

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 638 324**

51 Int. Cl.:

C40B 40/08 (2006.01)

C40B 70/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **16.02.2010** **PCT/US2010/024314**

87 Fecha y número de publicación internacional: **19.08.2010** **WO10094036**

96 Fecha de presentación y número de la solicitud europea: **16.02.2010** **E 10741877 (4)**

97 Fecha y número de publicación de la concesión europea: **03.05.2017** **EP 2396459**

54 Título: **Métodos de creación y examen de bibliotecas codificadas por ADN**

30 Prioridad:

13.02.2009 US 152508 P

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

19.10.2017

73 Titular/es:

X-CHEM, INC. (100.0%)
100 Beaver Street, Suite 101
Waltham, Massachusetts 02453, US

72 Inventor/es:

WAGNER, RICHARD, W.

74 Agente/Representante:

SERRAT VIÑAS, Sara

ES 2 638 324 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

Métodos de creación y examen de bibliotecas codificadas por ADN

5 Antecedentes de la invención

El creciente coste del descubrimiento de fármacos ha conducido a la búsqueda continua de nuevos métodos de examen de mayor espacio químico lo más económicos posible para encontrar moléculas con mayor potencia y de poca a ninguna toxicidad. Los enfoques de química combinatoria en la década de 1980 se anunciaron originalmente como métodos para superar el paradigma del descubrimiento de fármacos, pero fracasaron en gran medida debido a tamaños de biblioteca insuficientes y a métodos de desconvolución inadecuados. Recientemente, el uso de bibliotecas combinatorias con presentación de ADN de moléculas pequeñas ha creado un nuevo cambio de paradigma para el examen de compuestos principales terapéuticos.

Morgan *et al.* (publicación de solicitud de patente estadounidense n.º 2007/0224607, identifican los principales retos en el uso de enfoques combinatorios con presentación de ADN en el descubrimiento de fármacos: (1) la síntesis de bibliotecas de complejidad suficiente y (2) la identificación de moléculas que son activas en los exámenes usados. Además, Morgan *et al.* establecen que cuanto mayor es el grado de complejidad de una biblioteca, es decir, el número de estructuras distintas presentes en la biblioteca, mayor es la probabilidad de que la biblioteca contenga moléculas con la actividad de interés. Por tanto, la química empleada en la síntesis de la biblioteca debe poder producir grandes números de compuestos dentro de un intervalo de tiempo razonable. Este enfoque ha sido generalmente satisfactorio en la identificación de moléculas con diversos quimiotipos y alta afinidad. Sin embargo, han surgido varias cuestiones con respecto a la generación de bibliotecas de enorme complejidad y a la evaluación del producto de secuenciación en la escala que se ha descrito. Por ejemplo, la purificación de una biblioteca tras múltiples transformaciones químicas (por ejemplo, habitualmente 3 ó 4 etapas) y transformaciones biológicas (por ejemplo, ligamiento enzimático de etiquetas de ADN) es complicada y da como resultado una cantidad significativa de "ruido" en la biblioteca debido o bien a la síntesis incompleta de moléculas o bien a etiquetado erróneo durante la etapa de ligamiento. Además, la cantidad de secuenciación que se requiere para estudiar poblaciones seleccionadas es notable, requiriéndose habitualmente métodos de secuenciación de "siguiente generación". Esto último se debe al hecho de que se requieren esquemas de etiquetado genético sofisticados incluidos en la parte de ADN de la biblioteca, junto con algoritmos bioinformáticos para analizar el producto de secuenciación de "siguiente generación", para filtrar el ruido e identificar resultados positivos en la biblioteca. Como resultado, incluso con estas metodologías, la secuenciación todavía no ha avanzado lo suficiente como para reflejar completamente la diversidad de secuencias (representando tanto resultados positivos reales como "ruido") a partir de un examen dado.

La presentación de ADN de bibliotecas combinatorias de moléculas pequeñas se basa en la síntesis de división y combinación de múltiples etapas de la biblioteca, acoplada a la adición enzimática de etiquetas de ADN que codifican tanto para la etapa de síntesis como para el elemento estructural usado. Normalmente se llevan a cabo y se codifican varias (por ejemplo, 3 ó 4) etapas de síntesis, y estas incluyen posiciones de diversidad (descritas en el presente documento como A, B y C (figura 1)), tal como las formadas mediante acoplamiento de elementos estructurales con, por ejemplo, grupos funcionales de amina o carboxilato sobre una estructura de soporte química que presenta los elementos estructurales unidos en orientaciones definidas. Un ejemplo de una estructura de soporte (S) que se usa a menudo en las bibliotecas combinatorias es un resto de triazina, que puede derivatizarse ortogonalmente en tres posiciones alrededor de su estructura de anillo.

El procedimiento de formación de bibliotecas puede llevar mucho tiempo, los productos a menudo se purifican de manera ineficaz y el resultado es que pueden producirse reacciones desconocidas que crean moléculas no deseadas y/o desconocidas unidas al ADN. Además, la purificación incompleta de la biblioteca puede dar como resultado etiquetas que producen contaminación cruzada durante las etapas de ligamiento, dando como resultado etiquetado erróneo. El resultado final para el examen y secuenciación de resultados positivos a partir de la biblioteca es que tiene emplearse secuenciación paralela de manera masiva debido al "ruido" inherente tanto de los ADN que se unen a moléculas que no se desean (por ejemplo, productos sin reaccionar o secundarios) como que están etiquetados de manera errónea. Por tanto, se pierde la eficacia de la secuenciación.

En algunos casos, un oligonucleótido iniciador, a partir del cual se construye la biblioteca de moléculas pequeñas, contiene una región de unión a cebador para la amplificación mediante polimerasa (por ejemplo, PCR) en forma de un oligonucleótido bicatenario, cerrado covalentemente. Este constructo es muy problemático para realizar reacciones de polimerasa, debido a la dificultad de fusionar el dúplex y permitir que se una un oligonucleótido cebador e inicie la polimerización, lo que da como resultado una reacción ineficaz, reduciendo el rendimiento en de 10 a 1000 veces o más.

Existe la necesidad de un enfoque más gradual para examinar e identificar moléculas pequeñas que tengan mayor potencia y de poca a ninguna toxicidad.

65 Sumario de la invención

La presente divulgación presenta un método para crear y examinar bibliotecas codificadas por ADN simplificadas, debido al menor número de etapas de síntesis (por ejemplo, sin ligamiento enzimático o sin oligonucleótidos bicatenarios iniciadores cerrados covalentemente) y, por tanto, sustancialmente menos “ruido” durante el análisis de los oligómeros codificados (denominados en el presente documento “regiones identificadoras”). Por tanto, la secuenciación se vuelve mucho más eficaz, o alternativamente, se hace posible el análisis de microalineamiento, teniendo en cuenta los sesgos inherentes que pueden dificultar la interpretación de los datos que pueden introducirse mediante la amplificación de la región codificante. También se han identificado métodos para crear una mayor diversidad de reacciones químicas en lugar de las limitadas simplemente a condiciones acuosas para hacer que la biblioteca codificada por ADN sea más hidrófoba y soluble en disolventes orgánicos para etapas posteriores de la síntesis de la biblioteca. De este modo, pueden llevarse a cabo reacciones químicas con rendimiento potencialmente superior, una mayor diversidad de elementos estructurales y fidelidad mejorada de las reacciones químicas.

Por consiguiente, la presente divulgación presenta un método de etiquetado de bibliotecas químicas codificadas por ADN uniendo un primer grupo funcional de un ligador bifuncional a un oligonucleótido iniciador en el extremo 5' del oligonucleótido iniciador, en el que el oligonucleótido iniciador forma una estructura de horquilla, y uniendo un segundo grupo funcional del ligador bifuncional a un componente de la biblioteca química. El oligonucleótido iniciador puede incluir una primera región identificadora y una segunda región identificadora, de manera que la segunda región identificadora se hibrida con la primera región identificadora del oligonucleótido iniciador. La segunda región identificadora puede incluir una etiqueta fluorescente (por ejemplo, un fluoróforo o GFP) o etiqueta de biotina. Además, la segunda región identificadora no se amplifica antes del análisis tras una etapa de selección.

En otra realización, la divulgación presenta un método de creación de bibliotecas codificadas por ADN (a) creando un primer nodo de diversidad, (b) codificando el primer nodo de diversidad en recipientes independientes, (c) combinando el primer nodo de diversidad y (d) dividiendo el primer nodo de diversidad combinado en un segundo conjunto de recipientes independientes, en el que el primer nodo de diversidad reacciona para formar un segundo nodo de diversidad. En determinadas realizaciones, el segundo nodo de diversidad no se codifica ni se combina.

En otra realización, la presente divulgación presenta un método para crear bibliotecas usando reacciones químicas semiacuosas o no acuosas (por ejemplo, orgánicas) con rendimiento superior, una mayor diversidad de elementos estructurales y un mayor número de reacciones químicas que pueden usarse para crear más bibliotecas combinatorias con etiquetado de ADN que las logradas anteriormente.

En general, los métodos de la presente divulgación proporcionan un conjunto de bibliotecas que contienen, por ejemplo, una o dos posiciones de diversidad en una estructura de soporte química que puede generarse eficazmente con alto rendimiento, examinarse para identificar elementos estructurales individuales preferidos o combinaciones de elementos estructurales que residen, por ejemplo, en la una o dos posiciones de diversidad, y diversificarse iterativamente, por ejemplo, en una segunda, tercera y/o cuarta posición de diversidad para crear moléculas con propiedades mejoradas. Además, los métodos descritos en el presente documento permiten un análisis amplio y extenso de los compuestos seleccionados que tienen una propiedad biológica deseada lo que, a su vez, permite que se identifiquen compuestos relacionados con relaciones estructurales familiares (por ejemplo, relaciones de estructura-actividad).

Por “estructura de soporte” quiere decirse un resto químico que presenta nodo(s) de diversidad en una geometría especial particular. El/los nodo(s) de diversidad normalmente se une(n) a la estructura de soporte durante la síntesis de la biblioteca, pero en algunos casos un nodo de diversidad puede unirse a la estructura de soporte antes de la síntesis de la biblioteca (por ejemplo, adición de regiones identificadoras). En algunas realizaciones, la estructura de soporte se derivatiza de manera que puede desprotegerse ortogonalmente durante la síntesis de la biblioteca y hacerse reaccionar posteriormente con diferentes nodos de diversidad (por ejemplo, usando etiquetado identificador en cada etapa).

Por “región identificadora” quiere decirse la parte de etiqueta de ADN de la biblioteca que codifica para la adición de elemento estructural a la biblioteca.

Por “oligonucleótido iniciador” quiere decirse el oligonucleótido de partida para la síntesis de la biblioteca que también contiene un ligador unido covalentemente y un resto funcional para la adición de un nodo de diversidad o estructura de soporte. El oligonucleótido puede ser mono o bicatenario. El oligonucleótido puede consistir en bases naturales o modificadas.

Por “resto funcional” quiere decirse un resto químico que comprende uno o más elementos estructurales que pueden seleccionarse de cualquier molécula pequeña o diseñarse y construirse basándose en características deseadas, por ejemplo, de solubilidad, disponibilidad de donadores y aceptores de enlaces de hidrógeno, grados de libertad rotacionales de los enlaces, carga positiva, carga negativa y similares. El resto funcional debe ser compatible con la modificación química de manera que reacciona con la parte anterior. En determinadas realizaciones, el resto funcional puede hacerse reaccionar adicionalmente como una entidad bifuncional o trifuncional (o mayor). Los restos funcionales también pueden incluir elementos estructurales que se usan en cualquiera de las posiciones o nodos de

diversidad. En las tablas 1 y 2 se encuentran ejemplos de elementos estructurales y etiquetas de ADN codificante. Véase, por ejemplo, la publicación de solicitud de patente estadounidense n.º 2007/0224607.

Por "elemento estructural" quiere decirse una unidad estructural química que se une a otras unidades estructurales químicas o que puede unirse a otras unidades de este tipo. Cuando el resto funcional es polimérico u oligomérico, los elementos estructurales son las unidades monoméricas del polímero o el oligómero. Los elementos estructurales también pueden incluir una estructura de soporte (por ejemplo, un elemento estructural de estructura de soporte) a la que se unen, o pueden unirse, una o más estructuras adicionales (por ejemplo, elementos estructurales periféricos). Los elementos estructurales pueden ser cualquier compuesto químico que sea complementario (es decir, los elementos estructurales deben poder reaccionar entre sí para formar una estructura que comprende dos o más elementos estructurales). Normalmente, todos los elementos estructurales usados tendrán al menos dos grupos reactivos, aunque algunos de los elementos estructurales usados tendrán sólo un grupo reactivo cada uno. Los grupos reactivos en dos elementos estructurales diferentes deben ser complementarios, es decir, deben poder reaccionar entre sí para formar un enlace covalente.

Por "ligador" quiere decirse una molécula que une la parte de ácido nucleico de la biblioteca a las especies presentadas funcionales. Tales ligadores se conocen en la técnica, y los que pueden usarse durante la síntesis de la biblioteca incluyen, pero no se limitan a, 5'-O-dimetoxitritil-1',2'-didesoxirribosa-3'-[(2-cianoetil)-(N,N-diisopropil)]-fosforamidita; 1-[(2-cianoetil)-(N,N-diisopropil)]-fosforamidita de 9-O-dimetoxitritil-trietilenglicol; 3-(4,4'-dimetoxitritiloxi)propil-1-[(2-cianoetil)-(N,N-diisopropil)]-fosforamidita; y 1-[(2-cianoetil)-(N,N-diisopropil)]-fosforamidita de 18-O-dimetoxitritilhexaetilenglicol. Tales ligadores pueden añadirse en tándem entre sí en diferentes combinaciones para generar ligadores de diferentes longitudes deseadas. Por "ligador ramificado" quiere decirse una molécula que une la posición de ácido nucleico de la biblioteca a 2 o más especies funcionales idénticas de la biblioteca. Los ligadores ramificados se conocen bien en la técnica y los ejemplos pueden consistir en dobletes simétricos o asimétricos (1) y (2) o un triplete asimétrico (3). Véanse, por ejemplo, Newcome *et al.*, *Dendritic Molecules: Concepts, Synthesis, Perspectives*, VCH Publishers (1996); Boussif *et al.*, *Proc. Natl. Acad. Sci. USA* 92: 7297-7301 (1995); y Jansen *et al.*, *Science* 266: 1226 (1994).

Tal como se usa en el presente documento, el término "oligonucleótido" se refiere a un polímero de nucleótidos. El oligonucleótido puede incluir ADN o cualquier derivado del mismo conocido en la técnica que puede sintetizarse y usarse para el reconocimiento de pares de bases. El oligonucleótido no tiene que tener bases contiguas, sino que puede presentar restos de ligador intercalados. El polímero oligonucleotídico puede incluir nucleósidos naturales (por ejemplo, adenosina, timidina, guanosa, citidina, uridina, desoxiadensina, desoxitimidina, desoxiguanosina y desoxicitidina), análogos de nucleósido (por ejemplo, 2-aminoadenosina, 2-tiotimidina, inosina, pirrolo-pirimidina, 3-metiladenosina, C5-propinilcitidina, C5-propiniluridina, C5-bromouridina, C5-fluorouridina, C5-yodouridina, C5-metilcitidina, 7-desazaadenosina, 7-desazaguanosina, 8-oxoadenosina, 8-oxoguanosina, O(6)-metilguanina y 2-tiocitidina), bases modificadas químicamente, bases modificadas biológicamente (por ejemplo, bases metiladas), bases intercaladas, azúcares modificados (por ejemplo, 2'-fluororribosa, ribosa, 2'-desoxirribosa, arabinosa y hexosa) y/o grupos fosfato modificados (por ejemplo, enlaces fosforotioatos y 5'-N-fosforamidita).

Por "unido operativamente" quiere decirse que dos estructuras químicas se unen entre sí de tal manera que permanecen unidas a través de las diversas manipulaciones a las que se espera que se sometan. Normalmente, el resto funcional y el oligonucleótido codificante se unen covalentemente a través de un grupo de unión apropiado. Por ejemplo, el grupo de unión puede ser un resto bifuncional con un sitio de unión para el oligonucleótido codificante y un sitio de unión para el resto funcional.

Por "molécula pequeña" quiere decirse una molécula que tiene un peso molecular menor de aproximadamente 1000 Dalton. Las moléculas pequeñas pueden ser orgánicas o inorgánicas, y pueden aislarse, por ejemplo, de bibliotecas de compuestos o fuentes naturales, o pueden obtenerse mediante derivatización de compuestos conocidos.

Otras características y ventajas de la invención resultarán evidentes a partir de la siguiente descripción detallada, los dibujos, los ejemplos y las reivindicaciones.

Breve descripción de los dibujos

La figura 1 es una representación esquemática que ilustra las posiciones de diversidad A, B, y C.

La figura 2 es una representación esquemática de un elemento de biblioteca química codificada por ADN de modo 1, que muestra, en parte, el oligonucleótido iniciador, que incluye una estructura de horquilla complementaria en la región identificadora, que ha reaccionado con los nodos de diversidad A y B. Está añadiéndose la región identificadora para B. En esta figura, el nodo de diversidad "C" es la posición potencial para una posición de diversidad adicional que va a añadirse tras la adición de la región identificadora B.

La figura 3 es una representación esquemática de un elemento de biblioteca química codificada por ADN de modo 1, que muestra, en parte, el oligonucleótido iniciador, que incluye una secuencia en la región de bucle de la estructura

de horquilla que puede servir como región de unión a cebador para amplificación.

La figura 4 es una representación esquemática de un elemento de biblioteca química codificada por ADN de modo 1, que muestra, en parte, el oligonucleótido iniciador, que incluye una secuencia no complementaria en el extremo 3' de la molécula que puede servir para unir una segunda región identificadora o bien para la polimerización o bien para el ligamiento enzimático.

La figura 5 es una representación esquemática de un elemento de biblioteca química codificada por ADN de modo 1, que muestra, en parte, el oligonucleótido iniciador, en el que la región de bucle del oligonucleótido iniciador y al menos la región identificadora en el lado 3' de la región de bucle pueden servir para hibridarse con un nucleótido complementario que también contiene una segunda región identificadora.

La figura 6 es una representación esquemática de amplificación por PCR del modelo de horquilla, tal como se presenta en la figura 5.

La figura 7 es una representación esquemática de un elemento de biblioteca química codificada por ADN de modo 2, que muestra un oligonucleótido en horquilla que está cerrado covalentemente (por ejemplo, a través de una horquilla o químicamente) en el extremo distal al ligador.

La figura 8 es una representación esquemática de un elemento de biblioteca química codificada por ADN de modo 2, que muestra la inclusión de nodos de diversidad adicionales.

La figura 9 es una representación esquemática de un elemento de biblioteca química codificada por ADN de modo 2, que muestra las etapas para examinar bibliotecas y métodos para la desconvolución de las regiones identificadoras.

La figura 10 es una representación esquemática que muestra oligonucleótidos usados en la síntesis de la biblioteca. La parte anterior (HP) la sintetizó IDT DNA y se purificó mediante HPLC. Las flechas indican el sitio para la restricción por BbvCI (subrayado) o el digesto de formación de mella por Nb.BbvCI o Nt.BbvCI. También se muestran las secuencias de las etiquetas de ADN A1, B1 y C1 (cadenas superior e inferior), los cebadores de PCR en 5' y 3', y el extremo 3' de la HP.

La figura 11 es un gel electroforético (electroforesis en gel de TBE-urea (15%); sombreado UV en una placa de CCF) de la parte anterior en diferentes etapas de su síntesis. Se aciló la parte anterior HP (IDT DNA) mediante Fmoc-amino-PEG2000-NHS (JenKem Technology USA). El carril 1 es el oligonucleótido de HP (IDT DNA) (42 nucleótidos). El carril 2 es la HP acilada con Fmoc-amino-PEG2000-NHS. Tras la adición de Tris-HCl, se observó cierta desprotección de Fmoc. El carril 3 es la reacción en bruto con piperidina, que muestra la desprotección completa de Fmoc. El carril 4 es igual al carril 3 tras la desalinización en una columna de NAP-5 y la liofilización. (XC: xilencianol (migra como ADN de 60 nt); BPB: azul de bromofenol (migra como ADN de 15 nt)).

La figura 12 es una representación esquemática que muestra las etapas en la síntesis de la biblioteca modelo. Se conjugó DTAF con la parte anterior modificada por amino-PEG (HP-1) en la primera etapa. Tras esta etapa, se aciló adicionalmente una parte de HP-1-DTAF con pentilamino-biotina.

La figura 13A es un esquema del ligamiento de las etiquetas de ADN. La figura 13B ilustra un gel de agarosa al 4% de la biblioteca de HP-1-DTAF-biotina en diferentes etapas del ligamiento de etiqueta de ADN. M: marcador; carril 1: HP-1-DTAF-biotina; carril 2: 1 + etiqueta A solo; carril 3: 1 + etiquetas A, B y C, así como oligonucleótido en el extremo 3' ligado. La flecha indica fluorescencia verde brillante (DTAF). No se observa separación sustancial en el gel. La figura 13C ilustra amplificación por PCR (24 ciclos) de las reacciones de ligamiento. M: marcador (la banda inferior es 100); carril 1: amplificación por PCR de la banda fluorescente verde del carril 1 de la figura 14B (HP-1-DTAF-biotina + etiqueta A); carril 2: amplificación por PCR de la banda fluorescente verde del carril 2 de la figura 13B (HP-1-DTAF-biotina + las 3 etiquetas y el oligonucleótido en el extremo 3'); carril 3: amplificación por PCR de la reacción de ligamiento en bruto HP-1-DTAF-biotina + las 3 etiquetas; carril 4: sin control de molde.

La figura 14 es un conjunto de geles electroforéticos que muestra la purificación del compuesto modelo XChem y la selección de modelo (a través de una interacción de unión entre el resto de biotina del compuesto modelo XChem y estreptavidina). Los geles son geles de SDS NuPage al 4-12% con tampón de transferencia MES. Se exploraron los geles para determinar la fluorescencia verde usando un láser de 450 nm. La figura 14A es un gel que muestra las etapas de síntesis y purificación. Se mezclaron las muestras con tampón de carga y se sometieron a ebullición. M: marcador; carril 1: HP-1 + DTAF; carriles 2 y 2a: HP-1-DTAF + biotina (dos reacciones independientes); carriles 3-6 (etapas de purificación/selección de modelo usando perlas Dynal con estreptavidina): carril 3: fracciones no retenidas; carril 4: último lavado (lavado con agua a 80°C durante 10 minutos); carriles 5 y 5': elución con EDTA 25 mM a 90°C (1° y 2°); carriles 6 y 6': elución con EDTA 25 mM y NaOH a 5 mM a 90°C (1° y 2°). La figura 14B es un gel que muestra la unión de HP-1-DTAF-biotina ("biblioteca de 1") a estreptavidina. Se mezclaron las muestras con tampón de carga de gel y se cargaron directamente en el gel sin ebullición. Se incubaron las muestras, como en el gel de la figura 14A, con un exceso de estreptavidina en NaCl 50 mM/Tris HCl 10 mM, pH 7,0, durante 10 minutos. "S" indica la adición de estreptavidina. Se combinaron las muestras 5 y 6. Carril 1: HP-1-DTAF; carril 1S: HP-1-DTAF

+ estreptavidina; carril 2: HP-1-DTAF-biotina (desalinizado); carril 2S: HP-1-DTAF-biotina + estreptavidina; carril 4: último lavado (lavado con agua a 80°C durante 10 minutos); carril 4S: muestra de último lavado + estreptavidina; carril 5+6: muestras combinadas 5, 5', 6 y 6' (fracciones de elución de perlas de estreptavidina, HP-1-DTAF-biotina purificada y seleccionada); carril 5+6S': HP-1-DTAF-biotina purificada y seleccionada + estreptavidina. Obsérvese que no hay ninguna diferencia notable en la migración entre las diferentes etapas de síntesis de la "biblioteca de 1". La figura 14C es un gel de agarosa al 4% de la parte anterior (Trilink) HP-T, que se hace reaccionar con DTAF. Carril 1: marcador; carril 2: DTAF; carril 3: HP-T-DTAF. Panel izquierdo: visualización UV del gel (tinción con bromuro de etidio); panel derecho: mismo gel explorado para determinar la fluorescencia a la longitud de onda de excitación de 450 nm (verde, fluoresceína). La figura 14D es un gel de SDS NuPage al 4-12% con tampón de transferencia MES, que muestra la unión de HP-T-DTAF-biotina a estreptavidina. Se mezclaron las muestras con tampón de carga de gel y se cargaron directamente en el gel sin ebullición. Se incubaron las muestras, como en el gel de la figura 14A, con un exceso de estreptavidina en NaCl 50 mM/Tris HCl 10 mM, pH 7,0, durante 10 minutos. Carril 1: DTAF; carril 2: HP-T-DTAF; carril 3: HP-T-DTAF + estreptavidina; carril 4: HP-T-DTAF-biotina (desalinizado); carril 5: HP-T-DTAF-biotina + estreptavidina; carril 6: muestras combinadas 5, 5', 6 y 6' (fracciones de elución de perlas de estreptavidina, HP-1-DTAF-biotina purificada y seleccionada); carril 7: HP-1-DTAF-biotina purificada y seleccionada + estreptavidina.

La figura 15A es un esquema de la síntesis del constructo para el experimento de administración intracelular de ARN-P de T7. Se amplificó mediante PCR el clon de ADNbc de V_H para añadir un sitio de BsmI en el extremo 5' en el sentido de 5' del promotor de T7. Tras la digestión por restricción y la purificación, se ligó el constructo a HP-1-DTAF-R7 (parte anterior modificada con DTAF y péptido (-Arg-εAhx)₆-Arg). La figura 15B es un gel electroforético de la reacción de ligamiento. Los carriles 1 y 2 muestran diferentes muestras de HP-1 ligadas a V_H; el carril 3 muestra el producto de PCR de V_H no ligado; y M es el marcador. La figura 15C es un gel electroforético que muestra la validación para la actividad del promotor de T7. El gel muestra una reacción de T7 Megascript (Ambion, Inc.) usando muestras de los carriles 1-3 de la figura 15B.

La figura 16 es una electroforesis en gel de agarosa de las etapas en la síntesis de la biblioteca 10x10 La figura 16A es un gel de agarosa al 4% de parte anterior (Trilink) HP-T ligada con la etiqueta A. Carril 1: marcador; carril 2: HP-T; carril 3: etiqueta A apareada; carril 4: HP-T ligada con la etiqueta A; carril 5: HP-T ligada con la etiqueta A y desalinizada en la columna Zeba. La figura 16B es un gel de agarosa al 2% del ligamiento de HP-T-A con 12 etiquetas B diferentes. Carril M: marcador, carriles 1 y 9: HP-T-A; carriles 3, 4, 5, 6, 7, 8, 11, 12, 13, 14, 15 y 16: ligamiento de HP-T-A con las etiquetas B1 - B12. La figura 16C es un gel de agarosa al 4% de la biblioteca combinada (biblioteca B), con las etiquetas A y B1-B12 ligadas, tras la reacción con cloruro cianúrico y aminos B1-B12. Carril 1: Marcador; carril 2: HPT- A; carril 3: biblioteca-B combinada y desalinizada en columnas Zeba.

Descripción detallada de la invención

La presente divulgación presenta varios métodos para identificar uno o más compuestos que se unen a una diana biológica. Los métodos incluyen sintetizar una biblioteca de compuestos, en la que los compuestos contienen un resto funcional que tiene una o más posiciones de diversidad. El resto funcional de los compuestos está unido operativamente a un oligonucleótido iniciador que identifica la estructura del resto funcional. En resumen, el modo 1 proporciona varios métodos para conservar el carácter bicatenario del ADNbc durante la síntesis de la biblioteca, lo que es importante durante la etapa de reacción química, y puede usarse (tal como se muestra en las figuras 2-6) para generar hasta dos nodos de diversidad. El modo 2 (figuras 7-9) prevé un nodo de diversidad y usa un oligonucleótido en horquilla que está cerrado covalentemente (por ejemplo, a través de una horquilla o químicamente) en el extremo distal al ligador. El modo 3 proporciona métodos para crear bibliotecas con uno, dos, tres o más nodos de diversidad. Los modos 1, 2 y 3 se describen en detalle a continuación.

Modo 1

La presente divulgación presenta un método para identificar uno o más compuestos que se unen a una diana biológica. El método incluye sintetizar una biblioteca de compuestos, en la que los compuestos contienen un resto funcional que no tiene más de dos posiciones de diversidad. El resto funcional de los compuestos está unido operativamente a un oligonucleótido iniciador que identifica la estructura del resto funcional proporcionando una disolución que contiene A compuestos iniciadores.

El oligonucleótido iniciador incluye un ligador L (por ejemplo, polietilenglicol) con un número entero de uno o más, en el que los oligonucleótidos iniciadores contienen un resto funcional que incluye elementos estructurales A unidos a L y separados en A recipientes de reacción, en los que A es un número entero de dos o más, que está unido operativamente a un oligonucleótido iniciador que identifica los elementos estructurales A.

En algunas realizaciones, los elementos estructurales A pueden derivatizarse adicionalmente a través de un nodo común S. En otras realizaciones, A se transforma posteriormente con S, siendo S una molécula de estructura de soporte que permite la introducción de nodos de diversidad adicionales. En algunas realizaciones, A-S pueden examinarse directamente, representando un solo nodo de diversidad. En otras realizaciones, los recipientes de reacción con A-S (por ejemplo, que pueden incluir en primer lugar una purificación de A-S a partir de los materiales

de partida) se mezclan entre sí y se introducen alícuotas en B recipientes de reacción, en los que B es un número entero de uno o más, y se hacen reaccionar con uno de elementos estructurales B. A-S-B, todavía en B recipientes de reacción, se hace reaccionar en algunos casos con elementos estructurales C, donde C es un número entero de uno, se purifica y se somete a una reacción de polimerización o ligamiento usando cebadores B, en los que los cebadores B difieren en cuanto a la secuencia e identifican los elementos estructurales B.

En determinadas realizaciones, A-S puede ser un número entero de uno. En una realización, AS puede unirse directamente a oligonucleótidos iniciadores B, y tras la reacción de elementos estructurales B, se mezclan las B reacciones. En determinadas realizaciones, la mezcla de A-S-B, donde B representa el único nodo de diversidad, se examina directamente, representando un solo nodo de diversidad. En otras realizaciones, se introducen posteriormente alícuotas de la mezcla de A-S-B, donde B representa el único nodo de diversidad, en C recipientes de reacción, se hacen reaccionar con elementos estructurales C, y se someten a reacción de polimerización o ligamiento de segunda cadena usando cebadores C, en los que los cebadores C difieren en cuanto a la secuencia e identifican los elementos estructurales C.

En determinadas realizaciones, B puede ser un número entero de uno y A-S es mayor de uno, en cuyo caso se introducen alícuotas de A-S, ahora derivatizado con B, en C recipientes de reacción, se hacen reaccionar con elementos estructurales C y se someten a reacción de polimerización de segunda cadena usando cebadores C, en los que los cebadores C difieren en cuanto a la secuencia e identifican los elementos estructurales C. Esta estrategia general puede ampliarse para incluir nodos de diversidad adicionales (por ejemplo, D, E, F, etc.) de modo que el primer nodo de diversidad se hace reaccionar con elementos estructurales y/o S y se codifica por un oligonucleótido inicial, se mezcla, se introducen nuevas alícuotas en recipientes y entonces se derivatiza el nodo de diversidad posterior mediante elementos estructurales, lo que se codifica mediante el cebador usado para la reacción de polimerización o ligamiento.

En determinadas realizaciones, A puede ser un número entero de uno, B puede ser un número entero de uno y se usan oligonucleótidos iniciadores C. A-S-B, unido a oligonucleótidos iniciadores C, se forma en C recipientes de reacción, se hacen reaccionar con elementos estructurales C y se examinan directamente.

En determinadas realizaciones, S se hace reaccionar en primer lugar con el oligonucleótido iniciador, y A, B y/o C (por ejemplo, o D, E, F, etc.) se hacen reaccionar posteriormente.

En determinadas realizaciones, A, B o C (por ejemplo, o D, E, F, etc.) puede contener sitios para nodos de diversidad adicionales. Si este es el caso, entonces S puede usarse o necesitarse o no para introducir nodos de diversidad adicionales.

En una realización, el oligonucleótido iniciador incluye una estructura de horquilla complementaria en la región identificadora (figura 2). La región identificadora puede tener, por ejemplo, de 2 a 100 pares de bases de longitud, preferiblemente de 5 a 20 pares de bases de longitud, y lo más preferiblemente de 6 a 12 pares de bases de longitud. El oligonucleótido iniciador incluye además una secuencia en la región de bucle de la estructura de horquilla que puede servir como una región de unión a cebador para amplificación (figura 3), de manera que la región de unión a cebador tiene una temperatura de fusión mayor para su cebador complementario (por ejemplo, que puede incluir regiones identificadoras flanqueantes) que la región identificadora sola.

En una realización, la región de bucle puede incluir bases modificadas que pueden formar formaciones dúplex de afinidad mayor que las bases no modificadas, conociéndose tales bases modificadas en la técnica (figura 3). El oligonucleótido iniciador puede incluir además una secuencia no complementaria en el extremo 3' de la molécula que puede servir para unir una segunda región identificadora o bien para polimerización o bien para ligamiento enzimático (figura 4). En una realización, las cadenas pueden reticularse posteriormente, por ejemplo, usando psoraleno.

En otra realización, la región de bucle y al menos la región identificadora en el lado 3' de la región de bucle pueden servir para hibridarse con un nucleótido complementario que también contiene una segunda región identificadora (figura 5). En los casos en los que se usan muchos elementos estructurales y etiquetas correspondientes (por ejemplo, 100 etiquetas), puede emplearse una estrategia de mezclado y división durante la etapa de síntesis de oligonucleótidos para crear el número de etiquetas necesario. Tales estrategias de mezclado y división para la síntesis de ADN se conocen en la técnica. En una realización, las cadenas pueden reticularse posteriormente, por ejemplo, usando psoraleno. Los elementos de la biblioteca resultante pueden amplificarse mediante PCR tras la selección de entidades de unión frente a una(s) diana(s) de interés (figura 6).

Por ejemplo, puede hacerse reaccionar una parte anterior, que incluye un oligonucleótido iniciador, con un ligador y A, que incluye, por ejemplo, 1000 variantes diferentes. Por cada elemento estructural A, puede ligarse una etiqueta A de ADN o puede extenderse el cebador a la parte anterior. Estas reacciones pueden realizarse, por ejemplo, en una placa de 1000 pocillos o en 10 placas de 100. Todas las reacciones pueden combinarse, purificarse opcionalmente y dividirse en un segundo conjunto de placas. A continuación, puede realizarse el mismo procedimiento con elementos estructurales B, que también incluyen, por ejemplo, 1000 variantes diferentes. Puede

ligarse una etiqueta B de ADN a la parte anterior y pueden combinarse todas las reacciones. Una biblioteca de 1000 x 1000 combinaciones de A con B (es decir, 1.000.000 compuestos), puede etiquetarse mediante 1.000.000 combinaciones de etiquetas diferentes. Puede extenderse el mismo enfoque para añadir variantes C, D, E, etc. La biblioteca generada puede usarse entonces para identificar compuestos que se unen a la diana. La composición de los compuestos que se unen a la biblioteca puede evaluarse mediante PCR y secuenciación de las etiquetas de ADN para identificar los compuestos que se enriquecieron.

Modo 2

En otra realización (figura 7), el método incluye sintetizar una biblioteca de compuestos, en la que los compuestos contienen un resto funcional que no tiene más de dos posiciones de diversidad. El resto funcional de los compuestos está unido operativamente a un oligonucleótido iniciador, que contiene una secuencia genética única que identifica la estructura del resto funcional proporcionando una disolución que comprende compuestos iniciadores A, en la que L es un número entero de uno o más, donde los compuestos iniciadores incluyen un resto funcional que tiene elementos estructurales A separados en A recipientes de reacción, donde, por ejemplo, A es un número entero de dos o más, que está unido operativamente a un oligonucleótido inicial que identifica los elementos estructurales A. En algunas realizaciones, los elementos estructurales A se derivatizan previamente con un S común. En otras realizaciones, A se transforma posteriormente con S, siendo S una molécula de estructura de soporte que permite la introducción de nodos de diversidad adicionales. A continuación, los recipientes de reacción con A-S (que pueden incluir en primer lugar una purificación de A-S a partir de los materiales de partida) se mezclan entre sí y se introducen alícuotas en B recipientes de reacción, en los que B es un número entero de uno o más, y se hacen reaccionar con uno de elementos estructurales B. En algunas realizaciones se hace reaccionar A-S-B, todavía en B recipientes de reacción, con elementos estructurales C, donde C es un número entero de uno, se purifican y se mantienen separados en B recipientes para el examen. En algunas realizaciones, A-S es un número entero de uno. En una realización, A-S puede unirse directamente a oligonucleótidos iniciadores B y, tras la reacción de elementos estructurales B, se mezclan las B reacciones y se introducen alícuotas en C recipientes de reacción, se hacen reaccionar con elementos estructurales C y se mantienen separados en C recipientes para el examen. En otras realizaciones, B puede ser un número entero de uno y A-S es mayor de uno, en cuyo caso se introducen alícuotas de A-S, ahora derivatizado con B, en C recipientes de reacción, se hacen reaccionar con elementos estructurales C y se mantienen separados en C recipientes para el examen. Esta estrategia general puede ampliarse para incluir nodos de diversidad adicionales (por ejemplo, D, E, F, etc.) de modo que el primer nodo de diversidad se hace reaccionar con elementos estructurales y/o S y se codifica por un oligonucleótido iniciador, se mezcla, se introducen nuevas alícuotas en recipientes, y entonces se derivatiza el nodo de diversidad posterior mediante elementos estructurales y se mantiene en sus respectivos recipientes para el examen (figura 8).

Por ejemplo, tal como se describe en el modo 1, una parte anterior, que incluye un oligonucleótido iniciador, puede hacerse reaccionar con un ligador y elementos estructurales A, que incluyen, por ejemplo, 1000 variantes diferentes. Por cada elemento estructural A, puede ligarse una etiqueta A de ADN o puede extenderse el cebador a la parte anterior. Las reacciones pueden combinarse. A continuación, puede realizarse el mismo procedimiento con elementos estructurales B, pero no se añade una etiqueta de ADN para B. Puesto que B no se codifica, todas las reacciones de "B" pueden combinarse (por ejemplo, 1000 reacciones) y puede realizarse una etapa de selección para identificar todos los elementos estructurales A que producen el efecto de unión deseado con elementos estructurales B desconocidos. Entonces puede hacerse reaccionar una biblioteca de elementos estructurales A identificada en la etapa de selección (por ejemplo, 10 elementos estructurales A) con los mismos 1000 elementos estructurales B, dando como resultado un examen de 10.000 compuestos o menos. En esta ronda, pueden añadirse etiquetas de ADN para B y pueden identificarse los elementos estructurales B que producen el efecto de unión deseado en combinación con los, por ejemplo, 10 elementos estructurales A, dando como resultado una convolución gradual de una biblioteca inicial de, por ejemplo, 1.000.000 compuestos. Puede someterse a prueba un conjunto de estos compuestos finales de manera individual para identificar, por ejemplo, los mejores compuestos de unión, activadores o inhibidores.

Para evitar la combinación de todas las reacciones tras la síntesis de B, puede usarse un lector BIND (SRU Biosystems), por ejemplo, para monitorizar la unión en una superficie de sensor en formato de alto rendimiento (por ejemplo, placas de 384 pocillos y placas de 1536 pocillos). Por ejemplo, los elementos estructurales A pueden codificarse con etiquetas de ADN y los elementos estructurales B pueden codificarse por posición. Entonces pueden identificarse compuestos de unión usando un sensor BIND, secuenciación y análisis de microalineamiento o análisis de digesto de restricción de las etiquetas A. Este análisis permite la identificación de combinaciones de elementos estructurales A y B que producen las moléculas deseadas. Pueden usarse otros métodos para monitorizar la unión conocidos por los expertos en la técnica incluyendo, por ejemplo, ELISA.

Modos 1 y 2

El oligonucleótido iniciador de los modos 1 y 2 puede contener una estructura de horquilla, complementaria a la región identificadora. El oligonucleótido iniciador contiene además una secuencia en la región de bucle de la estructura de horquilla que puede servir como una región de unión a cebador para amplificación, de manera que la región de unión a cebador tiene una temperatura de fusión mayor para su cebador complementario (que puede

incluir regiones identificadoras flanqueantes) que la región identificadora sola.

En una realización, el oligonucleótido iniciador incluye una molécula de ligador que puede hacerse reaccionar funcionalmente con elementos estructurales. La molécula de ligador puede unirse directamente al extremo 5' del oligonucleótido a través de métodos conocidos en la técnica o puede incluirse dentro de la molécula, por ejemplo, en una base derivatizada (por ejemplo, la posición C5 de uridina), o el ligador puede colocarse en el medio del oligonucleótido usando técnicas convencionales conocidas en la técnica.

El oligonucleótido iniciador puede ser monocatenario o bicatenario. La formación de un oligonucleótido bicatenario puede lograrse a través de la formación de la horquilla del oligonucleótido o a través de reticulación usando, por ejemplo, un resto de psoraleno, tal como se conoce en la técnica.

El oligonucleótido iniciador puede contener dos regiones de unión a cebador (por ejemplo, para permitir una reacción de PCR) en cada lado de la región identificadora que codifica para el elemento estructural. Alternativamente, el oligonucleótido iniciador puede contener un sitio de unión a cebador en el extremo 5'. En otras realizaciones, el oligonucleótido iniciador es una horquilla, y la región de bucle forma un sitio de unión a cebador o el sitio de unión a cebador se introduce a través de hibridación de un oligonucleótido con la región identificadora en el lado 3' del bucle. Un oligonucleótido cebador, que contiene una región homóloga al extremo 3' del oligonucleótido iniciador y que porta una región de unión a cebador en su extremo 5' (por ejemplo, para permitir una reacción de PCR) puede hibridarse con el oligonucleótido iniciador, y puede contener una región identificadora que codifica para los elementos estructurales usados en una de las posiciones de diversidad. El oligonucleótido cebador puede contener información adicional, tal como una región de nucleótidos aleatorizados, por ejemplo, de 2 a 16 nucleótidos de longitud, que se incluye para el análisis bioinformático.

En una realización, el oligonucleótido iniciador no contiene un sitio de unión a cebador para PCR.

En otra realización, se pone en contacto la biblioteca de compuestos, o una parte de la misma, con una diana biológica en condiciones adecuadas para que al menos un miembro de la biblioteca de compuestos se una a la diana, seguido por la retirada de los miembros de la biblioteca que no se unen a la diana, y el análisis de la región o regiones identificadoras. Las dianas biológicas a modo de ejemplo incluyen, por ejemplo, enzimas (por ejemplo, cinasas, fosfatasas, metilasas, desmetilasas, proteasas y enzimas de reparación del ADN), proteínas implicadas en interacciones proteína:proteína (por ejemplo, ligandos para receptores), dianas de receptores (por ejemplo, GPCR y RTK), canales iónicos, bacterias, virus, parásitos, ADN, ARN, priones o hidratos de carbono).

En una realización, la biblioteca de compuestos, o una parte de la misma, se pone en contacto con una diana biológica en condiciones adecuadas para que al menos un miembro de la biblioteca de compuestos se una a la diana, seguido por eliminación de miembros de la biblioteca que no se unen a la diana, seguido por amplificación de la región identificadora mediante métodos conocidos en la técnica, y posteriormente análisis de la región o regiones identificadoras mediante métodos conocidos en la técnica.

En una realización, el método de amplificación de la región identificadora puede incluir, por ejemplo, reacción en cadena de la polimerasa (PCR), amplificación de cadena lineal (LCR), amplificación de círculo rodante (RCA) o cualquier otro método conocido en la técnica para amplificar secuencias de ácido nucleico.

En una realización adicional, la biblioteca de compuestos no se combina tras la etapa final de adición de elementos estructurales y las combinaciones se examinan individualmente para identificar compuesto(s) que se une(n) a una diana.

En otra realización, las moléculas que se unen a una diana no se someten a amplificación, sino que se analizan directamente. Los métodos de análisis incluyen, por ejemplo, análisis de microalineamiento o métodos basados en perlas para la desconvolución de las regiones identificadoras (figura 9). Las moléculas que se unen durante la etapa de examen también pueden detectarse mediante un biosensor de cristal fotónico sin etiquetas.

En una realización, el oligonucleótido iniciador y/o el oligonucleótido cebador contienen un resto funcional que permite su detección, por ejemplo, mediante etiquetas fluorescentes, puntos Q o biotina.

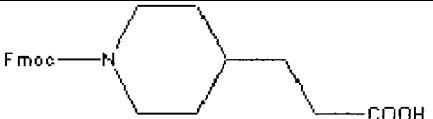
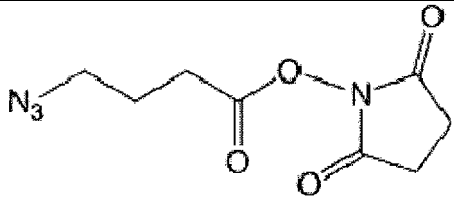
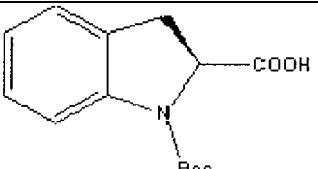
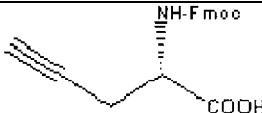
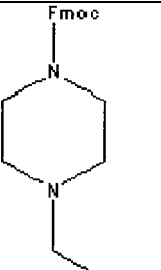
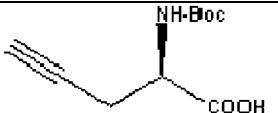
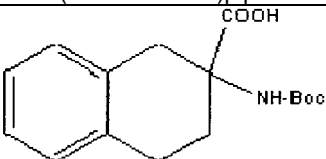
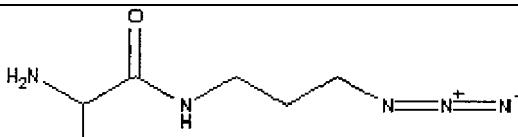
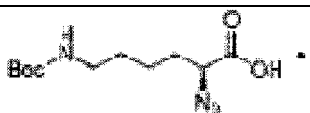
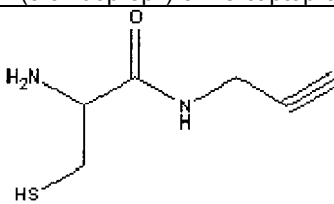
En una realización, el análisis de microalineamiento usa capacidad de detección avanzada, tal como, por ejemplo, cristales fotónicos de resonancia evanescente.

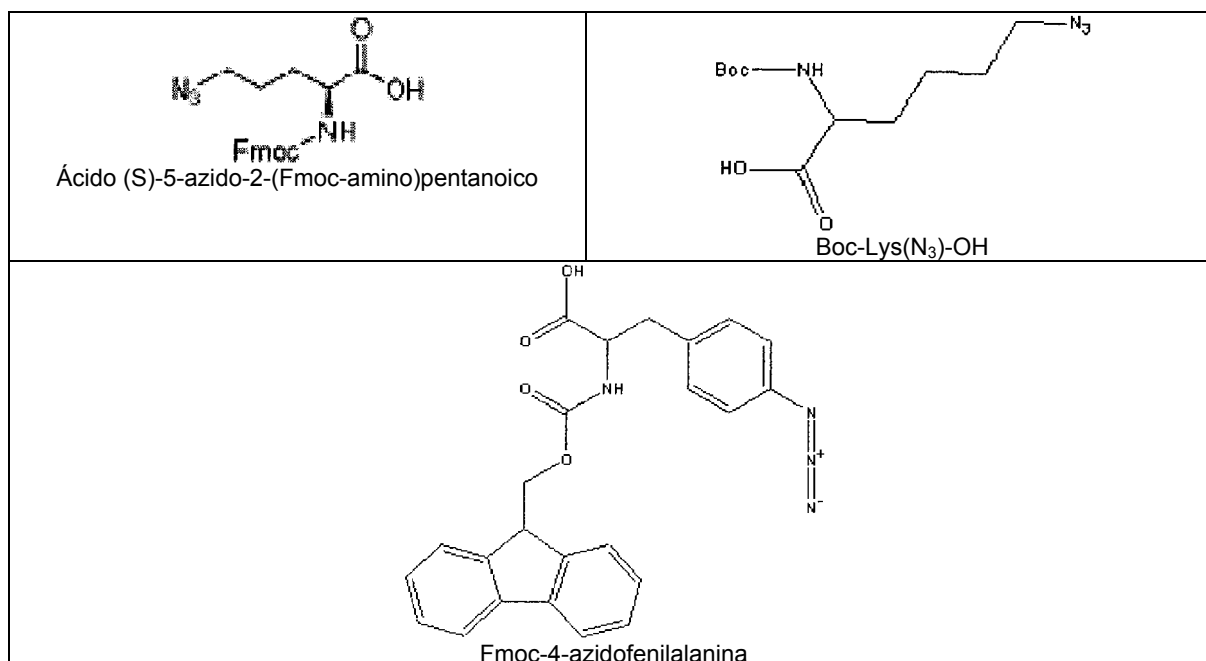
En una realización, el método de amplificación incluye formar una emulsión de agua en aceite para crear una pluralidad de microrreactores acuosos, en la que al menos uno de los microrreactores tiene al menos un miembro de una biblioteca de compuestos que se une a la diana, una única perla que puede unirse al oligonucleótido codificante del al menos un miembro de la biblioteca de compuestos que se une a la diana, y disolución de reacción de amplificación que contiene reactivos necesarios para realizar la amplificación de ácido nucleico, amplificar el oligonucleótido codificante en los microrreactores para formar copias amplificadas del oligonucleótido codificante, y unir las copias amplificadas del oligonucleótido codificante a las perlas en los microrreactores.

Una vez que se han identificado los elementos estructurales de la primera biblioteca que se unen a la diana de interés, puede prepararse una segunda biblioteca de manera iterativa, en la que se añaden uno o dos nodos de diversidad adicionales, y se crea la biblioteca y se muestrea la diversidad tal como se describe en el presente documento. Este procedimiento puede repetirse tantas veces como sea necesario para crear moléculas con propiedades moleculares y farmacéuticas deseadas.

Los elementos estructurales A a modo de ejemplo incluyen, por ejemplo, aminoácidos (no limitados a alfa-aminoácidos), reactivos de química clic (por ejemplo, azida o cadenas de alquino) con una amina, o un reactivo de tiol. La elección de elementos estructurales A depende, por ejemplo, de la naturaleza del grupo reactivo usado en el ligador, la naturaleza de un resto de estructura de soporte, y el disolvente usado para la síntesis química. Véase, por ejemplo, la tabla 1.

Tabla 1. Elementos estructurales en la posición A a modo de ejemplo

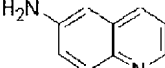
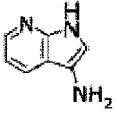
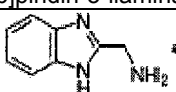
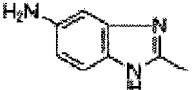
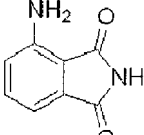
 Ácido 3-(1-Fmoc-piperidin-4-il)-propiónico	 Éster de N-hidroxisuccinimida del ácido 4-azido-butan-1-oico
 Ácido Boc-L-indolin-2-carboxílico	 Fmoc-L-propargilglicina
 Fmoc-(4-carboximetil)piperazina	 Boc-D-propargilglicin-DCHA
 Ácido Boc-2-amino-1,2,3,4-tetrahidronaftaleno-2-carboxílico	 2-Amino-N-(3-azidopropil)-3-mercaptopropionamida
 Ácido (S)-(-)-2-azido-6-(Boc-amino)hexanoico	 2-Amino-3-mercapto-N-(prop-2-inil)propionamida



En las tablas 2 y 3 se describen elementos estructurales B y C a modo de ejemplo, respectivamente. Puede introducirse un sitio de restricción, por ejemplo, en la posición B o C para el análisis del producto final y la selección mediante realización de PCR y digestión de restricción con una de las enzimas de restricción correspondientes.

5

Tabla 2. Ejemplos de elementos estructurales en la posición B y etiquetas de ADN codificante

Nombre químico y estructura	Sitio de restricción (enzima de restricción)	Cadena superior Cadena inferior
 6-Aminoquinolina (B1)	T/CCGGA (BspEI)	5'-Phos-CCTCCGGAGA (SEQ ID NO: 1) 5'-Phos-TCCGGAGGAC (SEQ ID NO: 2)
 3-Amino-7-azaindol,1H-pirrol[2,3-b]piridin-3-ilamina (B2)	GGC/GCC (SfoI)	5'-Phos-CCGGCGCCGA (SEQ ID NO: 3) 5'-Phos-GGCGCCGGAC (SEQ ID NO: 4)
 Diclorhidrato de 2-(aminometil)bencimidazol (B3)	GGTAC/C (KpnI)	5'-Phos-CCGGTACCGA (SEQ ID NO: 5) 5'-Phos-GGTACCGGAC (SEQ ID NO: 6)
 2-Metil-1H-bencimidazol-5-amina (B4)	CAC/GTG (PmlI)	5'-Phos-CCCACGTGGA (SEQ ID NO: 7) 5'-Phos-CACGTGGGAC (SEQ ID NO: 8)
 (Aminometil)ciclopropano (B5)	GAGCT/C (SacI)	5'-Phos-CCGAGCTCGA (SEQ ID NO: 9) 5'-Phos-GAGCTCGGAC (SEQ ID NO: 10)
 3-Aminoftalimida (B6)	G/GATCC (BamHI)	5'-Phos-CCGGATCCGA (SEQ ID NO: 11) 5'-Phos-GGATCCGGAC (SEQ ID NO: 12)

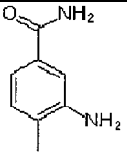
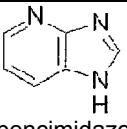
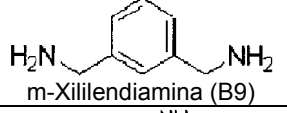
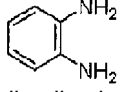
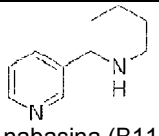
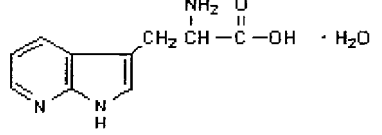
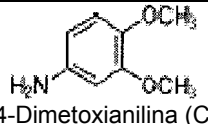
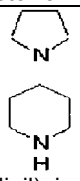
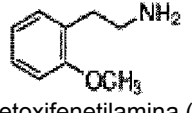
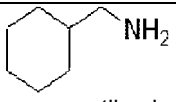
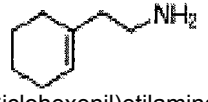
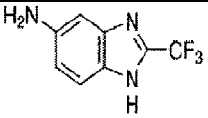
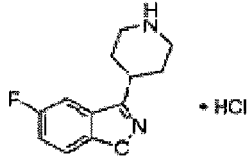
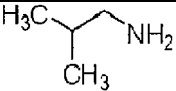
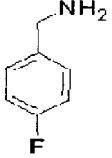
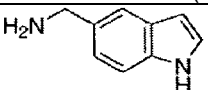
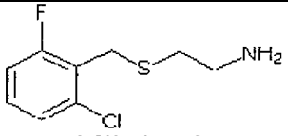
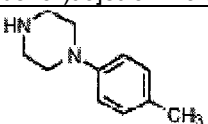
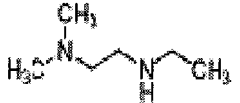
 3-Amino-4-metilbenzamida (B7)	AT/CGAT (BspDI)	5'-Phos-CCATCGATGA (SEQ ID NO: 13) 5'-Phos-ATCGATGGAC (SEQ ID NO: 14)
 4-Azabencimidazol (B8)	A/AGCTT (HindIII)	5'-Phos-CCAAGCTTGA (SEQ ID NO: 15) 5'-Phos-AAGCTTGGAC (SEQ ID NO: 16)
 m-Xililendiamina (B9)	A/GATCT (BglII)	5'-Phos-CCAGATCTGA (SEQ ID NO: 17) 5'-Phos-AGATCTGGAC (SEQ ID NO: 18)
 1,2-Fenilendiamina (B10)	G/AATTC (EcoRI)	5'-Phos-CCGAATTCGA (SEQ ID NO: 19) 5'-Phos-GAATTCGGAC (SEQ ID NO: 20)
 Anabasina (B11)	T/GATCA (BclI)	5'-Phos-CCTGATCAGA (SEQ ID NO: 21) 5'-Phos-TGATCAGGAC (SEQ ID NO: 22)
 Hidrato de DL-7-azatriptófano (B12)	CA/TATG (NdeI)	5'-Phos-CCCATATGGA (SEQ ID NO: 23) 5'-Phos-CATATGGGAC (SEQ ID NO: 24)

Tabla 3. Ejemplos de elementos estructurales en la posición C y etiquetas de ADN codificante

Nombre químico y estructura	Cadena superior Cadena inferior
 3,4-Dimetoxianilina (C1)	5'-Phos-GAACCTGCTT (SEQ ID NO: 25) 5'-Phos-GCAGGTTCTC (SEQ ID NO: 26)
 4-(1-Pirrolidinil)piperidina (C2)	5'-Phos-GAAGACGCTT (SEQ ID NO: 27) 5'-Phos-GCGTCTTCTC (SEQ ID NO: 28)
 2-Metoxifenetilamina (C3)	5'-Phos-GACCAGACTT (SEQ ID NO: 29) 5'-Phos-GTCTGGTCTC (SEQ ID NO: 30)
 Ciclohexanometilamina (C4)	5'-Phos-GACGACTCTT (SEQ ID NO: 31) 5'-Phos-GAGTCGTCTC (SEQ ID NO: 32)
 2-(1-Ciclohexenil)etilamina (C5)	5'-Phos-GACGCTTCTT (SEQ ID NO: 33) 5'-Phos-GAAGCGTCTC (SEQ ID NO: 34)

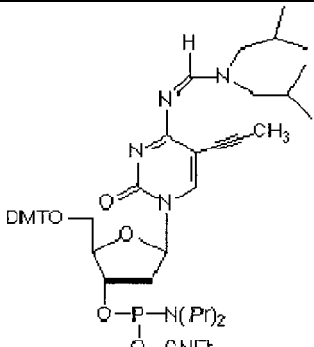
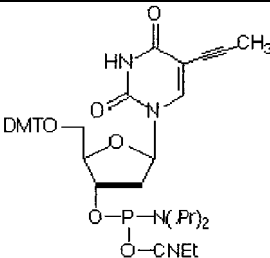
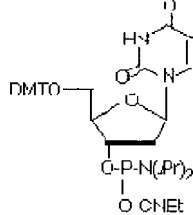
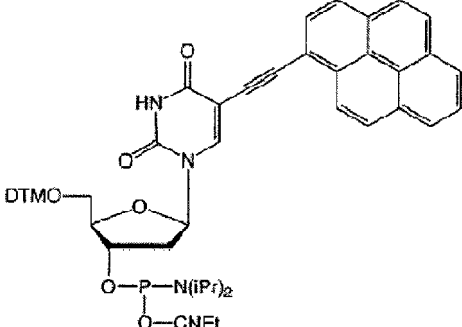
 5-Amino-2-(trifluorometil)bencimidazol (C6)	5'-Phos-GAGCAACCTT (SEQ ID NO: 35) 5'-Phos-GGTTGCTCTC (SEQ ID NO: 36)
 Clorhidrato de 5-fluoro-3-(4-piperidinil)-1,2-bencisoxazol (C7)	5'-Phos-GAGCCATCTT (SEQ ID NO: 37) 5'-Phos-GATGGCTCTC (SEQ ID NO: 38)
 Isobutilamina (C8)	5'-Phos-GCAACCACTT (SEQ ID NO: 39) 5'-Phos-GTGGTTGCTC (SEQ ID NO: 40)
 4-Fluorobencilamina (C9)	5'-Phos-GCACAGACTT (SEQ ID NO: 41) 5'-Phos-GTCTGTGCTC (SEQ ID NO: 42)
 5-(Aminometil)indol (C10)	5'-Phos-GCGATCACTT (SEQ ID NO: 43) 5'-Phos-GTGATCGCTC (SEQ ID NO: 44)
 2-[(2-cloro-6-fluorobencil)tio]etilamina (C11)	5'-Phos-GCGGTTACTT (SEQ ID NO: 45) 5'-Phos-GTAACCGCTC (SEQ ID NO: 46)
 1-(4-Metilfenil)piperazina (C12)	5'-Phos-GCATGACCTT (SEQ ID NO: 47) 5'-Phos-GGTCATGCTC (SEQ ID NO: 48)
 N,N-Dimetil-N'-etiletilendiamina (C13)	5'-Phos-GCGTACTCTT (SEQ ID NO: 49) 5'-Phos-GAGTACGCTC (SEQ ID NO: 50)

Modo 3

- 5 En cualquiera de los modos descritos en el presente documento (por ejemplo, los modos 1 y 2), el oligonucleótido de parte anterior puede modificarse para soportar solubilidad en condiciones semiacuosas o no acuosas (por ejemplo, orgánicas). La parte anterior incluye, en determinadas realizaciones, la región identificadora. En otras realizaciones, la parte anterior con ligador puede derivatizarse en primer lugar con un elemento estructural (por ejemplo, un resto funcional) o una estructura de soporte, y después se añade la secuencia identificadora.
- 10 Las bases de nucleótidos de la parte anterior pueden hacerse más hidrófobas modificando, por ejemplo, las posiciones C5 de las bases de T o C con cadenas alifáticas sin alterar significativamente su capacidad para formar enlaces de hidrógeno con sus bases complementarias. Véase, por ejemplo, la tabla 4 para ejemplos de bases modificadas. Además, el oligonucleótido de parte anterior puede intercalarse con modificaciones que fomentan la solubilidad en disolventes orgánicos. Por ejemplo, azobenceno-fosforamidita puede introducir un resto hidrófobo en el diseño de la parte anterior. Tales inserciones de amiditas hidrófobas en la parte anterior pueden producirse en cualquier lugar en la molécula. Sin embargo, la inserción no puede interferir con el posterior etiquetado usando etiquetas de ADN adicionales durante la síntesis de la biblioteca o reacciones de PCR posteriores una vez completada una selección o análisis de microalineamiento, si se usa para la desconvolución de etiqueta. Tales
- 15

adiciones al diseño de la parte anterior descritas en el presente documento harán que la parte anterior sea soluble, por ejemplo, en disolvente orgánico al 15%, 25%, 30%, 50%, 75%, 90%, 95%, 98%, 99% o al 100%. Por tanto, la adición de residuos hidrófobos en el diseño de la parte anterior permite una solubilidad mejorada en condiciones semiacuosas o no acuosas (por ejemplo, orgánicas), al tiempo que se hace que la parte anterior sea compatible para el etiquetado de ácido nucleico. Además, las etiquetas de ADN que se introducen posteriormente en la biblioteca también pueden modificarse en la posición C5 de bases de T o C de manera que también hacen que la biblioteca sea más hidrófoba y soluble en disolventes orgánicos para etapas posteriores de síntesis de la biblioteca.

Tabla 4. Bases de nucleótidos modificadas a modo de ejemplo

 3'-[(2-cianoetil)-(N,N-diisopropil)]-fosforamidita de 5'-dimetoxitritil-N4-diisobutylaminometiliden-5-(1-propinil)-2'-desoxicidina	 3'-[(2-cianoetil)-(N,N-diisopropil)]-fosforamidita de 5'-dimetoxitritil-5-(1-propinil)-2'-desoxiuridina
 3'-[(2-cianoetil)-(N,N-diisopropil)]-fosforamidita de 5'-dimetoxitritil-5-fluoro-2'-desoxiuridina	 3'-[(2-cianoetil)-(N,N-diisopropil)]-fosforamidita de 5'-dimetoxitritil-5-(piren-1-il-etinil)-2'-desoxiuridina

La molécula de ligador entre la parte anterior y la biblioteca de moléculas pequeñas puede variarse para aumentar la solubilidad de la parte anterior en disolvente orgánico. Hay una amplia variedad de ligadores comercialmente disponibles que pueden acoplar la parte anterior con la biblioteca de moléculas pequeñas. Los ligadores se seleccionan empíricamente para un diseño de biblioteca de moléculas pequeñas dado (estructuras de soporte y elementos estructurales) de manera que la biblioteca puede sintetizarse en disolvente orgánico, por ejemplo, disolvente orgánico al 15%, 25%, 30%, 50%, 75%, 90 %, 95%, 98%, 99% o al 100%. El ligador puede variarse usando reacciones modelo antes de la síntesis de la biblioteca para seleccionar la longitud de cadena apropiada que solubiliza la parte anterior en disolvente orgánico. Tales ligadores pueden incluir ligadores, por ejemplo, con una longitud de cadena de alquilo aumentada, unidades de polietilenglicol aumentadas, especies ramificadas con cargas positivas (para neutralizar las cargas de fosfato negativas en la parte anterior) o cantidades aumentadas de hidrofobia (por ejemplo, adición de estructuras cíclicas de benceno).

La molécula de ligador puede proporcionar un separador apropiado entre el ADN de parte anterior y un miembro de una biblioteca química. Por ejemplo, pueden usarse ligadores bifuncionales. En determinadas realizaciones, los ligadores bifuncionales pueden incluir, por ejemplo, tres partes. La parte 1 puede ser un grupo reactivo, que forma un enlace covalente con ADN, tal como, por ejemplo, un ácido carboxílico, preferiblemente activado mediante un éster de N-hidroxisuccinimida (NHS) para reaccionar con un grupo amino en el ADN (por ejemplo, dT modificada con amino), una amidita para modificar el extremo 5' o 3' de una parte anterior de ADN monocatenario (logrado mediante química de oligonucleótidos convencional), pares de química de clic (cicloadición de azida y alquino en presencia de catalizador de Cu(I)) o grupos reactivos con tiol. La parte 2 también puede ser un grupo reactivo, que forma un enlace covalente con la biblioteca química, o bien un elemento estructural en la posición A o bien un resto de estructura de soporte. Un grupo reactivo de este tipo puede ser, por ejemplo, una amina, un tiol, una azida o un alquino para reacciones basadas en agua u otros múltiples grupos reactivos para las reacciones basadas en disolvente orgánico. La parte 3 puede ser un separador químicamente inerte de longitud variable, introducido entre

las partes 1 y 2. Un separador de este tipo puede ser una cadena de unidades de etilenglicol (por ejemplo, PEG de diferentes longitudes), un alcano, un alqueno, cadena poliénica o cadena peptídica. El ligador puede contener ramificaciones o inserciones con restos hidrófobos (tal como, por ejemplo, anillos de benceno) para mejorar la solubilidad de la parte anterior en disolventes orgánicos, así como restos fluorescentes (por ejemplo fluoresceína o Cy-3) usados con fines de detección de biblioteca.

Los ejemplos de ligadores comercialmente disponibles incluyen, por ejemplo, ligadores amino-carboxílicos (por ejemplo, péptidos (por ejemplo, Z-Gly-Gly-Gly-Osu o Z-Gly-Gly-Gly-Gly-Gly-Osu), PEG (por ejemplo, Fmoc-amino-PEG2000-NHS o amino-PEG(12-24)-NHS) o cadenas de ácido de alcano (por ejemplo, ácido Boc-ε-aminocaproico-Osu)), ligadores de química de clic (por ejemplo, péptidos (por ejemplo, azidohomalanina-Gly-Gly-Gly-Osu o propargilglicina-Gly-Gly-Gly-Osu), PEG (por ejemplo, azido-PEG-NHS) o cadenas de ácido de alcano (por ejemplo, ácido 5-azidopentanoico, (S)-2-(azidometil)-1-Boc-pirrolidina o éster de N-hidroxisuccinimida del ácido 4-azido-butan-1-oico)), ligadores reactivos con tiol (por ejemplo, PEG (por ejemplo, SM(PEG)_n NHS-PEG-maleimida), cadenas de alcano (por ejemplo, ácido 3-(piridin-2-ildisulfanil)-propionico-Osu o 6-(3'-[2-piridilditio]propionamido)hexanoato de sulfosuccinimidilo)), amiditas para síntesis de oligonucleótidos (por ejemplo, modificadores de amino (por ejemplo, 6-(trifluoroacetilamino)-hexil-(2-cianoetil)-(N,N-diisopropil)-fosforamidita), modificadores de tiol (por ejemplo, S-tritil-6-mercaptohexil-1-[(2-cianoetil)-(N,N-diisopropil)]-fosforamidita, o modificadores de química de clic (por ejemplo, 6-hexin-1-il-(2-cianoetil)-(N,N-diisopropil)-fosforamidita, 3-dimetoxitritiloxi-2-(3-(3-propargiloxipropanamido)propanamido)propil-1-O-succinoilo, CPG de alquilamino de cadena larga, o éster de N-hidroxisuccinimida del ácido 4-azido-butan-1-oico)).

Los residuos hidrófobos en el diseño de la parte anterior pueden variarse con el diseño del ligador para facilitar la síntesis de la biblioteca en disolventes orgánicos. Por ejemplo, la combinación de parte anterior y ligador se diseña para tener residuos apropiados en los que el coeficiente de octanol:agua (Poet) es, por ejemplo, de desde 1,0 hasta 2,5.

Ejemplos

Se pretende que los siguientes ejemplos ilustren la invención. No se pretende que limiten la invención de ninguna manera.

Ejemplo 1. Preparación de la parte anterior (variante 1)

IDT DNA sintetizó una parte anterior de oligonucleótido fosforilado (oligo HP) que tenía la siguiente secuencia y la purificó mediante HPLC.

5'-(fosfato)TCCTGGCTGAGGCGAGAGTT(dT-C6-NH)
TTCTCTCGCCTCAGCCAGGACC-3' (SEQ ID NO: 51)

El oligonucleótido se pliega para dar una horquilla (figura 10) con una proyección y contiene un sitio de escisión (CCTCAGC) para la enzima de restricción BbvCI o versiones de mellado de esta enzima Nb.BbvCI o Nt.BbvCI, que pueden escindir o bien la cadena superior o bien la inferior (New England BioLabs). En el centro del bucle de horquilla, se inserta el dT modificado con C5-amino de cadena lateral (dT-C6-NH; "C6" se refiere a un ligador de 6 carbonos), que se usó para el acoplamiento del ligador de amino-PEG (PEG2000, aproximadamente 45 unidades de etilenglicol). IDT DNA sintetizó las cadenas superior e inferior de las etiquetas de ADN A, B, y C y las purificó y se purificaron mediante desalación convencional. IDT DNA sintetizó oligonucleótidos más largos, tales como los cebadores de PCR y extremo 3', y se purificaron mediante HPLC.

Se disolvieron diez nanomoles del oligonucleótido HP en 50 µl de agua. Se disolvió un exceso molar de 20 veces de éster de Fmoc-amino-PEG2000-carboxil-NHS (JenKem Technology USA) en 50 µl de dimetilformamida (DMF) y se añadió a la disolución de oligonucleótido en 2 porciones a lo largo del transcurso de 2 horas a temperatura ambiente (composición de disolvente final del 50% de DMF/el 50% de agua). Posteriormente, se añadieron 60 µl de Tris HCl 1 M, pH 7,0 (concentración final de 200 mM) para extinguir el exceso de ésteres de NHS, y se incubó la disolución durante 30 minutos adicionales a temperatura ambiente. Se diluyó la mezcla de reacción resultante hasta 500 µl con agua y se desalinizó haciéndola pasar a través de una columna NAP-5 (Sephadex-25, GE Healthcare).

Se liofilizó el material resultante y se disolvió en 100 µl agua. Se añadieron 20 µl de piperidina (hasta una concentración final del 20%) y se incubaron durante 2 horas a temperatura ambiente. Se formó un precipitado turbio debido a la desprotección de la amina y liberación del grupo Fmoc insoluble en agua. Después se filtró la reacción a través de filtros de centrifugadora de 0,2 µm (Millipore) y se precipitó usando acetato de sodio 300 mM mediante la adición de 3 volúmenes de etanol. Se encontró que la forma protegida con Fmoc del oligonucleótido modificado era soluble en etanol e isopropanol. Debido a la alta eficacia de acoplamiento, se usó la parte anterior resultante (HP-1) sin purificación adicional (figura 11).

Ejemplo 2. Preparación de la parte anterior (variante 2)

Se preparó una parte anterior completa (HP-1) que tenía la siguiente secuencia por Trilink, Inc., siguiendo un procedimiento similar al descrito anteriormente y se purificó mediante RP-HPLC.

5 5'-(fosfato)TCCTGGCTGAGGCGAGAGTT(dT-C6-NH)(X)
TTCTCTCGCCTCAGCCAGGACC-3' (SEQ ID NO: 52)

donde X representa amino-PEG2000.

10 Ejemplo 3. Síntesis de un miembro de biblioteca modelo

Etapa 1: Acoplamiento de DTAF

15 Con el fin de preparar una "biblioteca de 1", se acopló un compuesto modelo, 5-(4,6-diclorotriazinil-aminofluoresceína) (DTAF; Anaspec) (figura 12), al grupo amino de HP-1. DTAF representa estructuralmente una estructura de soporte de triclorotriazina con un compuesto de amino acoplado. Para formar una biblioteca, pueden derivatizarse estructuras de soporte de triclorotriazina con una diversidad de elementos estructurales en cada una de las tres posiciones de cloro. DTAF también proporciona una etiqueta fluorescente para la biblioteca modelo. Se configuró la reacción (10 µl) de la siguiente manera. A 5 µl de HP-1 400 µM disueltos en agua, se les añadieron 2 µl de tampón borato 750 mM, pH 9,5, y 1 µl de DMF. Se disolvió DTAF en DMF hasta 50 mM y se añadieron 2 µl a la reacción. Las concentraciones finales de HP-1 y DTAF fueron de 200 µM y 10 mM, respectivamente, generando por tanto un exceso de 50 veces de DTAF. La concentración de DMF final fue del 30%. Se observó que HP-1 permanecía soluble en hasta el 90% de DMF, demostrando que era soluble en un disolvente orgánico, por ejemplo, DMF. Se dejó avanzar la reacción a 4°C durante 16-20 horas. Después se diluyó la mezcla de reacción con agua hasta 30-50 µl y se desalinizó en una columna de centrifugación Zeba (Pierce). No se completó ninguna purificación adicional en este punto.

Etapa 2: Acoplamiento de amino-biotina

30 Tras acoplar DTAF a HP-1, todavía está disponible para la modificación un grupo reactivo más en la molécula de estructura de soporte. Se eligió un análogo de amino-biotina, EZ-Link Pentylamine-Biotin (Pierce), para acoplarlo en esta posición con el fin de generar un compuesto de unión modelo (figura 12). Se configuró la reacción de la siguiente manera. 20 µl de la mezcla de reacción contenían aproximadamente 200 pmol de HP-1-DTAF (etapa 1) disueltos en tampón borato 150 mM, pH 9,5, y 10 nmol de pentilamina-biotina. Se dejó avanzar la reacción durante 4-12 horas a 75°C. Después se purificó la reacción desalando en una columna de centrifugación Zeba, tal como se describió anteriormente.

Etapa 3: Ligamiento de las etiquetas de ADN a HP-1-DTAF-biotina

40 IDT DNA sintetizó etiquetas de ADN fosforiladas (región de cebador en el extremo 3' y cebadores de PCR 5' y 3'). Las secuencias de oligonucleótidos (figura 13) son las siguientes.

Etiqueta de ADN A1 (superior): 5'-phos-GGAGGACTGT (SEQ ID NO: 53)

45 Etiqueta de ADN A1 (inferior): 5'-phos-AGTCCTCCGG (SEQ ID NO: 54)

Etiqueta de ADN B1 (superior): 5'-phos-CAGACGACGA (SEQ ID NO: 55)

50 Etiqueta de ADN B1 (inferior): 5'-phos-GTCGTCTGAC (SEQ ID NO: 56)

Etiqueta de ADN C1 (superior): 5'-phos-CGATGCTCTT (SEQ ID NO: 57)

Etiqueta de ADN C1 (inferior): 5'-phos-GAGCATCGTC (SEQ ID NO: 58)

55 Extremo 3' (superior): 5'-phos-GCTGTGCAGGTAGAGTGC-3' (SEQ ID NO: 59)

Extremo 3' (inferior): 5'-AACGACACGTCCATCTCACG (SEQ ID NO: 60)

Cebador de PCR 5': 5'-CTCTCGCCTCAGCCAGGA (SEQ ID NO: 61)

60 Cebador de PCR 3': 5'-GCACTCTACCTGCACAGC (SEQ ID NO: 62)

Se disolvieron cantidades equivalentes de pares superior e inferior de etiquetas y oligonucleótidos del extremo 3' en agua y se aparearon mediante calentamiento hasta 85°C y reducción gradual hasta 4°C en tampón de NaCl 200 mM, Tris HCl 50 mM, pH 7,0.

65

En primer lugar, se ligó la etiqueta A1 bicatenaria a la parte anterior. La reacción de ligamiento (20 μ l) contenía 2,5 μ M de HP-1-DTAF-biotina y 2,5 μ M de etiqueta A1 bicatenaria en 1x tampón de ADN ligasa de T4 y 60 unidades Weiss de ADN ligasa de T4 (New England BioLabs). Se incubó la reacción a 16°C durante 16 horas. El producto resultante no se resolvió en ninguno de los geles sometidos a prueba, incluyendo diferentes porcentajes de TBE-urea, NativePage, SDS-PAGE, o gel de agarosa E al 2% y al 4% (Invitrogen, Inc.). La movilidad del oligonucleótido, modificado con ligador de PEG y DTAF-biotina, se determinó principalmente por la presencia de estos grupos en vez de por el propio ADN (datos no mostrados). Para someter a prueba la eficacia del ligamiento, se ligaron todas las etiquetas y los oligonucleótidos del extremo 3' y se realizaron ensayos de PCR del constructo resultante para confirmar la eficacia de ligamiento. La reacción de ligamiento (70 μ l) contenía: 2,5 μ M de HP-1-DTAF-biotina; 2,5 μ M de cada una de las etiquetas de ADN bicatenarias apareadas (A1, B1 y C1), así como la etiqueta de extremo 3'; 1x tampón de ADN ligasa de T4; y 210 unidades Weiss de ADN ligasa de T4. Se incubó la reacción a 16°C durante 20 horas.

Se cargó la mezcla de reacción en un gel de agarosa al 4% y se extrajo la banda fluorescente del gel. Se usó este material para la amplificación por PCR de 24 ciclos de prueba usando cebadores 5' y 3' tal como se describió anteriormente. Los resultados se resumen en la figura 13.

Etapla 4: Purificación de HP-1-DTAF-biotina con perlas de estreptavidina y reacción con estreptavidina

La purificación de HP-1-DTAF-biotina con perlas magnéticas Dynal de estreptavidina (SA) M-280 (Invitrogen) sirve como modelo para la selección por afinidad para la biblioteca marcada con etiquetas de ADN químicas. Se equilibraron previamente perlas de SA en 2x tampón PBS que contenía Triton X-100 al 0,05%. Se cargaron 50 pmol de HP-1-DTAF-biotina en 25 μ l de las perlas de SA previamente lavadas durante 15 minutos a temperatura ambiente con volteo. Se recogieron las fracciones no retenidas y se lavaron las perlas 3 veces durante 30 minutos con 1 ml del mismo tampón. Se realizó un lavado final a 80°C durante 10 minutos con 30 μ l agua (recogido). Se eluyeron las perlas con 30 μ l de EDTA 25 mM y NaOH 5 mM durante 10 minutos a 90°C, y se neutralizó inmediatamente el eluyente añadiendo 3 μ l de Tris HCl 1 M, pH 7,0.

Para el experimento de unión a estreptavidina, se incubaron 5 μ l de las muestras de elución con un exceso de estreptavidina en NaCl 50 mM/Tris HCl 10 mM, pH 7,0, durante 10 minutos. Se mezclaron las muestras con tampón de carga en gel sin ebullición y se resolvieron en un gel de SDS NuPage al 4-12% (Invitrogen) usando tampón de transferencia MES. Los resultados se resumen en la figura 14.

Ejemplo 4. Acoplamiento de péptido H(-Arg- ϵ Ahx)₆-Arg-OH a HP-1-DTAF

Se eligió un péptido rico en arginina R7, H(-Arg- ϵ Ahx)₆-Arg-OH (Bachem), para usarlo como otra modificación para el último grupo reactivo en la estructura de soporte de triazina. Esto es un péptido permeable a la membrana celular de ácido arginina-aminohexanoico usado para la administración intracelular de compuestos. Se estableció la reacción de manera similar a las condiciones de reacción descritas anteriormente: 20 μ l de reacción contenían aproximadamente 200 pmol de HP-1-DTAF (etapa 1) disueltos en tampón borato 150 mM, pH 9,5, y 10 nmol de péptido R7. En estas condiciones, las cadenas laterales de las argininas no reaccionan, y la única amina reactiva en el péptido es el extremo N-terminal. Se dejó avanzar la reacción durante 12 horas a 75°C y después se purificó mediante desalación en una columna de centrifugación Zeba.

Ejemplo 5. Constructo de ADN para experimento de detección de administración intracelular de ARN-P de T7

Se preparó el constructo de ADN usado para el experimento de administración intracelular de "biblioteca de 1" química a partir de un producto de PCR de un clon individual de ADN de V_H de ~400 pb que presentaba una región de promotor de T7 en el extremo 5' y una región Cmu constante de anticuerpo corta cerca del extremo 3' de la molécula. Con el fin de unir el constructo de ADN a la parte anterior modificada de la biblioteca química modelo, se añadió un sitio de restricción de BsmI en el sentido de 5' de la región de promotor de T7 mediante amplificación por PCR del clon. El digesto de restricción de BsmI produjo una proyección de GG en 3', que permitió el ligamiento a la parte anterior (proyección de CC en 3'). IDT DNA, Inc. sintetizó el cebador en 5' con sitio de BsmI (subrayado).

5'-GGATGCCGAATGCCCTAATACGACTCACTATAGGG-
ACAATTACTATTTACAATTACA (SEQ ID NO: 63)

Tras la amplificación por PCR, se purificó el constructo de ADN usando un kit de purificación por PCR (Invitrogen), y se digirió el ADN resultante con 250 U de BsmI (New England BioLabs) a 65°C en tampón NEB 4 durante 2 horas. Se purificó el ADN en un gel de agarosa al 2%. La reacción de ligamiento (30 μ l) contenía 2 pmol de cada constructo de ADN de V_H, digerido con BsmI, así como HP-1-DTAF-R7 (péptido de ácido arginina-aminohexanoico) en 1x tampón de ADN ligasa de T4 y 60 unidades Weiss de ADN ligasa de T4 (New England BioLabs). Se incubó la

reacción a 16°C durante 20 horas. Debido a la alta eficacia del ligamiento, el material se usó adicionalmente para el experimento de administración intracelular/ARN-P de T7 sin purificación adicional. Los resultados se resumen en la figura 15.

5 Ejemplo 6. Síntesis de biblioteca 10x10

Etapas 1. Ligamiento de la etiqueta A a la parte anterior HP-T

10 En esta biblioteca a modo de ejemplo, sólo se usan las posiciones B y C. Se liga una etiqueta A a HP-T. La etiqueta tiene la siguiente secuencia:

Etiqueta de ADN A1 (superior): 5'-phos-GGAGGACTGT (SEQ ID NO: 64)

Etiqueta de ADN A1 (inferior): 5'-phos-AGTCCTCCGG (SEQ ID NO: 65)

15 Se mezclaron 30 nmol de HP-T con 45 nmol de cada uno de los oligonucleótidos de etiqueta A1 superior y de etiqueta A1 inferior en 1x tampón de ADN ligasa de T4 y se aparearon mediante calentamiento hasta 95°C durante 1 minuto, seguido por enfriamiento hasta 4°C a 0,2°C/segundo. Después se llevó la muestra hasta 16°C. Se añadieron 300 unidades Weiss de ADN ligasa de T4 y se dejaron incubar las muestras durante 16-20 horas a 16°C. 20 Tras el ligamiento, se desalinizó HP-T-A usando una columna Zeba (Pierce). Véase, por ejemplo, la figura 16A.

Etapas 2. Ligamiento de etiquetas B1-B12 y etiquetas C

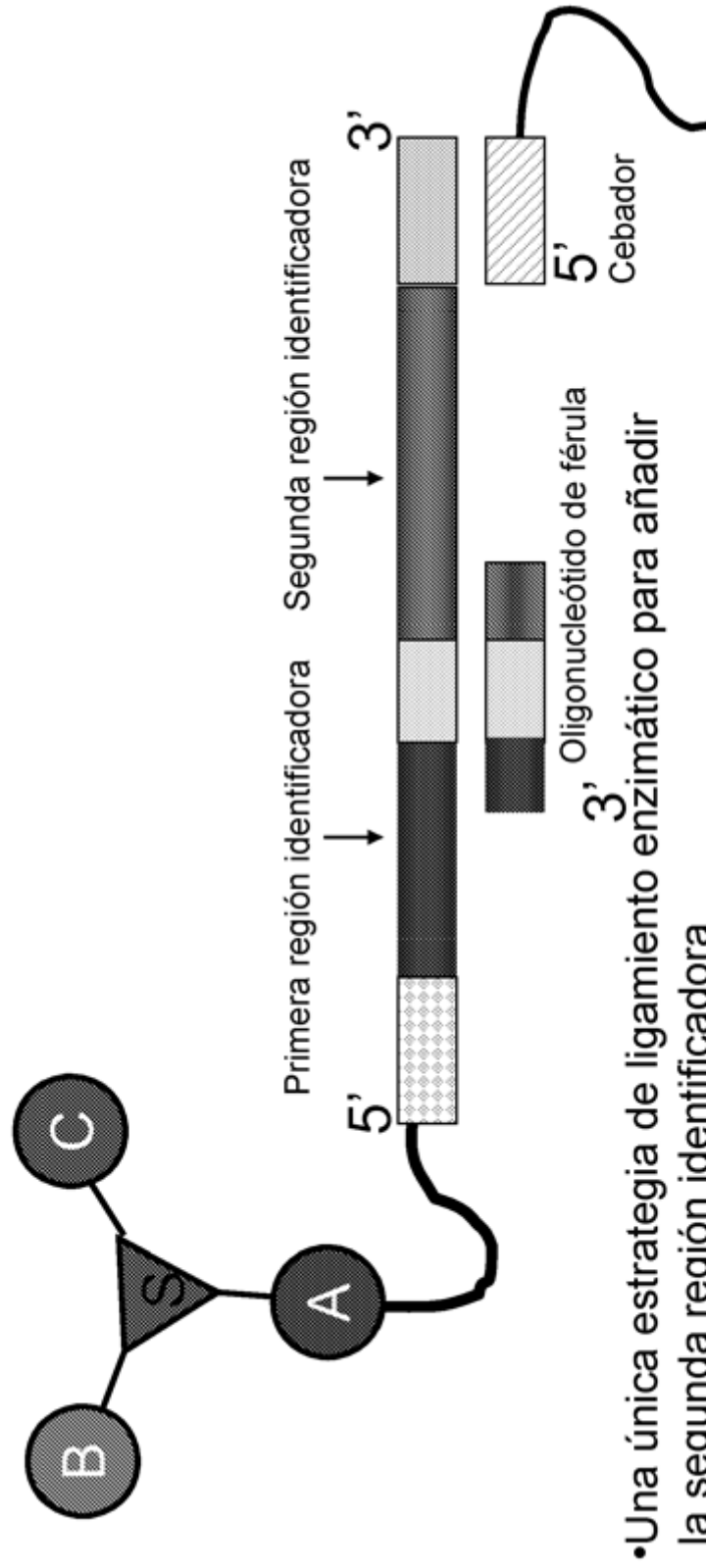
25 Se establecieron doce reacciones de ligamiento de manera similar a las reacciones de ligamiento descritas anteriormente. En cada uno de 12 tubos, se añadieron 5 nmol de pares de oligonucleótidos superior e inferior de B1-B12 a 1x tampón de ADN ligasa de T4 y se aparearon tal como se describió anteriormente. Se disolvió HP-T-A en 1x tampón de ADN ligasa de T4. Se añadieron alícuotas de 2,5 nmol de HP-T-A en estos 12 tubos. Se añadieron 30 unidades Weiss de ADN ligasa de T4 a cada tubo y se dejaron avanzar las reacciones durante 20 horas a 16°C. 30 Tras la incubación, se desalinizó individualmente cada mezcla de reacción en una columna de centrifugación Zeba de 0,5 ml, equilibrada con tampón borato 150 mM, pH 9,0. A cada tubo se le añadió un exceso de 20x de cloruro cianúrico (50 nmol) disuelto en acetonitrilo, y se incubaron durante 1,5 horas a 4°C. Tras esta incubación, se añadió un exceso de 100x (250 nmol, es decir, exceso de 5x con respecto al cloruro cianúrico) de aminos B1-B12, disueltos en acetonitrilo o DMF, en correspondencia con las etiquetas B1-B12 ligadas. Se dejó avanzar la reacción con aminos durante 20 horas a 4°C. Tras esta reacción se combinó la biblioteca, se desalinizó dos veces en columnas 35 Zeba de 2 ml y se liofilizó. Véanse, por ejemplo, las figuras 16B y 16C.

Al igual que las reacciones anteriores, las etiquetas C y las aminos se añaden usando condiciones de reacción similares a las descritas anteriormente.

REIVINDICACIONES

1. Método de etiquetado de bibliotecas químicas codificadas por ADN, comprendiendo dicho método:
 - 5 (a) unir un primer grupo funcional de un ligador bifuncional a una parte anterior que comprende un oligonucleótido iniciador, en el que dicho oligonucleótido iniciador comprende una estructura de horquilla;
 - (b) unir un segundo grupo funcional del ligador bifuncional a un primer elemento estructural;
 - 10 (c) hacer reaccionar un segundo elemento estructural con dicho primer elemento estructural; y
 - (d) unir una región identificadora en el extremo 3' de dicha parte anterior, en el que la región identificadora codifica para dicho segundo elemento estructural;
 - 15 en el que la etapa (c) se realiza en disolvente orgánico.
2. Método según la reivindicación 1, en el que dicho ligador bifuncional comprende una cadena de unidades de etilenglicol, un alcano, un alqueno, cadena poliénica o cadena peptídica.
- 20 3. Método según la reivindicación 1, en el que la parte anterior comprende una región identificadora.
4. Método según la reivindicación 1, en el que el método comprende además unir una región identificadora en el extremo 3' de dicha parte anterior tras la etapa (b) y antes de la etapa (c).
- 25 5. Método según la reivindicación 1, en el que dicha parte anterior comprende un nucleótido de T o C que comprende una cadena alifática en la posición C5.
6. Método según la reivindicación 1, en el que dicha parte anterior comprende un azobenceno.
- 30 7. Método según la reivindicación 1, en el que dicha región identificadora de la etapa (d) comprende un nucleótido de T o C que comprende una cadena alifática en la posición C5.

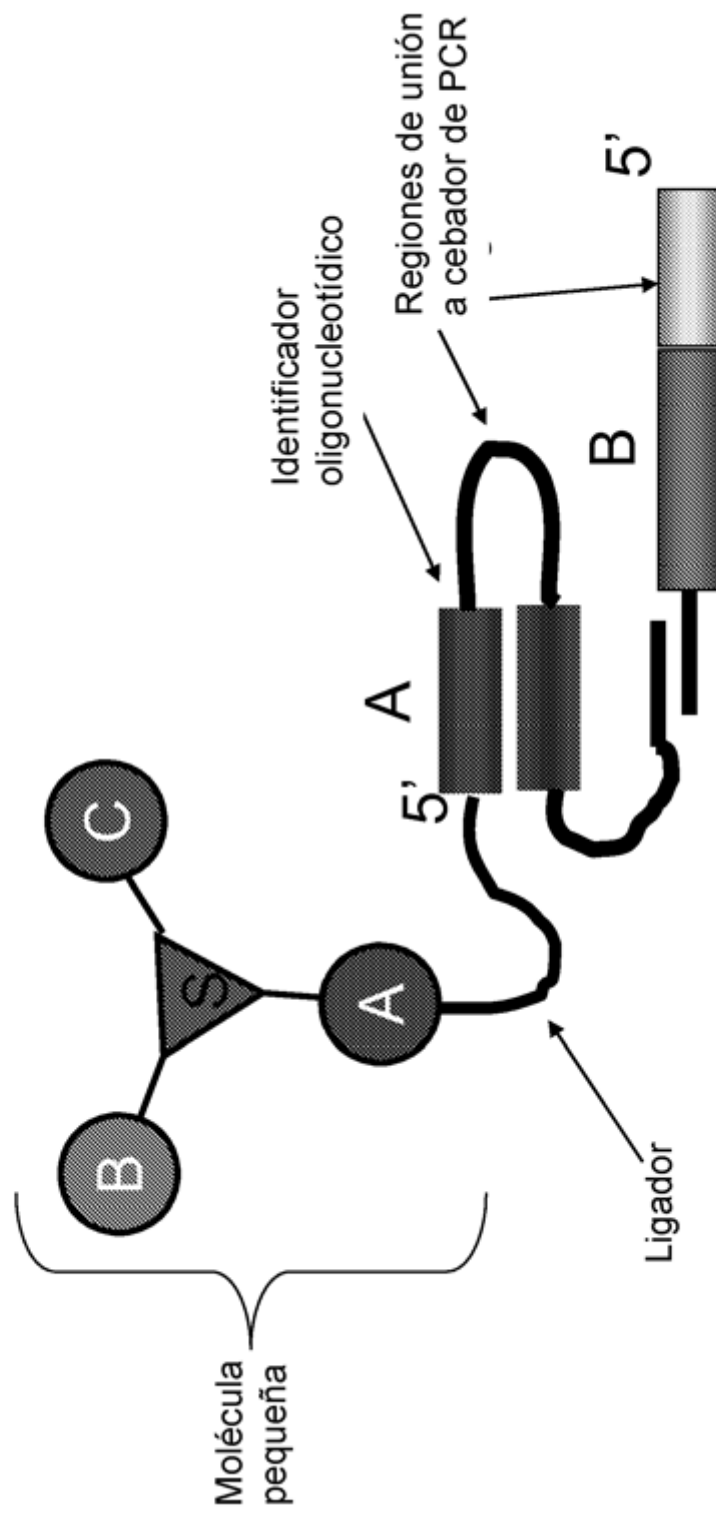
Figura 1



- Una única estrategia de ligamiento enzimático para añadir la segunda región identificadora
- Posibilidad de que las bases expuestas se dañen durante la síntesis química; buscar minimizar mediante enfoques bicatenarios

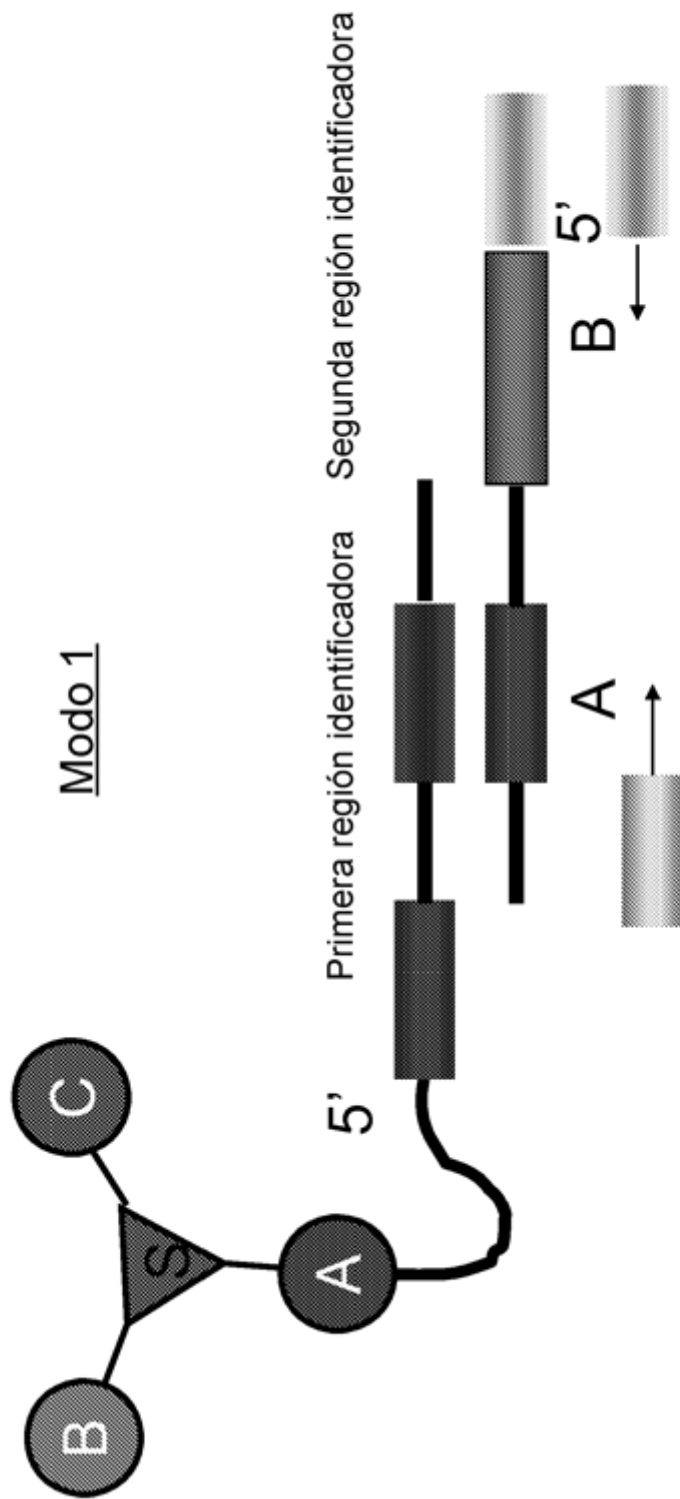
Figura 2

Modo 1



Polimerización para añadir la segunda región identificadora

Figura 4



Logrado mediante hibridación de etiquetas (no enzimáticas)
Puede introducir reticulación de psoraleno
o polimerización

Figura 5

Modo 1

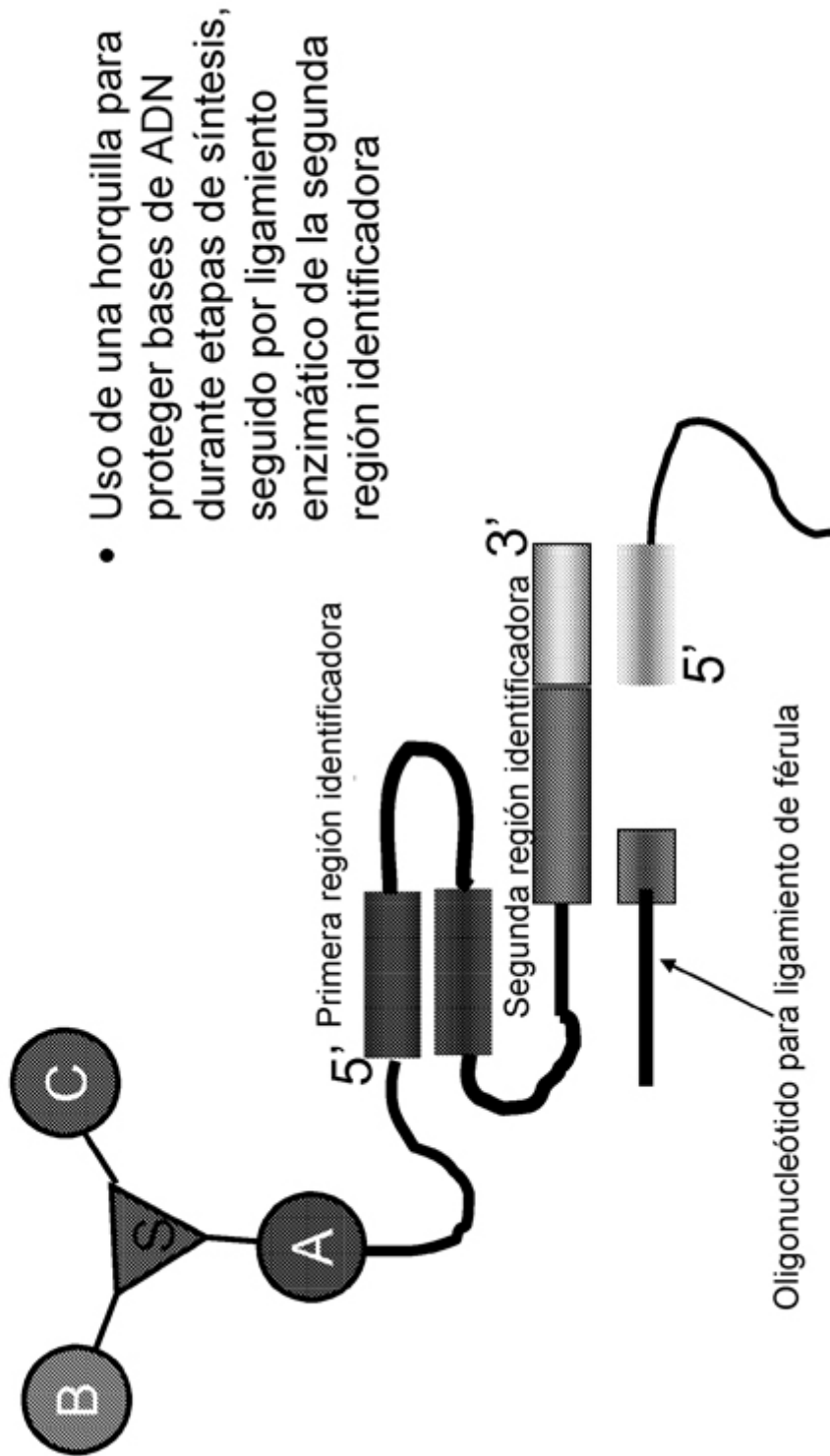


Figura 6

Modo 1

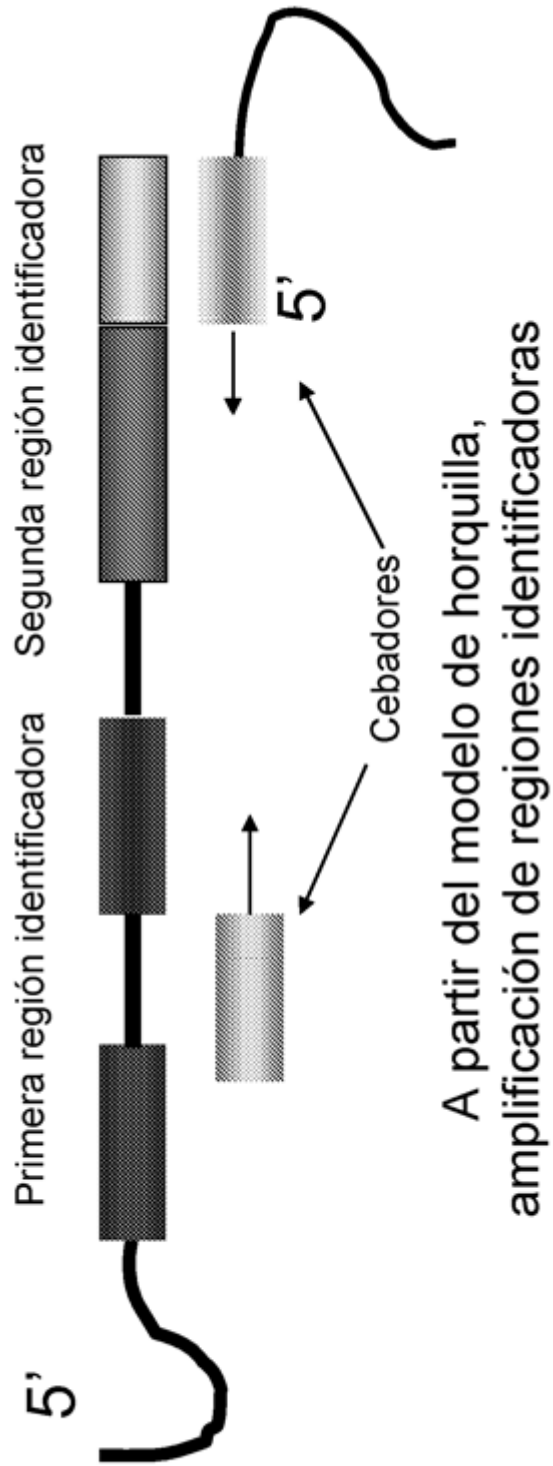
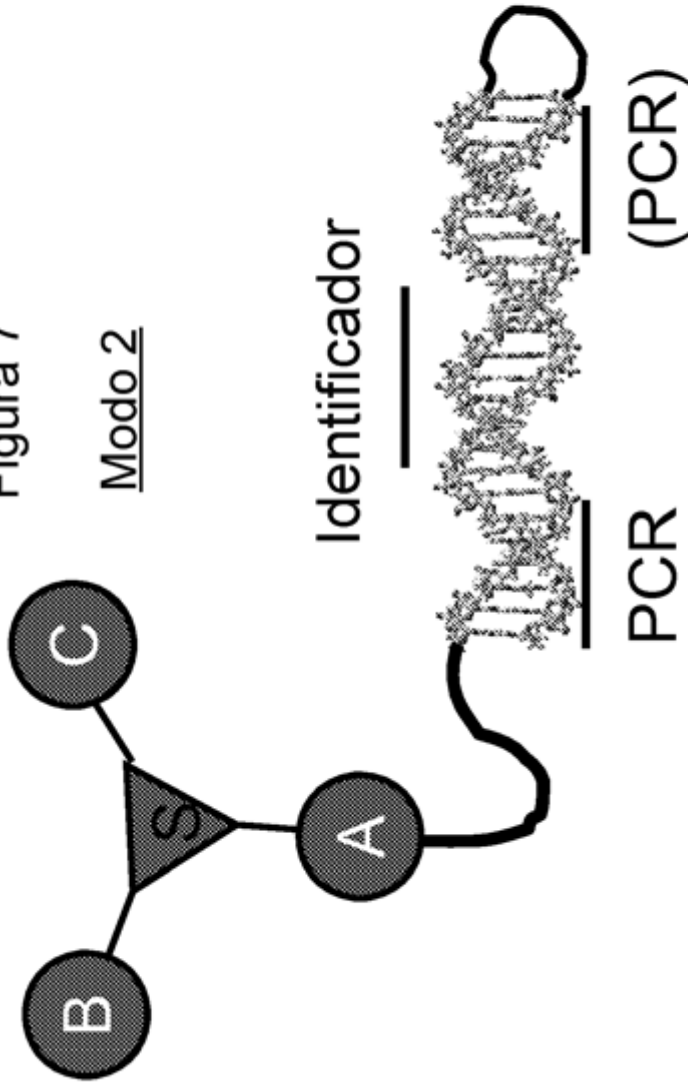


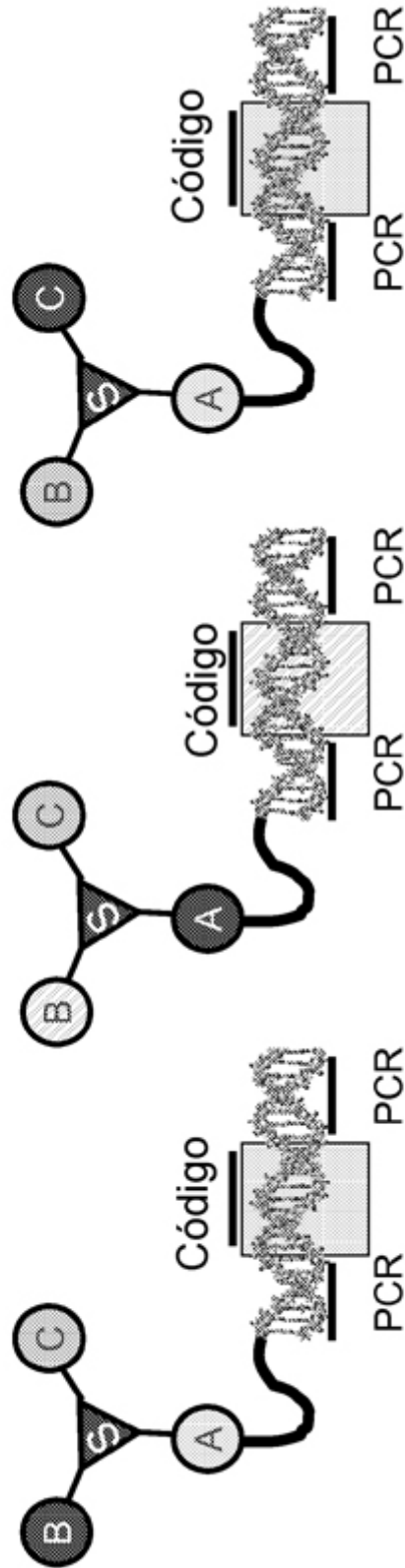
Figura 7

Modo 2

- Molécula pequeña anclada a ADNmc o ADNbc
 - el ADN puede ser una horquilla en un extremo opuesto
- Los ligadores deben variar en cuanto al tamaño y la constitución (normalmente de tipo PEG, de 20 a 200 átomos)
- La región identificadora puede ser de cualquier tamaño, pero se prefiere de 5 nt - 20 nt
- La región identificadora sólo codifica para A o B y, en algunos esquemas, S

Figura 8

Modo 2



- A tiene etiqueta
 - B es "metilo"
 - C no se combina
 - A x C interacciones (1000 x 1000 = 1.000.000)
- A es "metilo"
 - B tiene etiqueta
 - C no se combina
 - B x C interacciones
- A tiene etiqueta
 - B no se combina
 - C es "metilo"
 - A x B interacciones

•3.000.000 fragmentos combinatorios examinados

Figura 9

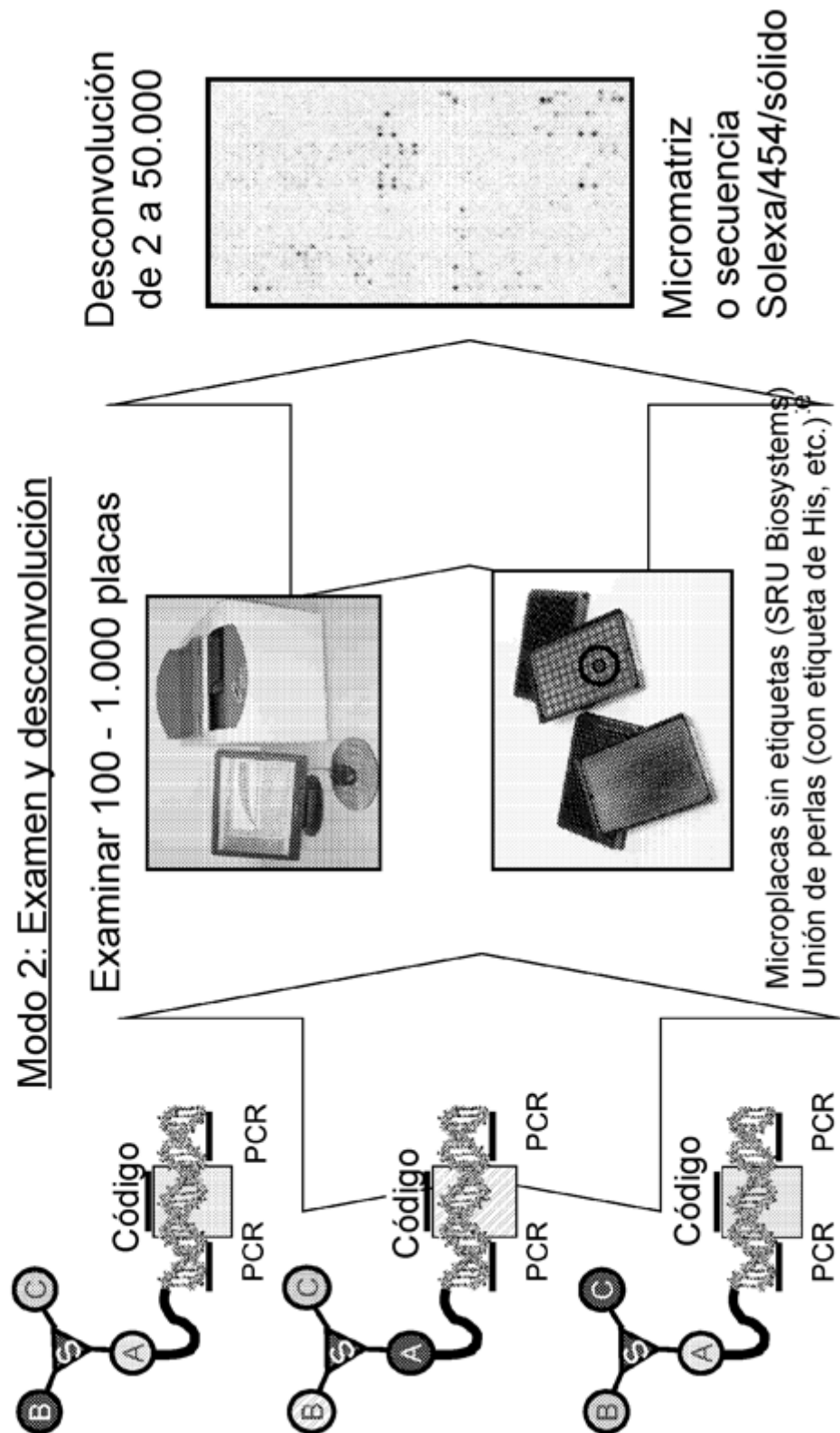


Figura 10

Parte anterior $\text{H}_2\text{N}-\text{T}$ T T CTC TCG CCTC AGC CA GGA CC-3'
T T GAG AGC GGAG TCG GT CCT (p)-5'
Cebador de PCR 5' 5'-CTC TCG CCTCAGC CA GGA- 3'

Etiqueta A1 5' XXXXXXXXGT 5' (p)GGAGGACT GT 3'
3' GGXXXXXXX 3' GG CCTCCTGA (p) 5'

Etiqueta B1 5' XXXXXXXXXGA 5'-(p)CAGACGAC GA 3'
3' CAXXXXXXXX 3' CA GTCTGCTG(p)-5'

Etiqueta C1 5' XXXXXXXXTT 5' (p)CGATGCTC TT 3'
3' CTXXXXXXX 3'CT GCTA CGAG (p) 5'

Extremo 3' 5'- (p) GCT GTG CAG GTA GAG TGC-3'
3'-AA CGA CAC GTC CAT CTC ACG-5'

Cebador de PCR 3' GCA CTC TAC CTG CAC AGC

Figura 11

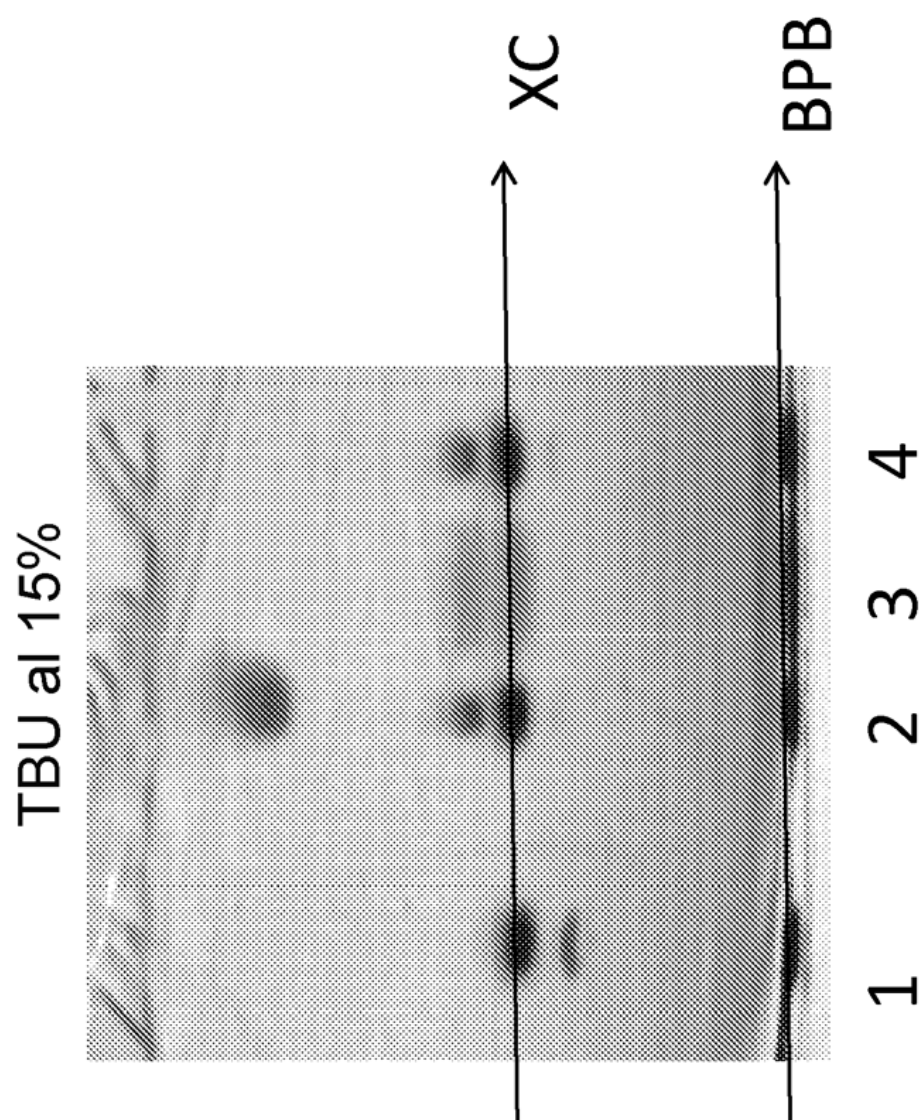


Figura 12

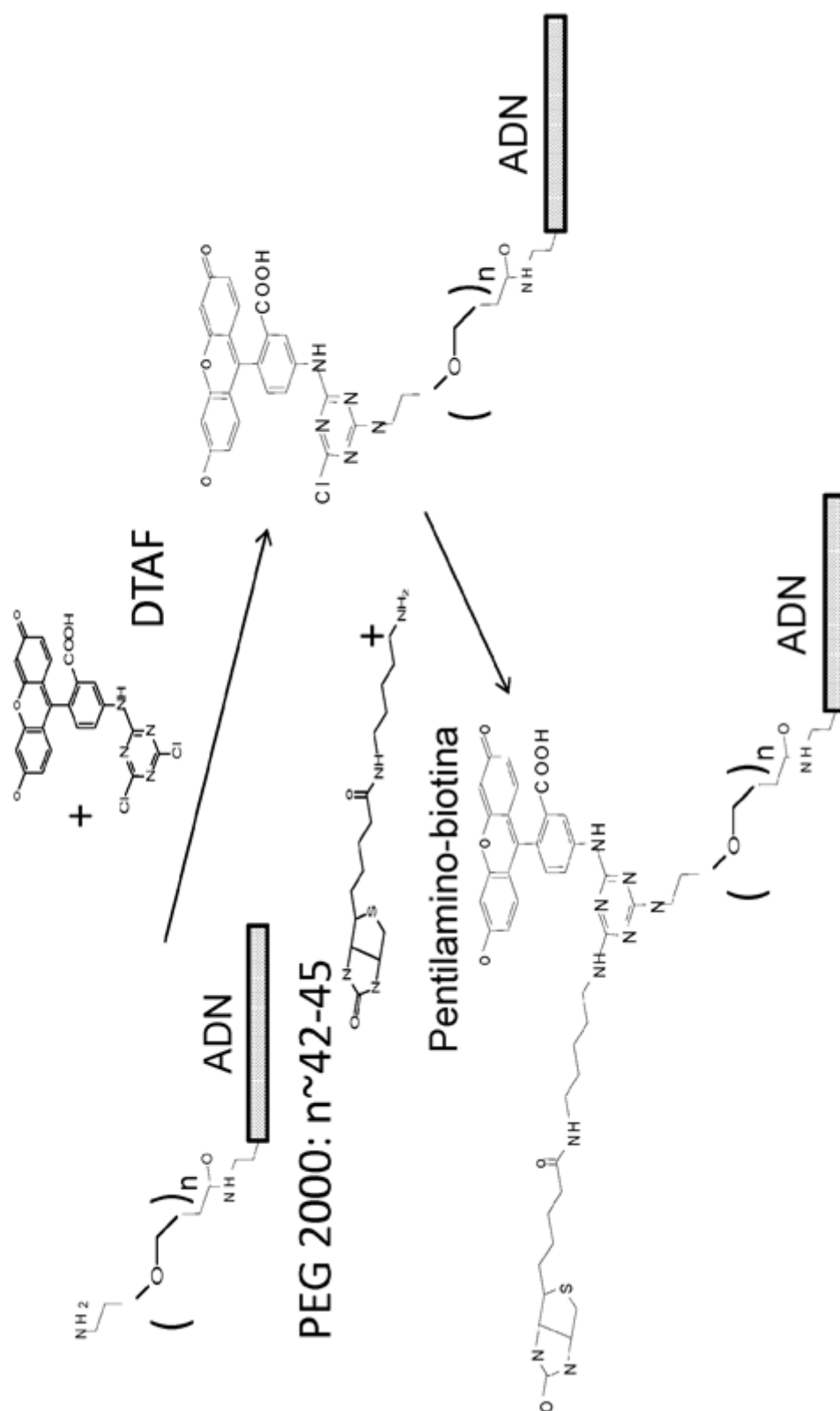
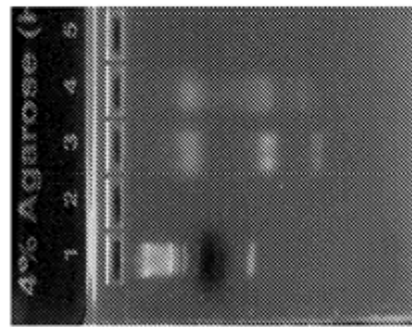


Figura 13

A. Ligamiento de etiquetas de ADN

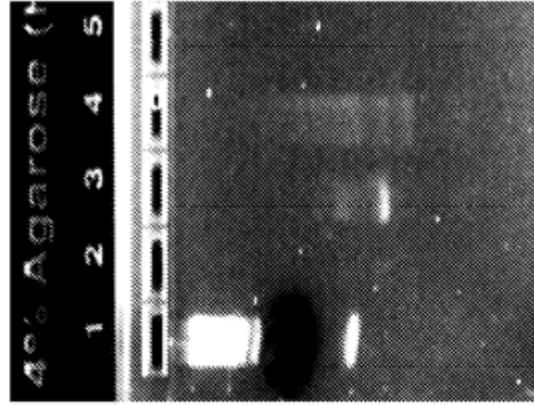
HP		A		B		C		extremo 3'	
H2N-T	T T	CTC	TCG	CCTC	AGC	CA	GGA	CC-3'	5'-(p) GCT GTG CAG GTA GAG TGC-3'
	T T	GAG	AGC	GGAG	TCG	GT	CCT	(p)-5'	3'-AA CGA CAC GTC CAT CTC ACG-5'

B



M 1 2 3

C



M 1 2 3 4

Figura 14

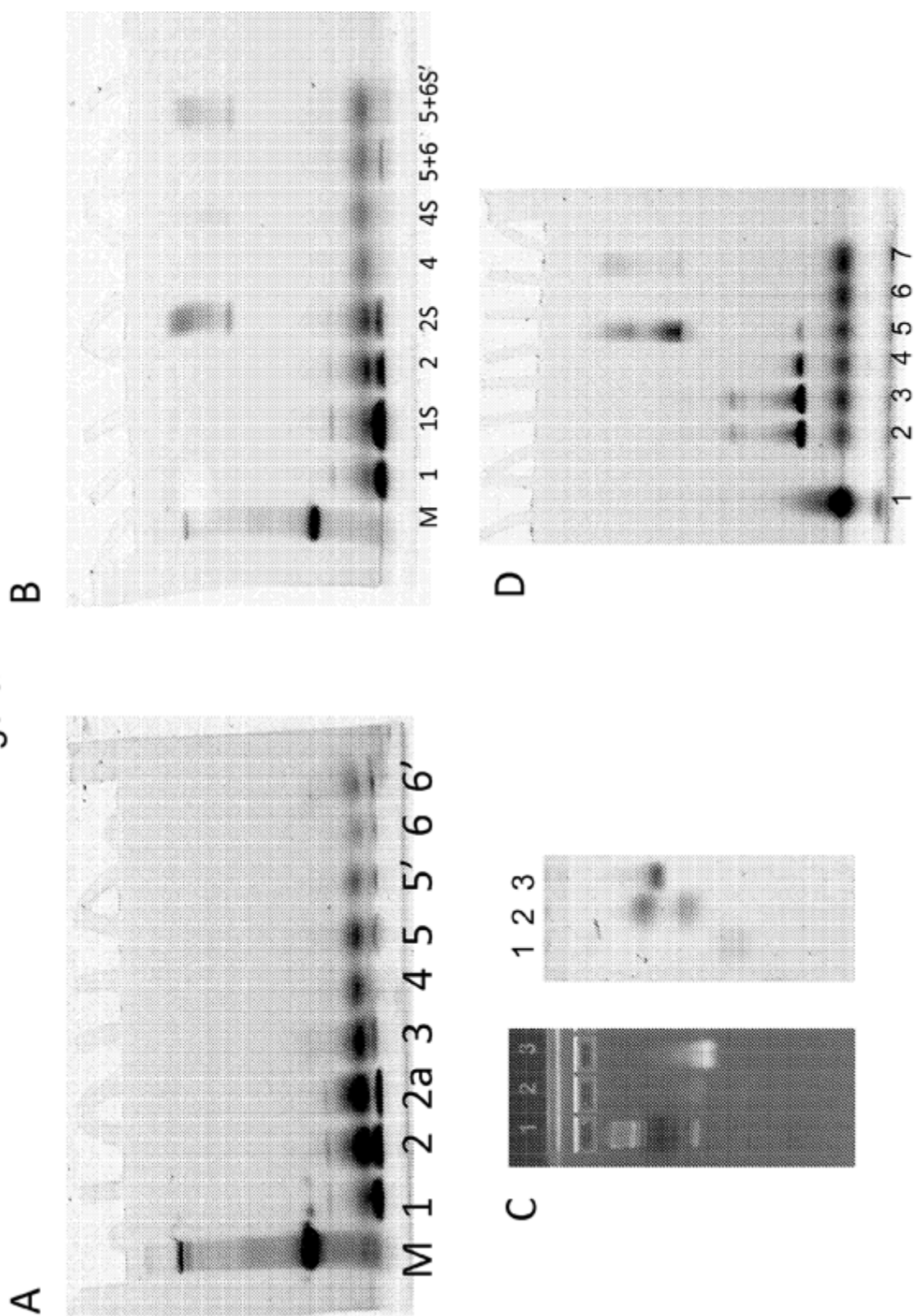


Figura 15

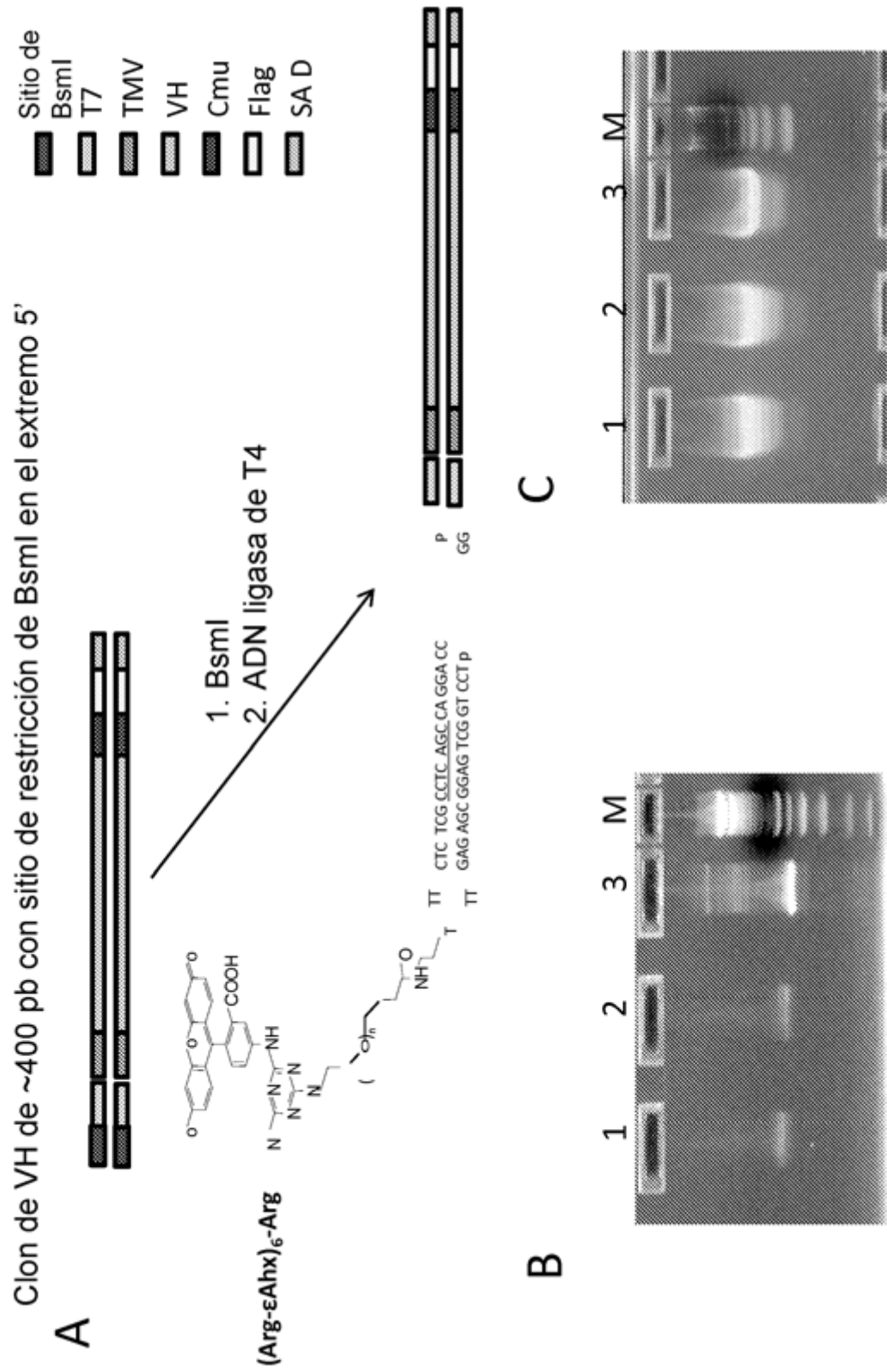


Figura 16

