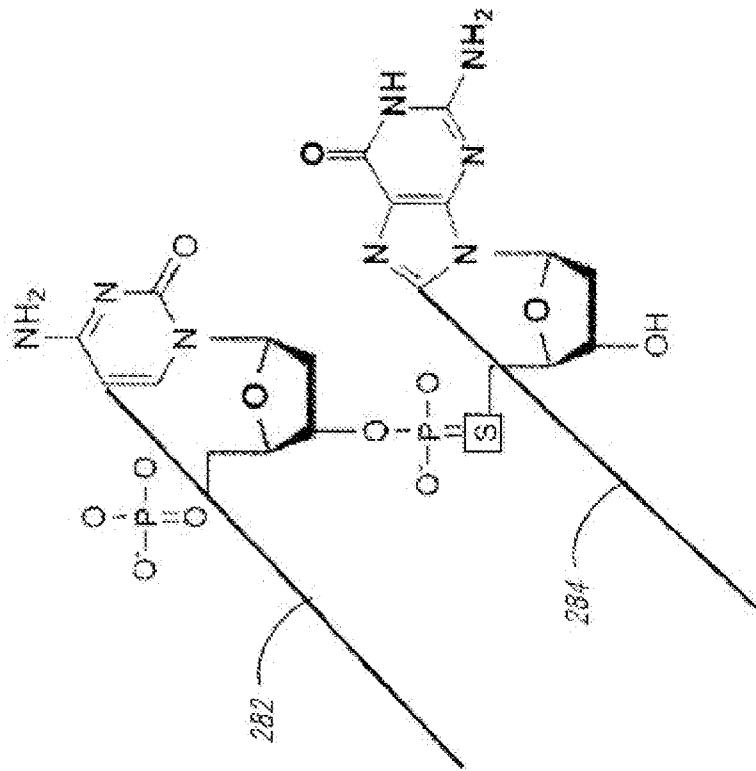


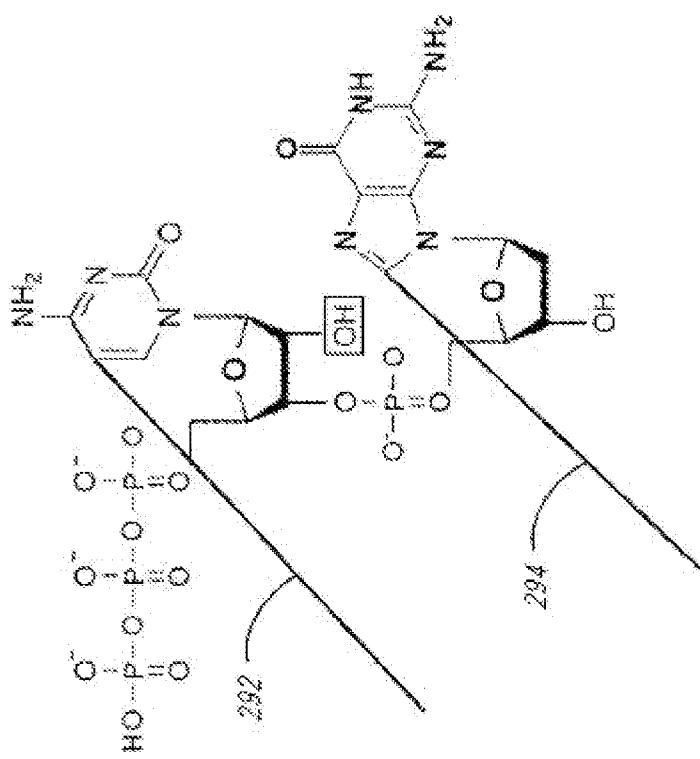
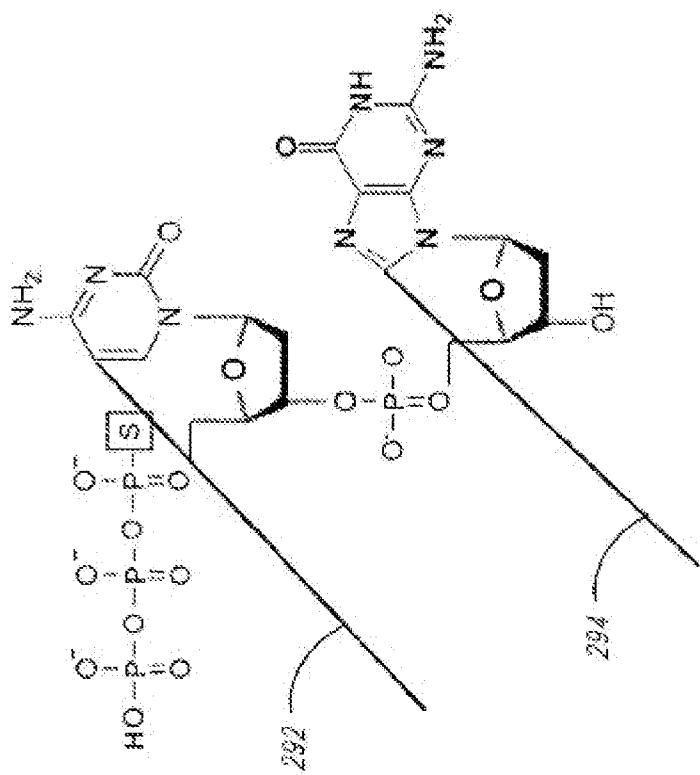
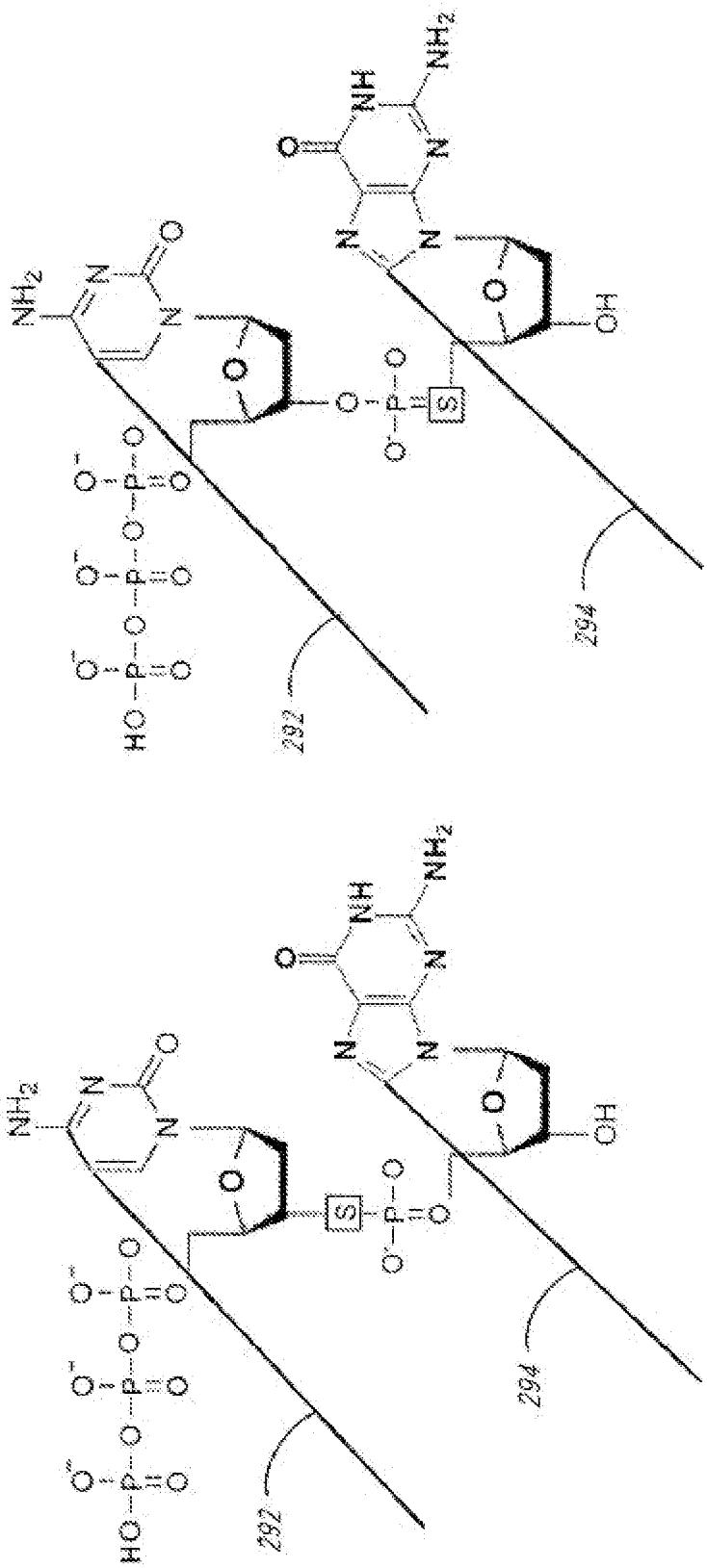
Fig. 29A*Fig. 29B*

Fig. 29C
Fig. 29D



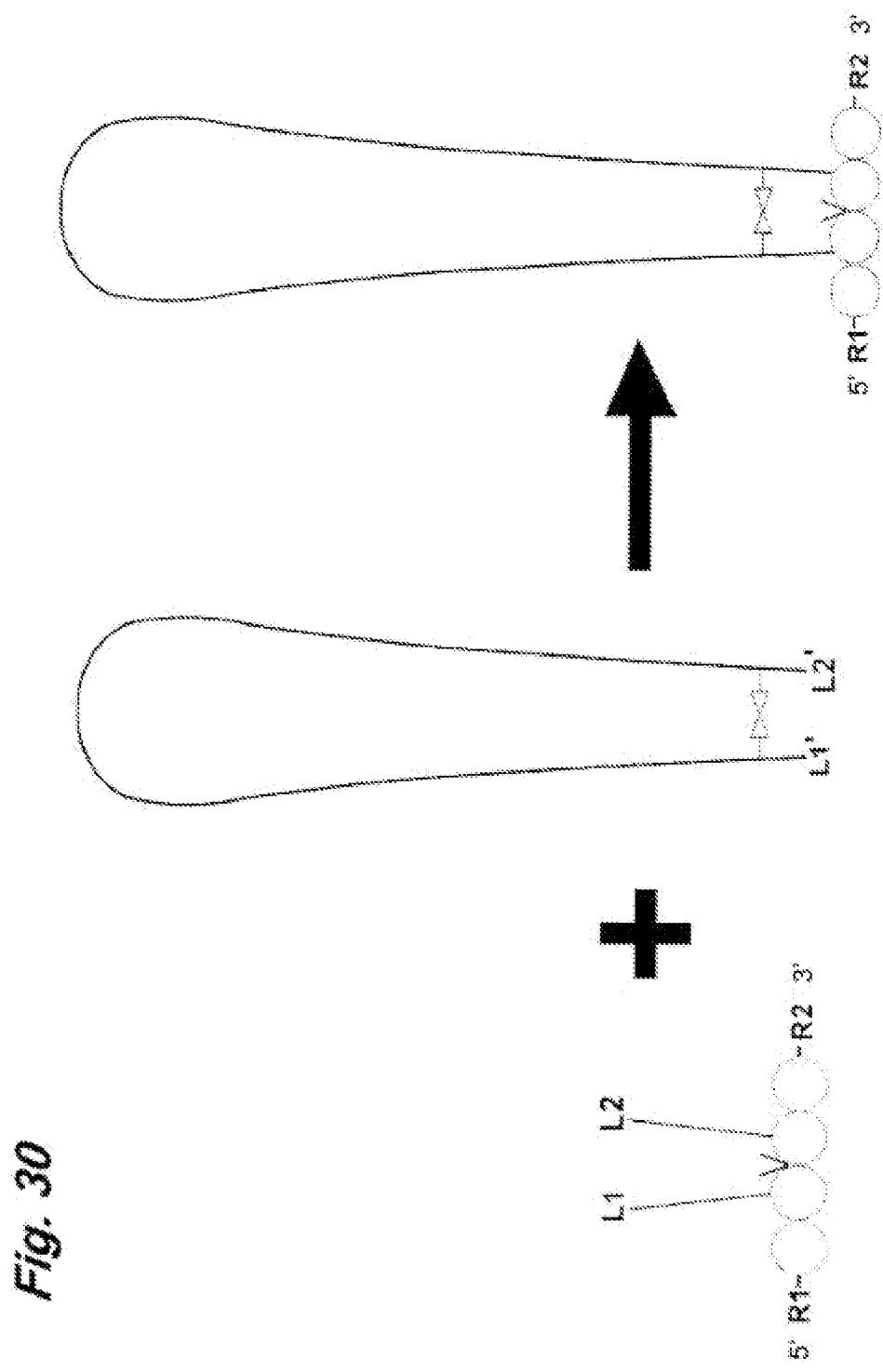


Fig. 30

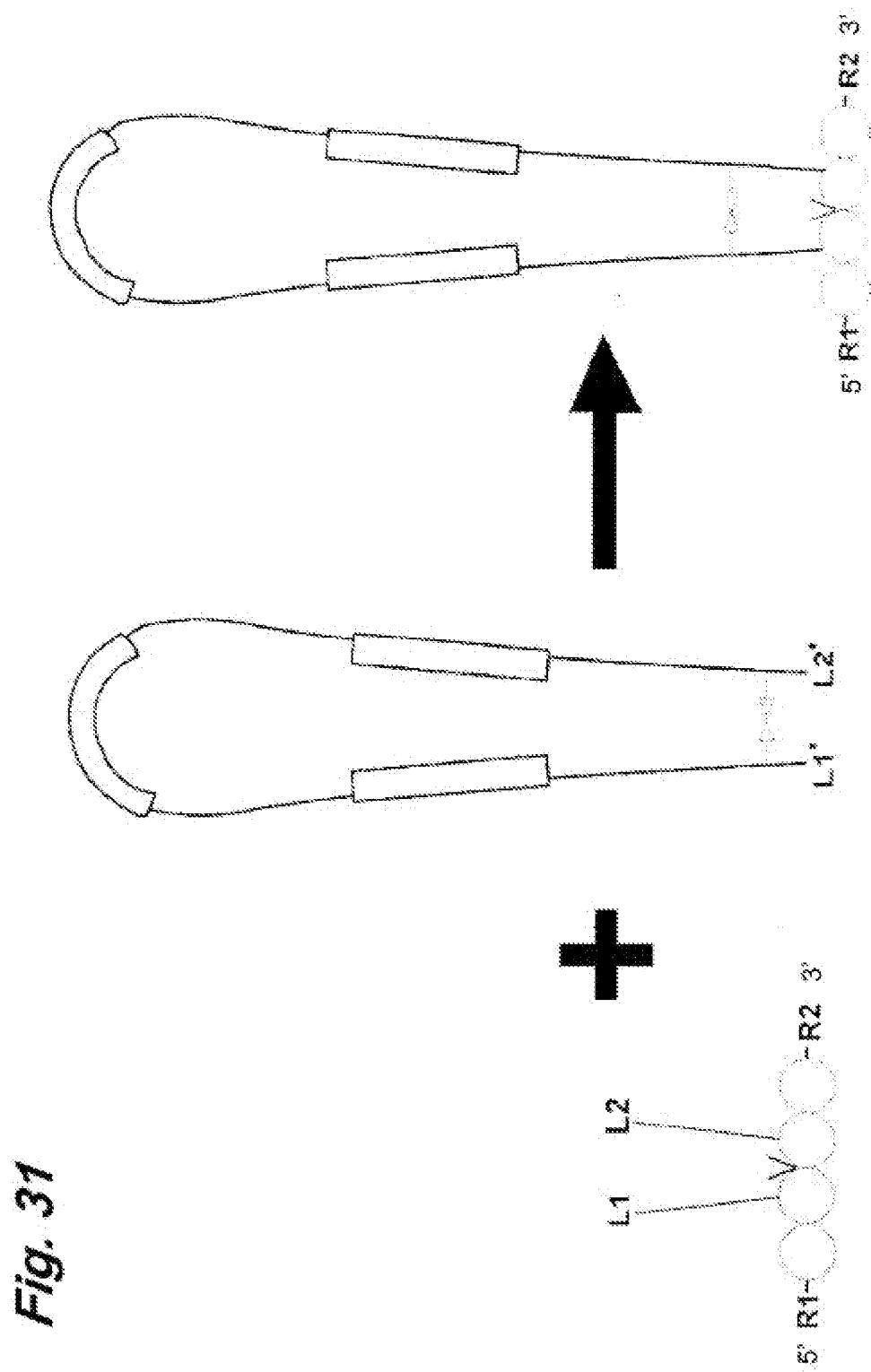


Fig. 31

Fig. 32C

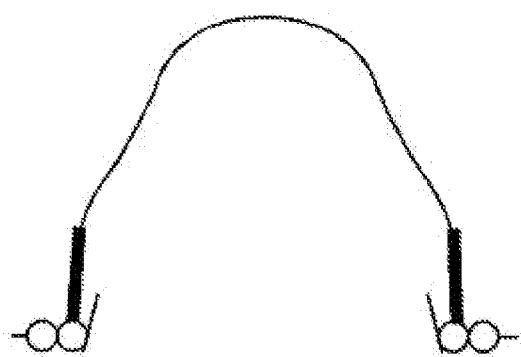


Fig. 32B



Fig. 32A

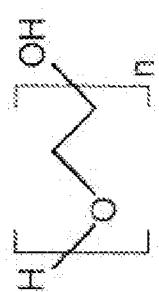


Fig. 33A

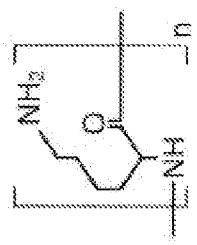


Fig. 33C

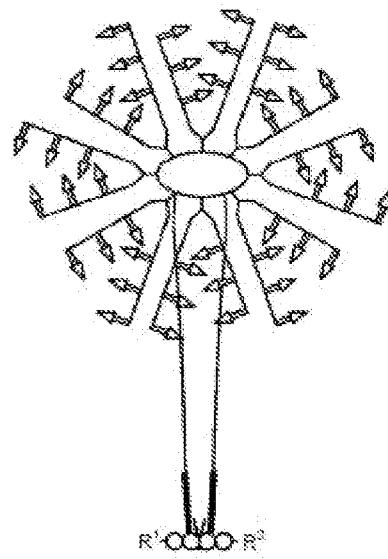


Fig. 330

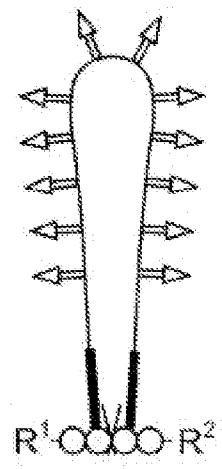
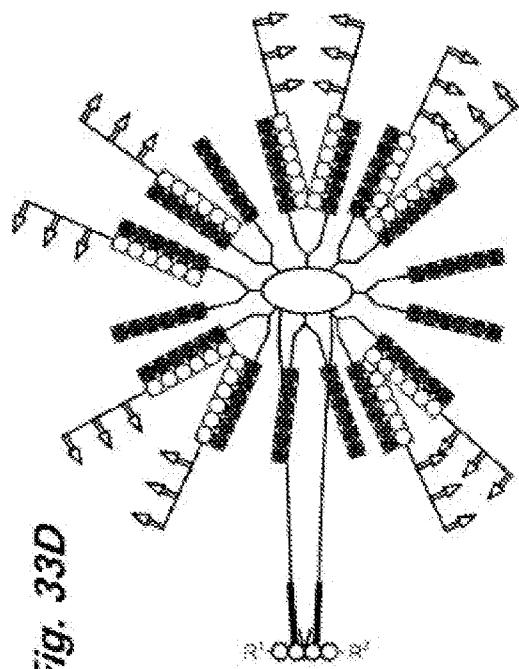
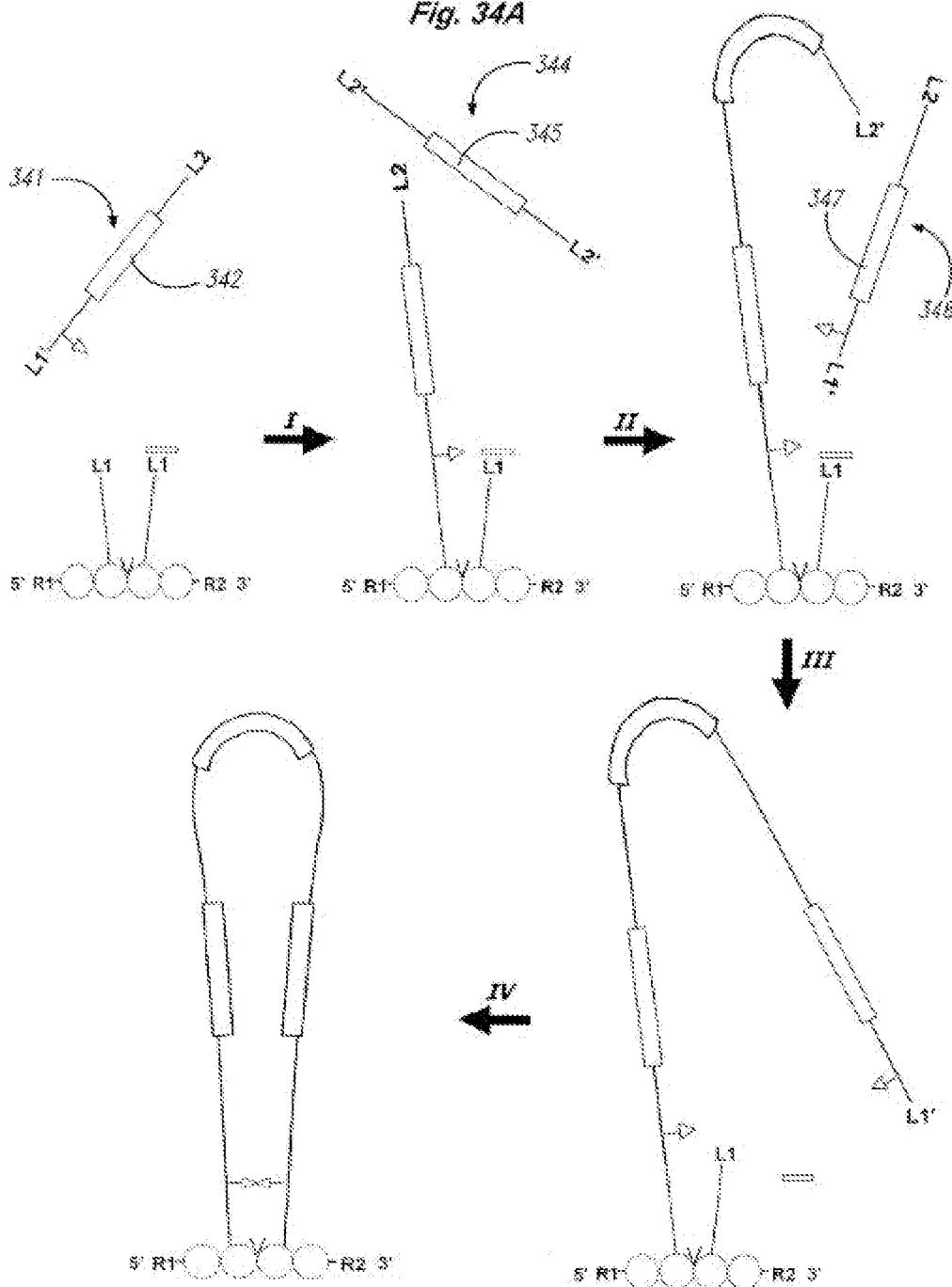


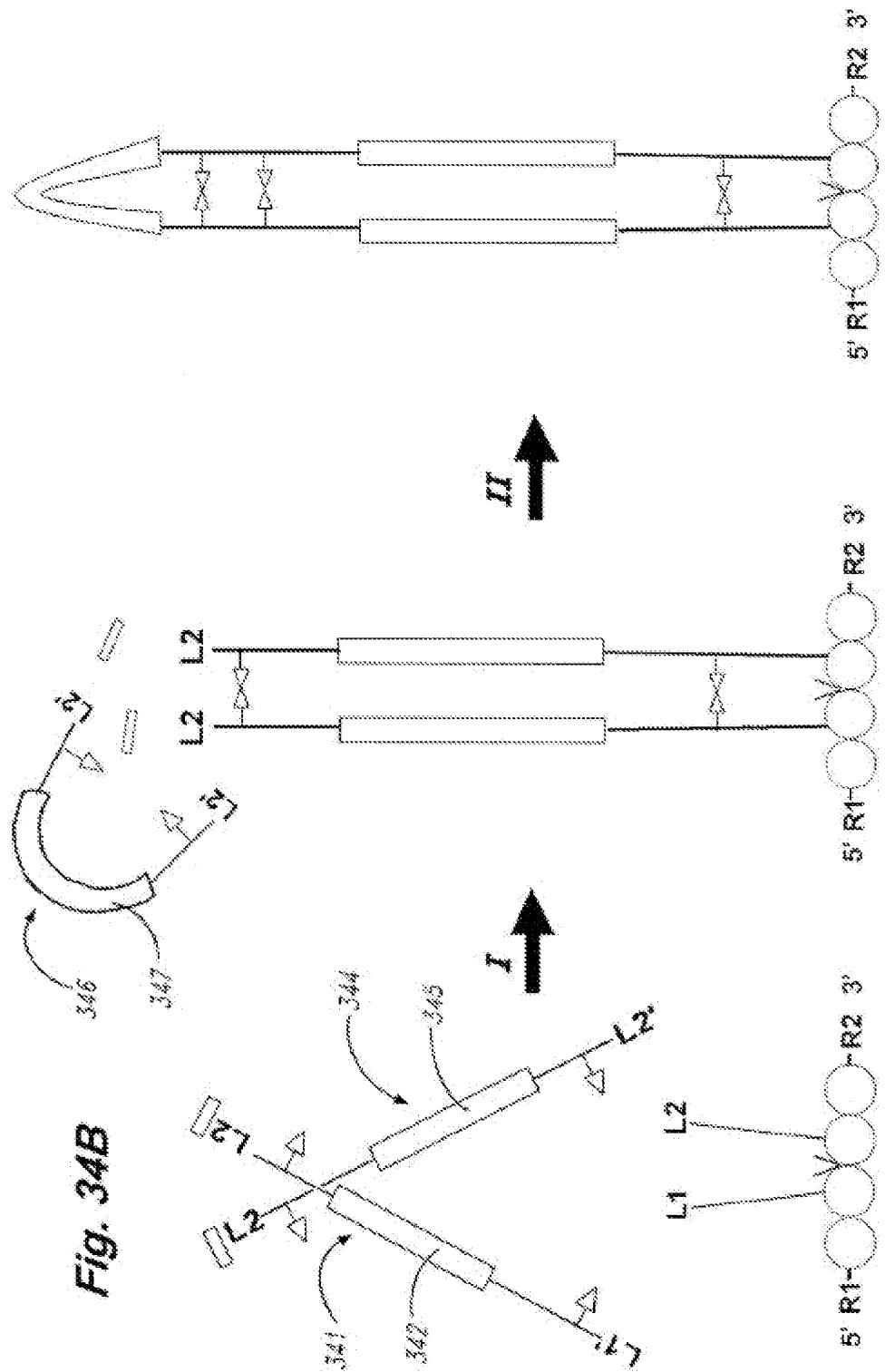
Fig. 33A



۷۰

Fig. 34A





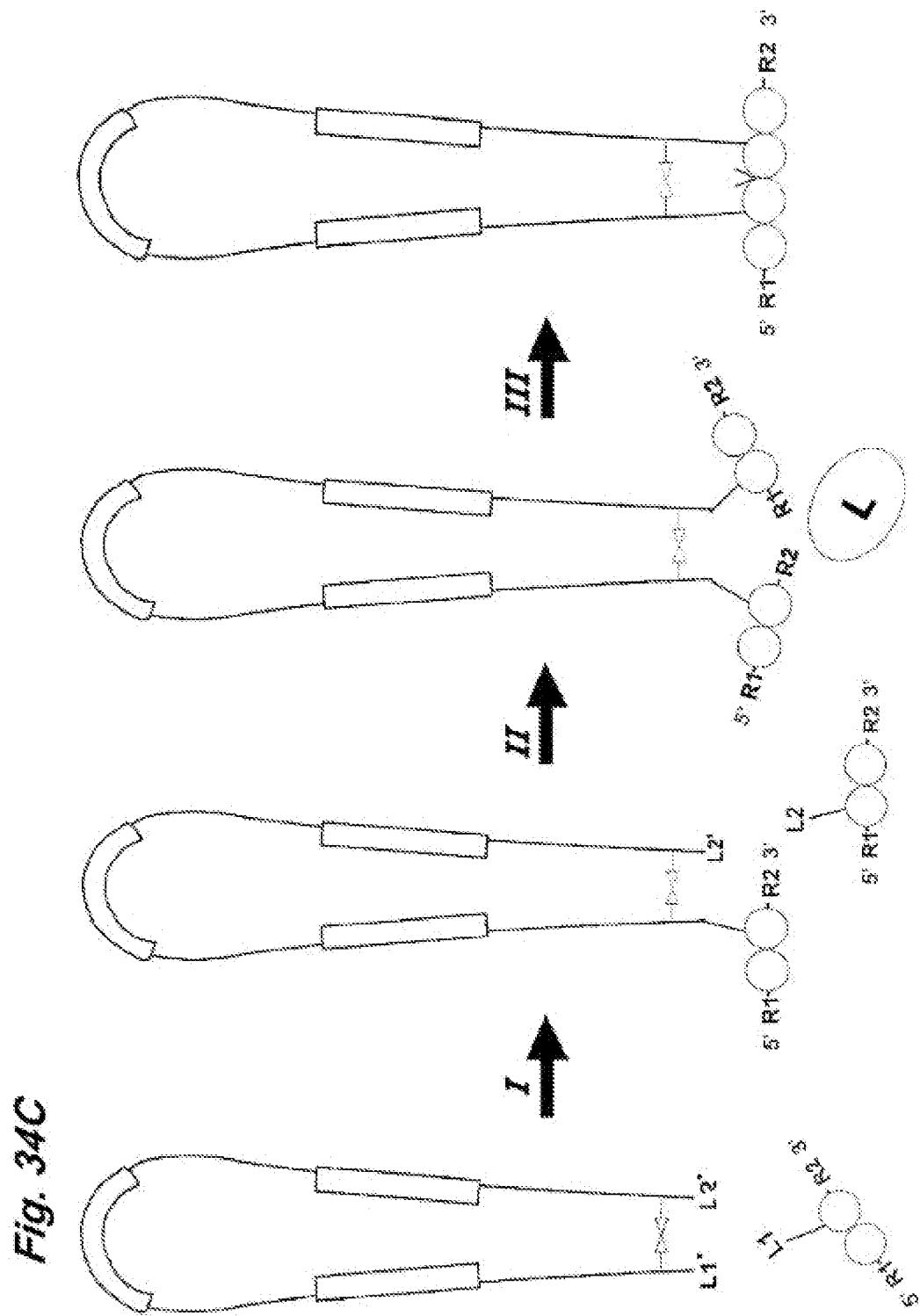


Fig. 34C

Fig. 35A

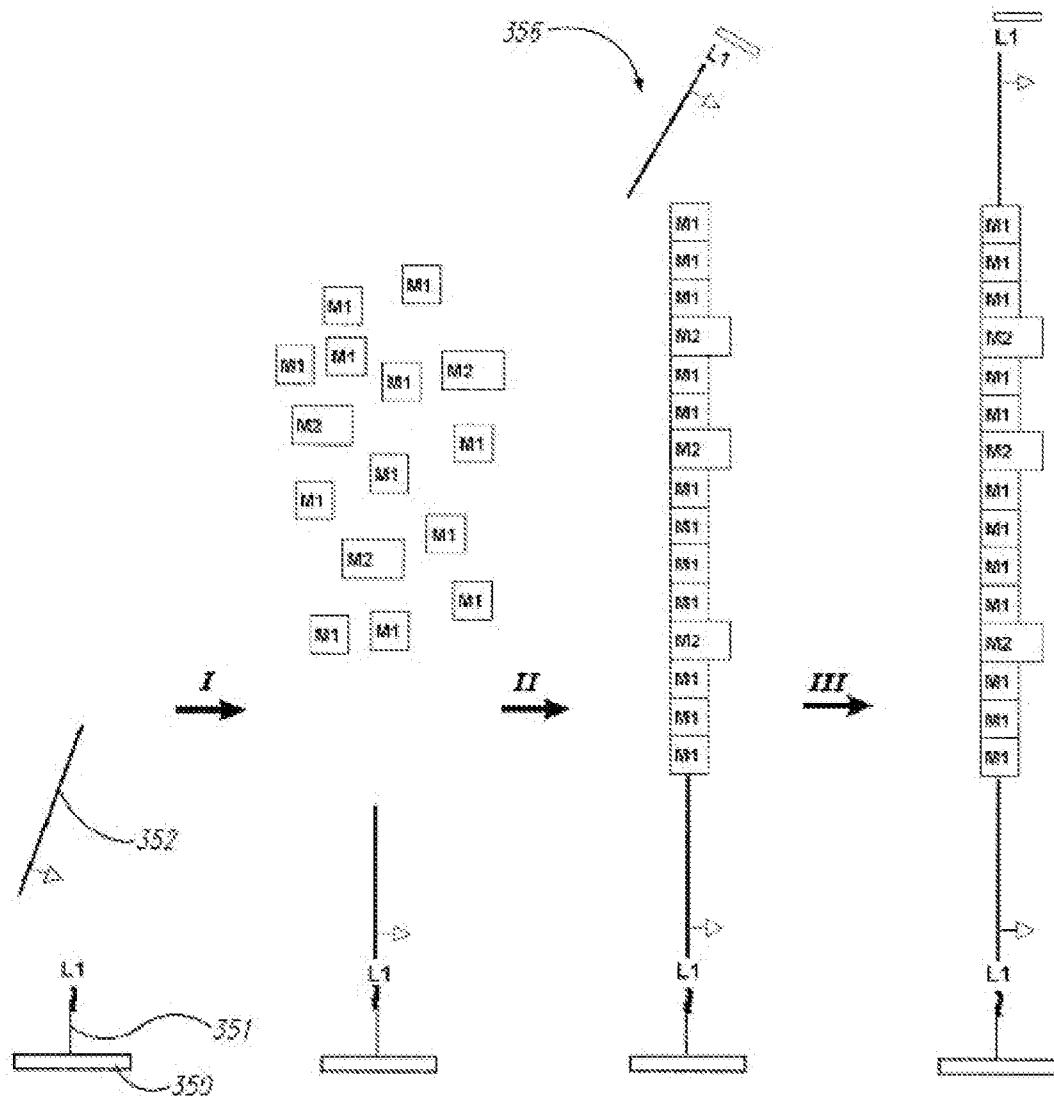


Fig. 35B

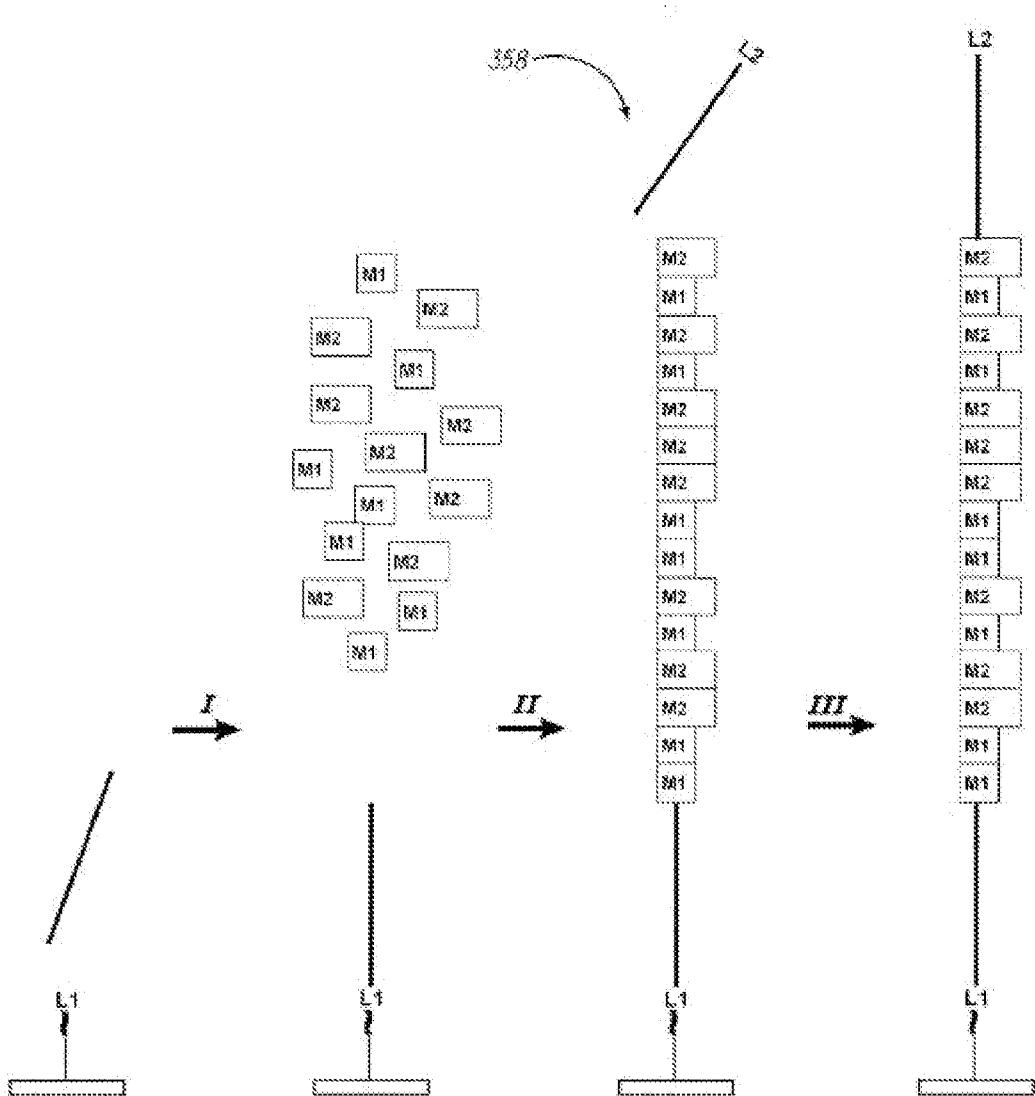


Fig. 36A

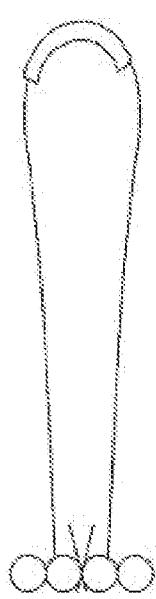


Fig. 36B

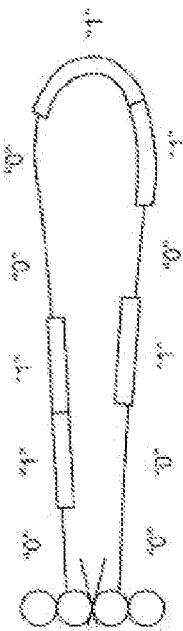


Fig. 36C

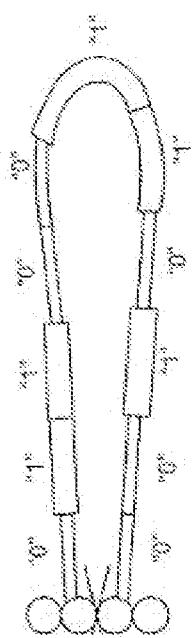


Fig. 36D

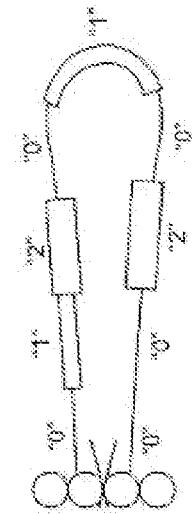


Fig. 36E

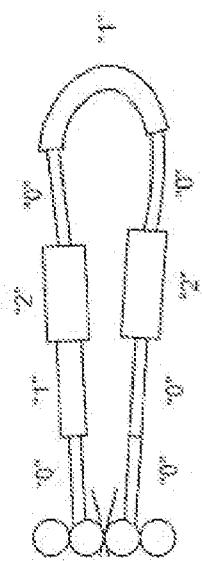


Fig. 36F

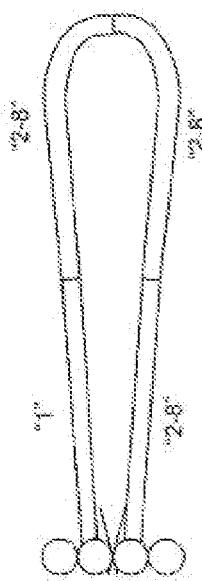


Fig. 36G

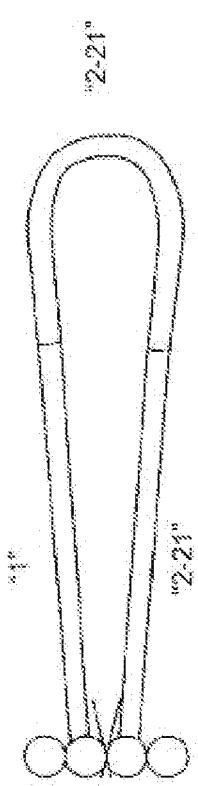


Fig. 36H

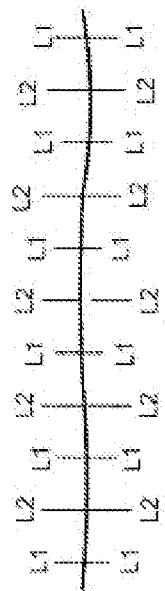


Fig. 36I

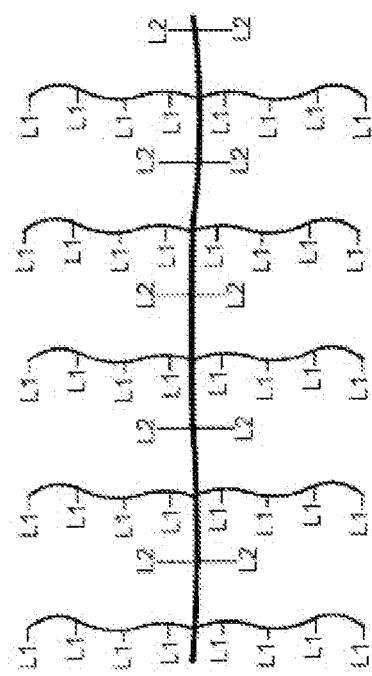


Fig. 37

Identidad de marcador	Masa (M)	Fluorescencia F		Conectores químicos
1	M1	ROJO	HIDRAZINA	-
2	M2	VERDE	AMINA	-
3	M3	AZUL	CARBOXILO	-
4	m1	rojo	hidrazina	-
5	m2	verde	amina	-
6	m3	azul	carboxilo	-
7	m1-m2	rojo-verde	hidrazina	-
8	m1-m3	rojo-azul	hidrazina	-
9	m1-M2	rojo-VERDE	hidrazina	-
10	m1-M3	rojo-AZUL	hidrazina	-
11	M1-m2	ROJO-verde	HIDRAZINA	CARBOXILO
12	M1-m3	ROJO-azul	HIDRAZINA	CARBOXILO
13	M2-m3	verde-azul	HIDRAZINA	CARBOXILO
14	M2-m3	VERDE-azul	HIDRAZINA	CARBOXILO
15	M2-M3	verde-Azul	HIDRAZINA	CARBOXILO
16	m1-m2-m3	rojo-verde-azul	HIDRAZINA	CARBOXILO
17	M1-m2-m3	ROJO-verde-azul	HIDRAZINA	CARBOXILO
18	m1 M2-m3	rojo-VERDE-azul	hidrazina	CARBOXILO
19	m1-m2-M3	rojo-verde-AZUL	hidrazina	CARBOXILO
20	M1-M2-m3	ROJO VERDE-azul	HIDRAZINA	CARBOXILO
21	m1-M2-M3	rojo-VERDE-AZUL	hidrazina	CARBOXILO
22	M1-m2-M3	ROJO-verde-AZUL	HIDRAZINA	CARBOXILO

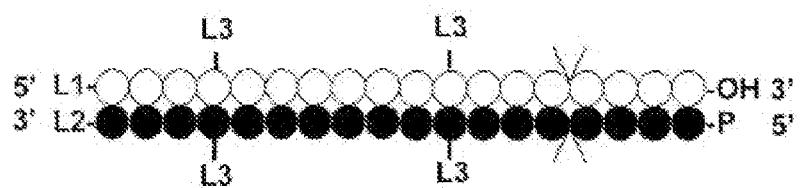
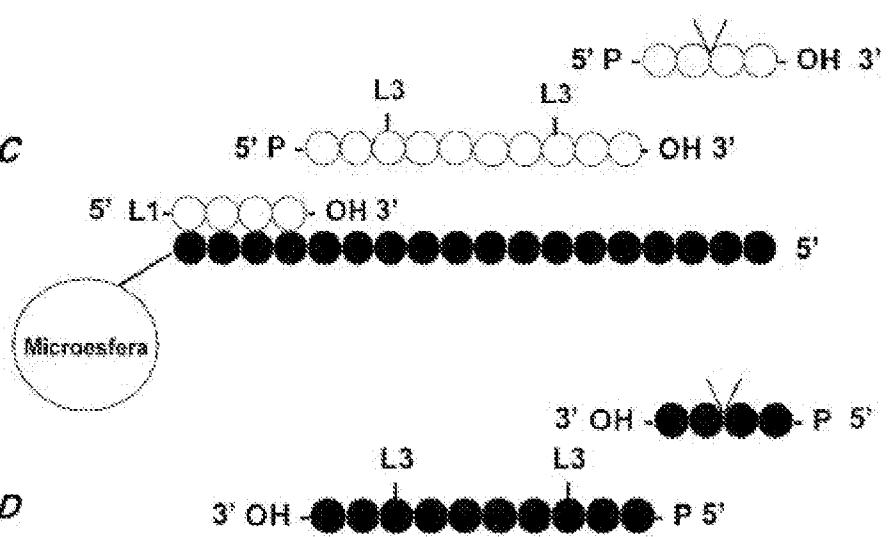
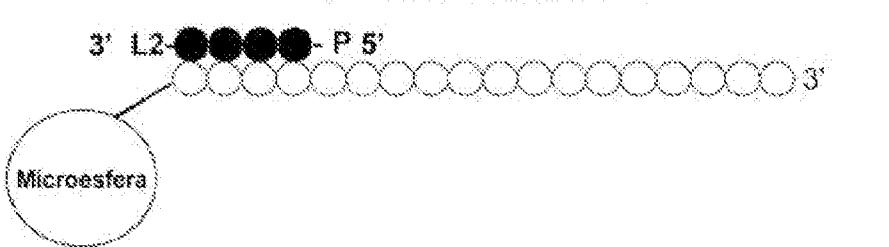
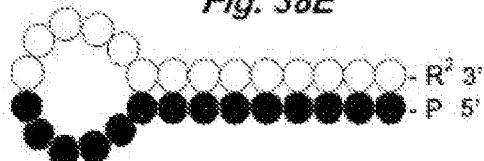
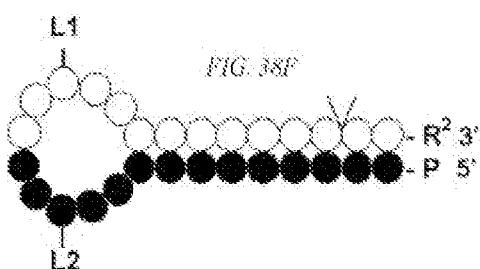
Fig. 38A*Fig. 38B**Fig. 38C**Fig. 38D**Fig. 38E**FIG. 38F*

Fig. 39

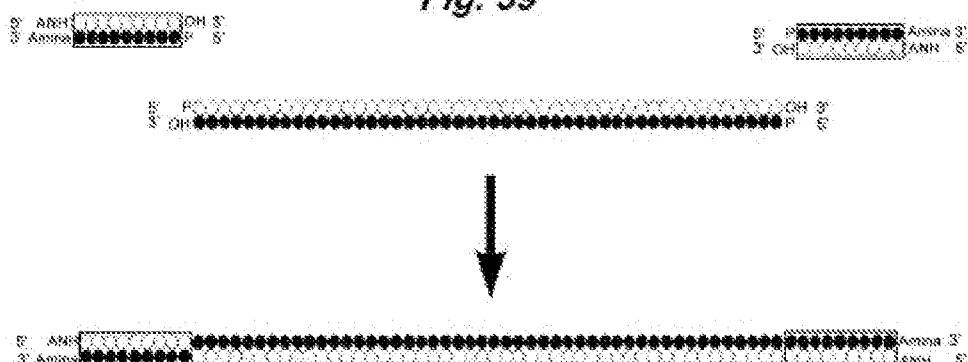


Fig. 40

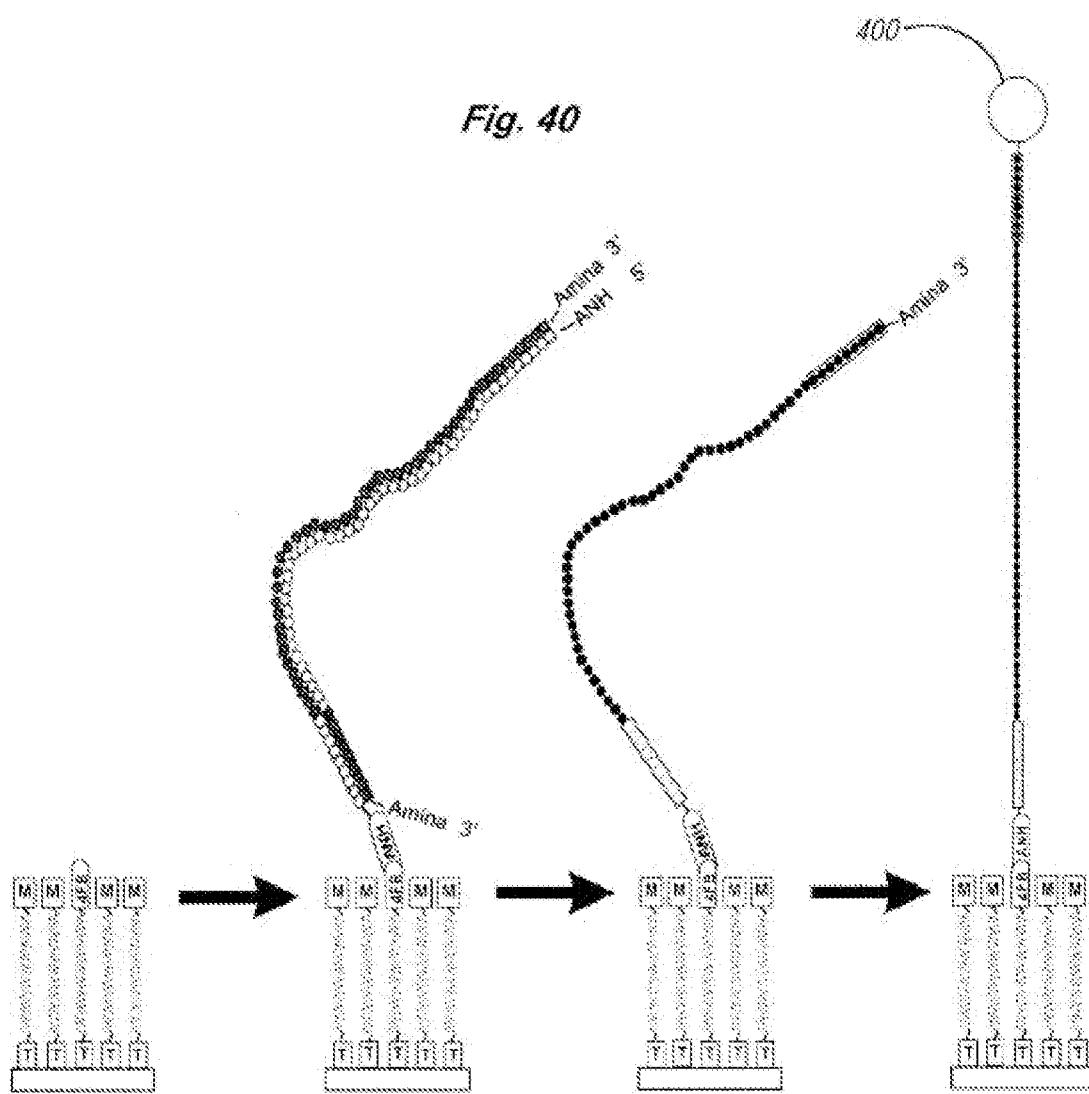


Fig. 41A

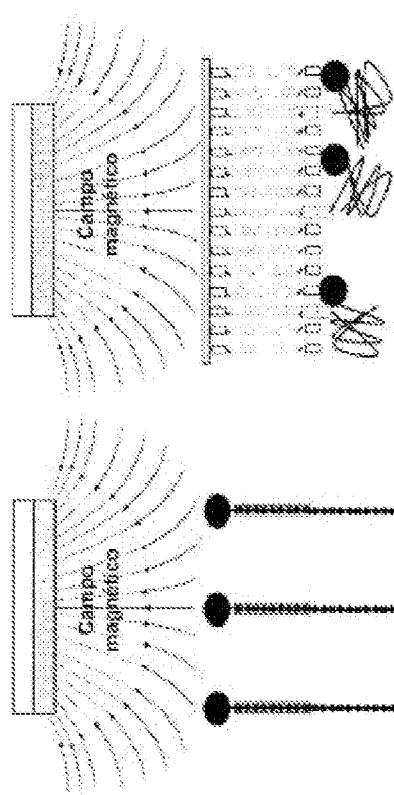


Fig. 41B

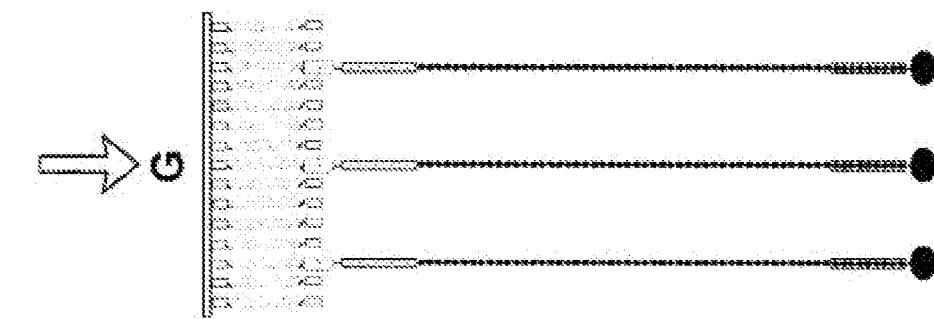


Fig. 41C

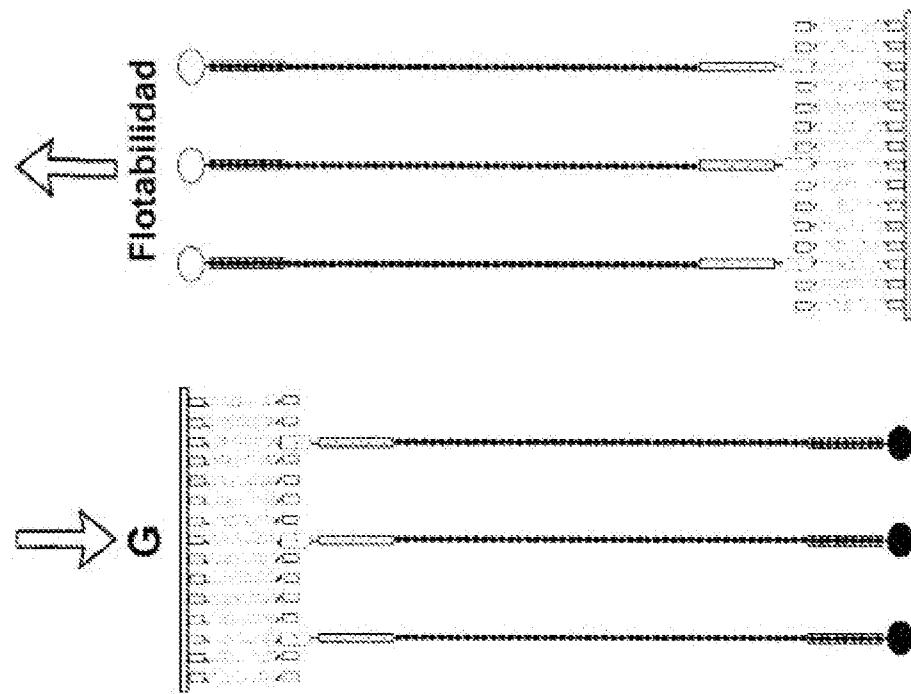


Fig. 41D

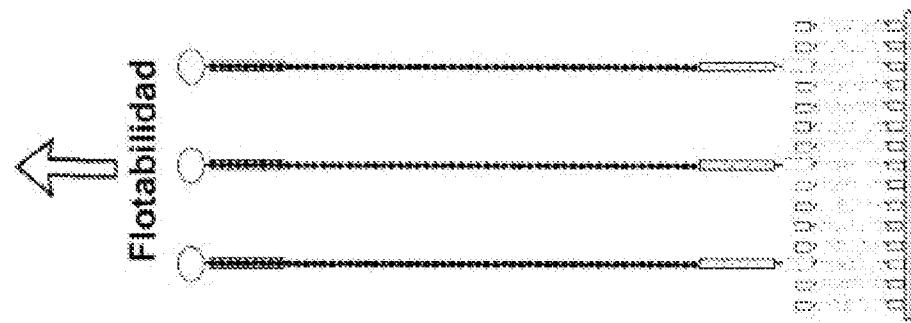


Fig. 41E

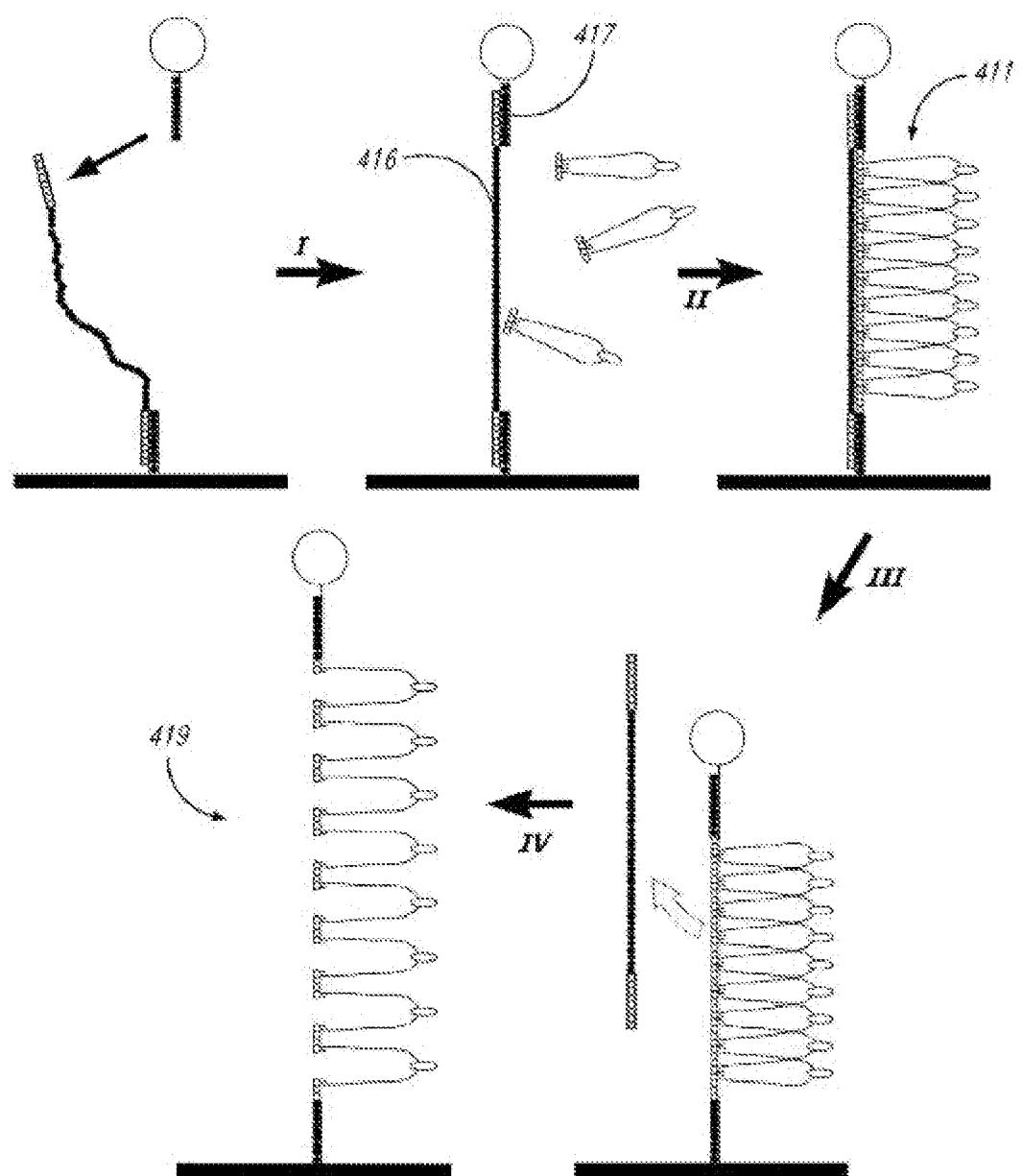


Fig. 42A

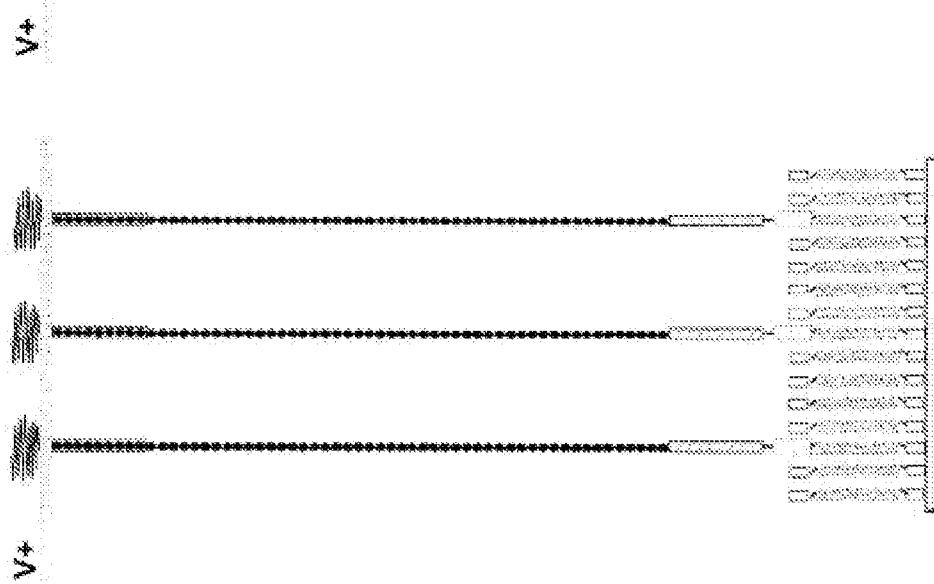


Fig. 42B

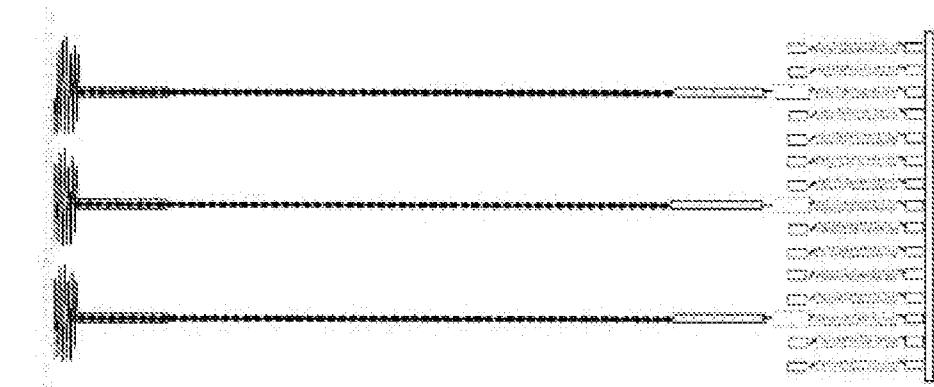
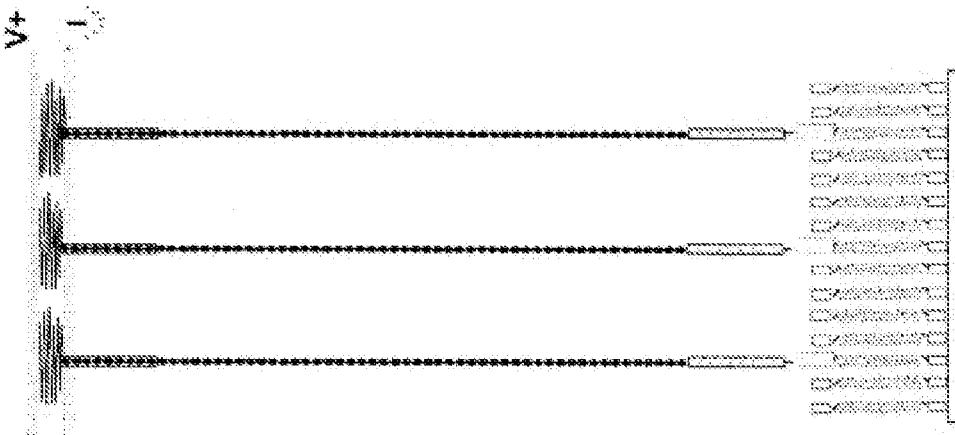
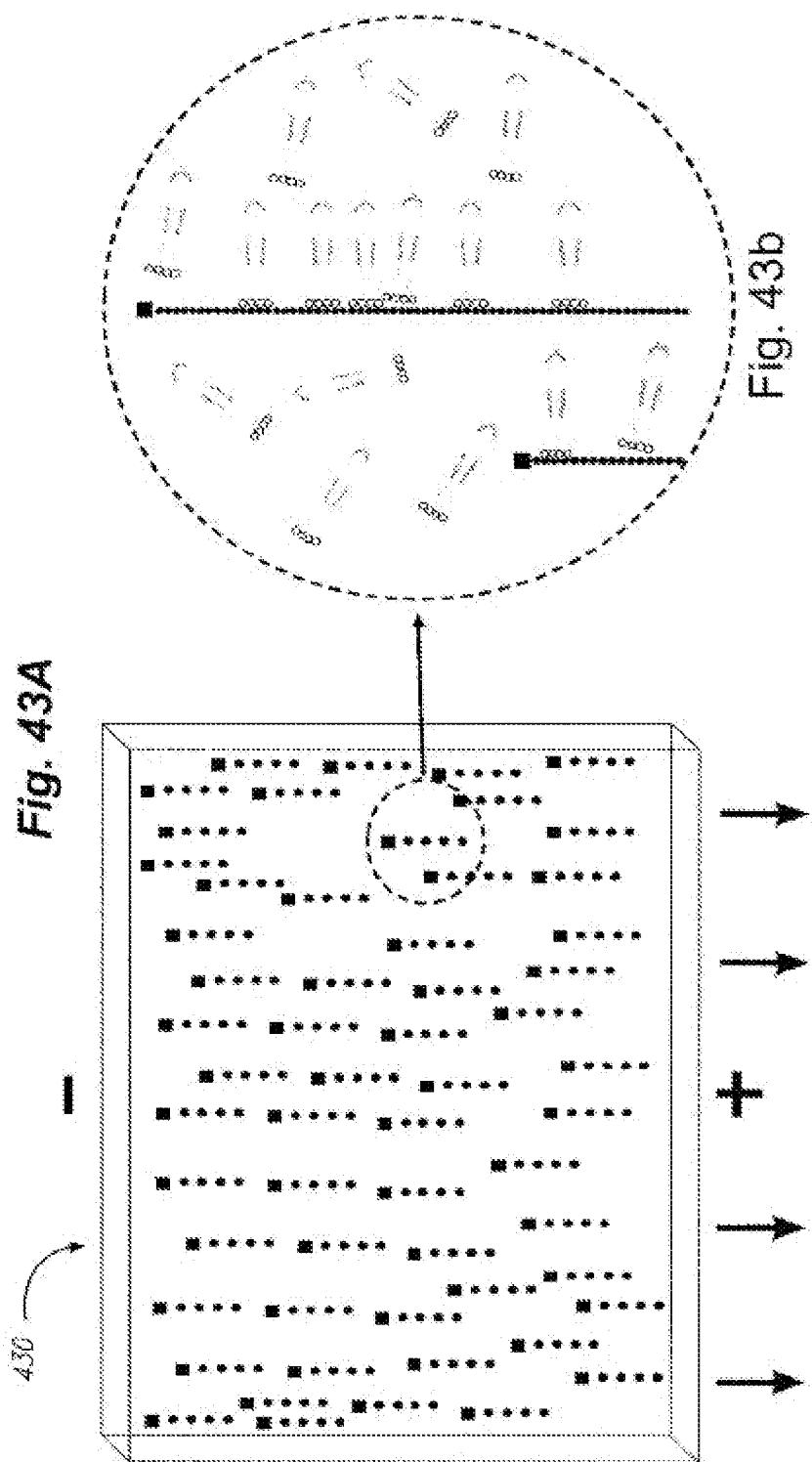


Fig. 42C





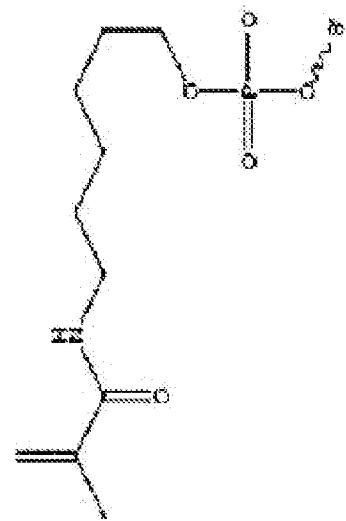


Fig. 43C

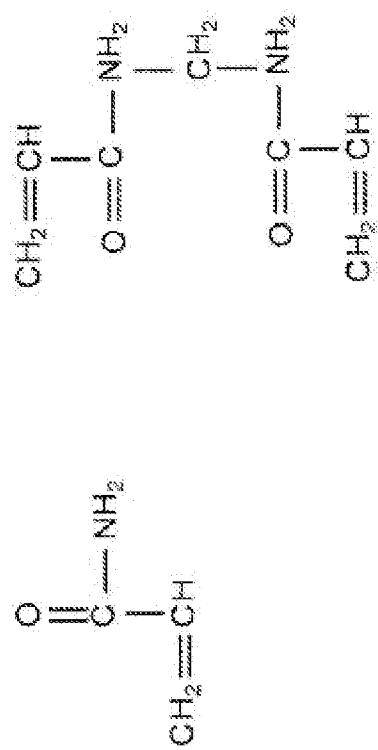


Fig. 43D

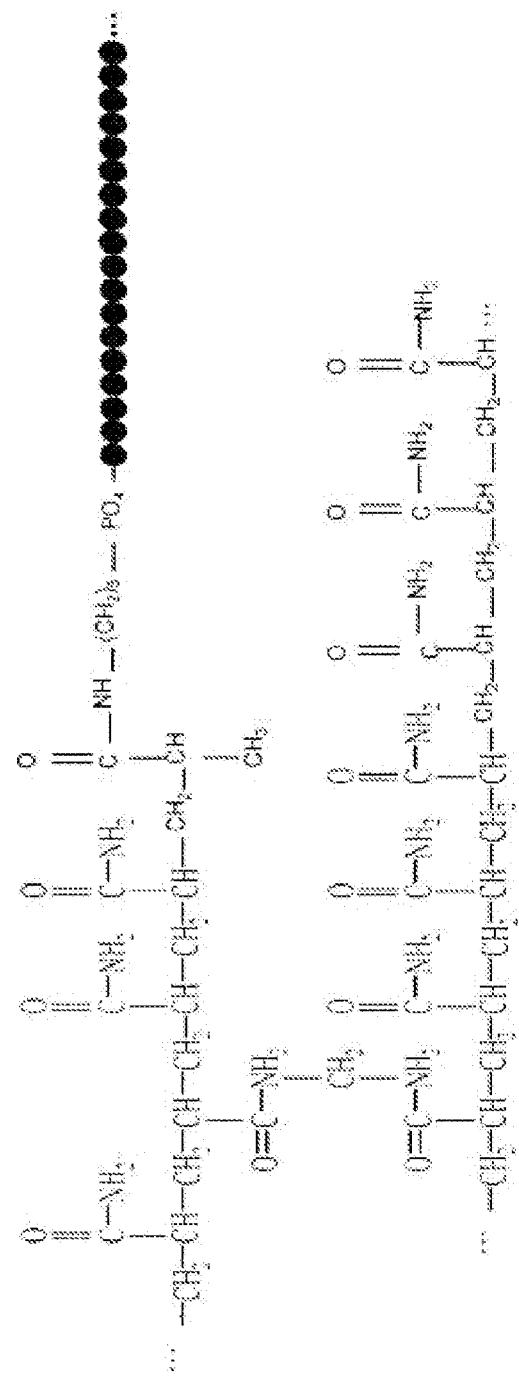


Fig. 44A

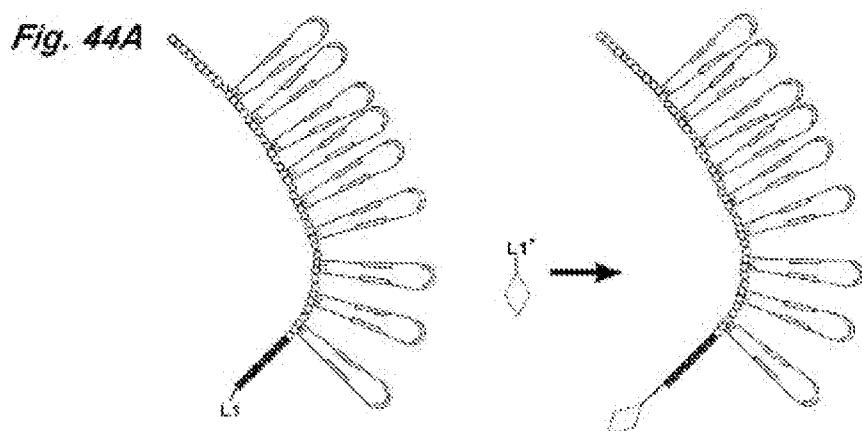


Fig. 44B

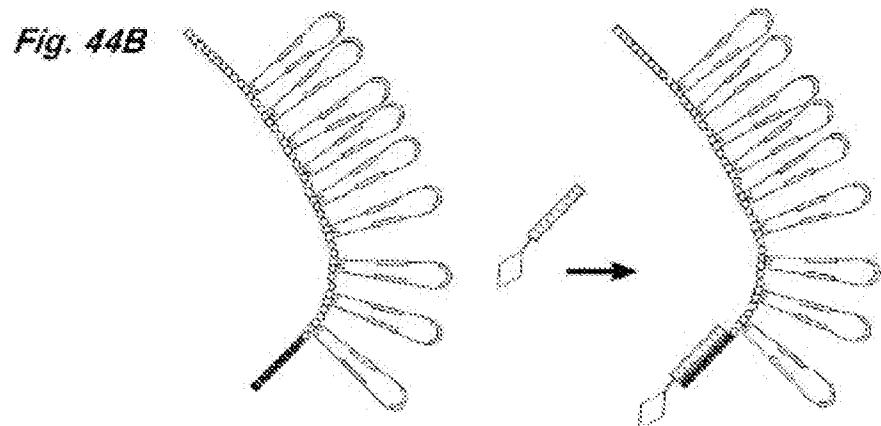


Fig. 44C

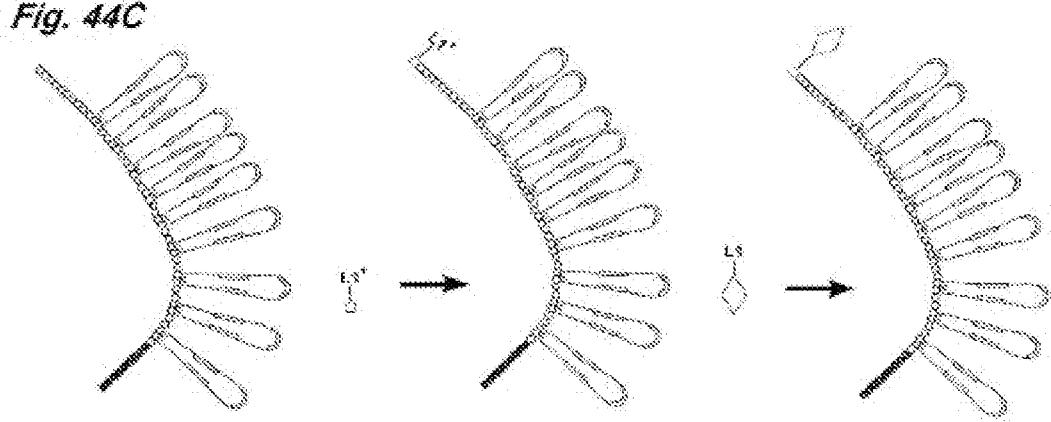


Fig. 45

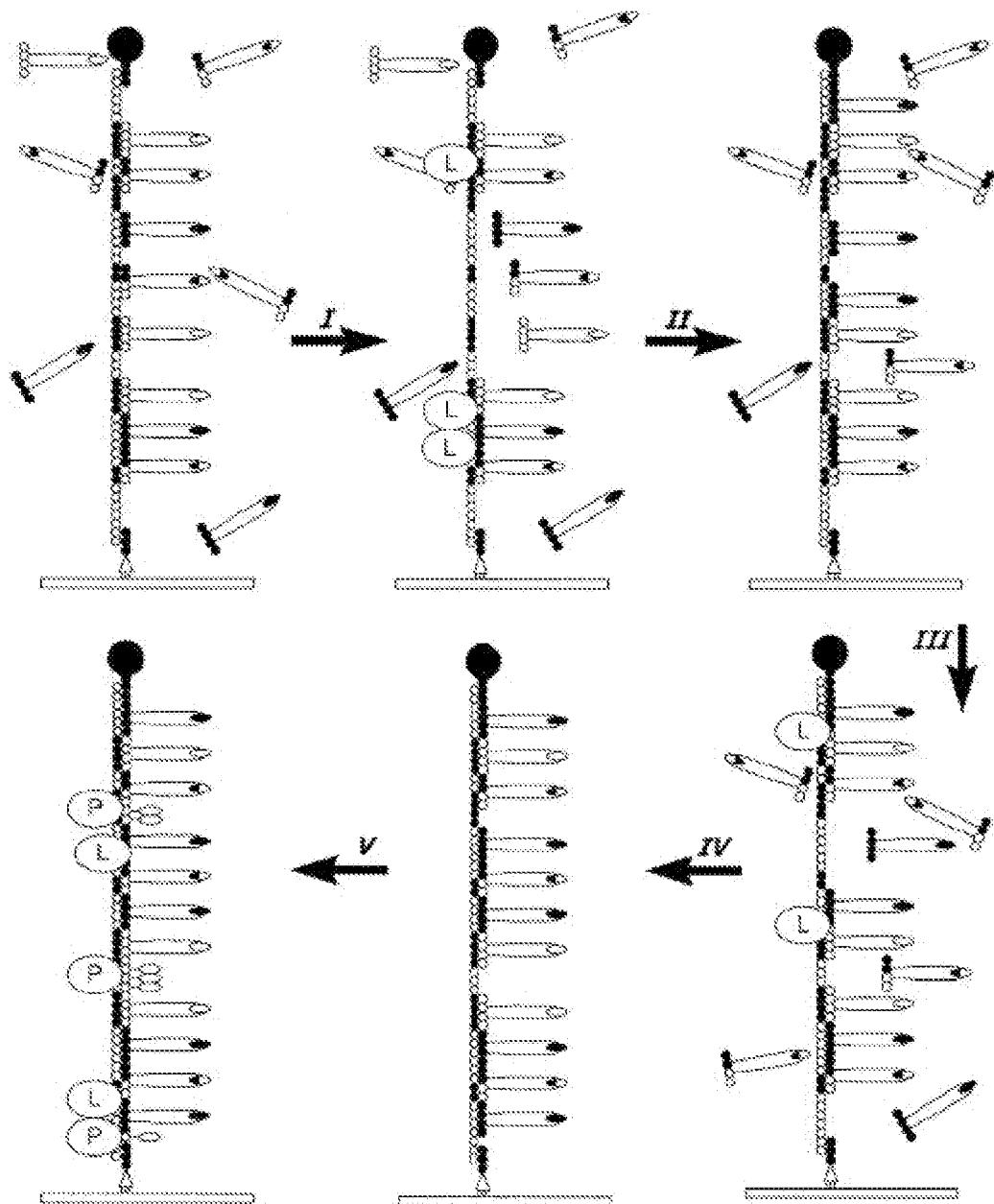


Fig. 46A

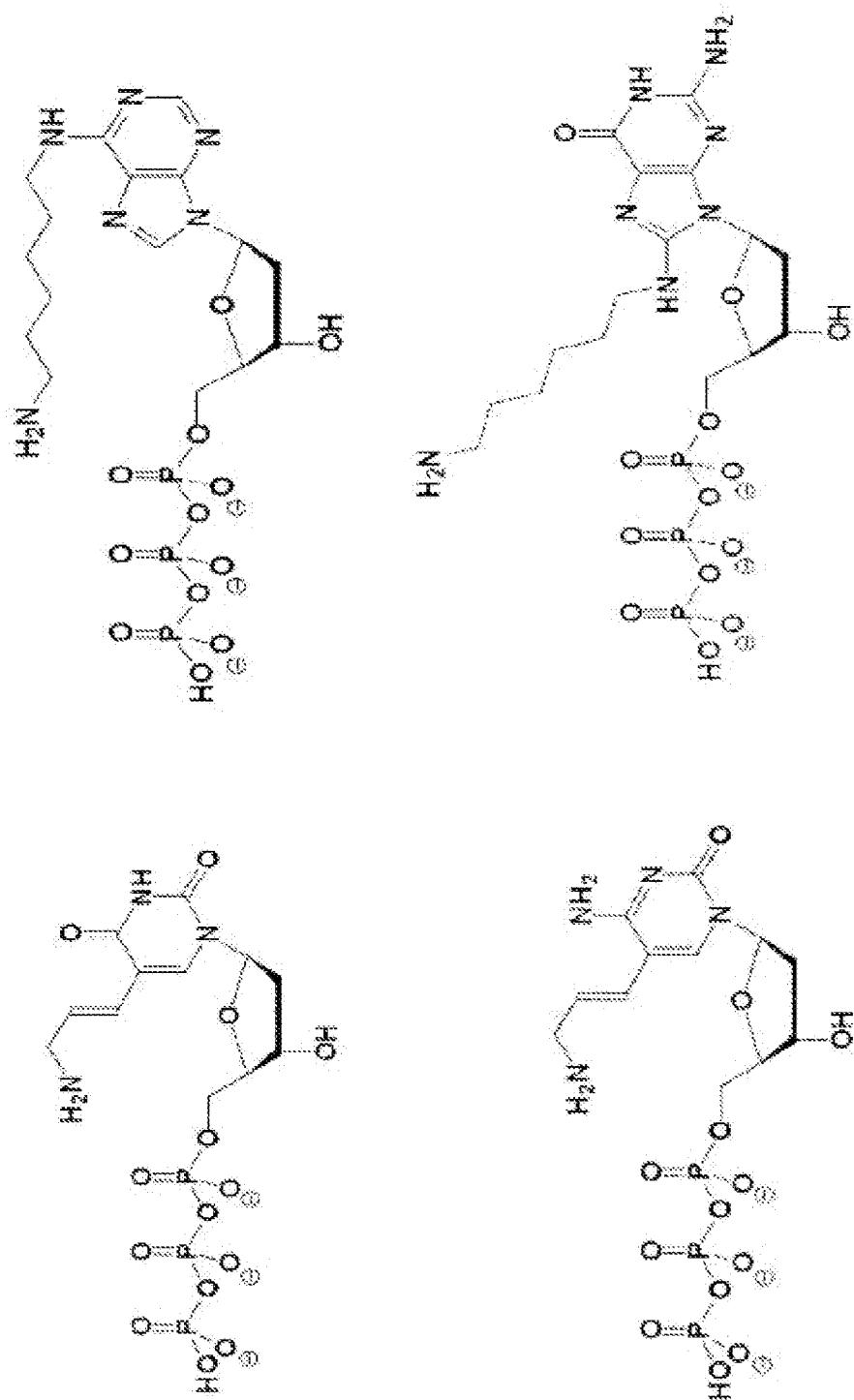


Fig. 46B

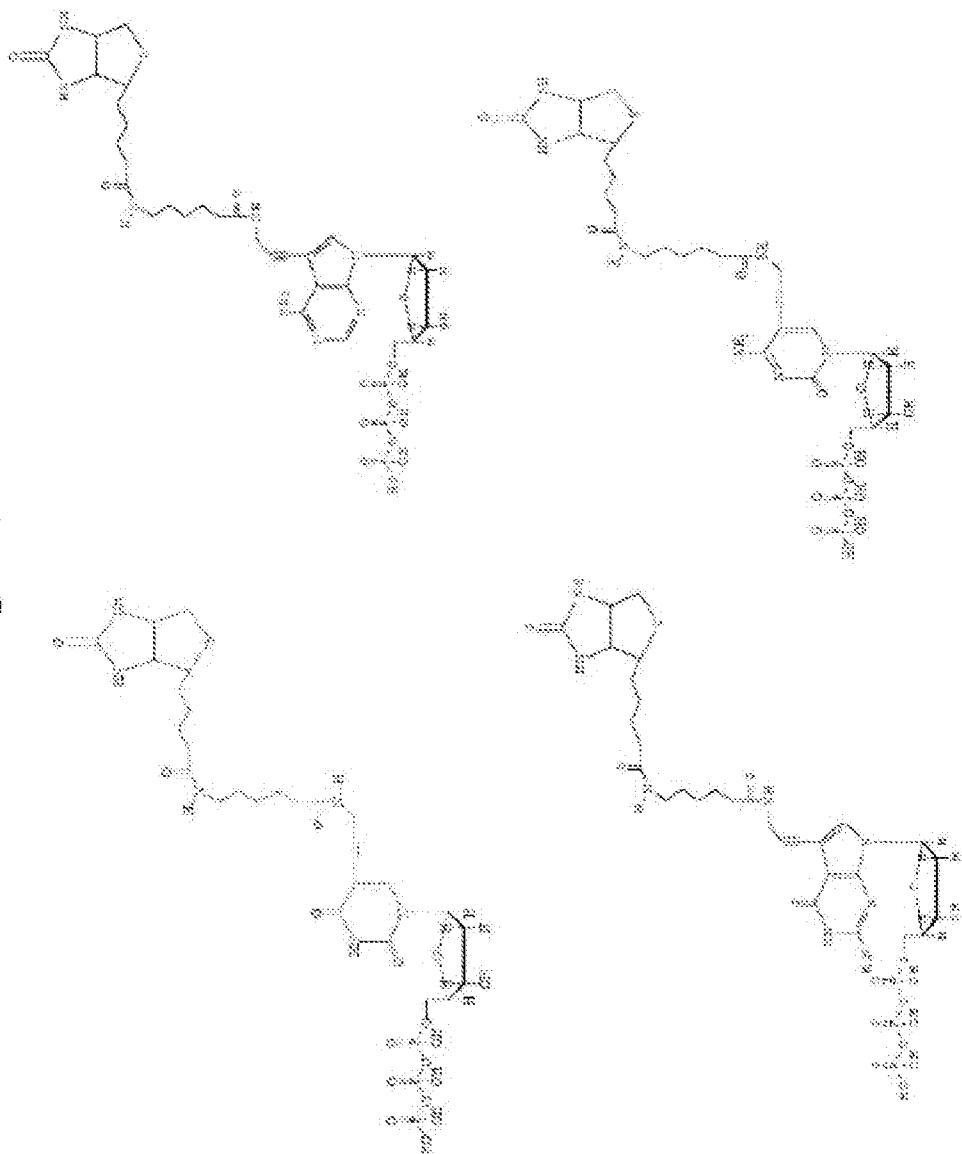


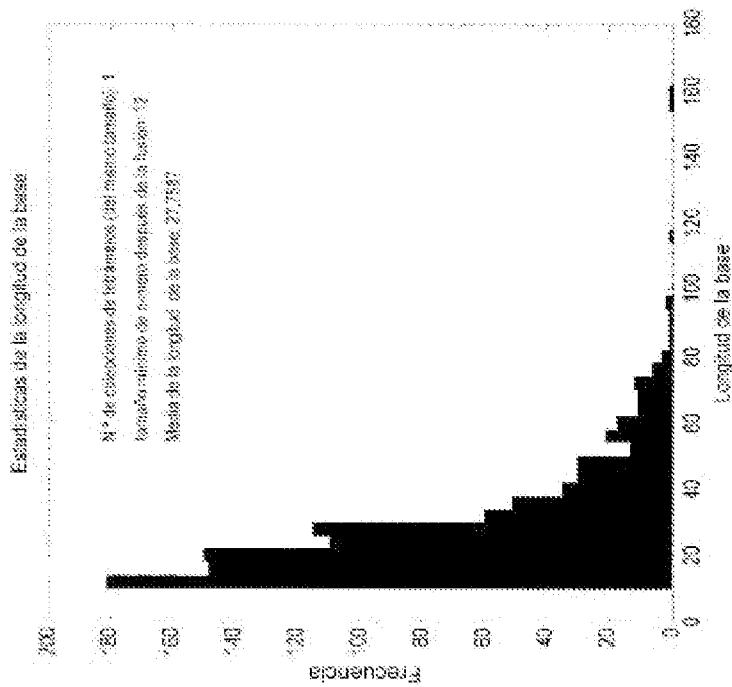
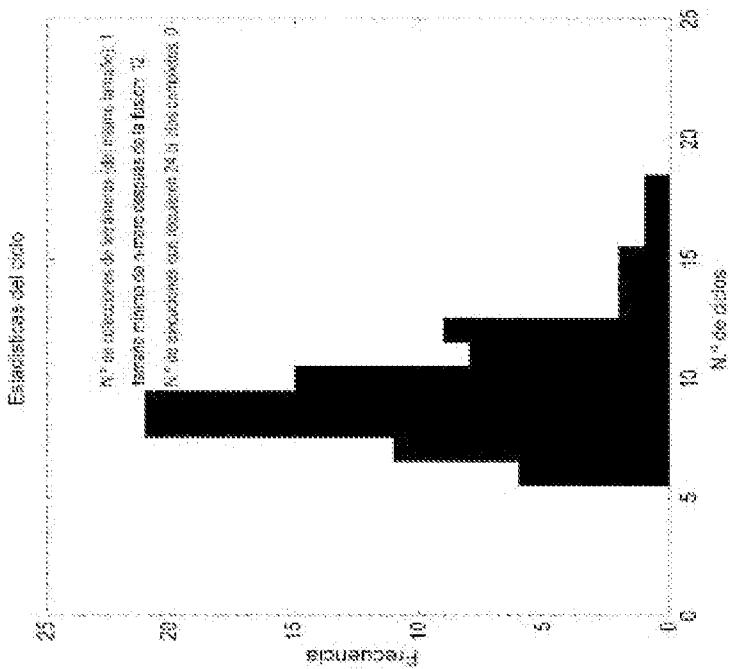
Fig. 47A*Fig. 47B*

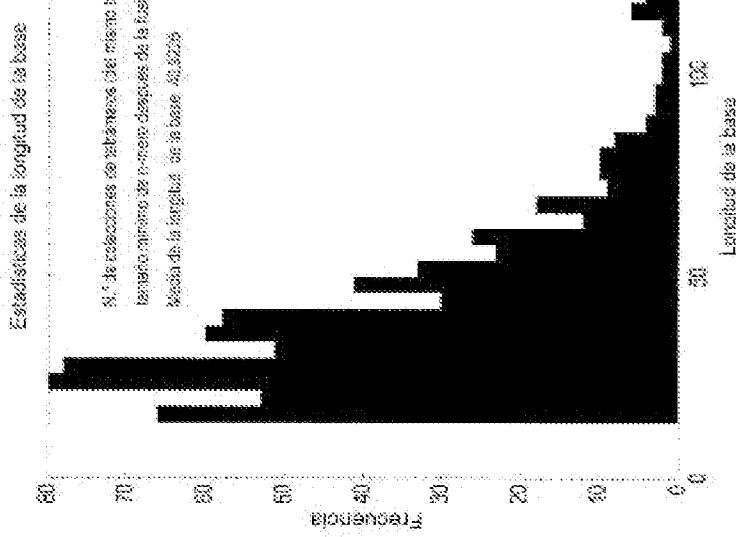
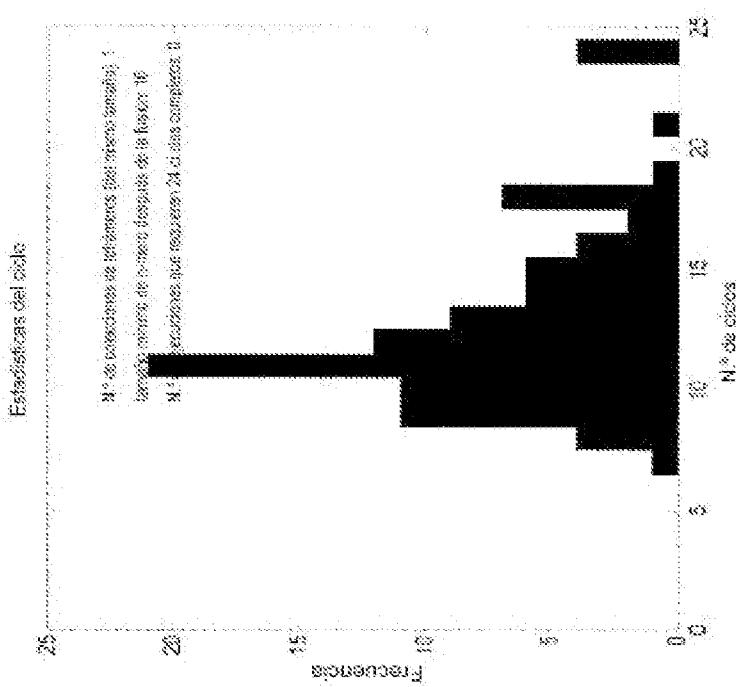
Fig. 48A*Fig. 48B*

Fig. 49

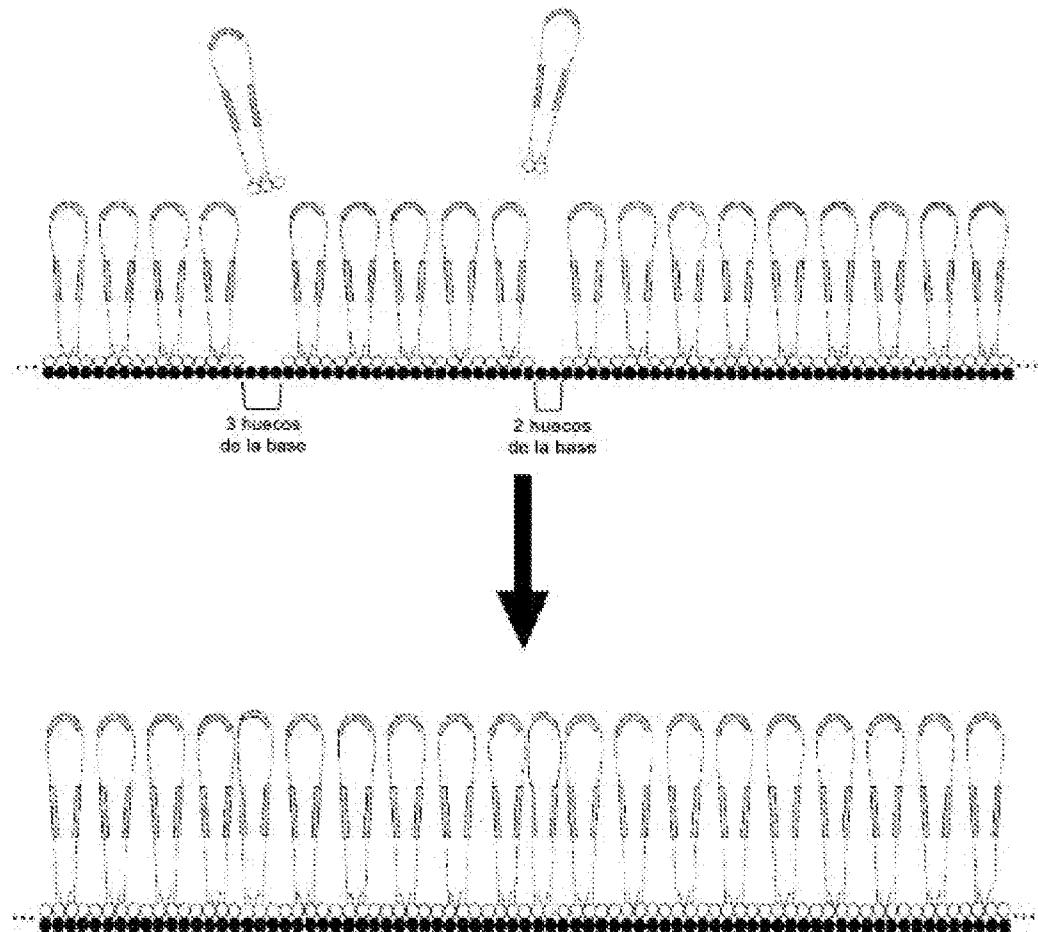


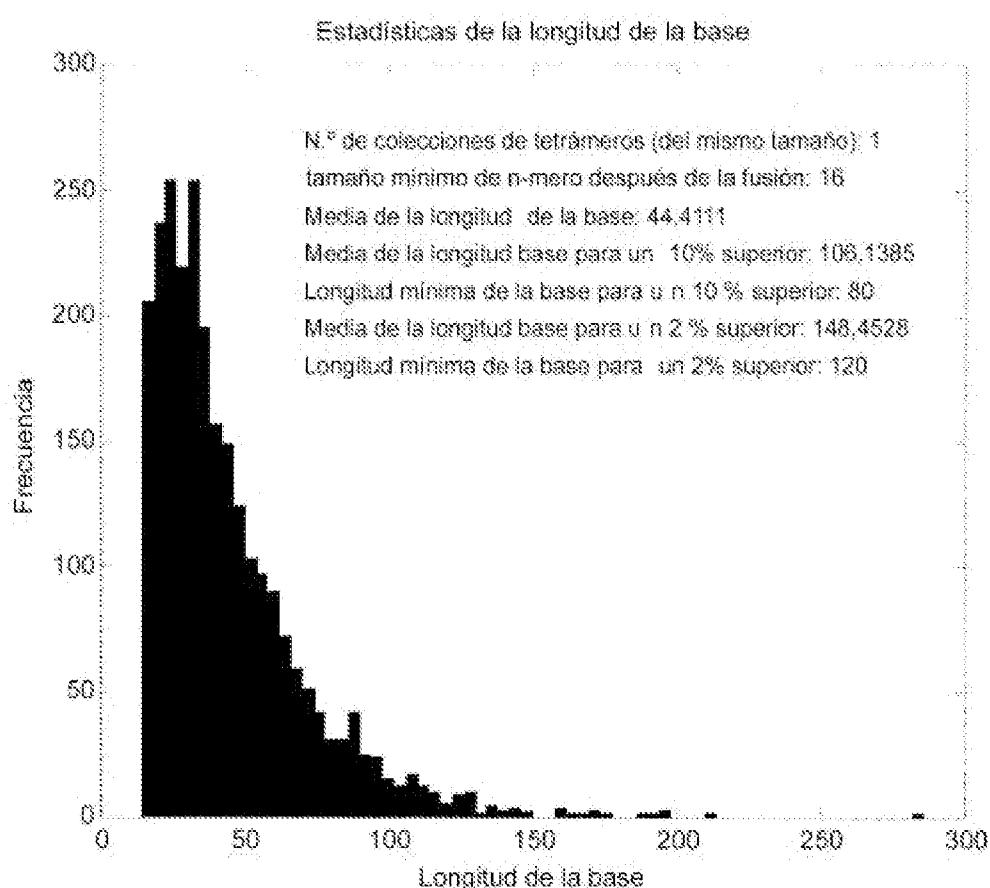
Fig. 50A

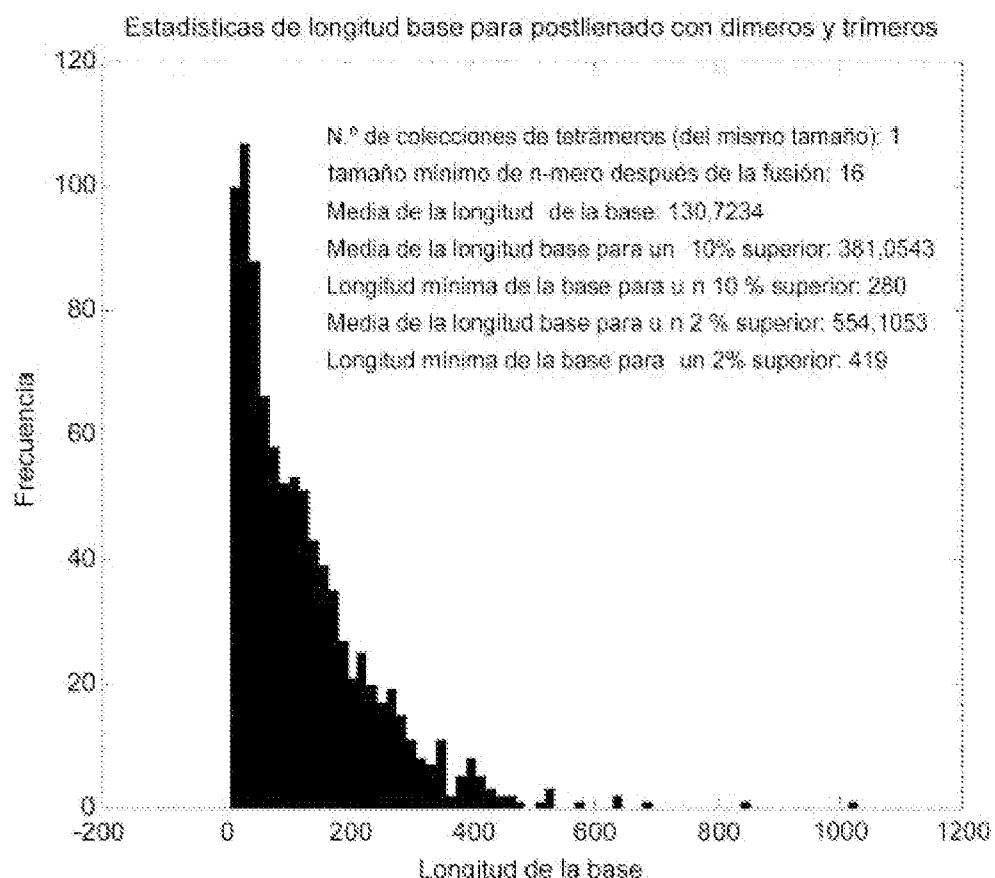
Fig. 50B

Fig. 51

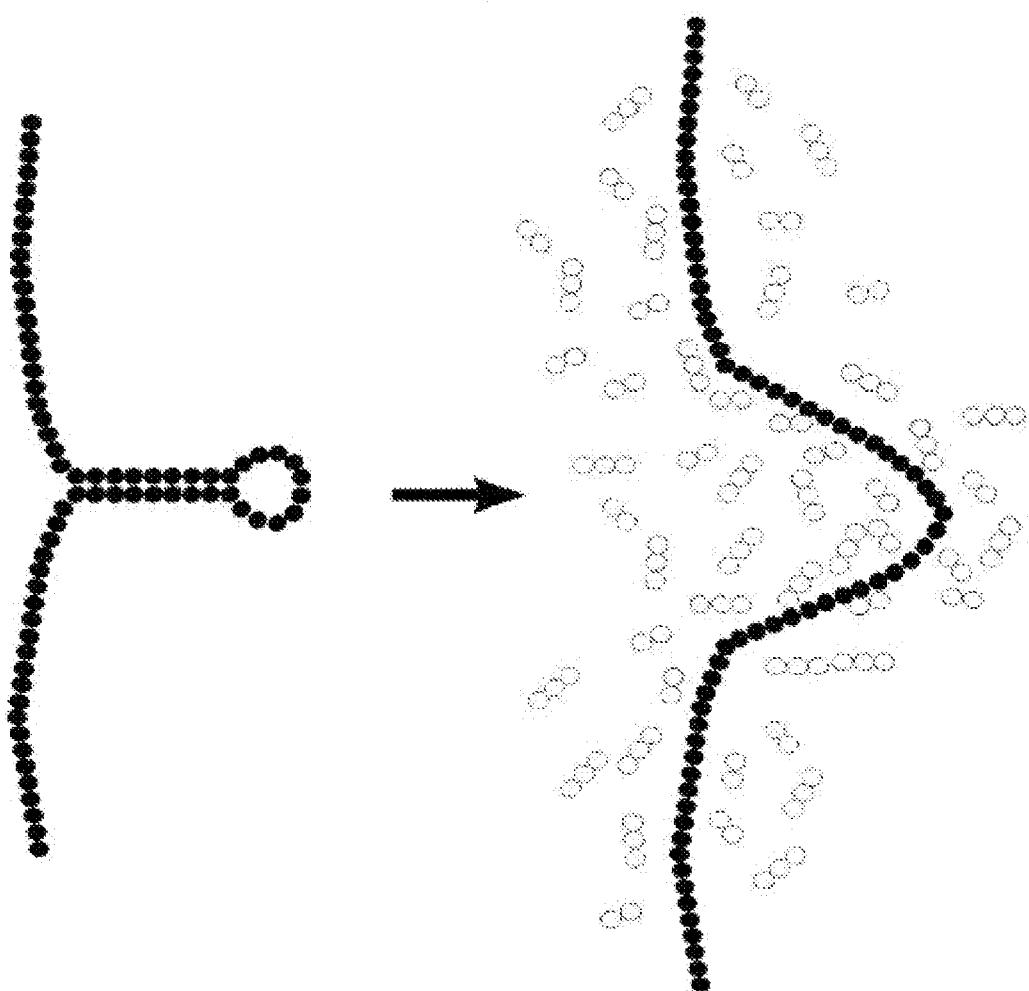


Fig. 52

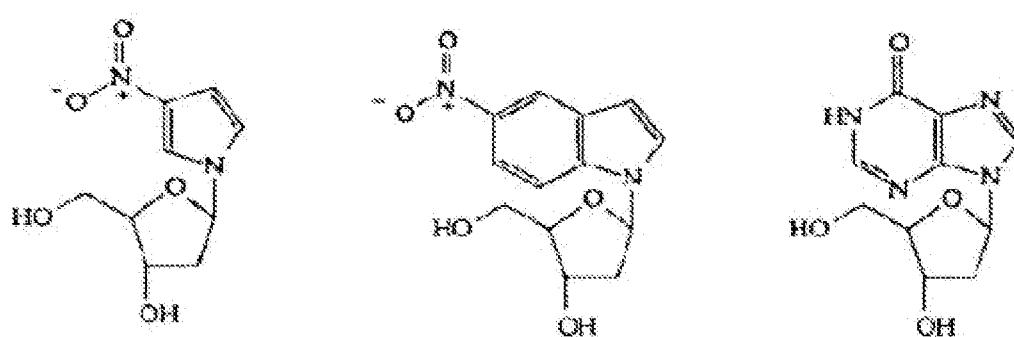
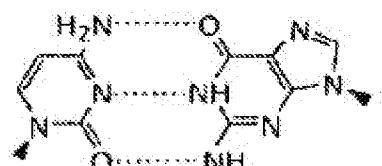
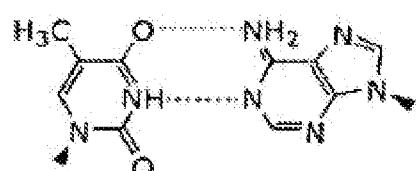


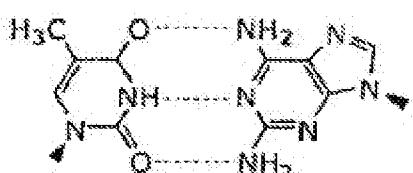
Fig. 53



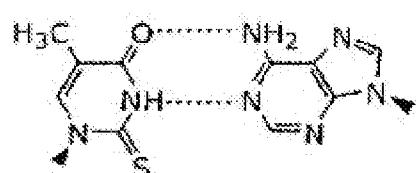
C - G



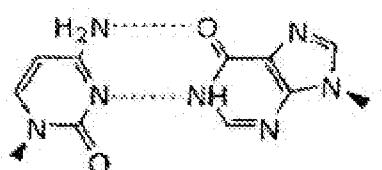
T (o U) - A



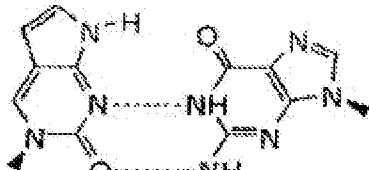
T (o U) - 2-amino A



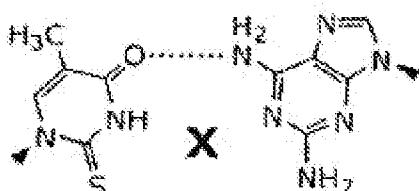
2-tio T (o U) - A



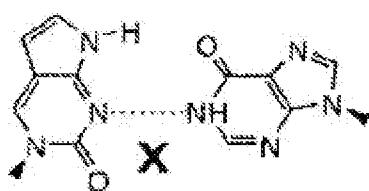
C - I



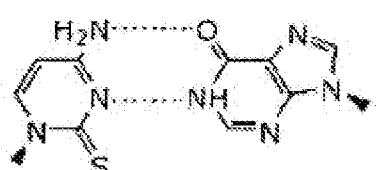
PirroloPir - G



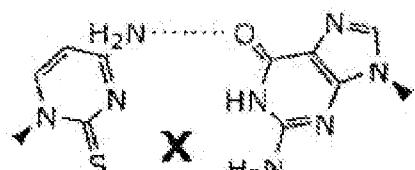
2-tio T (o U) - 2-amino A



PirroloPir - I



2-tio C - I



2-tio C - G

Fig. 54

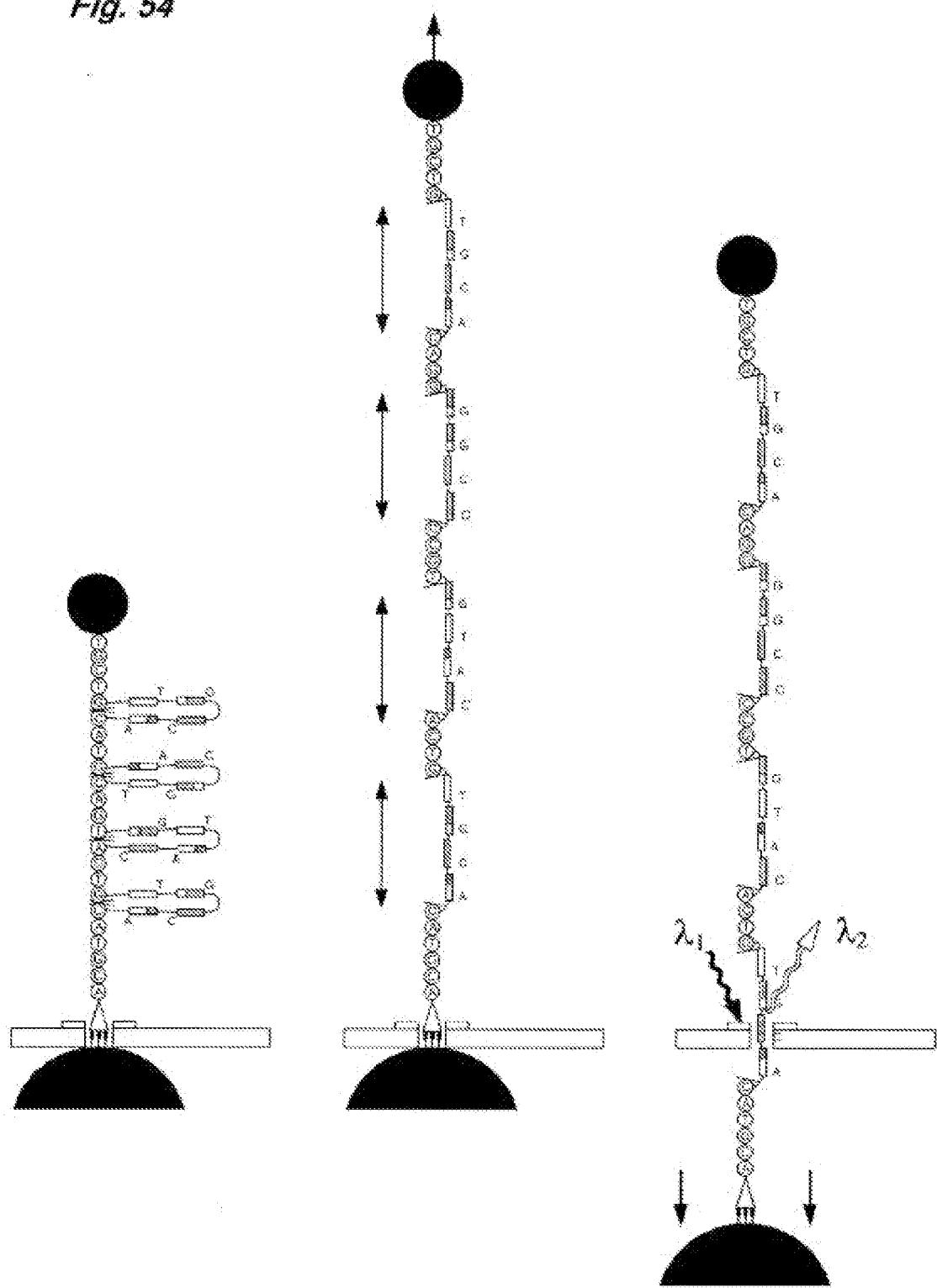


Fig. 55

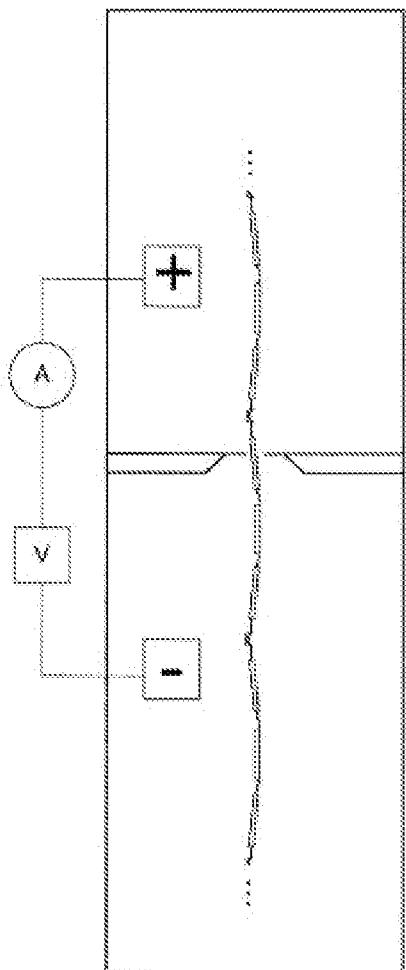


Fig. 56

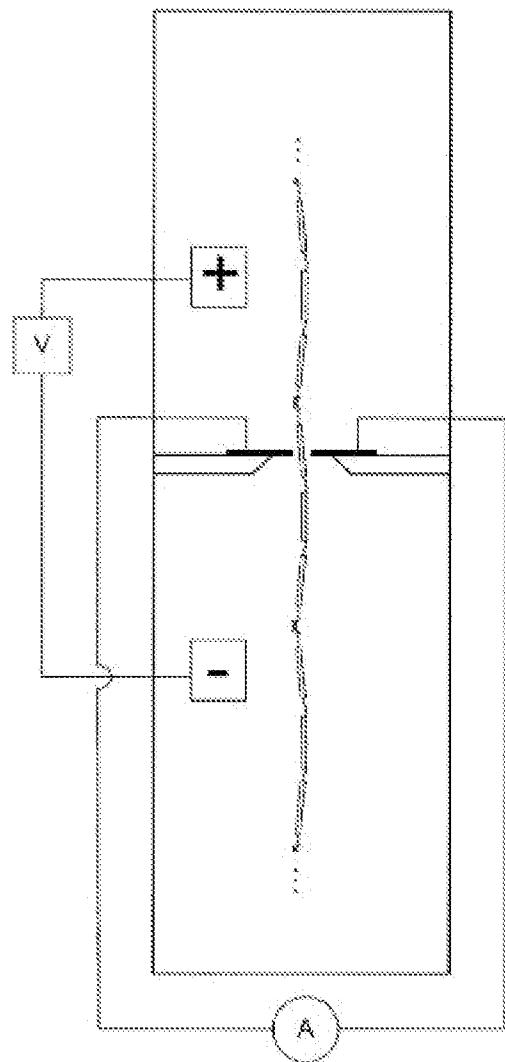


Fig. 57

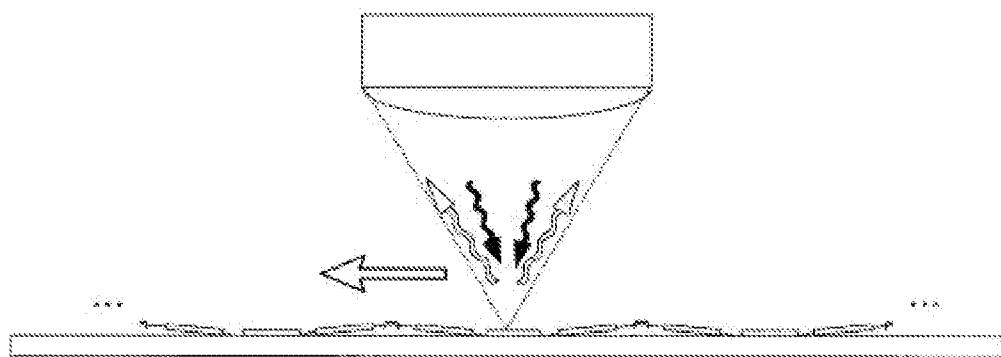


Fig. 58

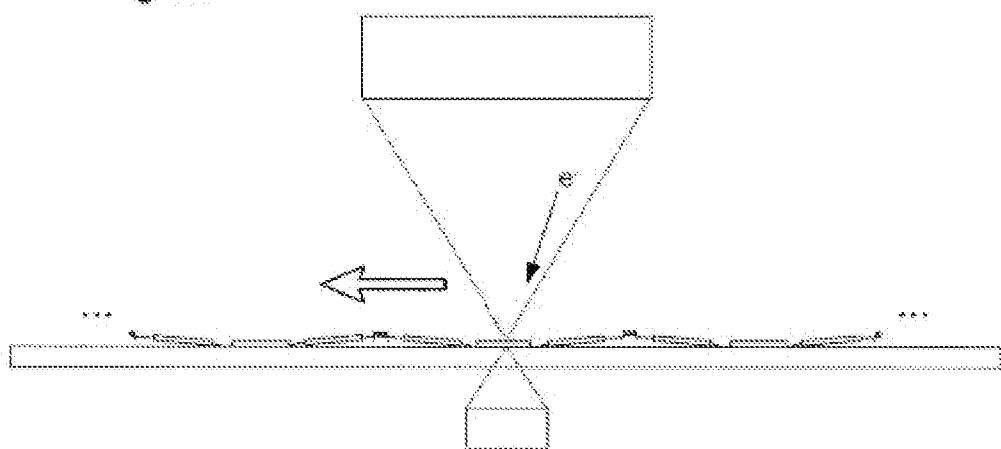


Fig. 59

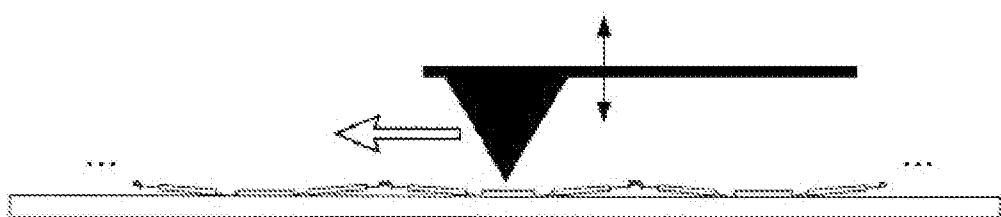


Fig. 60A

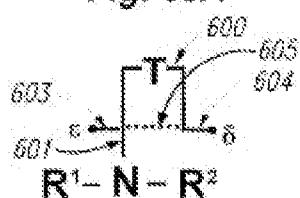


Fig. 60B

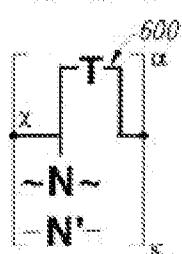


Fig. 60C

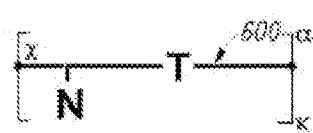


Fig. 60D

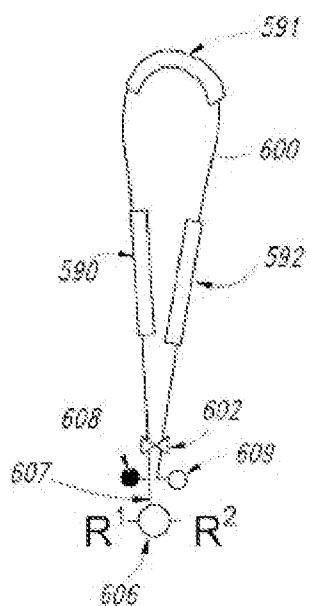


Fig. 60E

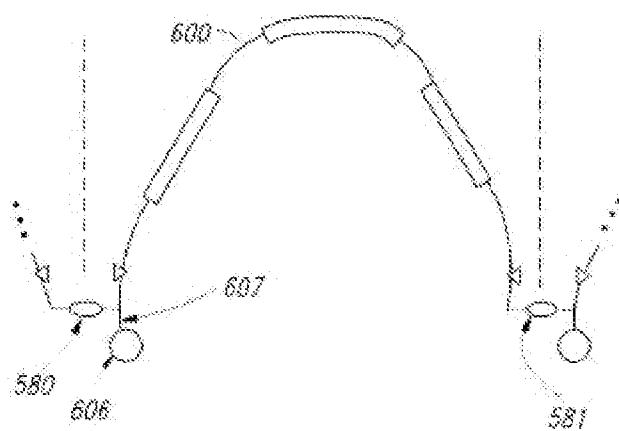


Fig. 6I

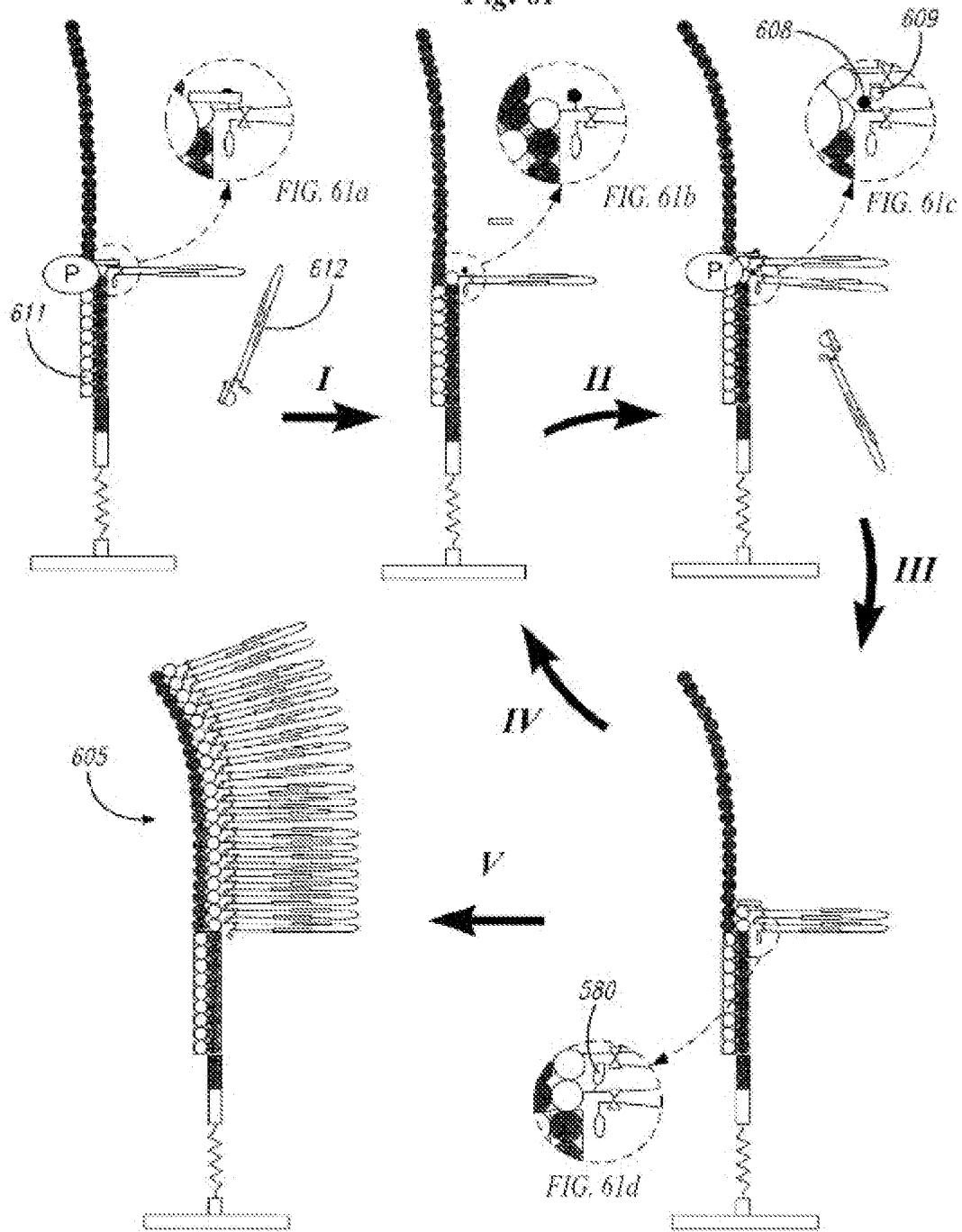


Fig. 62A

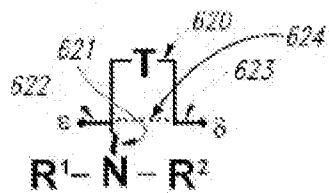


Fig. 62B

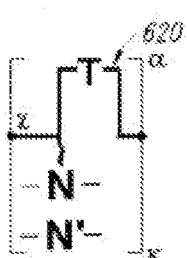


Fig. 62C

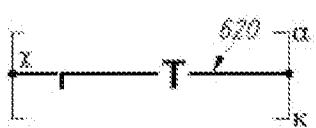


Fig. 62D

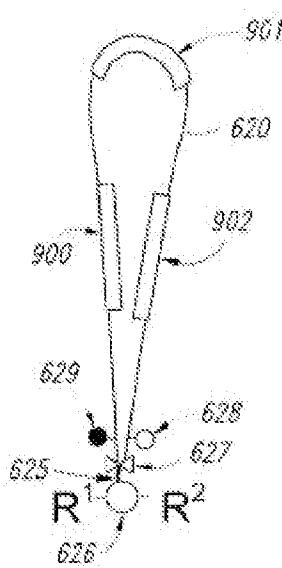


Fig. 62E

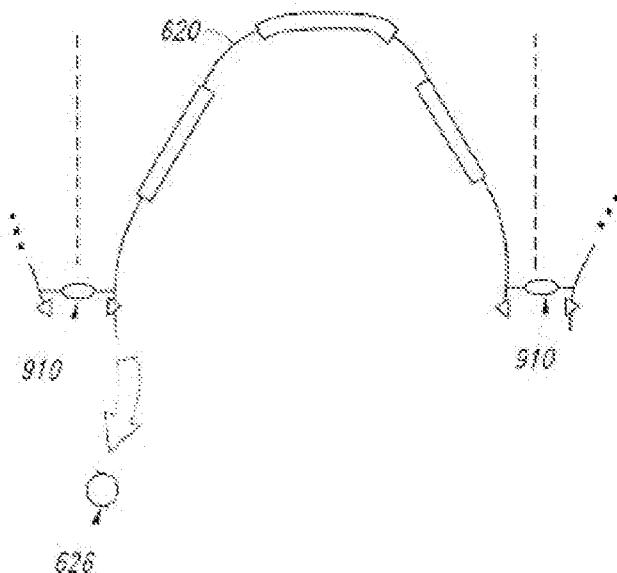


Fig. 63A

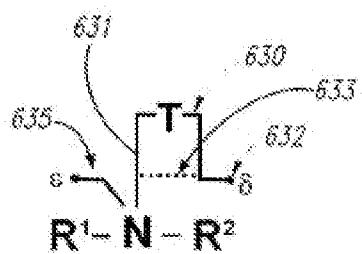


Fig. 63B

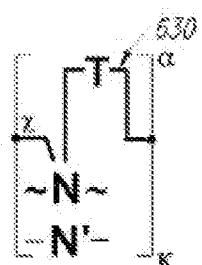


Fig. 63C

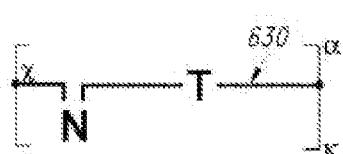


Fig. 63D

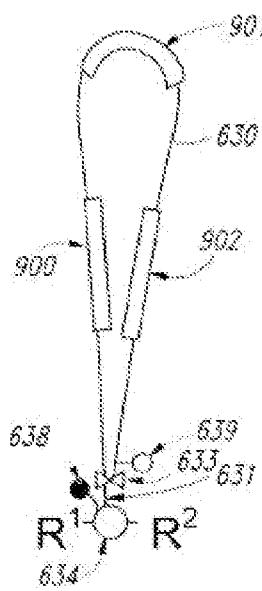


Fig. 63E

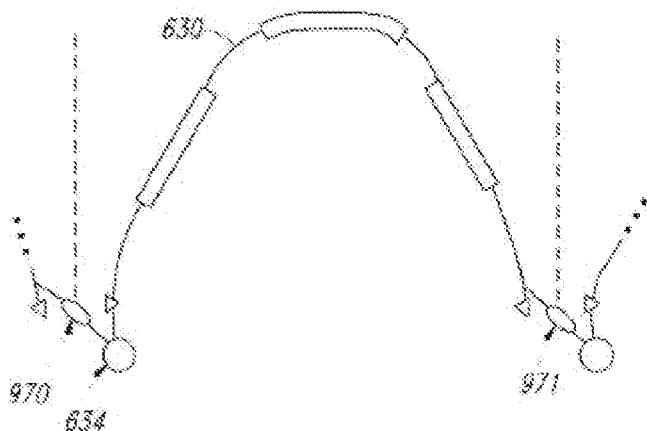


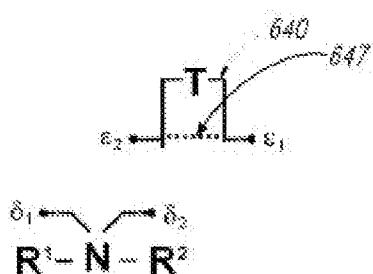
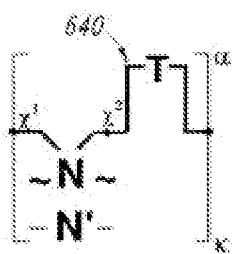
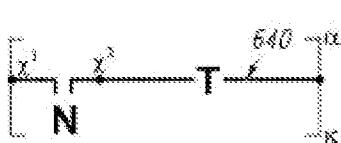
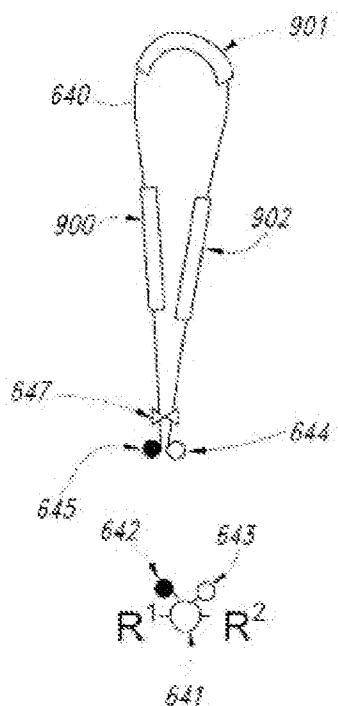
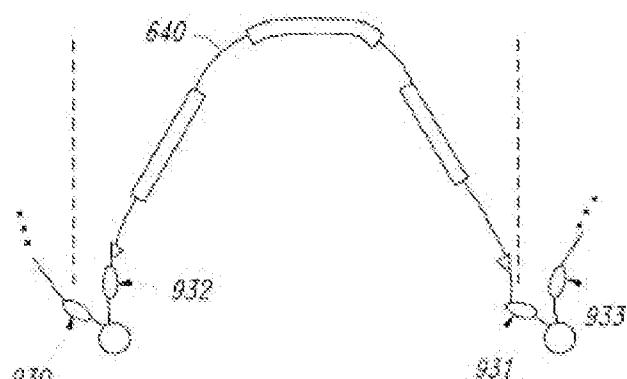
Fig. 64A*Fig. 64B**Fig. 64C**Fig. 64D**Fig. 64E*

Fig. 65

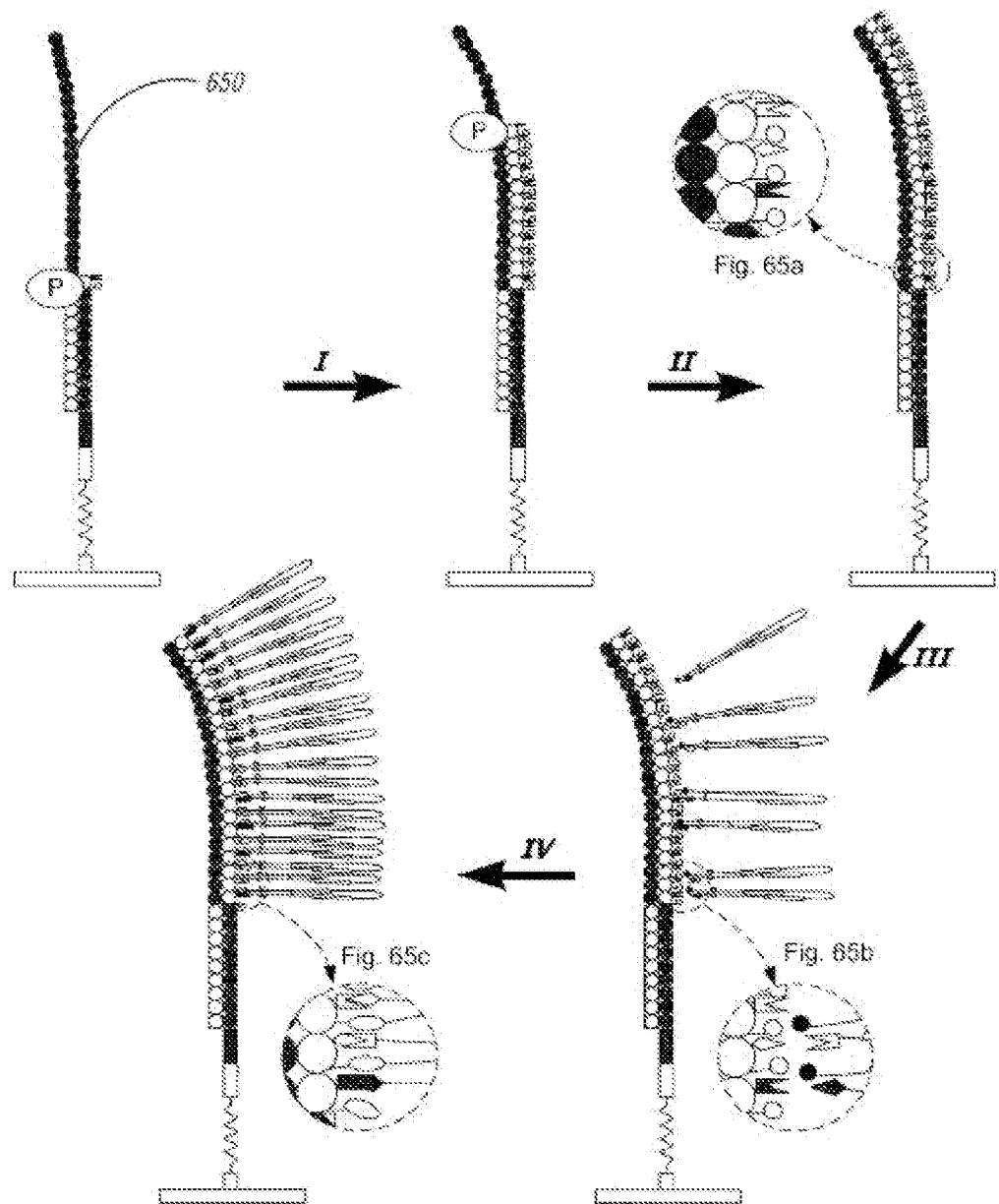


Fig. 66A

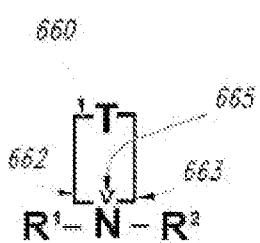


Fig. 66B

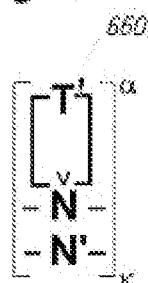


Fig. 66C

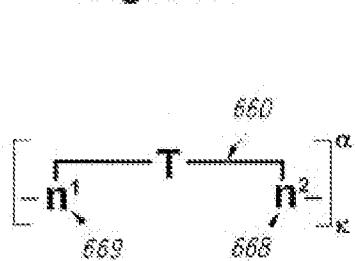


Fig. 66D

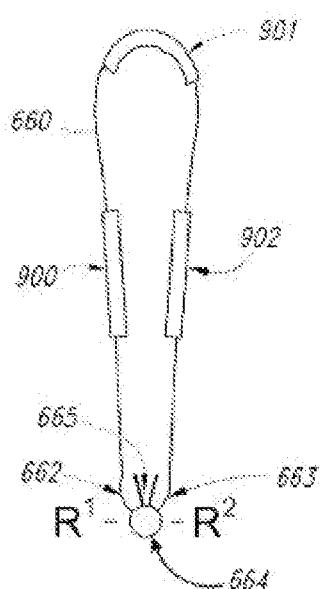


Fig. 66E

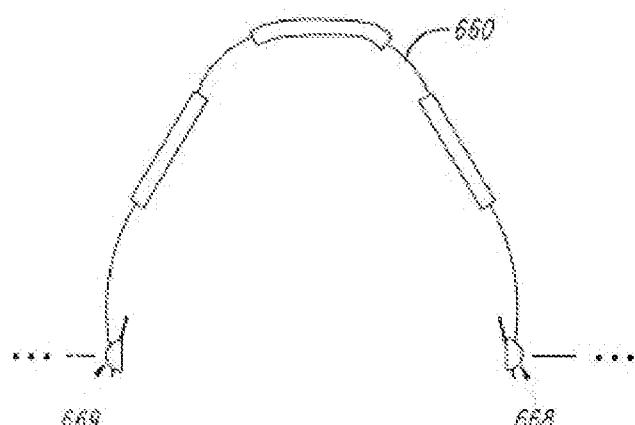


Fig. 67

