



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 319 634**

51 Int. Cl.:
G01N 33/543 (2006.01)
G01N 33/68 (2006.01)
C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **03721355 .0**
96 Fecha de presentación : **11.03.2003**
97 Número de publicación de la solicitud: **1502102**
97 Fecha de publicación de la solicitud: **02.02.2005**

54 Título: **Compuestos y métodos para analizar el proteoma.**

30 Prioridad: **11.03.2002 US 363433 P**

45 Fecha de publicación de la mención BOPI:
11.05.2009

45 Fecha de la publicación del folleto de la patente:
11.05.2009

73 Titular/es: **caprotec bioanalytics GmbH**
Volmerstr. 5
12489 Berlin, DE

72 Inventor/es: **Köster, Hubert;**
Shchepinov, Mikhail, S. y
Little, Daniel, P.

74 Agente: **Ungría López, Javier**

ES 2 319 634 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

Compuestos y métodos para analizar el proteoma.

5 **Campo**

El presente documento ofrece compuestos y métodos que utilizan los compuestos para analizar biomoléculas de manera específica y selectiva. En particular, los compuestos y métodos son de utilidad para analizar el proteoma.

10 **Antecedentes**

15 El proyecto Genoma Humano ha generado una secuencia bruta de los 3 millardos de pares de bases del genoma humano y ha revelado aproximadamente 35.000 genes. Se estudian las variaciones genéticas entre diferentes individuos y entre poblaciones para determinar la asociación con la predisposición a desarrollar enfermedades, o la correlación con la eficacia y/o los efectos secundarios de los medicamentos.

20 Una manifestación frecuente de las variaciones genéticas son los polimorfismos de nucleótido simple (SNPs en sus siglas en inglés). Se han desarrollado tecnologías para analizar los SNPs a escala industrial (por ejemplo, MassARRAY® y el sistema MassARRAY®, Sequenom, Inc., San Diego, CA) y en muestras combinadas para estudiar la frecuencia de SNPs en poblaciones de diferentes sexo, origen étnico, edad y estado de salud. El objetivo último de estos estudios es comprender la etiología de la enfermedad a nivel molecular (por ejemplo, basada en varianzas genéticas (farmacogenómica)), al objeto de desarrollar ensayos diagnósticos mano a mano con nuevos medicamentos más eficaces y exentos de efectos secundarios.

25 Sin embargo, no es suficiente el conocimiento de la asociación de un SNP (o SNPs) con una determinada enfermedad o efecto secundario medicamentoso. Adicionalmente, tampoco es suficiente establecer los perfiles de expresión diferencial de los ARN mensajeros en la comparación de muestras de tejido sano y enfermo tal como se lleva a cabo hoy en día con el uso de chips de ADN (por ejemplo, la tecnología GeneChip®, Affymetrix, Inc., Santa Clara, CA; la tecnología LifeArray®, Incyte Genomics, Inc., Palo Alto, CA). Esto obedece a que los ARNm no llevan a cabo
30 las actividades metabólicas en la célula, sino que éstas son responsabilidad de las proteínas traducidas a partir de los mismos y sus posteriores modificaciones post-traduccionales (alquilación, glicosilación, fosforilación, etc.).

35 El estudio de la proteómica comprende tanto el estudio de proteínas individuales como del modo en que estas proteínas actúan dentro de una vía bioquímica. La proteómica incluye también el estudio de las interacciones proteicas con respecto a cómo forman la arquitectura que constituyen las células vivas. En muchas enfermedades humanas tales como el cáncer, la enfermedad de Alzheimer, diabetes así como las respuestas del hospedador a enfermedades infecciosas, la determinación de las proteínas reguladoras de interacciones complejas, capaces de generar enfermedades, constituye una etapa crítica para encontrar un tratamiento efectivo. A menudo, se producen SNPs y otras mutaciones de ácidos nucleicos en los genes cuyos productos son proteínas tales como (1) hormonas relacionadas con el crecimiento,
40 (2) receptores de membrana para hormonas del crecimiento, (3) componentes de la vía de señalización transmembrana, (4) proteínas fijadoras de ADN que actúan sobre la transcripción, y la inactivación de genes supresores (por ejemplo, P53), que provocan la aparición de la enfermedad.

45 El documento WO 9859360 describe disposiciones para la proteómica en las que cada fila exhibe una propiedad diferente.

50 La única tecnología que se usa en la actualidad para analizar mezclas de proteínas complejas es la electroforesis sobre gel 2D y el posterior procesamiento de imágenes para identificar variaciones en el patrón (cambios estructurales) o intensidad de diversas manchas de proteínas. La electroforesis sobre gel 2D es un método laborioso y susceptible de errores, con una baja reproducibilidad y que adolece de la incapacidad para ser automatizado. Adicionalmente, la resolución de los geles 2D es insuficiente para mostrar todas las proteínas presentes en una mezcla. Por lo tanto, el análisis del proteoma requiere tecnologías que permitan aumentar hasta niveles industriales, con las características de un procedimiento industrial: gran exactitud, reproducibilidad y flexibilidad, en donde el procedimiento sea de alto rendimiento, automatizable y coste-eficaz. Un objeto de este documento es proporcionar estas tecnologías.

55 **Resumen**

60 Se proporcionan matrices de gradiente para el análisis de biomoléculas. De manera particular, se proporcionan matrices y métodos para analizar mezclas de proteínas complejas tales como el proteoma. Las matrices ofrecen reactivos bifuncionales que permiten la separación y aislamiento de mezclas de proteínas complejas. Igualmente, se ofrecen instrumentos automatizados para llevar a cabo los métodos.

65 Se proporcionan en este documento métodos, matrices de compuestos de captura (también llamados agentes de captura en este documento) para el análisis del proteoma a nivel industrial en un formato de alto rendimiento. Los métodos y matrices permiten clasificar mezclas complejas de biomoléculas. Además, permiten identificar las estructuras de proteínas predictivas o indicativas de fenotipos específicos, tales como estados patológicos, que eliminan, de este modo, la necesidad de análisis aleatorios de SNPs, el establecimiento de perfiles de expresión y métodos analíticos de proteínas. Las matrices y los métodos clasifican las mezclas complejas mediante la provisión de una variedad de

ES 2 319 634 T3

diferentes agentes de captura. Adicionalmente, se les puede utilizar para identificar “epítopes” estructurales que sirven como marcadores de estados patológicos específicos, estratificar poblaciones individuales en relación con fenotipos específicos, permitir una comprensión detallada de la función molecular subyacente de las proteínas, y ofrecer dianas para el desarrollo de medicamentos. La mayor comprensión de las proteínas diana permite el diseño de agentes terapéuticos de eficacia superior.

Se proporcionan matrices de compuestos de captura y métodos que utilizan los compuestos, de forma aislada o combinada, para capturar, separar y analizar biomoléculas que incluyen, sin limitación, mezclas de biomoléculas, incluidos biopolímeros y macromoléculas tales como proteínas, que comprenden proteínas individuales o de membrana. Las matrices contienen una pluralidad, típicamente al menos 10, 50, 100, 1000 o más compuestos de captura diferentes, en donde diferentes compuestos están localizados en cada locus.

De forma particular, las matrices son matrices de gradiente que, por lo general, son matrices direccionables de manera bidimensional de compuestos de captura. En estas realizaciones, los compuestos de captura que se definen en este documento son aquellos que contienen o son restos enlazados mediante X e Y a un soporte sólido. Por lo tanto, las matrices son una matriz bidimensional de restos X e Y sobre una superficie que presenta restos X e Y. La superficie es un soporte sólido; los restos X e Y están dispuestos en una matriz bidimensional de gradientes continuos de una o múltiples propiedades de X e Y; las propiedades del resto X están dispuestos en un gradiente a lo largo del eje X; las propiedades del resto Y se encuentran dispuestos en un gradiente a lo largo del eje Y; cada uno de los restos X se selecciona de manera independiente para unirse a las biomoléculas de forma covalente o con una afinidad suficientemente alta, de manera que los complejos resultantes de biomolécula/compuestos de captura sean estables bajo las condiciones de análisis de espectrometría de masa; y cada uno de los restos Y se selecciona, de forma independiente, para incrementar la selectividad de la unión por medio del resto X.

Cada resto X en cada fila (eje X) de loci puede diferir de manera gradual en cuanto a hidrofilia, lipofilia, carga, tamaño, especificidad al reactivo. Cada resto Y en cada columna (eje Y) de loci puede diferir de manera gradual en cuanto a hidrofilia, lipofilia, carga, tamaño especificidad al reactivo. En algunas realizaciones, los restos Y pueden estar presentes en cada locus con los restos X, o pueden estar unidos a cada resto X.

Por ejemplo, el resto X o Y puede ser un grupo azobenceno, y se puede crear un gradiente de hidrofilia aumentando (o reduciendo) la exposición a la luz en cada locus de la matriz. El resto X puede ser un grupo cargado y se puede crear un gradiente de carga por la exposición a un incremento de corriente o voltaje. Otras propiedades adicionales incluyen una especificidad aumentada o reducida de los grupos X o Y por los grupos NH_2 , SH, SS u OH.

Las matrices resultantes presentan una superficie con una pluralidad de loci (10, 50, 100, 500, 1000, 2000, 3000, 4000, 5000, 10.000 y más) que difieren en relación con una propiedad o pares de propiedades particulares, o con una pluralidad de propiedades en incrementos definidos. Una superficie de este tipo permite la captura de moléculas y partículas biológicas con diferentes afinidades por los restos X en loci discretos. Estas matrices de gradiente se pueden utilizar en métodos para clasificar mezclas complejas o sondear superficies de células y organelas, así como en otros métodos descritos en este documento o, por ejemplo, en la solicitud de EE.UU. en tramitación junto con la presente, de la Serie N° 10/197.954 o la solicitud de PCT Internacional N° PCT/US02/22821.

De este modo, las matrices están diseñadas para permitir sondear una mezcla de biomoléculas en virtud de la interacción de los compuestos de captura en la recolección con los componentes de una mezcla, bajo condiciones que mantienen intacta su configuración tridimensional. Cada locus de la matriz está diseñado 1) para unirse, covalentemente o a través de otra interacción química, con una elevada afinidad de fijación (k_a), de manera que la unión sea irreversible o estable bajo las condiciones de análisis de espectrometría de masa a menos de todas, típicamente a alrededor de 5 a 20 o más de las biomoléculas componentes de una mezcla, dependiendo de la complejidad y diversidad de la mezcla, bajo condiciones fisiológicas, incluidas las condiciones hidrófobas, y 2) para distinguir entre biomoléculas basándose en características topológicas.

Las matrices se utilizan en una variedad de métodos, pero están diseñadas especialmente para evaluar biomoléculas, tales como un biopolímero, que forman parte de mezclas de muestras biológicas. Las matrices se utilizan en métodos imparciales, jerarquizados de mayor a menor, que evalúan los cambios estructurales, incluidos los cambios estructurales post-traduccionales, y que, por ejemplo, se usan para comparar patrones, en particular patrones de proteínas post-traduccionales, en células enfermas frente a células sanas, procedentes generalmente de un mismo individuo. Las células que sirven como fuentes de biomoléculas pueden ser congeladas en un estado metabólico seleccionado o ser sincronizadas para permitir la comparación e identificación directas de biomoléculas específicas para el fenotipo tales como biomoléculas específicas de enfermedades, por lo general, proteínas.

Un locus en la matriz incluye un grupo X de reactividad química (llamado también en este documento función o funcionalidad), que determina una unión covalente o de alta afinidad (alta k_a), y una unión de mínima afinidad de otros tres grupos (también denominados en este documento funciones o funcionalidades). Los loci incluyen también una función de selectividad Y que modula la interacción de una biomolécula con la función de reactividad.

Por ejemplo, el grupo de reactividad (función de reactividad) incluye grupos que reaccionan o interactúan específicamente con funcionalidades en la superficie de una proteína, tales como los grupos hidroxilo, amina, amida, sulfuro y ácido carboxílico, o que reconocen áreas de superficie específicas tales como un anticuerpo, una lectina o un ligando

específico del receptor, o que interactúan con el sitio activo de las enzimas. Los expertos en la técnica pueden llevar a cabo una selección en una biblioteca de funcionalidades para lograr esta interacción. Aunque esta interacción puede ser altamente específica para la reacción, estos compuestos pueden reaccionar múltiples veces con la misma molécula de proteína, dependiendo del número de grupos funcionales asequibles en la superficie. La modificación de las condiciones de reacción permite la identificación de grupos funcionales asequibles en la superficie con diferente reactividad, permitiendo, de este modo, la identificación de uno o múltiples sitios altamente reactivos usados para separar una proteína individual de una mezcla. Las tecnologías disponibles no separan especies en la mezcla de reacción resultante. Las recolecciones y los compuestos aportados en este documento resuelven este problema a través de una segunda funcionalidad, el grupo de selectividad, que altera la unión de los grupos de reactividad a la biomolécula.

Las funciones de selectividad incluyen una variedad de grupos, así como el espaciado geométrico de la segunda funcionalidad, un oligonucleótido o análogo de oligonucleótido de cadena simple, no protegido o adecuadamente protegido. La funcionalidad selectiva puede estar separada del compuesto e incluir el soporte sólido o semisólido. La funcionalidad selectiva en esta realización puede ser la porosidad, hidrofobia, carga y otras propiedades químicas del material. Por ejemplo, las funciones de selectividad interactúan de forma no covalente con las proteínas diana para alterar la especificidad o la unión de la función de reactividad. Estas funciones incluyen grupos químicos y biomoléculas que pueden impedir estéricamente proteínas de un tamaño específico, compuestos o proteínas hidrófilos (por ejemplo, PEG y tritilos), compuestos o proteínas hidrófobos (por ejemplo, aromáticos polares, lípidos, glicolípidos, fosfotriésteres, oligosacáridos), grupos con carga positiva o negativa, grupos o biomoléculas que crean estructuras secundarias o terciarias definidas.

En la práctica, en una de las realizaciones, se hace contactar una matriz de gradiente con una mezcla de biomoléculas y las moléculas fijadas se evalúan usando, por ejemplo, espectrometría de masa, seguida de la aplicación opcional de marcaje tal como marcaje fluorescente, después de la ordenación para identificar las proteínas de escasa abundancia.

En este documento se ofrecen también métodos para el descubrimiento e identificación de proteínas, seleccionadas en base a un fenotipo definido. Los métodos permiten que las proteínas se unan a las moléculas diana bajo condiciones fisiológicas, mientras conservan la conformación secundaria y terciaria correcta de la diana. Los métodos se pueden llevar a cabo bajo condiciones fisiológicas y de otro tipo, que permiten la detección de proteínas biológicamente importantes, incluyendo proteínas de membrana, que se seleccionan en base a un fenotipo definido.

Las matrices de gradiente pueden estar compuestas por restos seleccionados para capturar proteínas diana o proteínas relacionadas con grupos capaces de imitar estructuras biológicas tales como la estructura nuclear y transmembranosa, membranas artificiales o paredes celulares intactas.

Las muestras de análisis incluyen cualquier biomolécula, en especial muestras que contienen proteínas tales como mezclas de proteínas, que incluyen, sin limitación, fuentes naturales y sintéticas. Las proteínas se pueden preparar por traducción de cromosomas aislados, genes, ADNc y bibliotecas genómicas. Las proteínas se pueden aislar de células y de otras fuentes. En ciertas realizaciones, los compuestos de captura proporcionados en este documento han sido diseñados para capturar de forma selectiva diferentes modificaciones post-traduccionales de la misma proteína (es decir, patrones de fosforilación (por ejemplo, oncogenes), glicosilación y otras modificaciones post-traduccionales).

Asimismo, se ofrecen otros métodos que emplean las recolecciones. En un método, las matrices se usan para distinguir entre diferentes conformaciones de una proteína y se pueden usar, por ejemplo, para la identificación fenotípica tal como para diagnóstico. Por ejemplo, en las enfermedades de agregación de proteínas, que son enfermedades que implican una proteína alterada de manera conformacional, tal como las enfermedades por amiloides, las recolecciones pueden distinguir entre la forma de la proteína que interviene en la enfermedad de la proteína normal y, por lo tanto, diagnosticar la enfermedad en una muestra.

Breve descripción de las figuras

Figura 1 muestra la hibridación, separación y análisis de espectro de masa de una mezcla de proteínas.

Figura 2 ofrece una representación esquemática de una realización del aparato proporcionado en este documento.

Figura 3 muestra ejemplos no limitantes de estructuras dendriméricas sobre las que están basados los compuestos ofrecidos en este documento.

Figura 4 ilustra la síntesis de enmascaramiento de un chip que posee parches de hidrofobia/hidrofilia variable.

Figura 5 muestra un ejemplo de chip que posee un gradiente bidimensional continuo de carga (eje Y) y de hidrofobia/hidrofilia (eje X).

Figura 6 ilustra un cambio diazo hidrófobo/hidrófilo ajustable por la luz.

Figura 7 muestra un sustrato bidimensional continuo que tiene un gradiente de cargas eléctricas en una dimensión, y un gradiente de grupos hidrófobos/hidrófilos, ajustables por un gradiente de luz, en la otra dimensión.

ES 2 319 634 T3

Figura 8 ilustra una proteína marcada con cuatro compuestos proporcionados en este documento, permitiendo de este modo la separación específica de la proteína.

Figura 9 muestra la hibridación incrementada y específica resultante del uso de dos o más marcas de oligonucleótidos.

Figura 10 ilustra las hibridaciones posibles con proteínas marcadas con oligonucleótidos dendriméricos.

Figura 11 muestra el marcaje de una única proteína con dos oligonucleótidos en una reacción.

Figura 12 muestra diversas estructuras dendriméricas para Z.

Figura 13 es un diagrama de flujo de la producción de proteínas recombinantes.

Figura 14 ilustra la producción de una biblioteca de ADNc cebado con el oligonucleótido dT adaptado.

Figura 15 muestra la producción de una biblioteca de ADNc específica para un motivo de secuencia adaptado.

Figura 16 muestra la producción de un ADNc específico para un gen adaptado.

Figura 17 ilustra la purificación de productos de amplificación a partir de una biblioteca molde.

Figura 18 muestra una biblioteca de ADNc cebada con oligonucleótido dT adaptado como molde universal para la amplificación de subpoblaciones de genes.

Figura 19 ilustra la reducción de la complejidad durante la amplificación por PCR.

Figura 20 muestra la unión de una molécula bifuncional a una superficie sólida.

Figura 21 muestra el análisis de proteínas purificadas a partir del rastreo de compuestos y la producción de anticuerpos.

Descripción detallada de las realizaciones

A. Definiciones

A menos que se defina de otra forma, todos los términos técnicos y científicos usados en este documento tienen el significado habitualmente comprendido por un experto en la técnica a la que pertenecen la o las invenciones. En el caso de que haya una pluralidad de definiciones para términos en este documento, prevalecerán las indicadas en esta sección. Cuando se hace referencia a una URL (Localizador Uniforme de Recursos, siglas en inglés) u otro sistema de identificación o dirección similar, se entenderá que tales sistemas de identificación pueden variar y la información particular disponible en Internet puede surgir y desaparecer, pero que es posible hallar información equivalente buscando en la red. La referencia a ello pone de manifiesto la disponibilidad y difusión pública de esta información.

Como se usa en este documento, un oligonucleótido significa una secuencia lineal de hasta aproximadamente 20, aproximadamente 50 o aproximadamente 100 nucleótidos, unidos por enlaces fosfodiéster. Por encima de esta longitud, se comienza a usar el término polinucleótido.

Como se usa en este documento, un análogo de oligonucleótido significa una secuencia lineal de hasta aproximadamente 20, aproximadamente 50 o aproximadamente 100 análogos de nucleótidos, o una secuencia lineal de hasta aproximadamente 20, aproximadamente 50, o aproximadamente 100 nucleótidos unidos por un enlace “estructural” diferente del enlace fosfodiéster, por ejemplo, un enlace fosfotriéster, un enlace fosforoamidato, un enlace fosforotioato, un enlace metilfosfonato diéster, un enlace tioéster, o un enlace peptídico (ácido nucleico peptídico).

Como se usa en este documento, ácido nucleico peptídico (PNA, siglas en inglés) hace referencia a análogos de ácido nucleico en los que el esqueleto de ribosa-fosfato está sustituido por uno unido por enlaces amida.

Como se usa en este documento, proteoma significa la totalidad de las proteínas presentes dentro de una célula.

Como se usa en este documento, una biomolécula es cualquier compuesto hallado en la naturaleza, o sus derivados. Las biomoléculas incluyen, pero no están limitadas a ellos, oligonucleótidos, oligonucleósidos, proteínas, péptidos, aminoácidos, lípidos, esteroides, ácidos nucleicos peptídicos (PNAs), oligosacáridos y monosacáridos.

Como se usa en este documento, MALDI-TOF se refiere a sistemas de espectrometría de masas en tándem a tiempos de vuelo con ionización por desorción láser asistida por matrices.

Como se usa en este documento, el término “condicionado” o “acondicionador”, cuando se utiliza en relación con una proteína, significa que el polipéptido ha sido modificado para reducir la energía láser necesaria para volatilizar la

proteína, para minimizar la probabilidad de fragmentación de la proteína, o para aumentar la resolución de un espectro de masa de la proteína o de los aminoácidos componentes. La resolución de un espectro de masa de una proteína se puede aumentar acondicionando la proteína antes de llevar a cabo la espectrometría de masa. El acondicionamiento se puede efectuar en cualquier etapa previa a la espectrometría de masa y se puede llevar a cabo mientras la proteína está inmovilizada. Por ejemplo, se puede acondicionar una proteína tratándola con un material intercambiador de cationes o intercambiador de aniones, que puede reducir la heterogeneidad de carga de la proteína para eliminar el ensanchamiento de los picos causado por la heterogeneidad del número de cationes (o aniones) unidos a las diversas proteínas de una población. En una realización, se lleva a cabo la eliminación de todos los cationes por intercambio iónico, excepto los iones H^+ y amonio. Haciendo contactar un polipéptido con un agente de alquilación tal como yoduro de alquilo, yodo-acetamida, yodo-etanol o 2,3-epoxi-1-propanol, es posible prevenir, por ejemplo, la formación de enlaces disulfuro en una proteína. De manera similar, las cadenas laterales de aminoácidos cargadas se pueden convertir de derivados exentos de carga empleando cloruro de trialkilsililo.

Dado que los compuestos de captura contienen porciones de proteínas y ácidos nucleicos, se contempla también el acondicionamiento apropiado para una o ambas porciones. Por consiguiente, en el análisis de la porción de ácido nucleico así como de la porción de proteína, resulta conveniente realizar una purificación previa para enriquecer las biomoléculas que se deben analizar, así como la eliminación de todos los cationes, por ejemplo, por intercambio iónico, excepto H^+ y amonio, o efectuar otro tratamiento de acondicionamiento para mejorar la resolución.

Por lo general, no es necesario el acondicionamiento de proteínas porque éstas son relativamente estables bajo condiciones ácidas y altamente enérgicas, de modo que las proteínas no requieren acondicionamiento para los análisis de espectrometría de masa. Sin embargo, en una realización dirigida a péptidos de menor longitud, existen medios para mejorar la resolución tales como la incorporación de aminoácidos modificados, que son más básicos que los correspondientes residuos no modificados. Esta modificación aumenta, en general, la estabilidad del polipéptido durante el análisis de espectrometría de masa. Igualmente, se puede usar la cromatografía de intercambio catiónico, así como procedimientos generales de lavado y purificación que eliminan proteínas y otros componentes de la mezcla de reacción de la proteína para incrementar la resolución del espectro resultante del análisis de espectrometría de masa de la proteína.

Como se usa en este documento, “matriz” se refiere al material con el que los conjugados de compuesto de captura-biomolécula se combinan para el análisis de espectrometría de masa MALDI. Se contempla cualquier material de matriz tal como ácidos sólidos, incluido el ácido 3-hidroxipicolínico, matrices líquidas tales como glicerol, conocidos por los expertos en la técnica para el análisis de ácidos nucleicos y/o proteínas. Dado que los conjugados de compuesto-biomolécula contienen ácidos nucleicos y proteínas, se puede utilizar una mezcla (óptima para ácidos nucleicos y proteínas) de moléculas de matriz.

Como se usa en este documento, macromolécula se refiere a cualquier molécula que tiene un peso molecular comprendido entre centenas y millones. Las macromoléculas incluyen, pero sin estar limitadas a ellos, péptidos, proteínas, nucleótidos, ácidos nucleicos, hidratos de carbono y otras moléculas similares, generalmente sintetizadas por organismos biológicos, pero se pueden preparar también de forma sintética o usando métodos de biología molecular recombinante.

Como se usa en este documento, el término “biopolímero” se refiere a una molécula biológica, incluidas macromoléculas, compuesta por dos o múltiples subunidades monómeras, o derivados de las mismas, unidas por un enlace o una macromolécula. Por ejemplo, un biopolímero puede ser un polinucleótido, un polipéptido, un hidrato de carbono, o un lípido, o derivados o combinaciones de los mismos, por ejemplo, una molécula de ácido nucleico que contiene una porción de ácido nucleico peptídico o una glicoproteína. Los métodos y recolecciones de los mismos, aunque han sido descritos en relación con biopolímeros, se pueden adaptar para usarlos con otros esquemas y ensayos sintéticos tales como las síntesis orgánicas de productos farmacéuticos, o inorgánicas y cualquier otra reacción o ensayo llevado a cabo sobre un soporte sólido, o en un pocillo en volúmenes de nanolitros o menor.

Como se usa en este documento, el término biomolécula incluye biopolímeros y macromoléculas, y todas las moléculas que se pueden aislar de organismos vivos y virus, incluidos, sin limitación, células, tejidos, priones, animales, plantas, virus, bacterias y otros organismos.

Como se usa en este documento, una partícula biológica se refiere a un virus tal como un vector viral o cápside viral, con o sin ácido nucleico empaquetado, fago, incluyendo un vector de fago o una cápside de fago, con o sin ácido nucleotídico encapsulado, una célula aislada, incluyendo células eucarióticas y procarióticas, o fragmentos de las mismas, un liposoma o agente micelar, u otra partícula de empaquetamiento, y otros materiales biológicos semejantes. A los efectos de este documento, las partículas biológicas incluyen moléculas que no se consideran típicamente macromoléculas porque, por lo general, no se sintetizan, pero que están derivadas de células y virus.

Como se usa en este documento, un medicamento se refiere a cualquier compuesto que es candidato para ser utilizado como agente terapéutico o como compuesto guía para diseñar un producto terapéutico, o que es un producto terapéutico conocido. Estos compuestos pueden ser moléculas de escaso tamaño, incluidas moléculas orgánicas pequeñas, péptidos, miméticos peptídicos, moléculas antisentido, anticuerpos, fragmentos de anticuerpos, anticuerpos recombinantes. Son especialmente interesantes los “medicamentos” que tienen propiedades de unión específicas, de modo que pueden ser usados como grupos de selectividad o para clasificar los compuestos de captura, ya se trate de

una funcionalidad de separación celular por *cell sorting* que se une a una diana en un soporte, o unido a un soporte sólido, en donde la funcionalidad de separación celular por *cell sorting* es la diana del medicamento.

5 Como se usa en este documento, la expresión “ácido nucleico” se refiere a polinucleótidos de cadena sencilla y/o doble tales como ácido desoxirribonucleico (ADN) y ácido ribonucleico (ARN), así como análogos o derivados de ARN o ADN. Una molécula de ácido nucleico es un polímero lineal de nucleótidos, unidos por enlaces fosfodiéster 3', 5'. En el ADN, o ácido desoxirribonucleico, el grupo glucídico es desoxirribosa y las bases de los nucleótidos son adenina, guanina, timina y citosina. El ARN, o ácido ribonucleico, tiene ribosa como azúcar y la timina está sustituida por uracilo. En la expresión “ácido nucleico” están incluidos también los análogos de ácidos nucleicos tales como el
10 ácido nucleico peptídico (PNA), ADN fosforotioato y otros análogos o derivados, o combinaciones de los mismos.

Como se usa en este documento, el término “polinucleótido” se refiere a un oligómero o polímero que contiene al menos dos nucleótidos o derivados nucleotídicos unidos, incluidos un ácido desoxirribonucleico (ADN), un ácido ribonucleico (ARN), un derivado de ADN o ARN, que contiene, por ejemplo, un análogo nucleotídico o un enlace “de esqueleto” diferente del enlace fosfodiéster, por ejemplo, un enlace fosfotriéster, un enlace fosforoamidato, un enlace metilfosfonato diéster, un enlace fosforotioato, un enlace tioéster, o un enlace peptídico (ácido nucleico peptídico). El término “oligonucleótido” se utiliza también en este documento como sinónimo de “polinucleótido”, aunque los expertos en la técnica reconocen que los oligonucleótidos, por ejemplo los cebadores de PCR, tienen una longitud generalmente menor que aproximadamente cincuenta a cien nucleótidos.
15
20

Los análogos de nucleótidos contenidos en un polinucleótido pueden ser, por ejemplo, nucleótidos modificados por masa, lo que permite la diferenciación de masa de los polinucleótidos; nucleótidos que contienen una marca detectable tal como una marca fluorescente, radiactiva, colorimétrica, luminiscente o quimioluminiscente, que permite la detección de un polinucleótido; o nucleótidos que contienen un grupo reactivo tal como un grupo biotina o tiol, que facilita la inmovilización del polinucleótido sobre un soporte sólido. Asimismo, un polinucleótido puede contener uno o múltiples enlaces fundamentales que pueden ser escindidos selectivamente, por ejemplo, de forma química, enzimática o fotolítica. Por ejemplo, un polinucleótido puede incluir uno o múltiples desoxirribonucleótidos, seguidos de uno o múltiples ribonucleótidos, que pueden ir seguidos por uno o múltiples desoxirribonucleótidos, en donde esta secuencia se puede escindir en la secuencia de ribonucleótidos por hidrólisis básica. Del mismo modo, un polinucleótido puede contener uno o múltiples enlaces que sean relativamente resistentes a la escisión, por ejemplo, un cebador oligonucleotídico quimérico, que puede incluir nucleótidos unidos por enlaces de ácidos nucleicos peptídicos y, al menos, un nucleótido en el extremo 3', que está unido por un enlace fosfodiéster, o semejante, y que es capaz de ser extendido por una polimerasa. Las secuencias de ácidos nucleicos peptídicos se pueden preparar usando métodos bien conocidos (véase, por ejemplo, Weiler *et al.* (1997) *Nucleic Acids Res.* 25:2792-2799).
25
30
35

Un polinucleótido puede ser una porción de una molécula de ácido nucleico mayor, por ejemplo, una porción de un gen, que puede contener una región polimórfica, o una porción de una región extragenética de un cromosoma, por ejemplo, una porción de una región de unidades de repetición de nucleótidos tales como un locus de microsatélite (STR, sigla en inglés), un locus de número variable de repeticiones en tándem (VNTR, sigla en inglés), un locus de microsatélite o un locus de minisatélite. Un polinucleótido puede ser de cadena sencilla o doble, incluyendo, por ejemplo, un híbrido de ADN-ARN, o puede tener una cadena triple o cuádruple. Cuando el polinucleótido es ADN de cadena doble, puede estar en una configuración A, B, L ó Z, y un polinucleótido simple puede contener combinaciones de tales configuraciones.
40

Como se usa en este documento, una “modificación de masa” con respecto a una biomolécula que se debe analizar por espectrometría de masa, se refiere a la inclusión de cambios de los átomos o grupos constituyentes que modifican el peso molecular de la molécula resultante en incrementos definidos, detectables por el análisis de espectrometría de masa. Las modificaciones de masa no representan marcas radiactivas, tales como las marcas de isótopos o los grupos fluorescentes u otras marcas habitualmente utilizadas para la detección por medios diferentes de la espectrometría de masa.
45
50

Como se usa en este documento, el término “polipéptido” significa al menos dos aminoácidos, o derivados de aminoácidos, incluidos aminoácidos modificados por masa y análogos de aminoácidos, unidos por un enlace peptídico y que puede ser un enlace peptídico modificado. Un polipéptido se puede traducir a partir de un polinucleótido, que puede incluir al menos una porción de una secuencia de codificación, o una porción de una secuencia nucleotídica que normalmente no se traduce debido, por ejemplo, a que está localizada en un marco de lectura diferente del marco de codificación, o a que es una secuencia intrón, una secuencia 3' ó 5' no traducida, o una secuencia reguladora tal como un promotor. Asimismo, un polipéptido puede ser sintetizado químicamente y se puede modificar por métodos químicos o enzimáticos tras la traducción o la síntesis química. Los términos “polipéptido”, “péptido” y “proteína” se utilizan esencialmente como sinónimos en este documento, aun cuando el experto en la técnica reconocerá que los péptidos contienen, por lo general, menos de aproximadamente cincuenta a cien residuos aminoácidos, y que a menudo las proteínas se obtienen a partir de una fuente natural y pueden contener, por ejemplo, modificaciones post-traduccionales. Un péptido se puede modificar después de la traducción, por ejemplo, por fosforilación (fosfoproteínas), glicosilación (glicoproteínas, proteoglicanos), que se pueden llevar a cabo en la célula o en una reacción *in vitro*.
55
60
65

Como se usa en este documento, el término “conjugado” se refiere a una unión estable, típicamente debida a una interacción química, incluyendo una fijación iónica y/o covalente. Entre los medios de conjugación se encuentra la interacción de estreptavidina o avidina con biotina; la interacción hidrófoba; la interacción magnética (por ejemplo,

ES 2 319 634 T3

con el uso de perlas magnéticas funcionalizadas tales como DYNABEADS, que son perlas magnéticas recubiertas con estreptavidina, comercializadas por Dynal, Inc., Great Neck, NY y Oslo, Noruega); las interacciones polares tales como las asociaciones de “humectación” entre dos superficies polares o entre oligo/polietilenglicol; la formación de un enlace covalente tal como un enlace amida, enlace disulfuro, enlace tioéter, o mediante agentes reticulantes; y a través de un enlazador ácido-lábil o foto-escindible.

Como se usa en este documento, “muestra” se refiere a una composición que contiene un material que debe ser detectado. A estos efectos, muestra se refiere a cualquier elemento que pueda contener una biomolécula. La muestra puede ser una muestra biológica tal como un fluido biológico o un tejido biológico obtenido de cualquier organismo o una célula, o de un organismo o una partícula viral o porciones de la misma. Los ejemplos de fluidos biológicos incluyen orina, sangre, plasma, suero, saliva, semen, heces, esputo, líquido cefalorraquídeo, lágrimas, moco, esperma, líquido amniótico y similares. Los tejidos biológicos son agregados de células, habitualmente de un tipo particular, junto con la sustancia intercelular que forman uno de los materiales estructurales de una estructura humana, animal, vegetal, bacteriana, fúngica o viral, incluidos los tejidos conectivo, epitelial, muscular y nervioso. Los ejemplos de tejidos biológicos incluyen también órganos, tumores, ganglios linfáticos, arterias y célula(s) individual(es).

De esta forma, las muestras incluyen muestras biológicas (por ejemplo, cualquier material obtenido de una fuente procedente de un ser vivo (por ejemplo, humano, animal, vegetal, bacteriano, fúngico, protozooario, viral). La muestra biológica puede estar en cualquier forma, incluidos materiales sólidos (por ejemplo, tejido, conglomerados celulares y biopsias, tejidos de cadáveres), y fluidos biológicos (por ejemplo, orina, sangre, saliva, líquido amniótico y enjuagues de la boca (que contienen células bucales)). En ciertas realizaciones, los materiales sólidos están combinados con un fluido. En realizaciones de este documento, la muestra para análisis por espectrometría de masa incluye muestras que contienen una mezcla de matriz usada para los análisis de espectrometría de masa y el complejo de compuesto de captura/biomolécula.

Como se usa en este documento, la expresión “soporte sólido” significa un material no gaseoso y no líquido, provisto de una superficie. Por lo tanto, un soporte sólido puede ser una superficie plana construida, por ejemplo, de vidrio, silicio, metal, plástico o un material compuesto; o puede adoptar la forma de una perla tal como gel de sílice, un vidrio de porosidad controlada, una perla magnética o de celulosa; o puede ser un alfiler, incluida una matriz de alfileres adecuada para la síntesis o análisis combinatorio.

Como se usa en este documento, una recolección se refiere a la combinación de, por lo general, 10, 50, 100, 500, 1000 o más miembros. En particular, una recolección se refiere a la combinación de los compuestos de captura descritos anteriormente.

Como se usa en este documento, una matriz se refiere a una recolección de elementos tales como los compuestos de captura. Una matriz direccionable es aquella en la que los miembros de la matriz son identificables, típicamente por su posición en un soporte de fase sólida, pero también gracias a un identificador o marca detectable. Por consiguiente, y por lo general, los miembros de una matriz están inmovilizados en loci identificables discretos en la superficie de una fase sólida. Una pluralidad de los compuestos está unida a un soporte tal como una matriz sobre la superficie de un soporte tal como un chip de silicio u otra superficie, generalmente por medio de la unión de la funcionalidad de separación celular por *cell sorting* con un grupo o compuesto en la superficie del soporte. El direccionamiento se puede lograr marcando electrónicamente cada miembro, por ejemplo, con una marca de radiofrecuencia (RF), mediante el uso de perlas codificadas por colores, u otras marcas codificadas por colores e identificables, o por medio del peso molecular. Por lo tanto, y por lo general, los miembros de una matriz están inmovilizados en loci discretos e identificables en la superficie de una fase sólida, o directa o indirectamente unidos o asociados por cualquier otro método con la marca identificable tal como fijados a una micro-esfera u otro soporte particulado (denominado perlas en este documento), y suspendido en solución o dispersado sobre una superficie.

Como se usa en este documento, una matriz en gradiente se refiere a una matriz de restos tales como X e Y, de manera que las propiedades de los miembros de la matriz varían gradualmente a lo largo de cada eje. X es un grupo o función de reactividad, tal como se ha definido anteriormente, e Y es un grupo o función de selectividad tal como se ha descrito anteriormente. Propiedades seleccionadas de X e Y se modifican de forma predeterminada.

Como se usa en este documento, “sustrato” se refiere a un soporte insoluble sobre el que está depositada la muestra y/o la matriz. El soporte se puede construir con prácticamente cualquier material insoluble o sólido. Por ejemplo, gel de sílice, vidrio (por ejemplo, vidrio de porosidad controlada (CGP, sigla en inglés)), nailon, resina Wang, resina Merrifield, dextrano reticulado con epíclorohidrina (por ejemplo, Sephadex[®]), agarosa (por ejemplo, Sepharose[®]), celulosa, perlas magnéticas, Dynabeads, una superficie metálica (por ejemplo, de acero, oro, plata, aluminio, silicio y cobre), un material plástico (por ejemplo, polietileno, polipropileno, poliamida, poliéster, difluoruro de polivinilideno (PVDF)). Ejemplos de sustratos incluyen, sin limitaciones, perlas (por ejemplo, gel de sílice, vidrio de porosidad controlada, magnéticas, dextrano reticulado con epíclorohidrina (por ejemplo, Sephadex[®]), agarosa (por ejemplo, Sepharose[®]), celulosa), capilares, soportes planos tales como filtros de fibra de vidrio, superficies de vidrio, superficies metálicas (acero, oro, plata, aluminio, cobre y silicio), materiales plásticos incluidas las placas o membranas multi-pocillos (por ejemplo, de polietileno, polipropileno, poliamida, difluoruro de polivinilideno), alfileres (por ejemplo, matrices de alfileres adecuadas para la síntesis o análisis combinatorio, o perlas en fosas de superficies planas tales como obleas (por ejemplo, obleas de silicio), con o sin placas de filtro. El soporte sólido adopta cualquier forma deseada incluidas, sin limitaciones, una perla, capilar, placa, membrana, oblea, peine, alfiler, oblea dotada de fosas,

ES 2 319 634 T3

una matriz de fosas o pocillos de nanolitro, así como otras geometrías y formas conocidas por los expertos en la técnica. Los soportes incluyen superficies planas diseñadas para recibir o fijar muestras en loci discretos. En una realización, las superficies planas incluyen las provistas de regiones hidrófobas que rodean loci hidrófilos para recibir, contener o fijar una muestra.

5

Los soportes pueden ser particulados o pueden estar en forma de una superficie continua tal como una placa o pocillo de microtitulación, una lámina de vidrio, un chip de silicio, una lámina de nitrocelulosa, una malla de nailon, u otros materiales de este tipo. Cuando se usa un material particulado, las partículas tienen, típicamente, al menos una dimensión dentro del intervalo de 5-10 mm o menor. Estas partículas, a las que se denomina en general “perlas” en este documento, son a menudo, pero no necesariamente, esféricas. Sin embargo, la referencia al término “perla” no limita la geometría de la matriz, que puede adoptar cualquier forma, incluidas formas aleatorias, como agujas, fibras y alargadas. También se contemplan “perlas”, en especial micro-esferas que son suficientemente pequeñas para ser usadas en la fase líquida. Las “perlas” pueden incluir componentes adicionales tales como partículas magnéticas o paramagnéticas (véase, por ejemplo, Dynabeads (Dynal, Oslo, Noruega) para la separación por el uso de imanes, con la condición de que los componentes adicionales no interfieran con los métodos y análisis de este documento.

10

15

Como se usa en este documento, “polimorfismo” hace referencia a la coexistencia de más de una forma de un gen o porción del mismo. Existen al menos dos formas diferentes en que puede estar presente la porción de un gen, por ejemplo, dos secuencias de nucleótidos diferentes, que se denomina “región polimórfica de un gen”. Una región polimórfica puede ser un único nucleótido, por ejemplo, un polimorfismo nucleotídico simple (SNP, sigla en inglés), cuya identidad difiere en distintos alelos. Una región polimórfica puede tener, también, una longitud de varios nucleótidos.

20

Como se usa en este documento, “gen polimórfico” se refiere a un gen que tiene al menos una región polimórfica.

25

Como se usa en este documento, “alelo”, que se utiliza de forma intercambiable con “variante alélica”, se refiere a formas alternativas de un gen o porciones del mismo. Los alelos ocupan el mismo locus o posición en cromosomas homólogos. Cuando el sujeto tiene dos alelos idénticos de un gen, se afirma que el sujeto es homocigoto para el gen o el alelo. Cuando un sujeto tiene dos alelos diferentes de un gen, se dice que el sujeto es heterocigoto para el gen. Los alelos de un gen específico pueden diferir entre sí en un único nucleótido o en varios nucleótidos, y pueden incluir sustituciones, deleciones e inserciones de nucleótidos. Igualmente, el alelo de un gen puede ser también una forma de un gen que contiene una mutación.

30

Como se usa en este documento, “alelo predominante” se refiere a un alelo que aparece representado con la máxima frecuencia de una población determinada. El o los alelos que están presentes con menor frecuencia se designan como variantes alélicas.

35

Como se usa en este documento, “asociado” se refiere a la coincidencia con el desarrollo o manifestación de una enfermedad, trastorno o fenotipo. La asociación puede deberse, pero no está limitada a ellos, a genes responsables de la economía del organismo, cuya alteración puede constituir la base para una diversidad de enfermedades y trastornos, los que forman parte de una vía que interviene en una enfermedad, trastorno o fenotipo específico, y los que contribuyen de forma indirecta a la manifestación de una enfermedad, trastorno o fenotipo.

40

Como se usa en este documento, el término “sujeto” se refiere a un organismo vivo tal como un mamífero, una planta, un hongo, un invertebrado, un pez, un insecto, un organismo patógeno tal como un virus o una bacteria, e incluye seres humanos y otros mamíferos.

45

Como se usa en este documento, la expresión “gen” o “gen recombinante” se refiere a una molécula de ácido nucleico que contiene un marco de lectura abierta e incluye al menos una secuencia de exón y (opcionalmente) de intrón. Un gen puede ser ARN o ADN. Los genes pueden incluir regiones previas y posteriores a la región de codificación.

50

Como se usa en este documento, “intrón” se refiere a un fragmento de ADN presente en un gen determinado, que se corta y empalma durante la maduración del ARNm.

55

Como se usa en este documento, “secuencia de nucleótidos complementaria a la secuencia de nucleótidos expuesta en SEQ ID NO: x” se refiere a la secuencia de nucleótidos de la cadena complementaria de una cadena de ácidos nucleicos que tiene la SEQ ID NO: x. La expresión “cadena complementaria” se utiliza de forma intercambiable con el término “complemento”. El complemento de una cadena de ácidos nucleicos puede ser el complemento de una cadena de codificación o el complemento de una cadena no codificadora. Cuando se hace referencia a ácidos nucleicos de doble cadena, el complemento de un ácido nucleico que tiene la SEQ ID NO: x se refiere a la cadena complementaria de la cadena que tiene la SEQ ID NO: x o a cualquier ácido nucleico que tenga la secuencia de nucleótidos de la cadena complementaria de SEQ ID NO: x. Cuando se hace referencia a un ácido nucleico de cadena sencilla que tiene la secuencia de nucleótidos SEQ ID NO: x, el complemento de este ácido nucleico es un ácido nucleico que tiene una secuencia de nucleótidos que es complementaria a la de SEQ ID NO: x.

60

65

Como se usa en este documento, la expresión “secuencia codificadora” se refiere a la porción de un gen que codifica un aminoácido que constituye un polipéptido o proteína.

ES 2 319 634 T3

Como se usa en este documento, la expresión “cadena con sentido” se refiere a la cadena de una molécula de ácido nucleico de doble cadena que posee la secuencia del ARNm que codifica la secuencia de aminoácidos codificada por la molécula de ácido nucleico de doble cadena.

5 Como se usa en este documento, la expresión “cadena antisentido” se refiere a esa cadena de una molécula de ácido nucleico de doble cadena que es el complemento de la secuencia del ARNm que codifica la secuencia de aminoácidos codificada por la molécula de ácido nucleico de doble cadena.

10 Como se usa en este documento, los aminoácidos que se encuentran en las diversas secuencias de aminoácidos que aparecen en este documento, se identifican de acuerdo con sus abreviaturas bien conocidas de tres o una letra. Los nucleótidos que existen en los distintos fragmentos de ADN se designan con la nomenclatura de una única letra usada habitualmente en la técnica (véase la Tabla 1).

15 Como se usa en este documento, residuo de aminoácido se refiere a un aminoácido formado tras la digestión química (hidrólisis) de un polipéptido en sus enlaces peptídicos. Los residuos de aminoácidos que se describen en este documento se encuentran, en determinadas realizaciones, en la forma isómera “L”. Los residuos en la forma isómera “D” pueden ser sustituidos por cualquier residuo aminoácido L, con la condición de que el polipéptido conserve la propiedad funcional deseada. NH₂ se refiere al grupo amino libre presente en el extremo amino terminal de un polipéptido. COOH se refiere al grupo carboxi libre presente en el extremo carboxilo terminal de un polipéptido. De acuerdo con la nomenclatura convencional de los polipéptidos, descrita en *J. Biol. Chem.*, 243:3552-59 (1969) y adoptada en la 37ª C.F.R. §§ 1.821-1.822, las abreviaturas de los residuos de aminoácidos se muestran en la Tabla siguiente:

TABLA 1

Tabla de correspondencias

SÍMBOLO		AMINOÁCIDO
1 LETRA	3 LETRAS	
Y	Tyr	Tirosina
G	Gly	Glicina
F	Phe	Fenilalanina
M	Met	Metionina
A	Ala	Alanina
S	Ser	Serina
I	Ile	Isoleucina
L	Leu	Leucina
T	Thr	Treonina
V	Val	Valina
P	Pro	Prolina
K	Lys	Lisina
H	His	Histidina
Q	Gln	Glutamina
E	Glu	Ácido glutámico
Z	Glx	Glu y/o Gln
W	Trp	Triptófano
R	Arg	Arginina
D	Asp	Ácido aspártico
N	Asn	Asparagina
B	Asx	Asn y/o Asp
C	Cys	Cisteína
X	Xaa	Desconocido u otro

65 Se debe señalar que todas las secuencias de residuos aminoácidos representadas en este documento mediante fórmulas tienen una orientación de izquierda a derecha en la dirección convencional de extremo amino terminal a extremo carboxilo terminal. Adicionalmente, la expresión “residuo (de) aminoácido” está definida en su sentido más

ES 2 319 634 T3

amplio para incluir los aminoácidos enumerados en la Tabla de Correspondencias, y los aminoácidos modificados y poco frecuentes, tales como los que se relacionan en 37 C.F.R. §§ 1.821-1.822. Además, se debe destacar que un guión situado al comienzo o al final de una secuencia de residuos aminoácidos indica un péptido unido a una secuencia adicional de uno o múltiples residuos aminoácidos, o a un grupo amino-terminal tal como NH₂ o a un grupo carboxilo-terminal tal como COOH.

En un péptido o proteína, los expertos en la técnica conocen sustituciones conservativas adecuadas de aminoácidos que, por lo general, se pueden llevar a cabo sin alterar la actividad biológica de la molécula resultante. Los expertos en la técnica reconocerán que, en general, las sustituciones de aminoácidos aislados en regiones no esenciales de un polipéptido no alteran sustancialmente la actividad biológica (véase, por ejemplo, Watson *et al. Molecular Biology of the Gene*, 4ª edición, 1987, The Benjamin/Cummings Pub. Co., pág. 224).

Dichas sustituciones se pueden llevar a cabo de acuerdo con las expuestas en la siguiente Tabla 2:

TABLA 2

Ala (A)	Gly; Ser
Arg (R)	Lys
Asn (N)	Gln; His
Asp (D)	Glu
Cys (C)	Ser
Gln (Q)	Asn
Glu (E)	Asp
Gly (G)	Ala; Pro
His (H)	Asn; Gln
Ile (I)	Leu; Val
Leu (L)	Ile; Val
Lys (K)	Arg; Gln
Met (M)	Leu; Tyr; Ile
Phe (F)	Met; Leu; Tyr
Ser (S)	Thr
Thr (T)	Ser
Trp (W)	Tyr
Tyr (Y)	Trp; Phe
Val (V)	Ile; Leu

Asimismo, es posible llevar a cabo otras sustituciones, que se pueden determinar empíricamente o de acuerdo con sustituciones conservativas conocidas.

Como se usa en este documento, un ADN u homólogo de ácido nucleico se refiere a un ácido nucleico que incluye una secuencia de nucleótidos conservada preseleccionada, tal como una secuencia que codifica un polipéptido terapéutico. Por la expresión “sustancialmente homólogo” se entiende la existencia de una homología de al menos 80%, al menos 90%, al menos 95% con el mismo, o un porcentaje menor de homología o identidad y una actividad o función biológica conservada.

Con frecuencia, los términos “homología” e “identidad” se usan de forma intercambiable. En este sentido, la homología o identidad porcentual se puede determinar, por ejemplo, comparando la información de secuencia usando el programa informático GAP. El programa GAP utiliza el método de alineación de Needleman y Wunsch (*J. Mol. Biol.* 48:443 (1970), según la revisión de Smith y Waterman (*Adv. Appl. Math.* 2:482 (1981)). En pocas palabras, el programa GAP define similitud como el número de símbolos alineados (por ejemplo, nucleótidos o aminoácidos) que son similares, dividido por el número total de símbolos en la más corta de las dos secuencias. Los parámetros por defecto para el programa GAP pueden incluir: (1) una matriz de comparación unaria (que contiene un valor de 1 para identidades y 0 para ausencia de identidades) y la matriz ponderada de comparación de Gribkov y Burgess, *Nucl. Acids Res.* 14:6745 (1986), según la describen Schwartz y Dayhoff, editores, *ATLAS OF PROTEIN SEQUENCE AND STRUCTURE*, National Biomedical Research Foundation, págs. 353-358 (1979); (2) una penalización de 3,0 para cada separación y una penalización adicional de 0,10 por cada símbolo en cada separación; y (3) ausencia de penalización para las separaciones finales.

Es posible determinar si dos moléculas de ácido nucleico cualquiera tienen secuencias “idénticas” en al menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ó 99% usando algoritmos informáticos conocidos tales como el programa “FAST A”, utilizando, por ejemplo, los parámetros por defecto como en Pearson y Lipman, *Proc. Natl. Acad. Sci.*

ES 2 319 634 T3

USA 85:2444 (1988). De forma alternativa, se puede usar la función BLAST de la base de datos de información del Centro Nacional de Biotecnología para determinar la identidad.

En general, las secuencias están alineadas de manera que se alcanza la coincidencia de máximo orden. La “identidad” tiene, por sí misma, un significado reconocido en la técnica, que se puede calcular usando métodos ya publicados. (Véanse, por ejemplo, *Computational Molecular Biology*, Lesk, A.M., ed. Oxford University Press, Nueva York, 1988; *Biocomputing: Informatic and Genome Projects*, Smith, D.W., ed., Academic Press, Nueva York, 1993; *Computer Analysis of Sequence Data, Part I*, Griffin, A.M., y Griffin, H.G., eds., Humana Press, Nueva Jersey, 1994; *Sequence Analysis Primer*, Gribskov, M. y Devereux, J., eds., M Stockton Press, Nueva York, 1991). Aunque existe una serie de métodos para medir la identidad entre dos secuencias de polinucleótidos o polipéptidos, el término “identidad” es bien conocido para los expertos en la técnica (Carillo, H. y Lipton, D., *SIAM J Applied Math* 48:1073 (1988)). Los métodos para determinar identidad y similitud están codificados en los programas informáticos. Los métodos de los programas informáticos para determinar la identidad y similitud entre dos secuencias incluyen, pero no están limitados a ellos, el paquete de programa GCG (Devereux, J. *et al*, *Nucleic Acids Research* 12(I):387 (1984); BLASTP, BLASTIN, FASTA (Atschul, S.F. *et al*, *J Molec Biol.* 215:403 (1990)).

Por lo tanto, tal como se usa en este documento, el término “identidad” representa una comparación entre un polipéptido o polinucleótido de ensayo y uno de referencia. Por ejemplo, un polipéptido de ensayo se puede definir como cualquier polipéptido que es idéntico en 90% o más a un polipéptido de referencia.

Como se usa en este documento, la expresión “idéntico en al menos 90% a” se refiere a identidades porcentuales desde 90 hasta 99,99 con respecto a los polipéptidos de referencia. La identidad a un nivel de 90% o más es indicativa del hecho de que, suponiendo con fines de ejemplificación que se comparan una longitud de polipéptido de ensayo y de referencia de 100 aminoácidos, no más de 10% (por ejemplo, 10 de cada 100) aminoácidos del polipéptido de ensayo difiere de los del polipéptido de referencia. Se pueden realizar comparaciones similares entre polinucleótidos de ensayo y de referencia. Estas diferencias se pueden representar como mutaciones puntuales, distribuidas aleatoriamente a lo largo de toda la longitud de una secuencia de aminoácidos, o pueden estar agrupadas en una o múltiples localizaciones de longitudes variables hasta el máximo permisible, por ejemplo, una diferencia de 10/100 aminoácidos (identidad de aproximadamente 90%). Las diferencias se definen como sustituciones o deleciones de ácidos nucleicos o aminoácidos.

Como se usa en este documento, la estringencia de hibridación en la determinación de desajustes porcentuales es la siguiente:

- 1) Estringencia elevada: 0,1 x SSPE, 0,1% SDS, 65°C
- 2) Estringencia media: 0,2 x SSPE, 0,1% SDS, 50°C
- 3) Estringencia baja: 1,0 x SSPE, 0,1% SDS, 50°C

Los expertos en la técnica saben que la etapa de lavado ayuda a seleccionar híbridos estables y conocen, también, los ingredientes de SSPE (véase, por ejemplo, Sambrook, E.F., Fritsch, T. Maniatis en: *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory Press (1989), vol. 3, pág. B.13; véanse, igualmente, diversos catálogos que describen soluciones de laboratorio usadas habitualmente). SSPE es solución tampón fosfato de NaCl 0,18 a pH 7,4. Además, los expertos en la técnica reconocerán que la estabilidad de los híbridos está determinada por T_m , que es una función de la concentración del ión sodio y la temperatura ($T_m = 81,5^\circ\text{C} - 16,6(\log_{10}[\text{Na}^+]) + 0,41 (\% \text{ G} + \text{C}) - 600/l$), de manera que los únicos parámetros críticos para la estabilidad del híbrido bajo las condiciones de lavado son la concentración del ión sodio en el SSPE (o SSC) y la temperatura.

Se entiende que es posible lograr estringencias equivalentes usando soluciones tampón, sales y temperaturas alternativas. A modo de ejemplo y no de limitación, los procedimientos que usan condiciones de baja estringencia son los siguientes (véase también Shilo y Weinberg, *Proc. Natl. Acad. Sci. USA* 78:6789-6792 (1981)): Durante 6 horas se pretratan filtros que contienen ADN a 40°C en una solución que contiene formamida al 35%, 5X SSC, Tris-HCl 50 mM (pH 7,5), EDTA 5 mM, PVP al 0,1%, Ficoll al 0,1%, BSA al 1% y 500 µg/ml de ADN desnaturalizado de esperma de salmón (10X SSC es cloruro sódico 1,5 M y citrato sódico 0,15 M, ajustado a pH 7).

Las hibridaciones se llevan a cabo en la misma solución, con las modificaciones siguientes: se utilizan PVP al 0,02%, Ficoll al 0,02%, BSA al 0,2%, 100 µg/ml de ADN de esperma de salmón, sulfato de dextrano al 10% (peso/vol) y 5-20 X 10⁶ cpm de sonda marcada con P³². Los filtros se incuban en la mezcla de hibridación durante 18-20 horas a 40°C y, a continuación, se lavan durante 1,5 horas a 55°C en una solución que contiene 2X SSC, Tris-HCl 25 mM (pH 7,4), EDTA 5 mM y SDS al 0,1%. La solución de lavado se sustituye por solución recién preparada y se incuba durante 1,5 horas adicionales a 60°C. Los filtros se secan con papel secante y se exponen para autorradiografía. Si es necesario, los filtros se lavan por tercera vez a 65-68°C y se exponen nuevamente a la película. Se conocen en la técnica otras condiciones de estringencia baja que se pueden usar (por ejemplo, según se emplean en hibridaciones de especies cruzadas).

A modo de ejemplo y no de limitación, los procedimientos que usan condiciones de estringencia moderada incluyen, por ejemplo, sin estar limitados a ellos, los siguientes: Se pretratan filtros que contienen ADN durante 6 horas a

ES 2 319 634 T3

55°C en una solución que contiene 6X SSC, 5X solución de DEHart, SDS al 0,5% y 100 µg/ml de ADN desnaturado de esperma de salmón. Las hibridaciones se llevan a cabo en la misma solución y se usa una sonda de 5-10 X 10⁶ cpm marcada con P³². Los filtros se incuban en mezcla de hibridación durante 18-20 horas a 55°C y, a continuación, se lavan dos veces durante 30 minutos a 60°C en una solución que contiene °X SSC y SDS al 0,1%. Los filtros se secan sobre papel secante y se exponen para autorradiografía. En la técnica se conocen bien otras condiciones de estringencia moderada que se pueden usar. El lavado de los filtros se efectúa a 37°C durante 1 hora en una solución que contiene 2X SSC, SDS al 0,1%.

A modo de ejemplo y no de limitación, los procedimientos que usan condiciones de alta estringencia son los siguientes: Se lleva a cabo una pre-hibridación de filtros que contienen ADN durante 8 horas hasta durante una noche a 65°C en una solución tampón compuesta por 6X SSC, Tris-HCl 50 mM (pH 7,5), EDTA 1 mM, PVP al 0,02%, Ficoll al 0,02%, BSA al 0,02%, y 500 µg/ml de ADN desnaturado de esperma de salmón. Los filtros se hibridan durante 48 horas a 65°C en mezcla de pre-hibridación que contiene 100 µg/ml de ADN desnaturado de esperma de salmón y una sonda de 5-20 X 10⁶ cpm marcada con P³². El lavado de los filtros se efectúa a 37°C durante 1 hora en una solución que contiene 2X SSC, PVP al 0,01%, Ficoll al 0,01% y BSA al 0,01%. A continuación, se realiza un lavado en 0,1X SSC a 50°C durante 45 minutos antes de la autorradiografía. En la técnica, se conocen bien otras condiciones de estringencia alta que se pueden usar.

La expresión “sustancialmente idéntico” o “sustancialmente homólogo o similar” varía con el contexto, tal como lo entenderán los expertos en la técnica pertinente y, por lo general, significa una identidad de al menos 60% o 70%, preferentemente significa una identidad de al menos 80%, 85% o, más preferentemente, de al menos 90% y, de forma especialmente preferida, una identidad de al menos 95%.

Se debe entender que los compuestos ofrecidos en este documento pueden contener centros quirales. Estos centros quirales pueden ser de configuración (R) o (S), o ser una mezcla de las mismas. De esta forma, los compuestos ofrecidos en este documento pueden ser enantioméricamente puros, o ser mezclas estereoisómeras o diastereoisómeras. En el caso de residuos aminoácidos, tales residuos pueden ser de la forma L o D. En una realización, la configuración de los residuos aminoácidos de origen natural es L.

Como se usa en este documento, sustancialmente puro significa suficientemente homogéneo para aparecer libre de impurezas fácilmente detectables por los métodos convencionales de análisis tales como cromatografía de capa fina (TLC, sigla en inglés), electroforesis sobre gel, cromatografía líquida de alto rendimiento (HPLC, sigla en inglés) y espectrometría de masa (MS), utilizados por los expertos en la técnica para evaluar dicha pureza, o suficientemente puro, de manera que una purificación adicional no alteraría de forma detectable las propiedades físicas y químicas, tales como las actividades biológica y enzimática, de la sustancia. Los expertos en la técnica conocen métodos de purificación de los compuestos para obtener compuestos sustancialmente puros desde el punto de vista químico. No obstante, un compuesto sustancialmente puro desde el punto de vista químico puede ser una mezcla de estereoisómeros. En tales casos, una purificación adicional podría incrementar la actividad específica del compuesto.

Como se usa en este documento, un enlace o resto susceptible de escisión hace referencia a un enlace o resto que se escinde o es escindible bajo condiciones específicas de tipo químico, enzimático o fotolítico. Cuando no se especifica en el documento, dicho enlace es escindible bajo las condiciones del análisis MALDI-MS, tal como un láser UV o IR.

Como se usa en este documento, un resto “escindible selectivamente” es un resto que puede ser escindido de forma selectiva sin afectar o alterar la composición de las otras porciones del compuesto de interés. Por ejemplo, un resto L escindible de los compuestos descritos en este documento es uno que puede ser escindido por medios químicos, enzimáticos, fotolíticos o de otro tipo, sin afectar o alterar la composición (por ejemplo, la composición química) de la biomolécula conjugada, incluida una proteína. Restos “no escindibles” son aquellos que no pueden ser escindidos selectivamente sin afectar o alterar la composición de las porciones restantes del compuesto de interés.

Como se usa en este documento, unir con afinidad alta significa una unión que tiene una constante de asociación k_a de al menos 10⁹ y, por lo general, 10¹⁰, 10¹¹ litros/mol o mayor, o una K_{eq} de 10⁹, 10¹⁰, 10¹¹, 10¹² o mayor. A los efectos del presente documento, las uniones de alta afinidad formadas por los grupos de reactividad son aquellas que son estables al láser (UV e IR) usado en los análisis MALDI-MS.

Como se usa en este documento, “alquilo”, “alquenilo” y “alquinilo”, si no se especifica, contienen desde 1 hasta 20 carbonos, o 1 a 16 carbonos, y son cadenas de carbono lineales o ramificadas. Las cadenas de carbono alquenilo tienen 2 hasta 20 carbonos y, en ciertas realizaciones, contienen 1 hasta 8 dobles enlaces. Las cadenas de carbono alquinilo tienen desde 2 hasta 20 carbonos y, en una realización, contienen 1 hasta 8 triples enlaces. Las cadenas de carbono alquinilo de 2 a 16 carbonos, en ciertas realizaciones, contienen 1 hasta 5 triples enlaces. Ejemplos de grupos alquilo, alquenilo y alquinilo incluyen, sin limitaciones, metilo, etilo, propilo, isopropilo, isobutilo, n-butilo, sec-butilo, terc-butilo, isopentilo, neopentilo, terc-pentilo e isohexilo. Los grupos alquilo, alquenilo y alquinilo, salvo que se especifique lo contrario, pueden estar opcionalmente sustituidos con uno o múltiples grupos, incluidos sustituyentes de grupos alquilo que pueden ser iguales o diferentes.

Como se usa en este documento, “alquilo inferior”, “alquenilo inferior” y “alquinilo inferior” se refieren a cadenas de carbono que tienen menos de aproximadamente 6 carbonos.

ES 2 319 634 T3

Como se usa en este documento, “alqu(en)(in)ilo” se refiere a un grupo alquilo que contiene al menos un doble enlace y al menos un triple enlace.

5 Como se usa en este documento, un “sustituyente de grupo alquilo” incluye, sin limitaciones, halo, haloalquilo, incluidos halo-alquilo inferior, arilo, hidroxilo, alcoxi, ariloxi, alquiloxi, alquiltio, ariltio, aralquilo, aralquiltio, carboxialcoxicarbonilo, oxo y cicloalquilo.

10 Como se usa en este documento, “arilo” se refiere a grupos aromáticos que contienen desde 5 hasta 20 átomos de carbono, y pueden ser un sistema de anillo mono-, multicíclico o fusionado. Los grupos arilo incluyen, sin limitaciones, fenilo, naftilo, bifenilo, fluorenilo y otros, que pueden estar no sustituidos o sustituidos con uno o múltiples sustituyentes.

15 Como se usa en este documento, “arilo” se refiere también a grupos que contienen arilo, incluidos, sin limitaciones, grupos ariloxi, ariltio, arilcarbonilo y arilamino.

20 Como se usa en este documento, “sustituyente de grupo arilo” incluye, sin limitaciones, alquilo, alqueno, alquino, cicloalquilo, cicloalquil-alquilo, arilo, heteroarilo sustituido opcionalmente con 1 o múltiples, incluidos 1 a 3 sustituyentes seleccionados de halo, haloalquilo y alquilo, aralquilo, heteroaralquilo, alqueno que contiene 1 a 2 dobles enlaces, alquino que contiene 1 a 2 triples enlaces, grupos alqu(en)(in)ilo, halo, pseudo-halo, ciano, hidroxilo, haloalquilo y polihaloalquilo, incluido halo-alquilo inferior, especialmente trifluorometilo, formilo, alquilcarbonilo, arilcarbonilo que está sustituido opcionalmente con 1 o múltiples incluidos 1 a 3 sustituyentes seleccionados de halo, haloalquilo y alquilo, heteroaril-carbonilo, carboxi, alcoxi-carbonilo, ariloxi-carbonilo, aminocarbonilo, alquil-amino-carbonilo, dialquil-amino-carbonilo, aril-amino-carbonilo, diaril-amino-carbonilo, aralquil-amino-carbonilo, alcoxi, ariloxi, perfluoroalcoxi, alqueno, alquino, arilalcoxi, aminoalquilo, alquil-aminoalquilo, dialquil-aminoalquilo, aril-aminoalquilo, amino, alquilamino, dialquilamino, arilamino, alquil-arilamino, alquil-carbonilamino, aril-carbonilamino, azido, nitro, mercapto, alquiltio, ariltio, perfluoroalquiltio, tiociano, isotiociano, alquilsulfinilo, alquilsulfonilo, arilsulfinilo, arilsulfonilo, aminosulfonilo, alquil-aminosulfonilo, dialquil-aminosulfonilo y aril-aminosulfonilo.

30 Como se usa en este documento, “aralquilo” se refiere a un grupo alquilo en el cual uno de los átomos de hidrógeno del alquilo está sustituido con un grupo arilo.

Como se usa en este documento, “heteroaralquilo” se refiere a un grupo alquilo, en el que uno de los átomos de hidrógeno del alquilo está sustituido con un grupo heteroarilo.

35 Como se usa en este documento, “cicloalquilo” se refiere a un sistema de anillo mono- o multicíclico saturado, en una realización, de 3 a 10 átomos de carbono, o 3 hasta 6 átomos de carbono; cicloalqueno y cicloalquino se refieren a sistemas de anillo mono- o multicíclicos que incluyen, respectivamente, al menos un doble enlace y al menos un triple enlace. Los grupos cicloalqueno y cicloalquino pueden contener, en una realización, 3 a 10 átomos de carbono, en tanto que, en otras realizaciones, el grupo cicloalqueno contiene 4 hasta 7 átomos de carbono, y los grupos cicloalquino, en otras realizaciones, contienen 8 a 10 átomos de carbono. Los sistemas de anillo de los grupos cicloalquilo, cicloalqueno y cicloalquino pueden estar compuestos por un anillo o dos o más anillos que pueden estar unidos entre sí de forma fusionada, mediante puentes o espiro-conectada y, opcionalmente, pueden estar sustituidos con uno o múltiples sustituyentes del grupo alquilo.

45 “Cicloalqu(en)(in)ilo” se refiere a un grupo cicloalquilo que contiene al menos un doble enlace y al menos un triple enlace.

50 Como se usa en este documento, “heteroarilo” se refiere a un sistema de anillo monocíclico o multicíclico, en una realización de aproximadamente 5 hasta aproximadamente 15 miembros, en donde uno o más, ó 1 hasta 3 de los átomos del sistema de anillo es un heteroátomo que es un elemento diferente del carbono, por ejemplo, átomos de nitrógeno, oxígeno y azufre. El heteroarilo puede estar sustituido opcionalmente con uno o múltiples, incluidos 1 a 3 sustituyentes del grupo arilo. El grupo heteroarilo puede estar fusionado opcionalmente con un anillo benceno. Ejemplos de grupos heteroarilo incluyen, sin limitaciones, pirroles, porfirinas, furanos, tiofenos, selenofenos, pirazoles, imidazoles, triazoles, tetrazoles, oxazoles, oxadiazoles, tiazoles, tiadiazoles, indoles, carbazoles, benzofuranos, benzotiofenos, indazoles, benzimidazoles, benzotriazoles, benzoxatriazoles, benzotiazoles, benzoselenozoles, benzotiadiazoles, benzoselenadiazoles, purinas, piridinas, piridazinas, pirimidinas, pirazinas, triazinas, quinolinas, acridinas, isoquinolinas, cinolinas, ftalazinas, quinazolininas, quinoxalinas, fenazinas, fenantrolinas, imidazolinilo, pirrolidinilo, pirimidinilo, tetrazolilo, tienilo, piridilo, pirrolilo, N-metilpirrolilo, quinolinilo e isoquinolinilo.

60 Como se usa en este documento, “heteroarilo” se refiere también a grupos que contienen un heteroarilo, incluidos, sin limitaciones, heteroariloxi, heteroariltio, heteroaril-carbonilo y heteroaril-amino.

65 Como se usa en este documento, “heterocíclico” se refiere a un sistema de anillo monocíclico o multicíclico, en una realización de 3 a 10 miembros, en otra realización de 4 a 7 miembros, incluidos 5 a 6 miembros, en donde uno o más, incluidos 1 a 3 de los átomos del sistema de anillo es un heteroátomo, que es un elemento diferente de carbono, por ejemplo, átomos de nitrógeno, oxígeno y azufre. El heterocíclico puede estar sustituido opcionalmente con uno o múltiples, o 1 a 3 sustituyentes del grupo arilo. En ciertas realizaciones, los sustituyentes del grupo heterocíclico incluyen hidroxilo, amino, alcoxi que contiene 1 a 4 átomos de carbono, halo-alquilo inferior, incluido trihalo-metilo, tales

ES 2 319 634 T3

como trifluorometilo, y halógeno. Como se usa en este documento, el término heterociclo puede incluir la referencia a heteroarilo.

5 Como se usa en este documento, la nomenclatura alquilo, alcoxi, carbonilo, etc. se usan de la forma generalmente aceptada por los expertos en esta técnica. Por ejemplo, como se usa en este documento, alquilo se refiere a cadenas de carbono saturadas que contienen uno o múltiples carbonos; la cadena puede ser lineal o ramificada, o incluir porciones cíclicas o ser cíclica.

10 Cuando no se especifica el número de cualquier sustituyente determinado (por ejemplo, "haloalquilo"), puede haber presentes uno o múltiples sustituyentes. Por ejemplo, "haloalquilo" puede incluir uno o múltiples halógenos iguales o diferentes. Como ejemplo adicional, "alcoxi-(C₁₋₃)-fenilo" puede incluir uno o múltiples grupo alcoxi iguales o diferentes que contienen uno, dos o tres carbonos.

15 Cuando los sustituyentes mencionados tales como carboxi, o los sustituyentes representados por variables tales como W se incluyen por separado, encerrados entre paréntesis, pero no muestran un sufijo fuera del paréntesis que indique un valor numérico, y va seguido de sustituyentes que no se hallan entre paréntesis, por ejemplo, "alquil-(C₁₋₄)(W)(carboxi)", "W" y "carboxi" están directamente unidos al alquilo-(C₁₋₄).

20 Como se usa en este documento, "halógeno" o "haluro" se refiere a F, Cl, Br o I.

25 Como se usa en este documento, los pseudo-haluros son compuestos que se comportan sustancialmente de manera similar a los haluros. Estos compuestos pueden ser usados y tratados de la misma forma que los haluros (X⁻, en donde X es un halógeno tal como Cl o Br). Los pseudo-haluros incluyen, sin limitaciones, cianuro, cianato, isocianato, tiocianato, isotiocianato, selenocianato, trifluorometoxi y azida.

30 Como se usa en este documento, "haloalquilo" se refiere a un radical alquilo inferior en el que uno o múltiples átomos de hidrógeno están sustituidos con halógeno, incluidos, sin limitaciones, clorometilo, trifluorometilo, 1-cloro-2-fluoroetilo y similares.

35 Como se usa en este documento, "haloalcoxi" se refiere a RO-, en donde R es un grupo haloalquilo.

40 Como se usa en este documento, "sulfinilo" o "tionilo" se refiere a -S(O)-. Como se usa en este documento, "sulfonilo" o "sulfurilo" se refiere a -S(O)₂-. Como se usa en este documento, "sulfo" se refiere a -S(O)₂O-.

45 Como se usa en este documento, "carboxi" se refiere a un radical bivalente -C(O)O-.

Como se usa en este documento, "aminocarbonilo" se refiere a -C(O)NH₂.

50 Como se usa en este documento, "alquil-aminocarbonilo" se refiere a -C(O)NHR, en donde R es hidrógeno o alquilo, incluido alquilo inferior.

55 Como se usa en este documento, "dialquil-aminocarbonilo" se refiere, en este caso, a -C(O)NR'R, en donde R' y R se seleccionan, independientemente, de hidrógeno o alquilo, incluido alquilo inferior.

60 Como se usa en este documento, "carboxamida" se refiere a grupos de la fórmula -NR'COR.

65 Como se usa en este documento, "diaril-aminocarbonilo" se refiere a -C(O)NRR', en donde R y R' se seleccionan, independientemente, de arilo, incluido arilo inferior tal como fenilo.

70 Como se usa en este documento, "aralquil-aminocarbonilo" se refiere a -C(O)NRR', en donde una de R y R' es arilo, incluido arilo inferior tal como fenilo, y la otra de R y R' es alquilo, incluido alquilo inferior.

75 Como se usa en este documento, "aril-aminocarbonilo" se refiere a -C(O)NHR, en donde R es arilo, incluido arilo inferior tal como fenilo.

80 Como se usa en este documento, "alcoxicarbonilo" se refiere a -C(O)OR, en donde R es alquilo, incluido alquilo inferior.

85 Como se usa en este documento, "ariloxi-carbonilo" se refiere a -C(O)OR, en donde R arilo, incluido arilo inferior tal como fenilo.

90 Como se usa en este documento, "alcoxi" y "alquiltio" se refieren a RO- y RS-, en donde R es alquilo, incluido alquilo inferior.

95 Como se usa en este documento, "ariloxi" y "ariltio" se refieren a RO- y RS-, en donde R es arilo, incluido arilo inferior tal como fenilo.

ES 2 319 634 T3

Como se usa en este documento, "alquileno" se refiere a un grupo hidrocarburo lineal, ramificado o cíclico, en una realización, lineal o ramificado, alifático bivalente que, en ciertas realizaciones, tiene desde 1 hasta aproximadamente 20 átomos de carbono y, en otras realizaciones, 1 a 12 carbonos, incluido alquileno inferior. El grupo alquileno está sustituido opcionalmente con uno o múltiples "sustituyentes del grupo alquilo". Opcionalmente, a lo largo del grupo alquileno pueden estar insertados uno o múltiples átomos de oxígeno, azufre o nitrógeno sustituido o no sustituido, en donde el sustituyente de nitrógeno es alquilo, como se ha descrito anteriormente. Ejemplos de grupos alquileno incluyen metileno (-CH₂-), etileno (-CH₂CH₂-), propileno (-CH₂CH₂CH₂-), ciclohexileno (-C₆H₁₀-), metilendioxi (-O-CH₂-O-) y etilendioxi (-O-(CH₂)₂-O-). La expresión "etileno inferior" se refiere a grupos alquileno que tienen 1 a 6 carbonos. En ciertas realizaciones, los grupos alquileno son alquileno inferior, incluido alquileno de 1 a 3 átomos de carbono.

Como se usa en este documento, "alquenileno" se refiere a un grupo hidrocarburo lineal, ramificado o cíclico, en una realización, lineal o ramificado, alifático y divalente, que en ciertas realizaciones tiene desde 2 hasta 20 átomos de carbono y al menos un doble enlace, y en otras realizaciones desde 1 hasta 12 carbonos, incluido alquenileno inferior. El grupo alquenileno está sustituido opcionalmente con uno o múltiples "sustituyentes del grupo alquilo". Opcionalmente, a lo largo del grupo alquenileno pueden estar insertados uno o múltiples átomos de oxígeno, azufre o nitrógeno sustituido o no sustituido, en donde el sustituyente de nitrógeno es alquilo, como se ha descrito anteriormente. Ejemplos de grupos alquenileno incluyen -CH=CH-CH=CH- y -CH=CH-CH₂-. La expresión "alquenileno inferior" se refiere a grupos alquenileno que tienen 2 a 6 carbonos. En ciertas realizaciones, los grupos alquenileno son alquenileno inferior, incluido alquenileno de 3 a 4 átomos de carbono.

Como se usa en este documento, "alquinileno" se refiere a un grupo hidrocarburo lineal, ramificado o cíclico, en una realización, lineal o ramificado, alifático y divalente, que en ciertas realizaciones tiene desde 2 hasta 20 átomos de carbono y al menos un triple enlace, y en otras realizaciones desde 1 hasta 12 carbonos, incluido alquinileno inferior. El grupo alquinileno está sustituido opcionalmente con uno o múltiples "sustituyentes del grupo alquilo". Opcionalmente, a lo largo del grupo alquinileno pueden estar insertados uno o múltiples átomos de oxígeno, azufre o nitrógeno sustituido o no sustituido, en donde el sustituyente de nitrógeno es alquilo, como se ha descrito anteriormente. Los ejemplos de grupos alquinileno incluyen -C≡C-C≡C-, -C≡C-CH₂-. La expresión "alquinileno inferior" se refiere a grupos alquinileno que tienen 2 a 6 carbonos. En ciertas realizaciones, los grupos alquinileno son alquinileno inferior, incluido alquinileno de 3 a 4 átomos de carbono.

Como se usa en este documento, "alqu(en)(in)ileno" se refiere a un grupo hidrocarburo lineal, ramificado o cíclico, en una realización lineal o ramificado, alifático y divalente que, en ciertas realizaciones tiene desde 2 hasta 20 átomos de carbono y al menos un triple enlace y al menos un doble enlace; en otras realizaciones, 1 hasta 12 carbonos, incluido alqu(en)(in)ileno inferior. El grupo alqu(en)(in)ileno está sustituido opcionalmente con uno o múltiples "sustituyentes del grupo alquilo". Opcionalmente, a lo largo del grupo alqu(en)(in)ileno pueden estar insertados uno o múltiples átomos de oxígeno, azufre o nitrógeno sustituido o no sustituido, en donde el sustituyente de nitrógeno es alquilo, tal como se ha descrito anteriormente. Los ejemplos de grupos alqu(en)(in)ileno incluyen -C=C(CH₂)_n-C≡C-, en donde n es 1 ó 2. La expresión "alqu(en)(in)ileno inferior" se refiere a grupos alqu(en)(in)ileno que tienen hasta 6 carbonos. En ciertas realizaciones, los grupos alqu(en)(in)ileno son alqu(en)(in)ileno inferior, incluyendo un alqu(en)(in)ileno de 4 átomos de carbono.

Como se usa en este documento, "arileno" se refiere a un grupo aromático divalente, monocíclico o policíclico, en una realización monocíclico, que en ciertas realizaciones tiene desde 5 hasta aproximadamente 20 átomos de carbono y al menos un anillo aromático; en otras realizaciones, 5 a 12 carbonos, incluido arileno inferior. El grupo arileno está sustituido opcionalmente con uno o múltiples "sustituyentes del grupo alquilo". Opcionalmente, a lo largo del grupo arileno pueden estar insertados uno o múltiples átomos de oxígeno, azufre o nitrógeno sustituido o no sustituido, en donde el sustituyente de nitrógeno es alquilo, tal como se ha descrito anteriormente. Ejemplos de grupos arileno incluyen 1,2-, 1,3- y 1,4-fenileno. La expresión "arileno inferior" se refiere a grupos arileno que tienen 5 ó 6 carbonos. En ciertas realizaciones, los grupos arileno son arileno inferior.

Como se usa en este documento, "heteroarileno" se refiere a un sistema de anillo aromático divalente, monocíclico o policíclico, en una realización de aproximadamente 5 hasta aproximadamente 15 miembros, en donde uno o múltiples, o 1 a 3 de los átomos en el sistema de anillo es un heteroátomo, que es un elemento diferente de carbono, por ejemplo, átomos de nitrógeno, oxígeno y azufre. El grupo heteroarileno puede estar sustituido opcionalmente con uno o múltiples, o 1 a 3 sustituyentes del grupo arilo.

Como se usa en este documento, "alquilideno" se refiere a un grupo divalente tal como =CR'R", que está unido a un átomo de otro grupo, formando un doble enlace. Ejemplos de grupos alquilideno son metilideno (=CH₂) y etilideno (=CHCH₃). Como se usa en este documento, "aralquilideno" se refiere a un grupo alquilideno en el que R' o R" es un grupo arilo.

Como se usa en este documento, "amido" se refiere al grupo divalente -C(O)NH-. "Tioamido" se refiere al grupo divalente -C(S)NH-. "Oxiamido" se refiere al grupo divalente -OC(O)NH-. "Tiaamido" se refiere al grupo divalente -SC(O)NH-. "Ditiaamido" se refiere al grupo divalente -SC(S)NH-. "Ureido" se refiere al grupo divalente -HNC(O)NH-. "Tioureido" se refiere al grupo divalente -HNC(S)NH-.

Como se usa en este documento, "semicarbazida" se refiere a -NHC(O)NHNH-. "Carbazato" se refiere al grupo divalente -OC(O)NHNH-. "Isotiocarbazato" se refiere al grupo divalente -SC(O)NHNH-. "Tiocarbazato" se refiere al

ES 2 319 634 T3

grupo divalente -OC(S)NHNH-. “Sulfonil-hidrazida” se refiere al grupo -SO₂NHNH-. “Hidrazida” se refiere al grupo divalente -C(O)NHNH-. “Azo” se refiere al grupo divalente -N=N-. “Hidrazinilo” se refiere al grupo divalente -NH-NH-.

5 Como se usa en este documento, el término “aminoácido” se refiere a α -aminoácidos que son racémicos o de la configuración D o L. La designación “d” previa al nombre de un aminoácido (por ejemplo, dAla, dSer, dVal, etc.) se refiere al isómero D del aminoácido. La designación “dl” previa al nombre de un aminoácido (por ejemplo, dlAla) se refiere a una mezcla de los isómeros L y D del aminoácido.

10 Como se usa en este documento, cuando se especifica cualquier grupo particular tal como fenilo o piridilo, esto significa que el grupo está no sustituido o está sustituido. Cuando no se especifican, los sustituyentes son halo, haloalquilo inferior y alquilo inferior.

15 Como se usa en este documento, una enfermedad por proteína con alteraciones conformacionales (o una enfermedad de agregación de proteínas) se refiere a enfermedades asociadas con una proteína o polipéptido que tiene una conformación asociada a una enfermedad. Los métodos y recolecciones ofrecidos en este documento permiten la detección de un conformero asociado con una enfermedad que se debe detectar. Las enfermedades y las proteínas asociadas que exhiben dos o más conformaciones diferentes, en las que al menos una conformación es una proteína con una alteración conformacional incluyen, sin limitaciones, enfermedades amiloides y otras enfermedades neurodegenerativas conocidas por el experto en la técnica y expuestas más adelante.

20 Como se usa en este documento, separación celular por *cell sorting* se refiere a un ensayo en el que las células se separan y recuperan de la suspensión, en base a propiedades medidas en el análisis de citometría de flujo. La mayoría de los ensayos usados para análisis pueden servir como base para los experimentos de separación celular, con la condición evidente de que las puertas y regiones que definen la o las subpoblaciones que se deben separar no estén superpuestas. Los índices máximos de rendimiento son típicamente de 5000 células/segundo (18 x 10⁶ células/hora). La velocidad de recolección de la o las poblaciones separadas depende fundamentalmente del estado de las células y del porcentaje de reactividad.

30 Como se usa en este documento, las abreviaturas de todos los grupos protectores, aminoácidos y de otros compuestos son conformes, a menos que se indique lo contrario, a sus usos habituales, a las abreviaturas reconocidas y a la Comisión de Nomenclatura Bioquímica de la IUPAC-IUB (véase, *Biochem.* 1972, 11:942). Por ejemplo, DMF = *N,N*-dimetilformamida; DMAc = *N,N*-dimetilacetamida; THF = tetrahidrofurano; TRIS = tris(hidroximetil)aminoetano; SSPE = solución tampón de solución salina fisiológica-fosfato sódico-EDTA; EDTA = ácido etilendiamino-tetraacético; SDS = dodecilsulfato sódico.

B. Matrices de compuestos

40 Se proporcionan matrices de gradiente de compuestos de captura, que se unen selectivamente a las biomoléculas presentes en las muestras, tal como biomoléculas en especial, aunque no de forma exclusiva, un lisado celular o polipéptidos traducidos *in vitro* a partir de un lisado celular. Cada locus en la matriz puede fijarse a grupos o clases específicas de biopolímeros, y está diseñado para unirse de manera covalente o íntima (por ejemplo, suficiente para resistir un análisis de espectrometría de masa) con un subconjunto de todas las biomoléculas en la muestra. Por ejemplo, una muestra puede contener millares de miembros, por ejemplo, un lisado celular. Las matrices de compuestos 45 permiten una selectividad suficiente, de modo que, por ejemplo, alrededor de 10-20 de los componentes de la muestra se unen a cada miembro de la recolección. El número exacto es lo suficientemente reducido como para que los análisis convencionales los identifiquen, por lo general, en una sola etapa tal como una espectrometría de masa.

50 Las matrices permiten establecer un enfoque integral para el análisis del proteoma, incluido el de las proteínas con modificaciones post-traduccionales, y de otras biomoléculas. Los patrones de proteínas y otras biomoléculas constituyen el punto de partida para los análisis que usan estas matrices, en cualquier caso, más que los ácidos nucleicos y el genoma (de abajo arriba). Las matrices se pueden usar para evaluar los componentes biomoleculares de una muestra, por ejemplo, una muestra biológica, para identificar componentes específicos de un fenotipo particular, tal como un estado patológico, para identificar la función estructural, las vías bioquímicas y los mecanismos de acción. 55 Las matrices y sus métodos de uso permiten realizar un análisis imparcial y no sesgado de las biomoléculas, puesto que los métodos no necesariamente evalúan clases específicas de dianas, sino que, en su lugar, se detectan o identifican variaciones presentes en las muestras. Las matrices permiten separar los componentes de una mezcla compleja de biomoléculas (es decir, una mezcla de 50, 100, 500, 1000, 2000 y más) en loci discretos que contienen números reducidos, típicamente una reducción de 10%, 50% o mayor de la complejidad o, lo que es igual, aproximadamente 60 1 a 50 biomoléculas por locus en una matriz, de manera que es posible analizar los componentes en cada punto, por ejemplo, por análisis de espectrometría de masa solamente, o combinado con otros análisis. En algunas realizaciones, por ejemplo, para análisis fenotípicos, la homogeneidad de la muestra inicial, tal como células, puede ser importante. Para aportar homogeneidad, se comparan células con diferentes fenotipos, por ejemplo, enfermas frente a sanas, del mismo individuo. En este documento, se proporcionan los métodos para llevar a cabo esta acción.

65 Gracias a la estructura de los restos en los loci de las matrices, éstas pueden ser usadas para detectar cambios estructurales tales como los derivados del procesamiento post-traduccionales de proteínas, y se les puede utilizar para detectar cambios en las proteínas de membrana que intervienen en los procesos más fundamentales tales como la

ES 2 319 634 T3

transducción de señales, canales de iones, receptores para la interacción de ligandos y las interacciones entre células. Cuando las células enferman suelen producirse cambios asociados con la enfermedad, tales como transformaciones, en las proteínas de membrana.

5 Las matrices contienen conjuntos de restos en cada locus. Los restos en cada locus incluyen un grupo X y un grupo Y. Por lo general, los miembros de cada conjunto difieren en al menos un grupo de propiedad y, generalmente, en dos, tres o más entre los distintos loci. Como se señala en este documento, las diferencias comprenden un gradiente, típicamente un gradiente bidimensional, basado en cambios graduales de propiedades de X e Y en cada locus, y a través de los ejes X e Y en la matriz.

10 En la ejecución de los métodos, las matrices se hacen contactar con una muestra o componentes parcial o totalmente purificados de la misma para producir la unión de las biomoléculas con los compuestos de captura de la matriz. La matriz resultante se trata opcionalmente con un reactivo que escinde de forma específica los polímeros fijados, por ejemplo, una proteasa, y se somete a análisis, en especial un análisis de espectrometría de masa, para identificar los componentes de las biomoléculas unidas en cada locus. Una vez determinado el peso molecular de una biomolécula, tal como una proteína o porción de la misma, es posible identificar la biomolécula. Los métodos de identificación incluyen la comparación de los pesos moleculares con bases de datos, por ejemplo, bases de datos de proteínas que incluyen fragmentos de proteasa y sus pesos moleculares.

20 Los grupos X e Y son grupos funcionales que aportan reactividad, selectividad y propiedades de separación, dependiendo de la especificidad de separación y de análisis requerida (que depende de la complejidad de la mezcla que se debe analizar). Por lo general, los loci en las matrices incluyen al menos dos grupos funcionales (funciones), seleccionados de: una función de reactividad (X), que se une con los biopolímeros de forma covalente o con una k_a alta (por lo general, mayor que aproximadamente 10^9 , 10^{10} , 10^{12} litros/mol, y/o de manera que la unión es sustancialmente irreversible o estable bajo las condiciones de los análisis de espectrometría de masa tales como las condiciones de MALDI-MS); y una función de selectividad (Y) que, debido a sus interacciones no covalentes, altera y, por lo general, aumenta la especificidad de la función de reactividad.

30 En general, la función de reactividad es un grupo reactivo que interactúa específicamente, de manera típica, covalentemente o con una elevada afinidad de unión (k_a), con biomoléculas particulares tales como proteínas o porciones de las mismas; y la otra funcionalidad, es decir, las funciones de selectividad, aumentan típicamente la especificidad de la función de reactividad. Por lo general, la función reactiva interactúa de manera covalente con grupos en una biomolécula particular, tales como los grupos amina en la superficie de una proteína. La función de reactividad interactúa con las biomoléculas para formar un enlace covalente o un enlace no covalente que es estable bajo las condiciones de análisis, generalmente con una k_a mayor que 10^9 litros/mol, o mayor que 10^{10} litros/mol. Las condiciones de análisis incluyen, sin limitaciones, análisis de espectrometría de masa tal como espectrometría de masas en tándem a tiempos de vuelo con ionización por desorción láser asistida por matrices (MALDI-TOF). La función de selectividad influye sobre el tipo de biomoléculas que pueden interactuar con la función de reactividad a través de una interacción no covalente. La función de selectividad altera la especificidad para los grupos particulares, reduciendo por lo general el número de proteínas o biomoléculas unidas a un locus, de forma que se pueden separar, entonces, las proteínas por espectrometría de masa.

45 Entre los restos mencionados en este documento están incluidos los que pueden ser clasificados en al menos dos conjuntos: uno para reacciones en solución acuosa (por ejemplo, para la reacción con biomoléculas hidrófilas), y el otro para la reacción en disolventes orgánicos (por ejemplo, cloroformo) (por ejemplo, para la reacción de biomoléculas hidrófobas). Así, en ciertas realizaciones, las presentes matrices discriminan entre biomoléculas hidrófilas e hidrófobas, incluidas, sin limitaciones, las proteínas, y permiten el análisis de ambas clases de biomoléculas.

C. Componentes de las matrices

50 Se proporcionan matrices con loci que contienen restos unidos (también designados como agentes de captura). Las matrices incluyen un núcleo "Z", la superficie sólida, que presenta una o múltiples funciones de reactividad "X" y, opcionalmente, una o múltiples funciones de selectividad "Y", que están dispuestas en los loci. La manera particular en que se presentan las funciones sobre la superficie sólida (Z) es cuestión de lección de diseño, pero se selecciona de forma que los loci resultantes capturen biomoléculas, en especial proteínas, con una especificidad suficiente, de modo covalente o con enlaces suficientemente estables o afines, para permitir el análisis, por ejemplo, por espectrometría de masa, incluido el análisis de espectrometría de masa MALDI, de manera que al menos una porción de las biomoléculas fijadas permanece unida (generalmente, una afinidad de unión de 10^9 , 10^{10} , 10^{11} litros/mol o mayor, o una K_{eq} de 10^9 , 10^{10} , 10^{11} , 10^{12} o mayor).

60 X, la funcionalidad de reactividad, se selecciona como cualquier elemento capaz de formar un enlace covalente o un enlace de alta afinidad, que sea estable bajo las condiciones del análisis de espectrometría de masa, en particular el análisis MALDI. La funcionalidad de selectividad, Y, es un grupo que "vigila" la topología de la proteína n torno a los sitios de unión de la reactividad, y que actúa seleccionando grupos particulares de biomoléculas de entre las que tienen un grupo de reactividad capaces de formar un enlace covalente (o un enlace de alta afinidad). Por ejemplo, un grupo de selectividad puede provocar impedimento estérico, o permitir la unión específica a un epítopo, o cualquier paso intermedio. Puede ser un sustrato para un medicamento, un lípido o un péptido. Selecciona en entorno de los grupos con los que interactúa la función de reactividad. La funcionalidad de selectividad, Y, puede ser aquella por la que un

ES 2 319 634 T3

compuesto de captura forma un enlace covalente con una biomolécula en una mezcla, o interactúa con alta estabilidad, de manera que la afinidad de la unión del compuesto de captura con la biomolécula a través de la funcionalidad de reactividad, en presencia de la funcionalidad de selectividad, sea al menos 10 ó 100 veces mayor que en ausencia de la funcionalidad de selectividad.

5

Las funciones de reactividad (“X”) confieren a los compuestos la capacidad de unirse ya sea de forma covalente o con alta afinidad (mayor que 10^9 , por lo general, mayor que 10^{10} ó 10^{11} litros/mol, típicamente mayor que un anticuerpo monoclonal y típicamente, estable al análisis de espectrometría de masa tal como MALDI-MS) a una biomolécula, en particular proteínas, incluidos los grupos funcionales que existen en ellas, que incluyen grupos agregados de forma post-traduccional. Generalmente, la unión es covalente o tiene tal afinidad que es estable bajo las condiciones de análisis tales como de espectro de masa, incluido el análisis MALDI-TOF. En este documento figuran ejemplos de los grupos.

15

En los compuestos ofrecidos en este documento, X es un resto que fija una cadena lateral de aminoácidos.

X es un grupo que reacciona o interactúa con funcionalidades sobre la superficie de una proteína para formar enlaces covalentes o no covalentes con alta afinidad. Hay disponible una extensa selección de diferentes grupos funcionales para que X reaccione con una proteína. X puede actuar ya sea como un nucleófilo o un electrófilo para formar enlaces covalentes tras la reacción con los residuos aminoácidos en la superficie de una proteína. Ejemplos de reactivos que se unen covalentemente con cadenas laterales de aminoácidos incluyen, sin limitación, grupos protectores para restos hidroxilo, carboxilo, amino, amida y tiol, incluidos, por ejemplo, los descritos en T.W. Greene y P.G.M. Wuts, “Protective Groups in Organic Synthesis”, 3ª ed. (1999, Wiley Interscience); grupos foto-reativos, parejas de Diels Alder (es decir, un dieno en un lado y un doble enlace sencillo en el otro lado). Otros grupos de X incluyen los que se describen en la solicitud de EE.UU. en tramitación de la Serie No. 10/197.954 (véase la Figura 16).

25

Las funciones de selectividad (“Y”) sirve para modular la función de reactividad al reducir el número de grupos a los que se fijan las funciones de reactividad, tal como por impedimento estérico u otras interacciones. Es un grupo que modifica las propiedades estéricas y/o electrónicas (por ejemplo, efectos mesoméricos, inductivos), así como las propiedades de afinidad resultantes del compuesto de captura. Las funciones de selectividad incluyen cualquier grupo funcional que eleve la selectividad del grupo de reactividad, de modo que se una a menos biomoléculas diferentes que en ausencia de la función de selectividad, o se una con mayor afinidad a las biomoléculas que en su ausencia. En los compuestos de captura citados en este documento, Y puede variar ampliamente, dependiendo del objetivo perseguido en relación con el impedimento estérico y factores electrónicos, en la medida que se relacionan con la modulación de la reactividad del enlace escindible L, si está presente, y la funcionalidad reactiva X. Por ejemplo, se puede seleccionar una función de reactividad X para que se fije a los grupos amina en las proteínas; la función de selectividad se puede seleccionar para garantizar que se pueda acceder sólo a los grupos expuesto en la superficie. La función de selectividad es tal que los compuestos se unen o reaccionan con (a través de la función de reactividad) menos biomoléculas diferentes cuando forma parte de una molécula que cuando está ausente y/o los compuestos se unen con mayor especificidad y una afinidad más alta. La función de selectividad puede estar enlazada directamente con un compuesto o puede estar a través de un enlazador tal como CH_2CO_2 o $\text{CH}_2\text{-O-(CH}_2\text{)}_n\text{-O}$, en donde n es un número entero desde 1 hasta 12, o 1 a 6, o 2 a 4. Otros grupos de Y incluyen los descritos en la solicitud de EE.UU. en tramitación de la Serie No. 10/197.954 (véase la Figura 17).

45

En ciertas realizaciones, cada Y es, independientemente, un grupo que modifica las propiedades de afinidad y/o las propiedades estéricas y/o electrónicas (por ejemplo, efectos mesoméricos, inductivos) del compuesto de captura resultante. Por ejemplo, Y, en ciertas realizaciones, se selecciona de análogos e inhibidores de ATP; péptidos y análogos de péptidos; polietilenglicol (PEG); ésteres activados de aminoácidos, aislados o dentro de un péptido; citocromo C; y grupos tritilo hidrófilos.

50

En otras realizaciones, Y es un resto de molécula pequeña, un producto natural, un agonista o antagonista de proteína, un péptido o un anticuerpo. En otra realización, Y es un compuesto hidrófilo o proteína (por ejemplo, PEG o éter tritilo), un compuesto hidrófobo o proteína (por ejemplo, aromáticos polares, lípidos, glicolípidos, fosfolípidos, oligosacáridos), un grupo con carga positiva o negativa, una molécula de tamaño pequeño, un compuesto farmacéutico o una biomolécula que crea estructuras secundarias o terciarias definidas.

55

En la descripción siguiente se ofrece una descripción y discusión más detalladas de cada funcionalidad.

Ejemplos de compuestos y descripciones de X e Y

60

Se entiende que para la matrices de gradiente propuestas en este documento, “Z” es un soporte sólido al que se unen o fijan de cualquier otra forma, directa o indirectamente, X e Y (y cualquier otro grupo).

65

En otra realización, los compuestos usados en los métodos ofrecidos en este documento tienen la fórmula



ES 2 319 634 T3

en donde Q es un oligonucleótido o análogo de oligonucleótido de cadena sencilla, no protegido o adecuadamente protegido (por ejemplo, PNA) de hasta 50 bloques de construcción, que es capaz de hibridar con una molécula de ácido nucleico de cadena sencilla y complementaria para las bases;

5 X es un grupo funcional que interactúa con y/o reacciona con funcionalidades en la superficie de una biomolécula, incluida, sin limitaciones, una proteína, para formar enlaces covalentes o enlaces estables bajo las condiciones del análisis de espectrometría de masa, según se definen en este documento; e

10 Y es un grupo funcional que interactúa con y/o reacciona imponiendo una selectividad única por medio de la introducción de funcionalidades que interactúan de forma no covalente con las proteínas diana.

En otra realización, los compuestos que se usan en los métodos propuestos en este documento tienen la fórmula:



20 en donde Q es un oligonucleótido o análogo de oligonucleótido de cadena sencilla, no protegido o adecuadamente protegido (por ejemplo, ácido nucleico peptídico (PNA)) de hasta 50 bloques de construcción, que es capaz de hibridar con una molécula de ácido nucleico de cadena sencilla y complementaria para las bases;

Z es un soporte sólido;

25 X es un grupo funcional que interactúa con y/o reacciona con funcionalidades en la superficie de una biomolécula, incluida, sin limitaciones, una proteína, para formar enlaces covalentes o enlaces estables bajo las condiciones del análisis de espectrometría de masa; e

30 Y es un grupo funcional que interactúa con y/o reacciona imponiendo una selectividad única, por medio de la introducción de funcionalidades que interactúan de forma no covalente con las proteínas diana.

35 En otra realización, Y se selecciona de análogos e inhibidores de ATP; péptidos y análogos de péptidos; polietilenglicol (PEG); ésteres activados de aminoácidos, aislados o dentro de un péptido; citocromo C; y grupos tritilo hidrófilos.

En otra realización, los compuestos que se usan en los métodos propuestos en este documento tienen las fórmulas



45 en donde Q, Q', Z, X e Y son como se han definido anteriormente; m es un número entero de 1 hasta 100, en una realización, 1 hasta 10, en otra realización, 1 hasta 3, 4 ó 5; y n es un número entero desde 1 hasta 100, en una realización, 1 hasta 10, en otra realización, 1 hasta 3, 4 ó 5.

50 En otra realización, X es un medicamento farmacéutico. Estos compuestos pueden ser usados como moléculas de transporte de medicamentos. Los compuestos de estas realizaciones se pueden utilizar en el rastreo de medicamentos por medio de biomoléculas de captura que incluyen, sin limitaciones, proteínas que se unen al medicamento farmacéutico. Se identifican las mutaciones en las biomoléculas que interfieren con la fijación al medicamento farmacéutico, determinando de este modo posibles mecanismos de resistencias medicamentosas. Véase, por ejemplo, Hessler *et al.* (9-11 de noviembre, 2001) *Ninth Foresight Conference on Molecular Nanotechnology* (Abstract) (<http://www.foresight.org/Conferences/MNT9/Abstracts/Hessler/>).

60 En realizaciones adicionales, los compuestos que se utilizan en los métodos propuestos en este documento son los de las fórmulas anteriores, en las que Z es un soporte insoluble o un sustrato, incluida una perla, que incluye, sin limitaciones, perlas polímeras, magnéticas, coloreadas, etiquetadas con R_f, etc. que, en ciertas realizaciones, están unidas a X a través de un primer espaciador opcional, y un enlace escindible; y están unidas a Q por medio de un segundo espaciador opcional. En estas realizaciones, la densidad del polímero que se debe analizar y, por consiguiente, la intensidad de la señal del análisis subsiguiente, aumentan en relación con las realizaciones en las que Z es un grupo bivalente o multivalente. En estas realizaciones, se empleará en los métodos propuestos una matriz apropiada de oligonucleótidos monocatenarios o análogos de oligonucleótidos que son complementarios para Q.

65 En realizaciones adicionales de este documento, Z es como se ha definido anteriormente, y es un enlazador no escindible. Los compuestos de estas realizaciones son generalmente de utilidad en realizaciones en las que no se requiere la escisión del biopolímero del resto Q antes o durante el análisis del biopolímero, incluido el análisis de las

interacciones biopolímero-biopolímero tales como interacciones de proteína-proteína, o interacciones del biopolímero con moléculas pequeñas (por ejemplo, medicamento o candidato a medicamento), incluidas las interacciones entre proteínas y moléculas pequeñas (por ejemplo, medicamento o candidato a medicamento).

5 En todas las realizaciones de esta invención, los compuestos que se usan pueden ser clasificados en al menos dos grupos: uno para reacciones en solución acuosa (por ejemplo, para la reacción con biomoléculas hidrófilas), y el otro para reacciones en disolventes orgánicos (por ejemplo, cloroformo) (por ejemplo, para reacciones con biomoléculas hidrófobas). Así, en ciertas realizaciones, los compuestos propuestos en este documento discriminan entre biomoléculas hidrófilas e hidrófobas, incluyendo, sin limitaciones, proteínas, y permiten el análisis de ambas clases de biomoléculas.

Las variables Y, X y Z, con referencia a enlazadores y espaciadores, se describen de manera más detallada a continuación. Las siguientes descripciones son aplicables a todas las fórmulas anteriores.

15 2. El resto Z

A los efectos de este documento, Z es un soporte sólido. A continuación, se describen enlazadores y espaciadores para unir restos X e Y con Z.

20 a. Z es un resto escindible

En ciertas realizaciones de esta invención, en los compuestos que se usan en los métodos propuestos, Z es un resto que es escindible antes de o durante el análisis de la biomolécula, incluido el análisis espectral de masa, sin alterar la estructura química de la biomolécula, incluida, sin limitaciones, una proteína. En una realización, los métodos ofrecidos en este documento incluyen métodos de análisis espectral de masa de las biomoléculas, incluidas las proteínas, que se muestran en formato direccionable. En ciertas realizaciones, el formato es una matriz de oligonucleótidos de cadena sencilla que son complementarios a las porciones de oligonucleótidos, o porciones de análogos de oligonucleótidos (Q) de los compuestos. En estas realizaciones, Z es un grupo que es (i) estable bajo las condiciones de reacción necesarias para la reacción de los compuestos proporcionados en la invención con la biomolécula, por ejemplo, una proteína; (ii) estable bajo las condiciones para la hibridación del resto Q con los oligonucleótidos de cadena sencilla, y (iii), escindibles antes de o durante el análisis de la biomolécula.

En una realización, Z es un grupo foto-escindible que se escinde mediante un láser usado en la espectrometría de masa MALDI-TOF. En otra realización, Z es un grupo ácido-lábil que se escinde tras la aplicación de una matriz a los conjugados hibridados de compuesto-biomolécula, o por exposición a ácidos (por ejemplo, ácido trifluoroacético o clorhídrico) en forma de vapor o líquida, antes del análisis. En esta realización, la matriz mantiene la integridad espacial de la matriz, permitiendo el análisis direccionable de la matriz.

40 b. Z es un resto no escindible

En otras realizaciones de este documento, los compuestos que se utilizan en los métodos propuestos en esta invención tienen un resto Z que no es escindible bajo las condiciones usadas para el análisis de biomoléculas, incluyendo, sin limitaciones, espectrometría de masa tal como la espectrometría de masas en tándem a tiempos de vuelo con ionización por desorción láser asistida por matrices (MALDI-TOF). Los compuestos de estas realizaciones son útiles, por ejemplo, en los métodos propuestos en esta invención para determinar interacciones biomolécula-biomolécula, incluidas las de proteína-proteína, y para determinar las interacciones de biomolécula-molécula pequeña, incluyendo las de proteína-medicamento o proteína-candidato a medicamento. En estas realizaciones, no es necesario escindir el grupo Z para el análisis.

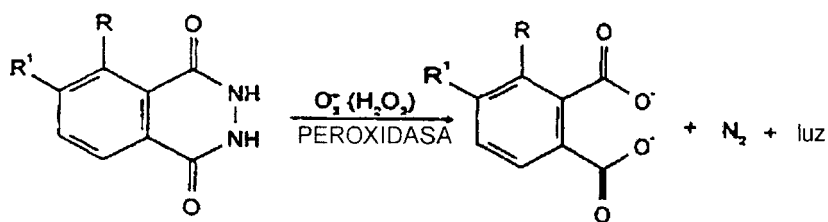
50 c. Restos Z divalentes o multivalentes

En una realización, Z es un grupo divalente o multivalente, escindible o no escindible, que contiene menos de 50, o menos de 20 miembros, y que se selecciona de alquileo de cadena lineal o ramificada, alquilenilo de cadena lineal o ramificada, alquilenilo de cadena lineal o ramificada, alquilenoxi de cadena lineal o ramificada, alquilentio de cadena lineal o ramificada, alquilenarbonilo de cadena lineal o ramificada, alquilenamino de cadena lineal o ramificada, cicloalquileo, cicloalquilenilo, cicloalquilenilo, cicloalquilenoxi, cicloalquilentio, cicloalquilen-carbonilo, cicloalquilen-amino, heterociclicileno, arileno, arilenoxi, arilentio, arilencarbonilo, arilenamino, heteroarileno, heteroarilenoxi, heteroarilentio, heteroarilencarbonilo, heteroarilenamino, oxi, tio, carbonilo, carboniloxi, éster, amino, amido, fosfino, óxido de fosfina, fosforamido, fosfinamido, sulfonamido, sulfonilo, sulfóxido, carbamato, ureido y sus combinaciones, y que está sustituido opcionalmente con uno o múltiples, incluyendo uno a cuatro, sustituyentes seleccionados, de forma independiente, de R¹⁵;

cada R¹⁵ es, independientemente, un grupo que modifica las propiedades estéricas y/o electrónicas (por ejemplo, efectos mesoméricos, inductivos) de Z; en una realización, R¹⁵ es un grupo que es un componente de un sistema luminiscente, incluyendo fluorescente, fosforescente, quimioluminiscente y bioluminiscente, o es un grupo que puede ser detectado en un ensayo colorimétrico.

Los grupos fluorescentes, colorimétricos y fosforescentes son bien conocidos por los expertos en la técnica (véanse, por ejemplo, la Patente de EE.UU. N° 6.247.337; Sapan *et al.* (1999) *Biotechnol. Appl. Biochem.* 29 (parte 2):99-108; Sittampalam *et al.* (1997), *Curr. Opin. Chem. Biol.* 1(3):384-91; Lakowicz, J.R., *Principles of Fluorescence Spectroscopy*, Nueva York: Plenum Press (1983); Herman, B., *Resonance Energy Transfer Microscopy*, en: *Fluorescence Microscopy of Living Cells in Culture*, Parte B, *Methods in Cell Biology*, vol. 30, ed. Taylor, D.L. y Wang, Y.-L., San Diego, Academic Press (1989), págs. 219-243; Turro, N.J., *Modern Molecular Photochemistry*, Menlo Park: Benjamin/Cummings Publishing Col., Inc. (1978), págs. 296-361 y Catálogo de Molecular Probes (1997), OR, EE.UU.). Los restos fluorescentes incluyen, sin estar limitados a ellos, 1- y 2-aminonaftaleno, p,p'-diaminoestilbenos, pirenos, sales de fenantridina cuaternaria, 9-aminoacridinas, p,p'-diaminobenzofenona iminas, antracenos, oxacarbocianina, merocianina, 3-aminoequilenina, perileno, bis-benzoxazol, bis-p-oxazolil-benceno, 1,2-benzofenazina, retinol, sales de bis-3-aminopiridinio, helebrigenina, tetraciclina, esterofenol, benzimidazol-fenilamina, 2-oxo-3-cromeno, indol, xanteno, 7-hidroxicumarina, fenoxazina, calicilato, estrofantidina, porfirinas, triarilmetanos y flavina. Los compuestos fluorescentes que tienen funcionalidades para unirse a un compuesto propuesto de esta invención, o que puede ser modificado para incorporar tales funcionalidades, incluyen, por ejemplo: cloruro de dansilo, fluoreceínas tales como 3,6-dihidroxi-9-fenil-xanthidrol; isotiocianato de rodamina; N-fenil-1-amino-8-sulfonato-naftaleno; N-fenil-2-amino-6-sulfonato-naftaleno; ácido 4-acetamido-4-isotiocianato-estilbeno-2,2'-disulfónico; ácido pireno-3-sulfónico; 6-sulfonato de 2-toluidino-naftaleno; 6-sulfonato de N-fenil-N-metil-2-aminonaftaleno; bromuro de etidio; estebrina; auromina-0,2-(9'-antroil)palmitato; dansil fosfatidil-etanolamina; N,N'-dioctadecil oxacarbocianina; N,N'-dihexil oxacarbocianina; merocianina, estearato de 4-(3'-pirenilo); d-3-aminodesoxi-equinelina; estearato de 12-(9'-antroilo); 2-metilantraceno; 9-vinilantraceno; 2,2'-(viniliden-p-fenil)bis-benzoxazol; p-bis(2-(4-metil-5-fenil-oxazolil))benceno; 6-dimetilamino-1,2-benzofenazina; retinol; diyoduro de bis-(3'aminopiridinio)-1,10-decadiilo; sulfonaftil-hidrazona de helibrienina; cloro-tetraciclina; N-(7-imetilamino-4-metil-2-oxo-3-cromenil)maleimida; N-(p-(2-benzimidazolil)-fenil)maleimida; N-(4-fluorantil)maleimida; ácido bis- (homovanílico); resazarina; 4-cloro-7-nitro-2,1,3-benzooxadiazol; merocianina 540; resorufina; rosa bengala; y 2,4-difenil-3(2H)-furanona. La compañía química SIGMA (San Luis, Mo.), Molecular Probes, R&D Systems (Minneapolis, Minn.), Pharmacia LKB Biotechnology (Piscataway, N.J.), CLONTECH Laboratories, Inc. (Palo Alto, CA.), Chem Genes Corp., Aldrich Chemical Company (Milwaukee, Wis.), Glen Research, Inc., GIBCO BRL Life Technologies, Inc. (Gaithersburg, Md.), Fluka Chemica-Biochemika Analytika (Fluka Chemie AG, Buchs, Suiza) y Applied Biosystems (Foster City, CA.), así como otras fuentes comerciales conocidas por el experto en la técnica comercializan numerosas etiquetas fluorescentes.

Los grupos quimioluminiscentes previstos para ser usados en este documento incluyen cualquier componente de sistemas generadores de luz, catalizados por una peroxidasa y que requieren un anión superóxido (O_2^-) (y/o peróxido de hidrógeno (H_2O_2)) (véase, por ejemplo, Musiani *et al.* (1998) *Histol. Histopathol.* 13(1):243-8). Los sistemas generadores de luz incluyen, pero sin estar limitados a ellos, luminol, isoluminol, fluoróforo de perioxalato, éster de actidinio, lucigenina, dioxietanos, ésteres oxalato, acridan, hemina, ésteres indoxilo incluidos los ésteres 3-O-indoxilo, derivados de naftaleno tales como hidrazida del ácido 7-dimetilamino-naftaleno-1,2-dicarbónico y análogos de ciperidín-luciferina, incluidos 2-metil-6-[p-metoxifenil]-3,7-dihidroimidazol-[1,2- α]-pirazin-3-ona, 2-metil-6-fenil-3,7-dihidroimidazol[1,2- α]-pirazin-3-ona, y 2-metil-6-[p-[2-3-carboxilato sódico-4-(6-hidroxi-3-xantenon-9-il)-fenil-tio-ureileno]etilen-oxi]fenil-3,7-dihidroimidazol[1,2- α]pirazin-3-ona. En otras realizaciones, los restos quimioluminiscentes previstos para ser usados en esta invención incluyen, sin limitaciones, luminol, isoluminol, N-(4-aminobutil)-N-etil-isoluminol (ABEI), N-(4-aminobutil)-N-metil-isoluminol (ABMI), que tienen las siguientes estructuras e intervienen en las siguientes reacciones:



en las que luminol está representado, cuando R es NH_2 y R^1 es H; isoluminol, cuando R es H y R^1 es NH_2 ; para ABEI ((6-[N-(4-aminobutil)-N-etilamino]-2,3-dihidroftalazina-1-4-diona), cuando R es H y R^1 es $C_2H_5-N-(CH_2)_4NH_2$; y para ABMI ((6-[N-(4-aminobutil)-N-metilamino]-2,3-dihidroftalazina-1-4-diona), cuando R es H y R^1 es $CH_3-N-(CH_2)_4NH_2$).

Los grupos bioluminiscentes para ser usados en esta invención incluyen parejas de luciferasa/luciferina, incluyendo luciferasa de luciérnaga [*Photinus pyralis*], el sistema *Aequorin* (es decir, la fotoproteína purificada de medusa, aequorina). Se han estudiado y caracterizado correctamente muchas luciferasas y sustratos, y se encuentran disponibles comercialmente (por ejemplo, la luciferasa de luciérnaga está comercializada por Sigma, San Luis, MO. y Boehringer Mannheim Biochemicals, Indianápolis, IN; luciferasa de luciérnaga producida de forma recombinante y otros reactivos basados en este gen o para ser usados con esta proteína están comercializados por Promega Corporation, Madison, WI; la fotoproteína aequorina luciferasa de medusa y la luciferasa de *Renilla* están disponibles

comercialmente por Sealite Sciences, Bogart, GA; celenterazina, el sustrato de origen natural para estas luciferasas, la comercializa Molecular Probes, Eugene, OR]. Otros sistemas bioluminiscentes incluyen sistemas de crustáceos, en especial de *Cypridina* (*Vargula*); los sistemas generadores de bioluminiscencia de insectos que incluyen luciérnagas, escarabajos de resorte y otros sistemas de insectos; sistemas bacterianos; sistemas generadores de bioluminiscencia de dinoflagelados; sistemas de moluscos tales como *Latia* y *Pholas*; gusanos de tierra y otros anélidos; gusanos brillantes; sistemas del gusano marino *Polichaeta*; el escarabajo del ferrocarril sudamericano; peces (es decir, los hallados en especies de *Aristostomias* tales como *A. scintillans* (véase, por ejemplo, O'Day *et al.* (1974) *Vision Res.* 14:545-550), *Pachystomias* y *Malacosteus*, tales como *M. niger*; los emisores azul/verde incluyen *cycithone*, *myctophids*, pez hacha (*agyropelecus*), *vinciguerria*, *howella*, *florenciella* y *Chauliodus*); y proteínas fluorescentes, incluidas las proteínas fluorescentes verdes (es decir, GFPs, incluidas las de *Renilla* y de *Ptilosarcus*), rojas y azules (es decir, BFPs, incluidas las de *Vibrio fischeri*, *Vibrio harveyi* o *Photobacterium phosphoreum*) (incluida la luciferasa de *Renilla mulleri*, luciferasa de especies de *Gaussia* y luciferasa de especies de *Pleuromamma*) y ficobilinas.

i. Restos Z divalentes o multivalentes escindibles

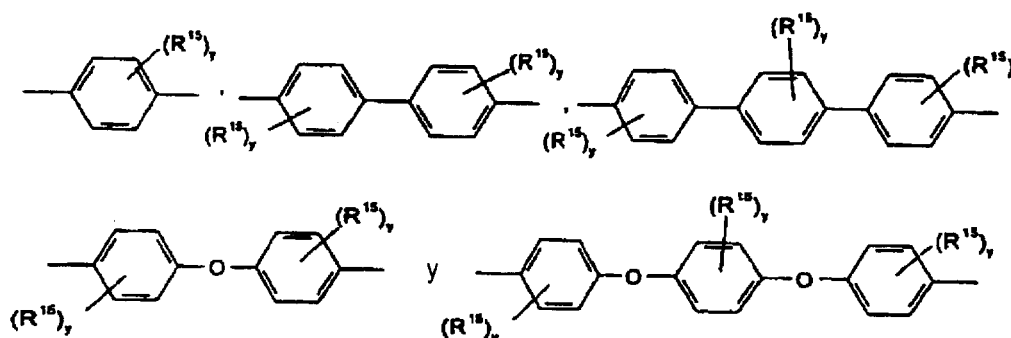
En una realización, Z es un resto divalente o multivalente escindible, y tiene la fórmula:



en la que S^1 y S^2 son restos espaciadores; t y b son, independientemente, 0 ó 1; M es un resto central que posee dos o más puntos de fijación (es decir, divalente o de valencia mayor); en ciertas realizaciones, presenta dos a seis puntos de fijación (es decir, divalente a hexavalente); en otras realizaciones, 2, 3, 4 ó 5 puntos de fijación (es decir, divalente, trivalente, tetravalente o pentavalente); R^{15} es un grupo funcional que modifica las propiedades estéricas y/o electrónicas (por ejemplo, efectos mesoméricos, inductivos) de M; a es 0 a 4, en ciertas realizaciones, 0, 1 ó 2; y L es un enlace que es escindible antes de o durante el análisis, incluido el análisis espectral de masa, de una biomolécula, sin alterar la estructura química de ésta, tal como una proteína.

En uno o en los dos lados del resto central M de los compuestos puede haber una región espaciadora S^1 y/o S^2 , reduciendo de este modo el impedimento estérico en reacciones con la superficie de grandes moléculas de proteínas (S^2), o facilitando la hibridación con secuencias complementarias (S^1). Para realizaciones en las que la proteína y la secuencia complementaria poseen un bajo impedimento estérico, puede no ser necesario un espaciador. En determinadas realizaciones, el impedimento estérico puede potenciar la selectividad. Esta selectividad aumentada se puede lograr por la presencia de uno o más sustituyentes voluminosos R^{15} , unido a M, o por la selección de las moléculas espaciadoras adecuadas para S^1 y/o S^2 .

En ciertas realizaciones, S^1 y S^2 se seleccionan, independientemente, de $-(CH_2)_r-$, $-(CH_2O)_r-$, $-(CH_2CH_2-O)_r-$, $-(NH(CH_2)_r-C(=O))_s-$, $-(NH-CH(R^{52})-C(=O))_s-$, $-(O-(CH)_r-C(=O))_s-$,



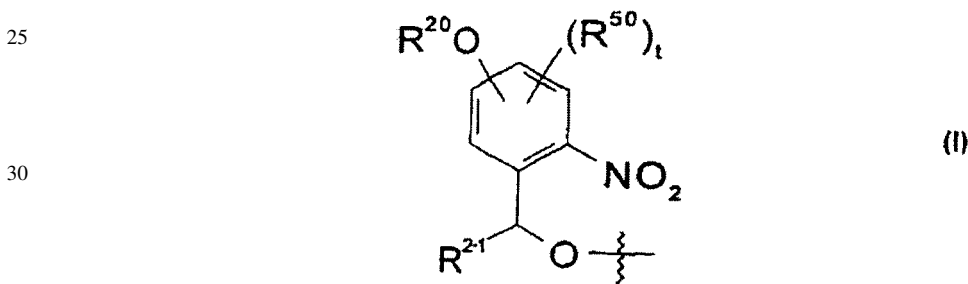
en donde R^{15} es como se ha definido anteriormente; r y s son, independientemente, un número entero desde 1 hasta 10; R^{52} es la cadena lateral de un α -aminoácido natural; e y es un número entero de 0 a 4. En una realización y es 0 ó 1.

En ciertas realizaciones, R^{15} es -H, -OH, $-OR^{51}$, -SH, $-SR^{61}$, $-NH_2$, $-NHR^{51}$, $-NR^{51}$, -F, -Cl, -Br, -I, $-SO_3H$, $-PO^{-2}_4$, $-CH_3$, $-CH_2CH_3$, $-CH(CH_3)_2$, o $-C(CH_3)_3$; en donde R^{51} es alquilo de cadena lineal o ramificada, alqueno de cadena lineal o ramificada, arilo, heteroarilo, cicloalquilo, heterociclilo, aralquilo de cadena lineal o ramificada, aralqueno de cadena lineal o ramificada, aralquinilo de cadena lineal o ramificada, heteroaralquilo de cadena lineal o ramificada, heteroaralqueno de cadena lineal o ramificada, heteroaralquinilo de cadena lineal o ramificada, cicloalquil-alquilo de cadena lineal o ramificada, heterociclil-alquilo de cadena lineal o ramificada, heterociclil-alqueno de cadena lineal o ramificada, o heterociclil-alquinilo de cadena lineal o ramificada.

En ciertas realizaciones, el grupo escindible L se escinde ya sea antes de o durante el análisis de la biomolécula, tal como una proteína. El análisis puede incluir el análisis espectral de masa, por ejemplo, el análisis espectral de masa MALDFI-TOF. El grupo escindible L se selecciona de manera que el grupo sea estable durante la conjugación a una biomolécula, hibridación del resto Q a una secuencia complementaria, y el lavado del híbrido; pero es susceptible de escisión bajo las condiciones de análisis de la biomolécula, que incluye, sin limitaciones, análisis espectral de masa, por ejemplo, el análisis espectral de masa MALDI-TOF. En ciertas realizaciones, el grupo escindible L puede ser un resto disulfuro, creado por la reacción de los compuestos en los que X = -SH, con la cadena lateral tior de los residuos cisteína en la superficie de las biomoléculas, que incluyen, sin limitaciones, proteínas. El enlace disulfuro resultante se puede escindir bajo diversas condiciones reductoras que incluyen, sin limitaciones, el tratamiento con ditioneitol y 2-mercaptoetanol.

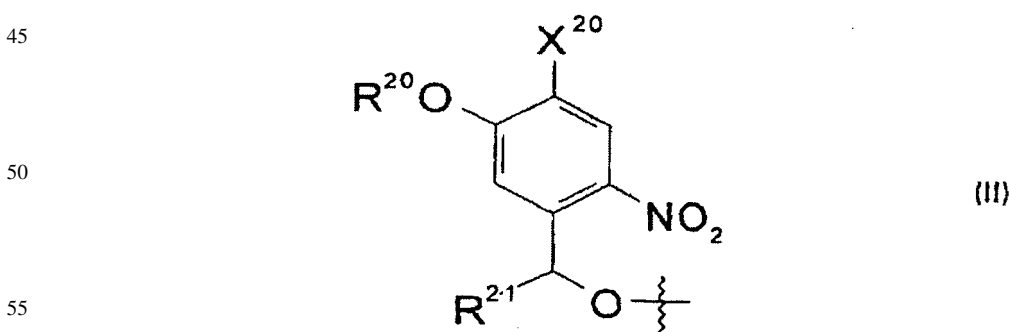
En otra realización, L es un grupo foto-escindible, que se puede escindir mediante un breve tratamiento con luz UV a la longitud de onda apropiada, tanto antes de cómo durante la espectrometría de masa. Se pueden utilizar grupos foto-escindibles, incluidos aquellos enlaces que pueden ser escindidos durante la espectrometría de masa MALDI-TOF por la acción de un rayo láser. Por ejemplo, el éter tritilo o aralquilo sustituido con orto-nitro, incluido el grupo bencilo, son susceptibles a la escisión del enlace inducida por láser durante la espectrometría de masa MALDI-TOF. Otros grupos foto-escindibles de utilidad incluyen, sin estar limitados a ellos, los grupos o-nitro-bencilo, fenacilo y nitro-fenilsulfonilo.

Otros grupos foto-escindibles que se pueden usar en esta invención incluyen los descritos en la Solicitud de Patente Internacional No. WO 98/20166. En una realización, los grupos foto-escindibles tienen la fórmula I:



35 en la que R²⁰ es ω-O-alkileno-; R²¹ se selecciona de hidrógeno, alquilo, arilo, alcocarbonilo, ariloxicarbonilo y carboxi; t es 0-3; y R⁵⁰ es alquilo, alcoxi, arilo o ariloxi. En una realización, Q está unido a R²⁰ a través de (S¹)_t-M(R¹⁵)_a-(S²)_b; y la biomolécula de interés se captura sobre el resto R²¹CH-O- a través de un derivado reactivo del oxígeno (por ejemplo, X).

40 En otra realización, los grupos foto-escindibles tienen la fórmula II:



55 en la que R²⁰ es ω-O-alkileno o alkileno; R²¹ se selecciona de hidrógeno, alquilo, arilo, alcocarbonilo, ariloxicarbonilo y carboxi; y X²⁰ es hidrógeno, alquilo u OR²¹. En una realización, Q está unido a R²⁰ a través de (S¹)_t-M(R¹⁵)_a-(S²)_b; y la biomolécula de interés se captura sobre el resto R²¹CH-O- a través de un derivado reactivo del oxígeno (por ejemplo, X).

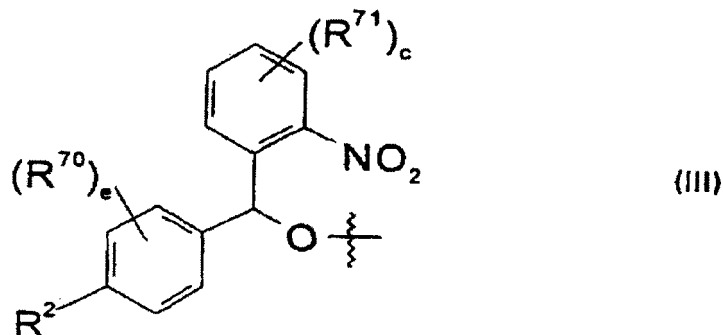
60 En realizaciones adicionales, R²⁰ es -O-(CH₂)₃- o metileno; R²¹ se selecciona de hidrógeno, metilo y carboxi; y X²⁰ es hidrógeno, metilo u OR²¹. En otra realización, R²¹ es metilo; y X²⁰ es hidrógeno. En ciertas realizaciones, R²⁰ es metileno; R²¹ es metilo; y X²⁰ es 3-(4,4'-dimetoxi-tritilo-oxi)propoxi.

En otra realización, los grupos foto-escindibles tienen la fórmula III:

5

10

15



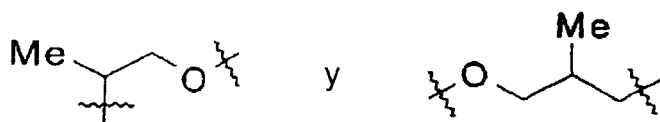
20

en la que R² se selecciona de ω-O-alquileo-O y ω-O-alquileo-, y está no sustituido o sustituido en la cadena alquileo con uno o múltiples grupos alquilo; c y e son, independientemente, 0-4; y R⁷⁰ y R⁷¹ son, independientemente, alquilo, alcoxi, arilo o ariloxi. En ciertas realizaciones, R² es ω-O-alquileo-, y está sustituido en la cadena alquileo con un grupo metilo. En una realización, Q está unido a R² a través de (S¹)_i-M(R¹⁵)_a-(S²)_b; y la biomolécula de interés se captura sobre el resto Ar₂CH-O- a través de un derivado reactivo del oxígeno (por ejemplo, X).

25

En realizaciones adicionales, R² se selecciona de 3-O-(CH₂)₃-O-, 4-O-(CH₂)₄-, 3-O-(CH₂)₃-, 2-O-CH₂CH₂-, -OCH₂-,

30



35

En otras realizaciones, c y e son 0.

40

Otros grupos escindibles L incluyen grupos sensibles al ácido en los que la escisión del enlace está determinada por la formación de un catión tras la exposición a ácidos suaves a fuertes. Para estos grupos ácido-lábiles, la escisión del grupo L se puede llevar a cabo antes de o durante el análisis, incluido el análisis espectrométrico de masa, por la acidez de las moléculas de la matriz, que se depositan como una capa fina sobre los híbridos, o aplicando un breve tratamiento de la matriz con un ácido, tal como vapor de ácido trifluoroacético. La exposición de un grupo tritilo al ácido acético o trifluoroacético provoca la escisión del enlace éter antes de o durante la espectrometría de masa MALDI-TOF.

45

La matriz de compuesto-biomolécula puede ser tratada con reactivos químicos, que incluyen, sin limitaciones, bromuro de cianógeno, o enzimáticos, que incluyen, sin limitaciones, tripsina, quimotripsina, una exopeptidasa (por ejemplo, aminopeptidasas y carboxipeptidasas) para efectuar la escisión. En este último caso, todos, excepto un fragmento peptídico, se mantendrán hibridados cuando la digestión es cuantitativa. La digestión parcial puede ser conveniente también para identificar y caracterizar proteínas tras la desorción de la matriz. Los fragmentos de proteína/péptido escindidos se someten a desorción, análisis y caracterización de acuerdo con sus respectivos pesos moleculares.

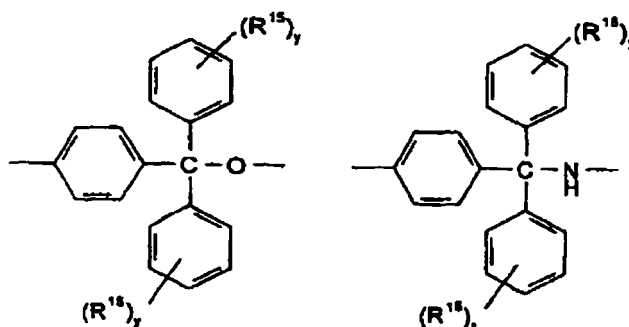
50

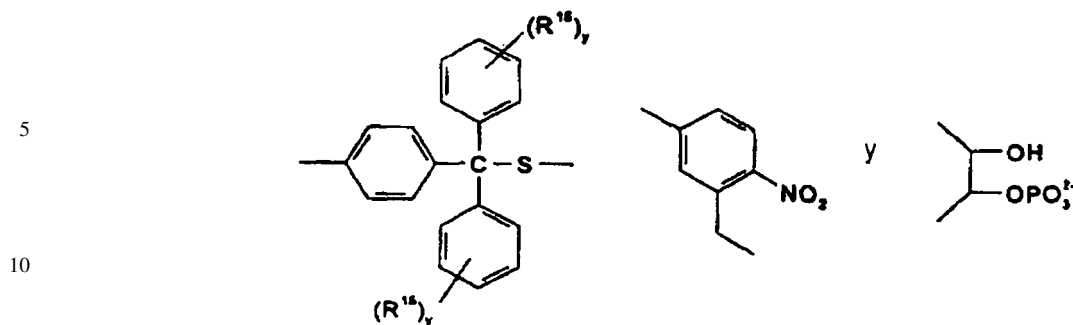
En ciertas realizaciones de la invención, L se selecciona de -S-S-, -O-P(=O)(OR⁵¹)-NH-, -O-C(=O)-,

55

60

65

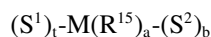




en las que R^{15} , R^{51} e y son como se han definido anteriormente. En ciertas realizaciones, R^{15} es -H, -OH, -OR⁵¹, -SH, -SR⁵¹, -NH₂, -NHR⁵¹, -N(R⁵¹)₂, -F, -Cl, -Br, -I, -SO₃H, -PO⁻²₄, -CH₃, -CH₂CH₃, -CH(CH₃)₂ o -C(CH₃)₃; en donde R^{51} es alquilo de cadena lineal o ramificada, alqueno de cadena lineal o ramificada, alquino de cadena lineal o ramificada, arilo, heteroarilo, cicloalquilo, heterocicilo, aralquilo de cadena lineal o ramificada, aralqueno de cadena lineal o ramificada, aralquino de cadena lineal o ramificada, heteroaralquilo de cadena lineal o ramificada, heteroaralqueno de cadena lineal o ramificada, heteroaralquino de cadena lineal o ramificada, cicloalquil-alquilo de cadena lineal o ramificada, cicloalquil-alqueno de cadena lineal o ramificada, cicloalquil-alquino de cadena lineal o ramificada, heterocicil-alquilo de cadena lineal o ramificada, heterocicil-alqueno de cadena lineal o ramificada, o heterocicil-alquino de cadena lineal o ramificada.

ii. Restos Z divalentes no escindibles

En otra realización, Z es un resto divalente no escindible y tiene la fórmula:



en la que S^1 , M, R^{15} , S^2 , t, a y b son como se han definido anteriormente.

e. Z es un soporte insoluble o un sustrato

Para las matrices de gradiente que se proponen en esta invención, Z puede ser el soporte o sustrato insoluble. En estas realizaciones, Z está unido a una pluralidad de restos X, o a una pluralidad de restos Y y X, o mezclas de X, Q y/o Y, y otros restos. En ciertas realizaciones, Z tiene decenas hasta centenares, millares, millones o más puntos de fijación. En estas realizaciones, el soporte o sustrato insoluble Z está basado en una superficie plana construida, por ejemplo, de vidrio, silicio, metal, plástico o un material compuesto; o puede adoptar la forma de una perla tal como de gel de sílice, de vidrio de porosidad controlada, o una perla magnética o de celulosa; o puede ser un alfiler, incluida una matriz de alfileres adecuada para la síntesis o el análisis combinatorio. Los sustratos se pueden fabricar prácticamente de cualquier material insoluble o sólido. Por ejemplo, gel de sílice, vidrio (por ejemplo, vidrio de porosidad controlada (CPG)), nailon, resina Wang, resina Merrifield, dextrano reticulado con epíclorohidrina (por ejemplo, Sephadex[®]), agarosa (por ejemplo, Sepharose[®]), celulosa, perlas magnéticas, Dynabeads[®], una superficie metálica (por ejemplo, acero, oro, plata, aluminio, silicio y cobre), un material plástico (por ejemplo, polietileno, polipropileno, poliamida, poliéster, difluoruro de polivinilideno (PVDF)). Ejemplos de sustratos incluyen, sin limitaciones, perlas (por ejemplo, gel de sílice, vidrio de porosidad controlada, magnéticas, dextrano reticulado con epíclorohidrina (por ejemplo, Sephadex[®]), agarosa (por ejemplo, Sepharose[®]), celulosa), capilares, soportes planos tales como filtros de fibra de vidrio, superficies de vidrio, superficies metálicas (acero, oro, plata, aluminio, cobre y silicio), materiales plásticos incluidas las placas o membranas multi-pocillos (por ejemplo, de polietileno, polipropileno, poliamida, difluoruro de polivinilideno (PVDF)). Ejemplos de sustratos incluyen, sin limitaciones, perlas (por ejemplo, de gel de sílice, de vidrio de porosidad controlada, magnéticas, de dextrano reticulado con epíclorohidrina (por ejemplo, Sephadex[®]), agarosa (por ejemplo, Sepharose[®]), celulosa), capilares, soportes planos tales como filtros de fibra de vidrio, superficies de vidrio, superficies metálicas (acero, oro, plata, aluminio, cobre y silicio), materiales plásticos, incluidas placas o membranas multi-pocillo (por ejemplo, de polietileno, polipropileno, poliamida, difluoruro de polivinilideno), alfileres (por ejemplo, matrices de alfileres adecuados para la síntesis o el análisis combinatorio, o perlas en cavidades de superficies planas tales como obleas (por ejemplo, obleas de sílice), con o sin placas. El soporte sólido muestra cualquier forma deseada, incluida, sin limitaciones, la de perla, capilar, placa, membrana, oblea, peine, oblea con cavidades, una matriz de cavidades o pocillos de nanolitro, y otras geometrías y formas conocidas por los expertos en la técnica. Los soportes incluyen superficies planas diseñadas para recibir o fijar muestras en loci discretos. En una realización, las superficies planas incluyen las que tienen regiones hidrófobas alrededor de loci hidrófilos para recibir, contener o fijar una muestra.

En una realización, los soportes o sustratos sólidos Z son perlas, incluidas, sin limitaciones, perlas polímeras, magnéticas, coloreadas, etiquetadas con R_f , etc. Las perlas pueden estar fabricadas de materiales hidrófobos que incluyen, sin limitaciones, polietireno, polietileno, polipropileno o teflón, o de materiales hidrófilos, que incluyen, sin limitaciones, celulosa, dextrano reticulado con epíclorohidrina (por ejemplo, Sephadex[®]), agarosa (por ejemplo, Sepharose[®]), poliácridamida, gel de sílice y vidrio de porosidad controlada.

ES 2 319 634 T3

En realizaciones adicionales, los restos Z de soporte o sustrato insoluble pueden poseer, opcionalmente, grupos espaciadores S¹ y/o S² o, en realizaciones en las que Z es un enlace escindible, L. Los restos S¹, S² y/o L están unidos a la superficie del soporte o sustrato insoluble.

- 5 En estas realizaciones, la densidad del biopolímero a analizar y, por lo tanto, la intensidad de la señal del análisis subsiguiente, está incrementada en relación con las realizaciones en las que Z es un grupo divalente o multivalente. En estas realizaciones, se empleará en los métodos propuestos una matriz apropiada de oligonucleótidos o análogos de oligonucleótidos monocatenarios, que son complementarios a Q.
- 10 En otra realización, Z es un soporte que, cuando está funcionalizado con Q, X e Y, simula la función de una membrana artificial, capturando biomoléculas que incluyen, sin limitaciones, proteínas. En una realización, Z es un soporte que, cuando está funcionalizado con Q, X e Y, simula la especificidad del interior de una membrana celular. En esta realización, el soporte es capaz de capturar proteínas y otras biomoléculas que, en otras circunstancias, los compuestos propuestos en esta invención serían incapaces de capturar, incluidas proteínas dentro de membranas celulares. En una
- 15 realización, los soportes funcionan bajo condiciones fisiológicas. Así, en las realizaciones anteriores, la elección de Z, Q, X e Y permite diseñar superficies y soportes que imitan membranas celulares y otras membranas biológicas. Cuando los compuestos propuestos en este documento actúan como una membrana artificial, se puede emplear la química polímera de dendrímeros para la síntesis controlada de membranas con dimensiones de poros y espesores de membrana consistentes, a través de la síntesis de copolímeros en bloque anfífilos, dendrímicos o híper-ramificados,
- 20 que pueden estar auto-ensamblados para formar membranas de película ultrafina sobre soportes porosos.

Matrices de gradiente

- En este documento, se proponen matrices de gradiente bidimensionales. En estas realizaciones, Z es un soporte
- 25 sólido y X e Y son restos (separados o unidos a Z) dispuestos sobre dicho soporte, de manera que la colección de loci presenta un gradiente de las propiedades seleccionadas. En cada locus de la matriz existen restos X e Y, dispuestos lado a lado o unidos, mientras que a lo largo de cada fila (es decir, eje X), cada resto Y es idéntico y se altera una propiedad seleccionada de X, por ejemplo, hidrofobia, lipofilia, carga, tamaño, especificidad u otra propiedad de cada
- 30 X de forma predeterminada a lo largo del eje X para ofrecer un gradiente de dicha propiedad. De manera similar, a lo largo de cada columna de loci en la matriz, se altera una propiedad de Y, por ejemplo, hidrofobia, carga, tamaño, especificidad u otra propiedad de cada Y de forma predeterminada a lo largo del eje Y para proporcionar un gradiente de dicha propiedad. Para cada matriz, la propiedad de los restos X que se ha alterado no es la misma que la propiedad que se ha alterado para el resto Y. De este modo, la matriz resultante presenta un gradiente bidimensional de dos (o
- 35 más) propiedades seleccionadas y muestra loci con diferentes afinidades. El número de loci en la matriz puede ser cualquiera, por ejemplo, 10, 50, 100, 500, 1.000, 10⁴ o más, o cualquier número predeterminado.

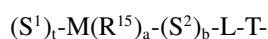
- Estos soportes sólidos pueden ser cualquier soporte sólido como los que se proponen en esta invención, e incluyen perlas, cuadrados y rectángulos planos y otras geometrías, pudiendo ser del tamaño deseado, por ejemplo, 1 μm, 10
- 40 μm, 50 μm, 100 μm, 1 mm, 1 cm, 10 cm y mayores, en donde las dimensiones se refiere a la dimensión máxima; a los efectos de este documento, el tamaño es adecuado para llevar a cabo análisis de espectrometría de masa, en particular análisis MALDI. En estas matrices, los restos X e Y se pueden establecer de manera que presenten un gradiente de propiedades seleccionadas.

- En la práctica, las matrices resultantes se analizan del modo descrito en la invención. Por ejemplo, las matrices
- 45 resultantes se ponen en contacto con una muestra, tal como una muestra que contiene proteínas, que se debe analizar, se añade la matriz y se analiza por espectrometría de masa la matriz resultante. A continuación, se describen ejemplos de gradientes.

- La Figura 5 muestra un soporte continuo con gradientes o monocapas liotrópicas, en las que las moléculas son
- 50 crecientemente hidrófobas/hidrófilas, cargadas (negativa o positivamente) y de tamaño creciente o decreciente. Se pueden utilizar otras propiedades bien conocidas por los expertos en la técnica, tales como propiedades químicas, incluida la especificidad de los reactivos para los grupos NH₂, SH, SS, OH, etc. En una realización, el gradiente a lo largo del eje X difiere del eje Y, por ejemplo, X = hidrofobia creciente, e Y = carga positiva creciente. En una
- 55 realización, se usan cambios de azobenceno (véase, por ejemplo, la Figura 6), en donde la hidrofilia se introduce cuando el azobenceno se expone a la luz. Se establece un gradiente de hidrofobia/hidrofilia a medida que aumenta la exposición a la luz. La carga se puede introducir usando electrodos que generan un gradiente de cargas electrónicas, véase, por ejemplo, la Figura 7. De manera alternativa, a lo largo de cada dirección (X e Y), habrá múltiples parches con diferentes afinidades. La mezcla de proteínas se aplica a la superficie del sustrato, con lo que las proteínas se
- 60 disponen de acuerdo con sus afinidades exclusivas. Seguidamente, la matriz se marca en cada zona y se analiza, por ejemplo, mediante MALDI-TOF.

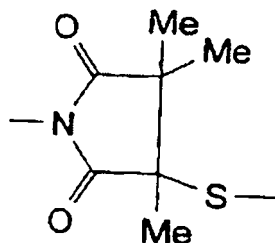
f. Restos Z de masa modificada

- En realizaciones adicionales, incluidas las realizaciones en las que Z incluye un resto escindible, Z puede incluir
- 65 una etiqueta modificadora de la masa. En ciertas realizaciones, la etiqueta modificadora de la masa está unida al enlazador L. En una realización, el resto Z de masa modificada tiene la fórmula:



ES 2 319 634 T3

en la que S¹, t, M, R¹⁵, a, S², b y L se seleccionan como se ha indicado anteriormente; y T es una etiqueta modificadora de la masa. Las etiquetas modificadoras de masa que se pueden usar en esta invención incluyen, sin limitaciones, grupos de la fórmula -X¹R¹⁰-, en la que X¹ es un grupo divalente tal como -O-, -O-C(O)-(CH₂)_y-C(O)O-, -NH-C(O)-, -C(O)-NH-, -NH-C(O)-(CH₂)_y-C(O)O-, -NH-C(S)-NH-, -O-P(o-alkyl)-O-, -O-SO₂-O-, -O-C(O)-CH₂-S-, -NH- y



10

15

y R¹⁰ es un grupo divalente que incluye -(CH₂CH₂O)_z-CH₂CH₂O-, -(CH₂CH₂O)_z-CH₂CH₂O-alkileno, alquileno, alquenileno, alquinileno, arileno, heteroarileno, -(CH₂)_z-CH₂-O-, -(CH₂)_z-CH₂-O-alkileno, -(CH₂CH₂NH)_z-CH₂CH₂NH-, -CH₂-CH(OH)-CH₂O-, -Si(R¹²)(R¹³)-, -CHF- y -CF₂-; en donde y es un número entero desde 1 hasta 20; z es un número entero desde 0 hasta 200; R¹¹ es la cadena lateral de un α-aminoácido; y R¹² y R¹³ se seleccionan, independientemente, de alquilo, arilo y aralquilo.

20

En otras realizaciones, -X¹R¹⁰- se selecciona de -S-S-, -S-, -(NH-(CH₂)_y-NH-C(O)-(CH₂)_y-C(O))_z-NH-(CH₂)_y-NH-C(O)-(CH₂)_y-C(O)O-, -(NH-(CH₂)_z-C(O))_z-NH-(CH₂)_y-C(O)-, -(NH-CH(R¹¹)-C(O))_z-NH-CH(R¹¹)-C(O)O-, y -(O-(CH₂)_y-C(O))_z-NH-(CH₂)_y-C(O)O-.

25

En las realizaciones anteriores, en donde R¹⁰ es un derivado de oligo/polietilenglicol, el incremento de modificación de masa es 44, es decir, se pueden generar cinco especies modificadoras de masa cambiando z desde 0 hasta 4, agregando de esta forma unidades de masa de 45 (z = 0), 89 (z = 1), 133 (z = 2), 177 (z = 3) y 221 (z = 4) a los compuestos. Los oligo/polietilenglicoles pueden estar también monoalquilados con un alquilo inferior tal como metilo, etilo, propilo, isopropilo, t-butilo y similares.

30

Otras etiquetas modificadoras de la masa incluyen, sin estar limitados a ellos, -CHF-, -CF₂-, -Si(CH₃)₂-, -Si(CH₃)(C₂H₅)- y -Si(C₂H₅)₂-. En otras realizaciones, las etiquetas modificadoras de masa incluyen homo- y heteropéptidos. Un ejemplo no limitante que genera especies de masa modificada, con un incremento de masa de 57, es una oligoglicina, que produce modificaciones de masa de, por ejemplo, 74 (y = 1, z = 0), 131 (y = 1, z = 2), 188 (y = 1, z = 3) o 245 (y = 1, z = 4). También se pueden usar oligoamidas, con las que se obtienen, por ejemplo, modificaciones de masa de 74 (y = 1, z = 0), 88 (y = 2, z = 0), 102 (y = 3, z = 0), 116 (y = 4, z = 0), etc. Los expertos en la técnica apreciarán que existen numerosas posibilidades adicionales a las que se ejemplifican en este documento, para introducir, de forma predeterminada, muchas etiquetas modificadoras de la masa diferentes a los compuestos citados.

35

40

En otras realizaciones, R¹⁵ y/o S² pueden estar funcionalizados con -X¹R¹⁰H o -X¹R¹⁰-alquilo, en donde X¹ y R¹⁰ son como se han definido anteriormente, para actuar como etiquetas modificadoras de masa.

45

3. El resto X

En los compuestos propuestos en este documento, X es una cadena lateral aminoácida de una proteína. Así, en ciertas realizaciones, X es un grupo que puede reaccionar o interactuar con funcionalidades situadas en la superficie de una proteína para formar enlaces covalentes o no covalentes. Existe una extensa selección de diferentes grupos funcionales para que X interactúe con una proteína. Por ejemplo, X puede actuar como un nucleófilo o un electrófilo para formar enlaces covalentes tras la reacción con los residuos aminoácidos en la superficie de una proteína. Los reactivos que se unen a las cadenas laterales de aminoácidos incluyen, sin limitaciones, grupos protectores para los restos hidroxilo, carboxilo, amino, amida y tiol, incluyendo los descritos en T.W. Greene y P.G.M. Wuts, "Protective Groups in Organic Synthesis", 3^a ed. (1999, Wiley Interscience). Estos grupos protectores pueden reaccionar con cadenas laterales de aminoácidos tales como hidroxilo (serina, treonina, tirosina); amino (lisina, arginina, histidina, prolina); amida (glutamina, asparagina); ácido carboxílico (ácido aspártico, ácido glutámico); y derivados de azufre (cisteína, Metionina), y son fácilmente adaptables para ser usados en estos compuestos como resto reactivo X.

55

Es preciso señalar que, además de la amplia gama de reactivos específicos para grupos conocidos por los expertos en la técnica, los reactivos conocidos en la química de productos naturales también pueden servir como base para X en la formación de enlaces covalentes.

60

De forma alternativa a como se describe en este documento, X es una tinción de purificación de proteínas, tales como acridina o azul de metileno, que exhibe una fuerte afinidad por ciertas proteínas.

65

X puede actuar como donante de electrones o como aceptor de electrones para formar enlaces no covalentes o un complejo, tal como un complejo de transferencia de carga, con una biomolécula, incluyendo, sin limitaciones,

ES 2 319 634 T3

una proteína. Estos reactivos incluyen a los que interactúan con intensidad y con alta especificidad con biomoléculas, incluidas, sin limitaciones, proteínas, sin formar enlaces covalentes a través de la interacción de superficies de afinidad complementaria. Por ejemplo, se describen complejos bien conocidos tales como biotina-estreptavidina, anticuerpo-antígeno, receptor-ligando, lectina-hidrato de carbono y otros tipos de reactivos. Adicionalmente, estos complejos muestran fuertes interacciones que son suficientemente estables para permitir un lavado exhaustivo de las biomoléculas no fijadas, incluidas, sin limitaciones, las proteínas, de las mezclas biológicas que forman complejo.

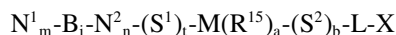
Si S^2 no está presente, la reactividad de X puede estar influida por una o múltiples funcionalidades sustituidas, por ejemplo, R^{15} en M. Se pueden usar los efectos electrónicos (por ejemplo, mesoméricos, inductivos) y/o estéricos para modular la reactividad de X y la estabilidad del enlace X-biomolécula resultante. En estas realizaciones, subconjuntos de mezclas biomoleculares que incluyen, sin limitaciones, mezclas de proteínas, pueden reaccionar y ser analizadas gracias a la modulación por R^{15} , lo que modifica las propiedades electrónicas o estéricas de X y, por lo tanto, aumenta la selectividad de la reacción de X con la biomolécula.

En estas realizaciones, X es un éster activo tal como $-C(=O)O-Ph-pNO_2$, $-C(=O)O-C_6F_5$ o $-C(=O)-O-(N-succinimidilo)$; un resto halo activo tal como un éter α -halo o un grupo La matriz de gradiente según la reivindicación 1, en la que la propiedad de los compuestos que se varía a lo largo del eje X es especificidad para un grupo NH_2 , SH, SS u OH.

carbonilo α -halo que incluye, sin limitaciones, $-OCH_2-I$, $-OCH_2-Br$, $-OCH_2-Cl$ y $-C(O)CH_2I$, $-C(O)CH_2Br$ y $-C(O)CH_2Cl$ grupos funcionales específicos para cadenas laterales de aminoácidos tales como maleimido (para cisteína), un complejo metálico, incluidos complejos de oro o mercurio (para cisteína o metionina), un epóxido o isotiocianato (para arginina o lisina); reactivos que se unen a sitios activos de enzimas que incluyen, sin limitaciones, análogos de estado de transición; ligandos que se unen a receptores tales como insulina; péptidos específicos que se unen a las superficies de las biomoléculas, incluyendo péptidos adhesivos; lectinas (por ejemplo, del tipo manosa, del tipo lactosa); anticuerpos, por ejemplo, contra péptidos fosforilados; antígenos, tales como una biblioteca de expresión en fago; haptenos; biotina; avidina; o estreptavidina. Otras realizaciones de X son bien conocidas para los expertos en la técnica, e incluyen las descritas en *Techniques in Protein Chemistry*, Vol. 1 (1989), T. Hugli ed. (Academic Press); *Techniques in Protein Chemistry*, Vol. 5 (1994), J.W. Crabb ed. (Academic Press); *Lundblad Techniques in Protein Modification* (1995) (CRC Press, Boca Ratón, FL); *Glazer et al. (1976) Chemical Modification of Proteins* (Holanda del Norte (Amsterdam) (American Elsevier, Nueva York); y *Hermanson (1996) Bioconjugate Techniques* (Academic Press, San Diego, CA).

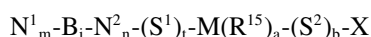
4. Realizaciones adicionales

En ciertas realizaciones, los compuestos propuestos en esta invención tienen la fórmula:



en la que N^1 , B, N^2 , S^1 , M, S^2 , L, X, m, i, n, t, a y b son como se han definido anteriormente.

En otras realizaciones, incluidas aquellas en las que Z no es un enlazador escindible, los compuestos propuestos en esta invención tienen la fórmula:



en la que N^1 , B, N^2 , S^1 , M, S^2 , X, m, i, n, t, a y b son como se han definido anteriormente.

D. Preparación de los Compuestos

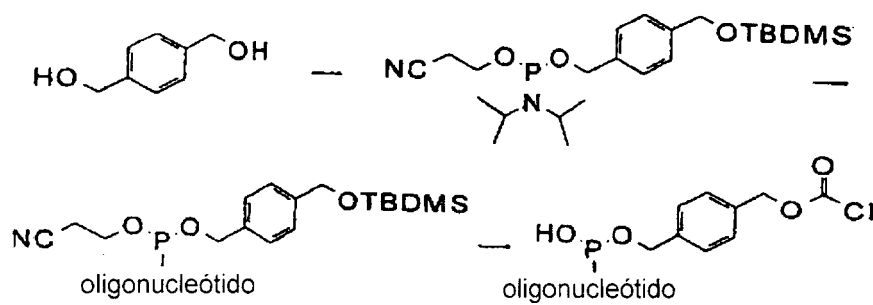
La preparación de los compuestos se describe más adelante. Cualquier compuesto o compuesto similar se puede sintetizar de acuerdo con un método que se analiza en general más abajo, o mediante modificaciones mínimas de los métodos, seleccionando los materiales de partida apropiados.

En general, los compuestos se pueden preparar a partir del resto central Z. En ciertas realizaciones, Z es $-(S^1)_t-M(R^{15})_a-(S^2)_b-L-$. En estas realizaciones, los compuestos se pueden preparar a partir del grupo M adecuadamente sustituido (por ejemplo, con uno o múltiples grupos R^{15}). $M(R^{15})_a$ se une opcionalmente con S^1 y/o S^2 , seguida por la unión al enlazador escindible L. De manera alternativa, el grupo L se une opcionalmente a S^2 , seguido por la reacción con $M(R^{15})_a$ y, opcionalmente, S^1 . A continuación, este grupo Z se derivatiza en su extremo S^1 (o $M(R^{15})_a$) para tener una funcionalidad para acoplarse con un oligonucleótido o análogo de oligonucleótido Q (por ejemplo, un grupo fosforamidita, H-fosfonato o triéster fosfórico). Por lo general, el grupo Q estará N-prottegido en las bases para evitar las reacciones de competición tras la introducción del resto X. En una realización, el grupo Z se hace reaccionar con una mezcla de todas las permutaciones posibles de un oligonucleótido o análogo de oligonucleótido Q (por ejemplo, permutaciones de 4' donde i es el número de nucleótidos o análogos de nucleótidos en B). El o los compuestos Q-Z resultantes se derivatizan entonces a través del extremo L para poseer un grupo X para reaccionar con una biomolécula tal como una proteína. Si se desea, se retiran entonces los grupos N-protectores tras la reacción del

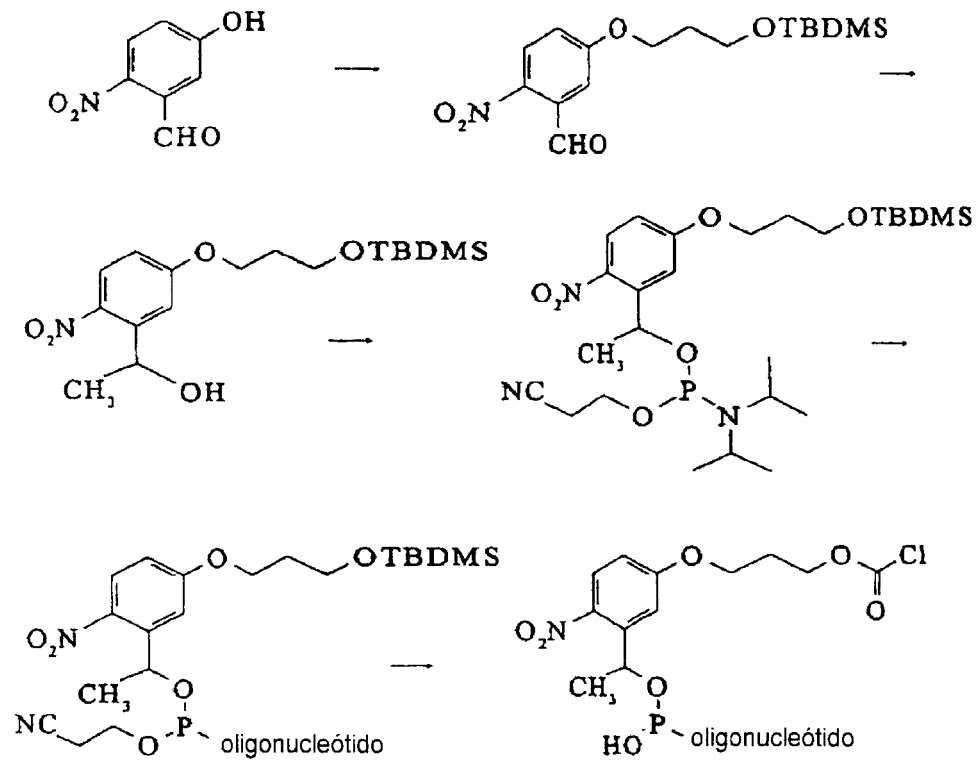
compuesto con una biomolécula, incluida una proteína. En otras realizaciones, Q se puede sintetizar sobre Z, incluyendo realizaciones sobre un soporte o sustrato insoluble tales como una perla. En una realización adicional, Q se pre-sintetiza mediante técnicas de estado sólido convencionales y, a continuación, se enlaza con M. De forma alternativa, Q se puede sintetizar gradualmente sobre el resto M.

Más adelante, se ofrecen ejemplos de síntesis de los compuestos propuestos en esta invención, que contienen enlazadores alcalino-lábiles y foto-escindibles. Un experto en la técnica tendría la capacidad para preparar otros compuestos dentro del alcance de esta descripción por la modificación rutinaria de los métodos expuestos más adelante, o por otros métodos bien conocidos por los expertos en la técnica.

Para la síntesis de un compuesto propuesto en este documento, que contiene un enlazador alcalino-lábil, 1,4-di(hidroximetil)benceno (es decir, M) se mono-protecte, por ejemplo, como el correspondiente éter mono-terc-butildimetilsililo. El alcohol libre remanente se derivatiza como la correspondiente 2-ciano-etil-N,N-diisopropil-fosforamidita mediante la reacción con 2-ciano-etil-N,N-diisopropil-clorofosforamidita. La reacción de esta amidita con un oligonucleótido (es decir, Q) va seguida de la eliminación del grupo protector para dar el alcohol correspondiente. La reacción con, por ejemplo, cloroformiato de triclorometilo proporciona el cloroformiato ilustrado (es decir, X).

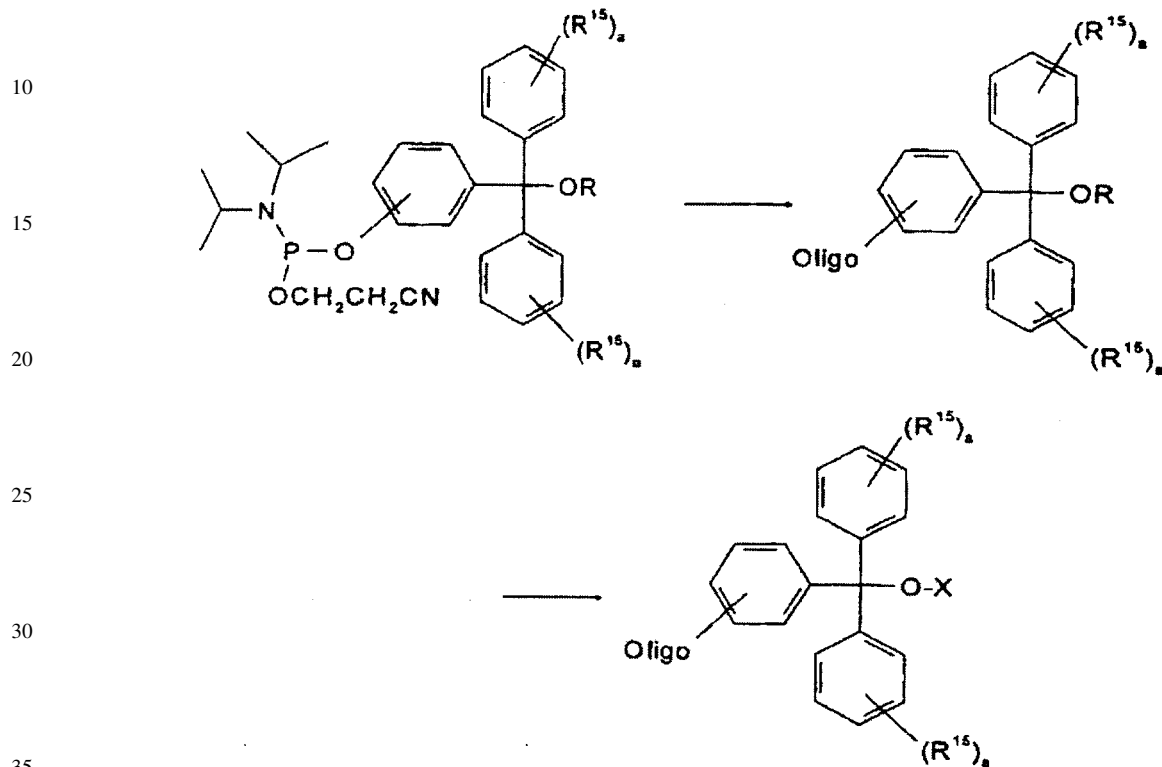


Para la síntesis de un compuesto propuesto en esta invención, que contiene un enlazador foto-escindible, 2-nitro-5-hidroxi-benzaldehído (es decir, un precursor de L), se hace reaccionar, por ejemplo, con 3-bromo-1-propanol para dar el correspondiente éter-alcohol. A continuación, se protege el alcohol, por ejemplo, como el correspondiente éter terc-butildimetilsililo. La reacción de este compuesto con trimetilaluminio da el correspondiente alcohol bencilico, que se derivatiza en forma de su fosforamidita, usando el procedimiento descrito anteriormente. La amidita se hace reaccionar con un oligonucleótido (es decir, Q) seguida de la eliminación del grupo protector y la derivatización del alcohol resultante en forma del cloroformiato correspondiente (es decir, X).



Para la síntesis de los compuestos propuestos en esta invención, que contienen un enlazador ácido-lábil, por ejemplo, un éter tritilo heterobifuncional, se hace reaccionar el éter tritilo de la fosforamidita requerido con el oligonucleótido o análogo de oligonucleótido Q, seguido de la desprotección del éter tritilo y la captura de una biomolécula, por ejemplo, una proteína, en el alcohol, a través de un derivado reactivo del alcohol (X), como se ha descrito anteriormente.

5



10

15

20

25

30

35

Los expertos en la técnica podrán apreciar que las síntesis anteriores de los compuestos propuestos en esta invención son solamente ejemplares. Es posible vislumbrar otras síntesis de los compuestos ofrecidos en este documento. Asimismo, un experto en la técnica podrá modificar las síntesis anteriores de manera convencional para sintetizar otros compuestos, dentro del alcance de la presente descripción.

40

E. Métodos de Utilización de las Matrices

45

1. Métodos generales

Las matrices que se proporcionan en este documento se pueden usar para el análisis, cuantificación, purificación y/o identificación de los componentes de mezclas de biomoléculas, incluidas, sin limitaciones, mezclas de proteínas. Para iniciar el procedimiento analítico, estas mezclas de pre-purifican de acuerdo con procedimientos convencionales. En una realización, las proteínas se aíslan de fluidos biológicos por lisis celular, seguida de métodos de precipitación (por ejemplo, sulfato de amonio) o degradación enzimática de los ácidos nucleicos y carbohidratos (si es preciso), y se retira el material de bajo peso molecular por tamización molecular. Las proteínas se pueden obtener también de bibliotecas de expresión. Se hacen reaccionar partes alícuotas de la mezcla de proteínas con los compuestos propuestos por la invención, que tienen diferentes funcionalidades X para segregar la mezcla en familias de proteínas separadas, según la reactividad seleccionada de X. La diversidad de B se selecciona en función de la complejidad de la mezcla de proteínas. Por lo tanto, existen conjuntos de compuestos que difieren en X y B, que pueden ser seleccionados para el análisis. En ciertas realizaciones, el análisis se lleva a cabo usando el menor número posible de reacciones necesarias para analizar por completo la mezcla. Así, en estas realizaciones, la selección de la diversidad de B y del número de grupos X de diferente reactividad será una función de la complejidad de la mezcla biomolecular que se debe analizar. La minimización de la diversidad de B y del número de grupos X permite un análisis completo de la mezcla con complejidad mínima.

50

55

60

La separación de proteínas de una mezcla compleja se consigue en virtud de los productos compuesto-proteína que se unen a una matriz de secuencia complementaria. El sobrenadante, que contiene los productos compuesto-proteína, se hace contactar con y se le permite hibridar con una matriz de secuencias complementarias. En una realización, se hibrida un soporte sólido plano, portador en localizaciones espacialmente distintas de una matriz de oligonucleótidos o análogos de oligonucleótidos que es complementaria al oligonucleótido o análogo de oligonucleótido N¹_m-B₁-N²_n con los productos compuesto-proteína.

65

ES 2 319 634 T3

En realizaciones en las que Z es un soporte o sustrato insoluble, tal como una perla, la separación de los productos compuesto-proteína en una matriz direccionable se puede lograr separando en una matriz de placas de micro-pocillos o de microtitulación, u otras matrices de micro-contenedores. En ciertas realizaciones, las placas de micro-pocillos o de microtitulación, o micro-contenedores, incluyen oligonucleótidos o análogos de oligonucleótidos monocatenarios que son complementarios al oligonucleótido o análogo de oligonucleótido Q.

Tras la reacción o complejación de los compuestos con las proteínas, se puede retirar cualquier exceso de compuestos agregando un reactivo diseñado para actuar como "agente de captura". Por ejemplo, se permite que una molécula pequeña biotinizada, que tiene una funcionalidad idéntica o similar a la que ha reaccionado con el X seleccionado, reacciona con el exceso de compuesto. La exposición de esta mezcla a estreptavidina, unida a una perla magnética, permite la retirada del exceso del compuesto.

La hibridación de los productos compuesto-proteína con una secuencia complementaria se lleva a cabo según condiciones convencionales (por ejemplo, en presencia de sales caotrópicas para equilibrar los valores T_m de los diversos híbridos). Se puede eliminar por lavado todo el material que no haya hibridado y analizar el material hibridado.

En otras realizaciones, se puede efectuar una combinación selectiva de los productos de diferentes reactivos que contienen el resto X (por ejemplo, grupos X reactivos con amino y tiol; grupos X reactivos con anticuerpos y amino; grupos X de anticuerpo y lectina, etc.) para el análisis combinado en un solo ensayo (por ejemplo, en un único chip).

La Figura 1 muestra un método para separar y analizar una mezcla compleja de proteínas, mediante el uso de la espectrometría de masa MALDI-TOF. La exposición de un compuesto descrito en este documento a una mezcla de biomoléculas, que incluyen, sin limitaciones, proteínas (P1 a P4), proporciona una matriz de compuesto-proteína (NA = resto de oligonucleótido o resto de análogo de oligonucleótido, L = enlazador escindible, P = proteína). La separación de la matriz se lleva a cabo por hibridación de la porción Q de la matriz con una secuencia complementaria unida al soporte, tal como un chip de oligonucleótidos. A continuación, se analizan las proteínas (P1 a P4) por espectrometría de masa MALDI-TOF.

Cuando la complejidad de una mezcla de biomoléculas que incluye, sin limitaciones, proteínas, es baja, se pueden aplicar métodos de cromatografía de afinidad o de filtración de afinidad para separar los productos compuesto-proteína de la mezcla de proteínas. Si las proteínas que se deben analizar fueron marcadas con radiactividad antes (o después) de la reacción con el compuesto, pero antes de la hibridación, se podrían detectar también estas proteínas en la matriz. De esta forma, se pueden detectar las posiciones portadoras de un híbrido antes de escanear la matriz con espectrometría de masa MALDI-TOF, y se reduce al mínimo el tiempo necesario para analizar la matriz. Se pueden aplicar espectrómetros de masa de diversos tipos para analizar las proteínas (por ejemplo, lineales o con reflexión, con o sin extracción retardada, con TOF, Q-TOF o analizador de la Transformada de Fourier con láseres de diferentes longitudes de onda y fases de muestra xy).

Los formatos de espectrómetros de masa para usar en esta invención son de ionización por desorción láser asistida por matrices (MALDI), electro-nebulización continua o pulsada (ES), ionización, nebulización de iones, termonebulización, o espectrometría de masa por impacto masivo en racimo, y un formato de detección tal como tiempo de vuelo (TOF) lineal, tiempo de vuelo de reflectrón, cuádruple simple, cuádruple múltiple, de sector magnético simple, de sector magnético múltiple, de transformada de Fourier, resonancia de ciclotrón iónico (ICR), trampa iónica y sus combinaciones, tales como la espectrometría MALDI-TOF. Por ejemplo, para la ES, las muestras, disueltas en agua o en una solución tampón volátil, se inyectan continua o discontinuamente en una interface de ionización de presión atmosférica (API) y, a continuación, se analiza la masa por un cuádrupolo. La generación de múltiples picos iónicos, que se puede obtener usando la espectrometría de masa ES, puede incrementar la exactitud de la determinación de masa. Es posible obtener una información todavía más detallada de la estructura específica usando una configuración cuádrupolo MS/MS.

Los expertos en la técnica conocen bien los métodos para llevar a cabo MALDI. También se conocen numerosos métodos para mejorar la resolución. Por ejemplo, la resolución en la espectrometría de masa MALDI-TOF se puede mejorar reduciendo el número de colisiones de alta energía durante la extracción de iones (véase, por ejemplo, Juhasz *et al.* (1996), *Analysis, Anal. Chem.* 68:941-946; véanse también, por ejemplo, en la Patente de EE.UU. No. 5.777.325, la Patente de EE.UU. No. 5.742.049, la Patente de EE.UU. No. 5.654.545, la Patente de EE.UU. No. 5.641.959, la Patente de EE.UU. No. 5.654.545, la Patente de EE.UU. No. 5.760.393, y la Patente de EE.UU. No. 5.760.393 las descripciones de los protocolos de MALDI y de extracción retardada).

En la espectrometría de masa MALDI-TOF se pueden utilizar diversos analizadores de masa, por ejemplo, configuraciones de los instrumentos de sector magnético/deflexión magnética en modo de cuádrupolo simple o triple (MS/MS), transformada de Fourier y tiempo de vuelo (TOF), incluido el tiempo de vuelo ortogonal (O-TOF), como son conocidas en la técnica de espectrometría de masa. Para el proceso de desorción/ionización, se pueden usar numerosas combinaciones de matriz/láser. También se pueden emplear configuraciones de trampa iónica y reflectrones.

MALDI-MS requiere la incorporación de la biomolécula en una matriz. Se ha llevado a cabo en polipéptidos y en ácidos nucleicos mezclados en una matriz sólida (es decir, cristalina). La matriz se selecciona de forma que absorba la radiación láser. En estos métodos, se utiliza un láser tal como láser UV o IR para incidir sobre la mezcla de biopolímero/matriz, que se cristaliza en el extremo de una sonda u otro soporte apropiado, efectuando de este modo la

ES 2 319 634 T3

desorción e ionización del biopolímero. Además, se ha llevado a cabo el procedimiento MALDI-MS en polipéptidos, glicerol y otros líquidos como matriz.

5 Es posible diseccionar selectivamente una mezcla compleja de proteínas y, al recopilar todos los datos, analizarla por completo mediante el uso de compuestos con diferentes funcionalidades X. Las proteínas presentes en una mezcla de origen biológico se pueden detectar porque todas las proteínas tienen funcionalidades reactivas presentes en su superficie. Si en cada posición de la matriz compuesto-proteína existe la misma proteína escindible bajo las mismas condiciones que L, o se agrega sin unión covalente al soporte sólido y sirve como estándar interno de peso molecular, se puede determinar la cantidad relativa de cada proteína (o péptido, si la matriz de proteína ha sido digerida enzimáticamente). Este procedimiento permite la detección de variaciones de las proteínas expresadas cuando se comparan tejidos de individuos sanos y enfermos, o cuando se compara el mismo tejido bajo diferentes condiciones fisiológicas (por ejemplo, estudios dependientes del tiempo). El procedimiento permite, igualmente, la detección de cambios de las proteínas expresadas cuando se comparan secciones diferentes de tejidos (por ejemplo, tumores), que se pueden obtener, por ejemplo, por biopsia láser.

15 Se pueden estudiar las interacciones proteína-proteína y proteína-molécula pequeña haciendo contactar la matriz compuesto-proteína con una mezcla de las moléculas de interés. En este caso, se usará un compuesto que no tenga un enlace L escindible, o que tenga un enlace L que sea estable bajo las condiciones de MS MALDI-TOF. El subsiguiente escaneo de la matriz con el espectrómetro de masa demuestra qué proteínas hibridadas de la matriz de proteínas han interactuado efectivamente con las mezclas de proteína o molécula pequeña de interés.

20 También es posible el análisis usando la conocida metodología de 2 híbridos, que se puede detectar por espectrometría de masa. Véanse, por ejemplo, las Patentes de EE.UU. Nos. 5.512.473, 5.580.721, 5.580.736, 5.955.280, 5.695.941. Véase también Brent *et al.* (1996), *Nucleic Acids Res.* 241(17):3341-3347.

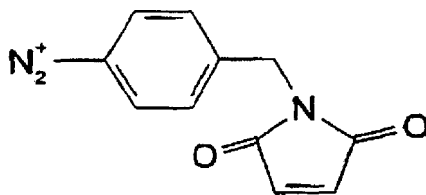
25 En las realizaciones anteriores, incluidas aquellas en las que Z contiene un enlace escindible, los compuestos pueden contener una etiqueta modificadora de masa. En estas realizaciones, la etiqueta modificadora de masa se usa para analizar las diferencias de estructura (por ejemplo, modificaciones de la cadena lateral tal como fosforilación o defosforilación) y/o los niveles de expresión de biopolímeros, proteínas incluidas. En una realización, se usan dos compuestos (o dos conjuntos de compuestos que tienen restos B permutados idénticos), que difieren solamente por la presencia o ausencia de una etiqueta modificadora de masa (o tienen dos etiquetas de masa con diferencias apropiadas de masa). Un compuesto (o un conjunto de compuestos) se hace reaccionar con tejido “sano” y el o los compuestos de masa modificada se hacen reaccionar con el tejido “enfermo”, bajo condiciones por lo demás idénticas. Se combinan y analizan las dos reacciones en un modo dúplex. Las diferencias de masa pondrán de manifiesto las proteínas con alteraciones estructurales o expresadas en cantidades diferentes en el tejido enfermo. Se pueden usar tres o más etiquetas modificadoras de masa en reacciones separadas, combinándolas para análisis en múltiplex para realizar un seguimiento de las diferencias durante las diferentes etapas del desarrollo de la enfermedad (es decir, etiqueta modificadora de masa 1 en el tiempo 1, etiqueta modificadora de masa 2 en el tiempo 2, etc.) o, de forma alternativa, para analizar diferentes secciones de tejido de un tejido enfermo, tal como una muestra tumoral.

40 En realizaciones adicionales, se puede lograr la selectividad en la reacción de los compuestos propuestos en esta invención con una mezcla de biopolímeros, incluidas proteínas, llevando a cabo las reacciones bajo control cinético y retirando partes alícuotas a diferentes intervalos de tiempo. De manera alternativa, se pueden realizar diferentes reacciones paralelas (todas diferentes en el resto B del grupo Q) y llevadas a cabo con diferentes relaciones estequiométricas o detenidas a diferentes intervalos de tiempo y analizadas por separado.

50 En realizaciones en las que los compuestos propuestos en esta invención poseen un grupo luminiscente o colorimétrico, el conjugado inmovilizado de compuesto-biomolécula se puede visualizar sobre el soporte insoluble antes del análisis. La visualización del conjugado ofrece información sobre las zonas en que la hibridado el conjugado (por ejemplo, para el subsiguiente análisis de espectrometría de masa MALDI-TOF). En ciertas realizaciones, con reactivos seleccionados, se puede determinar la cantidad de una proteína dada de experimentos separados (por ejemplo, sano frente a enfermo, tiempo 1 frente a tiempo 2, etc.), usando tinciones que se pueden diferenciar por espectrometría.

55 En otra realización, los métodos se llevan a cabo etiquetando los biopolímeros que se deben analizar, incluidas, sin limitaciones, las proteínas, con más de uno - en una realización, con tres a cinco - de los compuestos propuestos en este documento. Estos compuestos tendrían una funcionalidad diseñada para dianizar características químicas menores de las biomoléculas, en lugar de una característica macromolecular. Véase, por ejemplo, la Figura 8. Estas características químicas menores incluyen, pero sin estar limitadas a ellas, NH₂, SH, SS (después de retirar SH, se puede dianizar SS, por ejemplo, con oro), y OH. En un ejemplo no limitante, el OH fenólico de tirosina se captura de forma selectiva usando un diazocompuesto tal como una sal de aril-diazonio. En esta realización, la reacción se puede llevar a cabo en agua. Por ejemplo, se podría utilizar una sal funcionalizada de diazonio cuando la funcionalidad permite la captura subsiguiente de un compuesto propuesto en esta invención, proporcionando de este modo una biomolécula marcada con oligonucleótido. Una sal funcionalizada de diazonio de este tipo es:

65



5

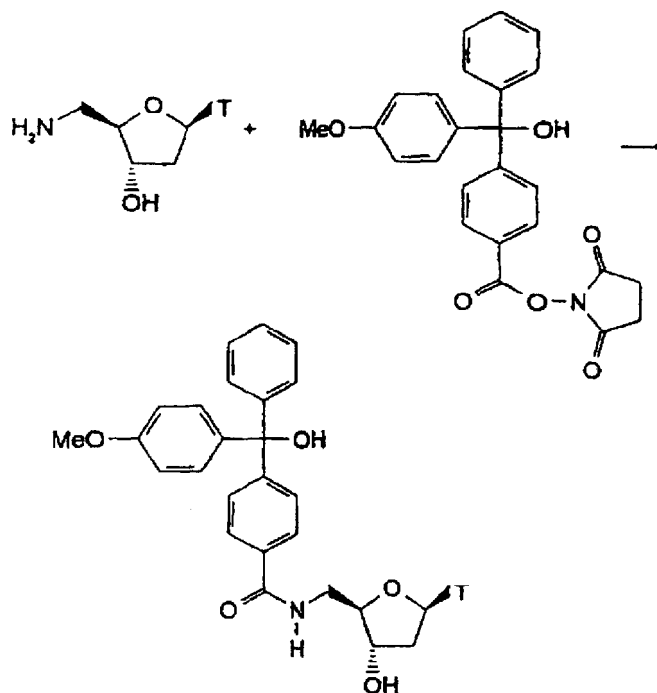
10 A continuación, la molécula modificada con este reactivo se marca con un oligonucleótido que tiene un residuo dieno. Los expertos en la técnica podrán apreciar que, en estas realizaciones, se pueden utilizar muchas parejas de reactivos además de dienófilo/dieno. En el caso de dienófilo/dieno, la reacción del dienófilo con dieno se puede llevar a cabo en presencia de otros muchos grupos funcionales, incluidos los oligonucleótidos activados con N-hidroxisuccinimido que reaccionan con un grupo NH_2 . De esta forma, estas dos reacciones específicas de marcado se pueden efectuar de manera simultánea (es decir, en una mezcla de reacción). Véase, por ejemplo, la Figura 11.

15

20 Subsiguientemente, las biomoléculas multi-etiquetadas se hibridan sobre una matriz de oligonucleótidos antisentido; en una realización, un chip que contiene una matriz de oligonucleótidos antisentido. En general, la especificidad de la hibridación aumenta usando compuestos de la invención en los que Z es un dendrímero. Véase, por ejemplo, la Figura 9. En una realización, Z es una estructura dendrítica que contiene hasta aproximadamente 6 ramas. En esta realización, los métodos propuestos en este documento permiten la separación entre biomoléculas marcadas con, por ejemplo, cinco oligo-etiquetas, en las que cuatro son similares y una es diferente. Véase, por ejemplo, la Figura 10.

25 En realizaciones en las que los compuestos utilizados en los métodos proporcionados en esta invención son insolubles o escasamente solubles en agua o soluciones tampón acuosas, se agregan disolventes orgánicos a las soluciones tampón para mejorar la solubilidad. En una realización, la relación solución tampón:disolvente orgánico es tal que no se produce la desnaturalización de la biomolécula. En otra realización, los disolventes orgánicos usados incluyen, sin limitaciones, acetonitrilo, formamida y piridina. En otra realización, la relación de solución tampón:disolvente orgánico es de aproximadamente 4:1. Para determinar si se necesita un disolvente orgánico, se mide la velocidad de reacción de los compuestos de la invención con una amina hidrosoluble tal como 5'aminotimidina. Por ejemplo, la siguiente reacción se efectúa en una variedad de mezclas disolventes, bien conocida por los expertos en la técnica para determinar las condiciones óptimas para el subsiguiente etiquetado y análisis biomoleculares:

30



35

40

45

50

55

60

2. Análisis de Fenotipo

65 Las matrices permiten un enfoque integral de arriba abajo del análisis del proteoma y otras biomoléculas. Como se ha señalado, las matrices y los métodos de uso proporcionan una forma imparcial de analizar biomoléculas, dado que los métodos no necesariamente evalúan clases específicas de dianas, sino que, más bien, detectan o identifican cambios en las muestras. Los cambios identificados incluyen variaciones estructurales que están relacionadas con las secuencias principales y las modificaciones, incluidas las modificaciones post-traduccionales. Además, puesto que los compuestos de captura pueden incluir una función de solubilidad, se les puede diseñar para reacciones bajo condiciones hidrófobas, permitiendo así el análisis de moléculas unidas a membrana y asociadas a membrana, en especial, proteínas.

ES 2 319 634 T3

Los problemas del análisis del proteoma surgen de la variación genética que no está relacionada con un fenotipo diana, variación del proteoma debido a diferencias tales como sexo, edad, estado metabólico, las mezclas complejas de células en los tejidos diana, y variaciones del estado del ciclo celular. De esta forma, para identificar o detectar cambios, por ejemplo, relacionados con una enfermedad, entre los componentes biomoleculares de tejidos y células, puede ser importante la homogeneidad de la muestra. Para proporcionar homogeneidad, se comparan células con diferentes fenotipos, tales como enfermas frente a sanas, del mismo individuo. Como consecuencia, las diferencias en los patrones de las biomoléculas se pueden atribuir a diferencias en el fenotipo y no a diferencias entre individuos. Por lo tanto, se pueden obtener muestras de un solo individuo, y se separan células con diferentes fenotipos, por ejemplo, sanas frente a enfermas o que responden frente a las que no lo hacen. Adicionalmente, las células se pueden sincronizar o congelar en un estado metabólico para reducir aún más las diferencias de fondo.

De este modo, se pueden usar matrices para identificar proteínas específicas para un fenotipo, o modificaciones de las mismas u otras biomoléculas específicas para el fenotipo y sus patrones. Esto se puede lograr comparando muestras de biomoléculas de células o tejidos con un fenotipo, con las células equivalentes de muestras de biomoléculas de células o tejidos con otro fenotipo. Se comparan los fenotipos en células del mismo individuo y tipo celular. En particular, se comparan células primarias, cultivos de células primarias y/o células sincronizadas. Se pueden identificar los patrones de unión de las biomoléculas de las células frente a los miembros del compuesto de captura de la colección, usándolos como una firma o perfil de un estado patológico o sano, u otros fenotipos. Se puede identificar la biomolécula unida particular, tal como una proteína, y también se pueden identificar proteínas y nuevas marcas asociadas a una enfermedad, tales como proteínas especiales o estructuras de las mismas. El Ejemplo 6 ofrece un ejemplo de realización en el que las células se separan. Véase también la Figura 19.

Los fenotipos de comparación incluyen, pero no están limitados a ellos:

1) Muestras de células o tejidos enfermos frente a sanos para identificar proteínas u otras biomoléculas asociadas con la enfermedad, o que son marcas de una enfermedad;

2) Muestras de pacientes que responden a un medicamento y que no lo hacen (es decir, 20-30% de pacientes con melanoma maligno responden al α -interferón, y otros no lo hacen), para identificar las biomoléculas indicativas de respuesta;

3) Muestras de células o tejidos con un perfil de toxicidad a medicamentos o condiciones ambientales, para identificar las biomoléculas asociadas con la respuesta o un marcador de la respuesta; y

4) Muestras de células o tejidos expuestos a cualquier condición o que exhiben cualquier fenotipo, con el fin de identificar biomoléculas, tales como proteínas, asociadas con la respuesta o fenotipo, o que son marcadoras de ello.

Por lo general, las muestras para cada fenotipo se obtienen del mismo organismo, por ejemplo, del mismo mamífero, de modo que las células son esencialmente comparables y cualquier variación debe reflejar un cambio debido al fenotipo y no al origen de las células. Las muestras se pueden obtener de células primarias (o tejidos). En todos los casos, las muestras se pueden obtener del mismo individuo, ya sea antes de la exposición o tratamiento, o de tejido sano y no patológico, al objeto de permitir la identificación de biomoléculas asociadas con el fenotipo.

Las células se pueden separar por cualquier método apropiado que permita la identificación de un fenotipo particular y, entonces, la separación de las células basada en él. Se puede utilizar cualquier método de separación tal como, por ejemplo, sembrado, sembrado negativo (en el que se capturan células no deseadas y las células deseadas se mantienen en el sobrenadante), en el que se recuperan células vivas. Estos métodos incluyen, pero no están limitados a ellos:

1) Citometría de flujo;

2) Captura específica;

3) Sembrado negativo, en el que se capturan las células no deseadas y las células diana permanecen en el sobrenadante, y se recuperan células vivas para el análisis; y

4) Microdissección por Captura con Láser (LCM) (Arcturus, Inc. Mountain View, CA).

De este modo, los criterios de separación incluyen, sin limitaciones, potencial de membrana, flujo iónico, actividad enzimática, marcadores de la superficie celular, marcadores de enfermedad, y otros criterios de este tipo que permiten la separación de células de un individuo basada en el fenotipo.

a) Ejemplos de Métodos de Separación

Microdissección por Captura con Láser (LCM) Arcturus, Inc. Mountain View, CA): utiliza una plataforma de microscopio combinada con un láser IR de baja energía para activar una película de captura plástica sobre las células de interés seleccionadas. A continuación, las células se levantan suavemente del tejido circundante. Este método impide que las células micro-disseccionadas o el tejido circundante absorban la radiación láser, con lo que se garantiza

la integridad del ARN, ADN y las proteínas preparadas a partir de las muestras micro-diseccionadas para el análisis posterior.

2) Citometría de flujo para separación

La citometría de flujo es un método, semejante a la microscopia fluorescente, en el que las mediciones se efectúan sobre partículas (células) en suspensión líquida, que fluyen de forma individual a través de un rayo láser enfocado, a una velocidad de hasta varios miles de partículas por segundo. Se mide la luz dispersada y la fluorescencia emitida por las partículas (células), se filtra, se digitaliza y se envía a un ordenador para su análisis. Típicamente, la citometría de flujo mide la unión de una sonda marcada con fluorocromo a las células, y la comparación de la fluorescencia resultante con respecto a la fluorescencia de base de células no teñidas. Las células se pueden separar usando una versión de la citometría de flujo, la separación por flujo, en la que las partículas (células) se separan y recuperan de una suspensión, basándose en propiedades medidas en el flujo. Las células que se recuperan a través de la separación de flujo son viables y se pueden recolectar bajo condiciones de esterilidad. Típicamente, se recuperan subpoblaciones puras por encima de 99,5% (véanse las Figuras 19a y 19b).

La citometría de flujo permite distinguir las células usando diversos parámetros, que incluyen características físicas y/o químicas asociadas con las células o propiedades de reactivos o sondas asociados a las células, en donde cualquiera de ellos se mide por sensores instrumentales. Separación: Para la identificación y clasificación primarias de las poblaciones celulares se usan las dispersiones hacia adelante y laterales de células vivas con respecto a las muertas. Los parámetros de dispersión se usan para excluir los desechos, las células muertas y los agregados no deseados. En una muestra de sangre periférica o de médula ósea, se pueden definir las poblaciones de linfocitos, monocitos y granulocitos, clasificándolas y analizándolas por separado, basándose en la dispersión hacia adelante y lateral. Las células recuperadas por separación de flujo son viables y pueden ser recogidas bajo condiciones de esterilidad. Las subpoblaciones recuperadas puras están típicamente por encima de 99,5%.

Los experimentos de separación celular frecuentes comprenden a menudo ensayos de inmunofluorescencia, es decir, tinción de células con anticuerpos conjugados con tinciones fluorescentes para la detección de antígenos. Además, la separación se puede llevar a cabo usando construcciones de GFR-informador para aislar poblaciones puras de células que expresan un gen/construcción determinado.

a. Fluorescencia

La medición del parámetro fluorescente permite la investigación de estructuras y funciones celulares basada en la tinción directa, reacciones con sondas marcadas con fluorocromo (por ejemplo, anticuerpos), o la expresión de proteínas fluorescentes. Las señales de fluorescencia se pueden medir como parámetros únicos y múltiples, correspondientes a diferentes longitudes de onda de excitación láser y emisión de fluorescencia. Cuando se utilizan simultáneamente diferentes fluorocromos, puede producirse un derrame de señales entre los canales de fluorescencia. Esto se corrige mediante la compensación. Determinadas combinaciones de fluorocromos no pueden ser usadas simultáneamente; los expertos en la técnica pueden identificar tales combinaciones.

b. Inmunofluorescencia

La inmunofluorescencia implica la tinción de células con anticuerpos conjugados a tinciones fluorescentes tales como FITC (fluoresceína), PE (ficoeritrina), APC (aloficocianina) y conjugados en tándem basados en PE (R670, CyChrome y otros). Los antígenos de la superficie celular son las dianas habituales de este ensayo, pero los anticuerpos pueden estar dirigidos también contra antígenos o citoquinas del citoplasma.

La tinción de ADN se utiliza principalmente para trazar el perfil del ciclo celular, o como método para medir la apoptosis. El yoduro de propidio (PI), la tinción de ADN más frecuentemente usada, no puede penetrar en las células vivas y se le puede usar, por lo tanto, en ensayos de viabilidad. Para los ensayos de ciclo celular o de apoptosis con empleo de PI, las células se deben fijar previamente para que tenga lugar la tinción (véase el protocolo). La cantidad relativa de tinción PI-ADN corresponde a la proporción de células en fases G0/G1, S y G2/M, en donde cantidades menores de tinción indican células apoptóticas/necróticas. La tinción con PI se puede llevar a cabo simultáneamente con ciertos fluorocromos tales como FITC y GFP en ensayos para la caracterización adicional de la apoptosis o la expresión de genes.

La Expresión de Genes y la Transfección se pueden medir de manera indirecta usando un gen informador en la construcción. Se pueden usar construcciones de tipo Proteína Verde Fluorescente (EGFP, proteínas fluorescentes rojas y azules) y β -galactosidasa, por ejemplo, para cuantificar las poblaciones de las células que expresan el gen/construcción. En la actualidad, hay disponibles mutantes de GFP que pueden ser excitados a frecuencias comunes, pero emiten fluorescencia a diferentes longitudes de onda. Esto permite la medición de la co-transfección, así como la detección simultánea de la expresión de genes y anticuerpos. Se deben incluir controles negativos (de fondo) apropiados para los experimentos en los que intervienen construcciones de tipo GFP. Los controles incluyen, por ejemplo, el mismo tipo de célula, usando el inserto génico menos la construcción de tipo GFP.

3) *Estudios Metabólicos y otros estudios*

La anexina-V se puede marcar con diversos fluorocromos para identificar células en estadios tempranos de apoptosis. CFSE se une a las membranas celulares y se distribuye uniformemente cuando se dividen las células. Se puede contar el número de divisiones que sufren las células en un período de tiempo. Se puede usar CFSE junto con ciertos fluorocromos para inmunofluorescencia. Se puede medir el flujo de calcio usando marcadores Indo-1. Esto se puede combinar con la tinción inmunofluorescente. Es posible efectuar ensayos de conjugación intercelular usando combinaciones de tinciones tales como calceína o hidroetidina.

b) *Sincronización de ciclos celulares*

Una vez obtenidas las células separadas pueden ser cultivadas y se pueden sincronizar o congelar en un estado metabólico determinado. Se refuerza así la capacidad para identificar biomoléculas específicas del fenotipo. Estas células se pueden separar por los métodos anteriores, incluida la citometría de flujo. Adicionalmente, las células en el mismo ciclo celular, le mismo estado metabólico u otro estado sincronizado se pueden separar en grupos, usando citometría de flujo (véase la Figura 19C).

Los ciclos celulares se pueden sincronizar o congelar por una variedad de métodos que incluyen, sin limitaciones, quelación celular de iones críticos por ejemplo, por la retirada de magnesio, cinc, manganeso, cobalto y/u otros iones que llevan a cabo funciones específicas, mediante EDTA u otros quelantes (véanse, por ejemplo, los EJEMPLOS). Otros métodos incluyen controlar distintas vías metabólicas o bioquímicas. La Figura 18 muestra ejemplos de puntos de regulación de mecanismos de control metabólico para la sincronización o "congelación" celular. El Control Metabólico para sincronizar células incluye, sin limitaciones, los siguientes procedimientos:

1) Control de la expresión génica;

2) Regulación de reacciones enzimáticas;

3) Control negativo: inhibición de retroalimentación o regresión del producto final y la inducción de enzimas son mecanismos de control negativo que conducen a un descenso de la transcripción de proteínas;

4) Control positivo: la represión catabólica se considera una forma de control positivo, porque determina un incremento de la transcripción de proteínas;

5) Control de traducción de proteínas individuales:

a) Los oligonucleótidos que hibridan con el sitio 5' del casquete inhiben la síntesis de proteínas mediante la inhibición de la interacción inicial entre el ARNm y la subunidad 40S de los ribosomas;

b) los oligonucleótidos que hibridan con la UTR 5' hasta el codón de inicio de la traducción, éste incluido, inhiben el barrido de la subunidad 40S (o 30S), o ensamblaje del ribosoma completo (80S para sistemas eucarióticos o 70S para sistemas bacterianos);

6) Control de la modificación post-traduccional;

7) Control de enzimas alostéricas, en donde el sitio activo se une al sustrato de la enzima y lo convierte en un producto. El sitio alostérico está ocupado por alguna molécula pequeña que no es un sustrato. Si la proteína es una enzima, cuando el sitio alostérico está ocupado, la enzima es inactiva, es decir, la molécula efectora reduce la actividad de la enzima. Algunas enzimas alostéricas de componentes múltiples tienen varios sitios ocupados por diversas moléculas efectoras que modulan la actividad sobre una serie de condiciones.

3. *Análisis de proteínas de baja abundancia*

Importantes marcadores y dianas asociados con enfermedades podrían ser proteínas de baja abundancia, que podrían escapar a la detección por espectrometría de masa. Para asegurar la detección, se puede llevar a cabo un primer experimento de expresión de un compuesto de captura. La matriz resultante de proteínas capturadas se hace reaccionar con una tinción no selectiva, tal como una tinción fluorescente, que destacará o hará visibles más proteínas en la matriz. La tinción puede proporcionar una estimación semi-cuantitativa de la cantidad de una proteína. Se puede determinar la cantidad de diferentes proteínas detectadas por la tinción y comparar, a continuación, el número detectado por el análisis de espectrometría de masa. Si se detectan más proteínas usando la tinción, se pueden repetir los experimentos usando un número inicial de células más alto, de modo que se puedan detectar las proteínas de baja abundancia, identificándolas por el análisis de espectrometría de masa.

Por ejemplo, las proteínas de mantenimiento tales como actina y otras semejantes se encuentran presentes en gran abundancia y pueden enmascarar las proteínas de baja abundancia. Los compuestos de captura u otros compuestos de purificación, diseñados para capturar o eliminar las proteínas o biomoléculas de gran abundancia de una mezcla, llevan a cabo esta acción antes de usar la recolección para evaluar los componentes de una mezcla. Una vez retiradas las proteínas de gran abundancia, las de baja abundancia alcanzan una concentración efectivamente mayor y pueden ser

detectadas. Por lo tanto, estos métodos tienen dos fases: una primera etapa para capturar los componentes abundantes de las mezclas de biomoléculas, tales como las actinas. Por ejemplo, se puede hacer contactar un lisado celular con moléculas de captura que incluyan un grupo de reactividad tal como una biotina u otra función de reactividad general asociada con un grupo de separación para eliminar dichas proteínas abundantes, y utilizar, entonces, una colección adecuada de compuestos de captura para identificar los compuestos de menor abundancia que permanecen en el lisado.

Asimismo, tal como se ha señalado anteriormente, se pueden diseñar compuestos de captura, por ejemplo, por una selección apropiada de W, para que interactúen con organelas intactas antes de degradarlas en células que han sido sometidas a un lisado suave o a un tratamiento semejante que permita el acceso a las organelas y a las membranas internas. A continuación, se pueden degradar las organelas capturadas, por ejemplo sobre una membrana artificial que puede ser incluida, tal como una bicapa lipídica o recubrimiento micelar, para capturar las proteínas de la organela y otras biomoléculas en un ambiente que conserva su estructura tridimensional. Se pueden analizar estas proteínas capturadas. Esto permite que los compuestos de captura interactúen con las proteínas y otras biomoléculas capturadas en su estructura terciaria nativa.

4. Monitorización de la conformación de proteínas como indicador de enfermedades

Las matrices y/o loci de las mismas se pueden utilizar para detectar o distinguir conformeros específicos de proteínas. Así, por ejemplo, si una conformación particular de una proteína se asocia con una enfermedad (o estado saludable), las matrices o los loci miembros de las mismas pueden detectar un conformero o distinguir conformeros sobre la base de un patrón de fijación a los compuestos de captura en una recolección. De este modo, las matrices y/o sus miembros se pueden utilizar para detectar enfermedades por proteínas de conformación alterada (o enfermedades por agregación de proteínas), en las que la proteína o polipéptido asociado con la enfermedad muestra una conformación asociada con la enfermedad. Los métodos y matrices propuestos en este documento permiten detectar un conformero asociado con una enfermedad que debe ser detectada. Estas enfermedades incluyen, sin limitaciones, enfermedades amiloides y enfermedades neurodegenerativas. Otras enfermedades y proteínas asociadas que exhiben dos o múltiples conformaciones diferentes en las que al menos una conformación se asocia con una enfermedad, se exponen en la tabla siguiente:

TABLA

Enfermedad	Proteína insoluble
Enfermedad de Alzheimer	APP, A β , σ 1-anti-quimiotripsina, tau, componente no-A β , presenelina 1, presenelina 2, apoE
Enfermedades por priones, incluida, sin limitaciones, enfermedad de Creutzfeldt-Jakob, "scrapie" (en ovejas), encefalopatía espongiforme bovina	PrP ^{Sc}
Esclerosis lateral amiotrófica (ELA)	Superóxido dismutasa (SOD) y neurofilamento
Enfermedad de Pick	Cuerpo de Pick
Enfermedad de Parkinson	α -sinucleína en cuerpos de Lewis
Demencia fronto-temporal	Tau en fibrillas
Diabetes Tipo II	Amilina
Mieloma múltiple	Cadena IgGL
Discrasias de células plasmáticas	
Polineuropatía amiloidótica familiar	Transtiretina
Carcinoma medular del tiroides	Procalcitonina
Insuficiencia renal crónica	β ₂ -microglobulina
Insuficiencia cardíaca congestiva	Factor natriurético auricular
Amiloidosis Cardíaca Senil y sistémica	Transtiretina
Inflamación crónica	Amiloide A sérico
Aterosclerosis	ApoAI
Amiloidosis familiar	Gelsolina
Enfermedad de Huntington	Huntington

Las matrices se pueden poner en contacto con una mezcla de los confórmers y se pueden identificar los miembros que se fijan o retienen cada forma, asociando de este modo un patrón con cada conformero. De manera alternativa, se pueden identificar los que se unen a un solo conformero, tal como el conformero asociado con una enfermedad, y se pueden usar sub-colecciones de una o múltiples matrices de este tipo como reactivo diagnóstico para la enfermedad.

5. Identificación de moléculas pequeñas e investigación de la interacción biomolécula-biomolécula

Las biomoléculas tales como proteínas se separan usando una interacción covalente o no covalente con compuestos de captura inmovilizados. Entonces, se pueden usar matrices unidas a las biomoléculas, tales como los procedentes de lisados celulares, para rastrear bibliotecas u otras mezclas de candidatos a medicamento, o para rastrear, adicionalmente, mezclas de biomoléculas para determinar lo que se une a las biomoléculas fijadas.

Los complejos de biomolécula de captura-biomolécula o los complejos de biomolécula-candidato a medicamento se pueden analizar para identificar las vías bioquímicas e identificar, asimismo, dianas para el candidato a medicamento.

Por ejemplo, se exponen interacciones proteína-proteína o proteína-biomolécula a compuestos de ensayo, típicamente moléculas pequeñas que incluyen moléculas pequeñas orgánicas, péptidos, péptido-miméticos, moléculas antisentido o ARNs, anticuerpos, fragmentos de anticuerpos, anticuerpos recombinantes y sintéticos y sus fragmentos, y otros compuestos de este tipo que pueden actuar como candidatos a medicamento o compuestos guía. Las moléculas pequeñas fijadas se identifican por espectrometría de masa u otros métodos analíticos.

F. Sistemas

En realizaciones adicionales, los compuestos y los métodos descritos en este documento han sido diseñados para ser situados en un sistema integrado que estandariza y automatiza las siguientes etapas de procedimientos:

- Aislamiento de biomoléculas de una fuente biológica, incluido el aislamiento de proteínas de lisados celulares (lisis, digestión enzimática, precipitación lavado);
- Opcionalmente, retirada del material de bajo peso molecular;
- Opcionalmente, determinación de partes alícuotas de la mezcla de biomoléculas, tal como una mezcla de proteínas;
- Reacción de la mezcla de biomoléculas, tal como una mezcla de proteínas, con compuestos de diferente reactividad química (X) y diversidad de secuencias (B) que se proponen en la invención; esta etapa se puede llevar a cabo en paralelo, usando partes alícuotas de la mezcla de biomoléculas;
- Opcionalmente, retirada del exceso de compuesto;
- Hibridación del conjugado compuesto-biomolécula tal como un conjugado de compuesto-proteína, a oligonucleótidos o análogos de oligonucleótidos de cadena sencilla que son complementarios al resto Q del compuesto; los oligonucleótidos o análogos de oligonucleótidos de cadena sencilla se presentan, opcionalmente, en formato de matriz y están inmovilizados, opcionalmente, en un soporte insoluble;
- Opcionalmente, subsiguiente tratamiento químico o enzimático de la matriz de proteínas;
- Análisis de la matriz de biomoléculas, que incluye, sin limitaciones, las etapas de (i) depósito de la matriz, y (ii) espectrometría de masa MALDI-TOF paso a paso, usando un espectrómetro de masa matricial (con o sin estándar interno, por ejemplo, en un chip, de peso molecular para calibración y cuantificación).

El sistema incluye las colecciones propuestas en este documento, opcionalmente, matrices de estas colecciones, software de control de los procedimientos de preparación de muestras y análisis instrumental, y para el análisis de los datos resultantes, e instrumentación, tal como un espectrómetro de masa, para el análisis de las biomoléculas. El sistema incluye otros dispositivos tales como instrumentos de cromatografía líquida, de modo que la mezcla de proteínas se separa al menos de forma parcial. El eluyente se recoge en una serie continua de partes alícuotas en, por ejemplo, placas de microtitulación, y cada parte alícuota se hace reaccionar con un compuesto de captura propuesto.

En reacciones múltiple, las partes alícuotas en cada pocillo pueden reaccionar simultáneamente con uno o múltiples de los compuestos propuestos en este documento que, por ejemplo, difieren en X (es decir, funcionalidad específica para amino, tiol, lectina), en donde cada uno de ellos tiene un resto de selectividad Y específico y diferenciador, y en el grupo Q. La cromatografía se puede llevar a cabo en medio acuoso u orgánico. Las mezclas de reacción resultantes se combinan y analizan directamente. Alternativamente, se efectúan reacciones secundarias subsiguientes o estudios de interacción molecular antes del análisis, incluidos análisis de espectrometría de masa.

ES 2 319 634 T3

Los sistemas propuestos en la invención pueden contener una línea de ensamblaje, tales como robots pipeteadores en las etapas xy y módulos de suministro de reactivos/lavado, unidos con un dispositivo central de separación y un espectrómetro de masa terminal para el análisis e interpretación de datos. Los sistemas pueden estar programados para llevar a cabo etapas del procedimiento (véase, por ejemplo, Fig. 2), por ejemplo:

- 1) Cultivos celulares (o muestras de tejido) que se aportan en placas de microtitulación (MTPs) con 1, 2, ...im pocillos. A cada pocillo se agregan soluciones para la lisis celular, liberando así las proteínas. En algunas realizaciones, se incluyen etapas apropiadas de lavado, así como la adición de enzimas para digerir ácidos nucleicos y otros componentes no proteicos. En realizaciones adicionales, en lugar de MTPs regulares, se usan MTPs con placas de filtro en el fondo de los pocillos. Los desechos celulares se eliminan por filtración o centrifugación. Se agrega una solución de acondicionamiento para el proceso de separación adecuado y se carga por separado el material de cada pocillo en el dispositivo de separación.
- 2) La separación utiliza diferentes principios de separación tales como carga, tamaño molecular, adsorción, intercambio de iones y principios de exclusión molecular. Dependiendo del tamaño de la muestra, se usan dimensiones apropiadas tales como cromatografía líquida de alto rendimiento (HPLC) con célula de micro-boro. En ciertas realizaciones, se utiliza un proceso de flujo continuo y se extraen partes alícuotas continuas del efluente en MTP 1, 2, ... n.
- 3) Reacción con Reactivos del Proteoma. A su vez, cada MTP se transfiere a una Estación de Reactivo del Proteoma que contiene 1, 2, ...m reactivos que difieren únicamente en la parte de secuencia de oligonucleótidos (es decir, Q) y/o en la naturaleza química de la funcionalidad que reacciona con las proteínas (es decir, X). Si hay más de una MTP procedente de una muestra de tejido, se añade entonces reactivo 1 al mismo pocillo de las correspondientes MTPs 1, 2, ... n, es decir, en el pocillo A1, reactivo 2 en el pocillo A2, etc. En las realizaciones en las que las MTPs tienen 96 pocillos ($i = 1-96$), se suministran 96 Reactivos del Proteoma diferentes (es decir, 96 compuestos propuestos en la invención, $m = 1-96$) a través de 96 boquillas diferentes desde la Estación de Reactivos del Proteoma, para evitar la contaminación cruzada.
- 4) Combinación: Se desactiva el exceso de Reactivo del Proteoma, se combinan partes alícuotas de cada pocillo, pertenecientes a una y misma muestra de tejido, y el material restante se almacena bajo condiciones que conserven intacta la estructura (y, si es necesario, la conformación) de las proteínas, actuando así como MTPs Maestras para experimentos subsiguientes.
- 5) Se elimina el exceso de Reactivo del Proteoma en la muestra combinada, usando, por ejemplo, el sistema biotina/estreptavidina con perlas magnéticas y, seguidamente, se concentra el sobrenadante y se le acondiciona para la hibridación.
- 6) Transferencia a un Chip de Oligonucleótidos. Después de la etapa de lavado para eliminar el material no hibridado y otros materiales de bajo peso molecular, se agrega una matriz. De forma alternativa, antes de la adición de la matriz, se efectúa una digestión con, por ejemplo tripsina y/o quimotripsina. Después de eliminar por lavado la enzima y los productos de digestión, se agrega la matriz.
- 7) Transferencia del chip al espectrómetro de masa. En una realización, se lleva a cabo una espectrometría de masa MALDI-TOF. También se pueden aplicar otras configuraciones de espectrometría de masa adecuadas para el análisis de proteínas. El espectrómetro de masa tiene una etapa xy y rastrea, de este modo, cada posición para su análisis. El Reactivo del Proteoma se puede diseñar de manera que la mayor parte del reactivo (incluida la parte que hibrida con la matriz del chip de oligonucleótidos) se escinda antes o durante la espectrometría de masa y sea detectada, por tanto, en la zona de bajo peso molecular del espectro, estando bien separada del péptido (en caso de digestión enzimática) o de las señales de peso molecular de la proteína en el espectro de masa.
- 8) Por último, se pueden procesar las señales de peso molecular para reducir el ruido, sustraer el fondo y otras etapas de este tipo. Los datos obtenidos se pueden archivar e interpretar. Los valores de peso molecular de las proteínas (o de los péptidos obtenidos tras la digestión enzimática) se asocian con la información de las secuencias de ADN humano y la información sobre la secuencia derivada de proteína, procedente de las regiones codificadoras de la proteína. La interacción con las bases de datos disponibles revelará si las proteínas y sus funciones son ya conocidas. Si la función es desconocida, la proteína se puede expresar a partir de la secuencia conocida de ADN a escala suficiente, usando métodos convencionales, para determinar su función y subsiguiente localización en una vía bioquímica, donde desempeña su función metabólica en un individuo sano, o en la vía patológica para un individuo enfermo.

Dado que las placas maestras que contienen partes alícuotas de las diferentes proteínas dentro de una muestra determinada de tejido se conservan y están disponibles, se pueden llevar a cabo experimentos subsiguientes de una forma ahora preseleccionada, por ejemplo, las proteínas se expresan en la superficie del chip para los estudios de interacción proteína-proteína (biomolécula) para la validación de dianas y/o para estudiar la interacción con bibliotecas combinatorias para la selección de candidatos a medicamento.

G. Bioinformática

Los datos brutos generados del análisis de espectrometría de masa de compuesto-especies de proteínas se procesan por sustracción del fondo, reducción del ruido, calibración de peso molecular y refinamiento de picos (por ejemplo, integración de picos). Los valores de peso molecular de las proteínas escindidas o de los productos de digestión se interpretan y comparan con bases de datos existentes de proteínas, para determinar si la proteína en cuestión es conocida y, en caso afirmativo, qué modificaciones presenta (glicosilada o no glicosilada, fosforilada o no fosforilada, etc.). Se componen, comparan e interpretan los diferentes conjuntos de experimentos correspondientes a un conjunto de compuestos. Por ejemplo, un conjunto de experimentos utiliza un conjunto de compuestos con un resto X y diferentes restos Q. Este grupo de experimentos proporciona datos para una parte del proteoma, puesto que no todas las proteínas del proteoma reaccionarán con un resto X determinado. La superposición de los datos de este grupo de experimentos con los de otros conjuntos de ensayos, con diferentes restos X, ofrece datos para el proteoma completo.

Se investigan grupos de experimentos que comparan tejidos de individuos sanos y enfermos, o de diferentes etapas fisiológicas o de desarrollo (por ejemplo, progresión tumoral, dependencia de tratamientos farmacológicos para controlar los resultados de una terapia, inmunorrespuestas a infecciones virales o bacterianas), o diferentes zonas del tejido (por ejemplo, de un tumor), y los datos finales se archivan.

Los ejemplos siguientes se incluyen únicamente con fines de ilustración y no pretenden limitar el alcance de la invención.

Ejemplo 1

Ejemplos para $N^1_m-B_i-N^2_n$

a. N^1 y N^2 son tetrámeros idénticos, B es un trímero

$N^1 = N^2$, $m = n = 4$, $i = 3$, B = 64 permutaciones de secuencia

GTGC ATG GTGC

AAG

ACG

AGG

TTG

CTG

GTG

...

...

...

GGG

b. N^1 y N^2 son tetrámeros no idénticos, B es un tetrámero

$N^1 \neq N^2$, $m = n = 4$, $i = 4$, B = 256 permutaciones de secuencia

GTCC ATCG CTAC

AACG

ACCG

AGCG

....

....

....

GGGG

ES 2 319 634 T3

c. N¹ es un heptámero, N² es un octámero, B es un octámero

$N^1 \neq N^2$, $m = 7$, $N = 8$, $l = 8$, $b = 65.536$ permutaciones de secuencia

5

<u>GCTGCCC</u>	<u>ATTCGTAC</u>	<u>GCCTGCCC</u>
N¹	B	N²

10

Ejemplo 2

Separación de proteínas en una matriz de ADN

20

Proteína $N^1_m - B_i - N^2_n - (S^1)_t - M(R^{15})_a - (S^2)_b - L - X$, en la que B es un trímero:
 $m = n = 4$, $i = 3$, $t = b = 1$; las secuencias subrayadas con N^1 y N^2

25

|| CTGC ATG GTGC - S₁ - M(R¹⁵)_a - S₂ - L - X - Proteína 1
 || - CACG TAC CACG

30

|| CTGC AAG GTGC - S₁ - M(R¹⁵)_a - S₂ - L - X - Proteína 2
 || - CACG TTC CACG

35

|| CTGC ACG GTGC - S₁ - M(R¹⁵)_a - S₂ - L - X - Proteína 3
 || - CACG TGC CACG

40

|| ...
 || ...
 || ...
 || CTGC GGG GTGC - S₁ - M(R¹⁵)_a - S₂ - L - X - Proteína 64
 || - CACG CCC CACG

Ejemplo 3

45

Síntesis de matrices basadas en vidrio

50

Q-Z-X
|
Y

55

Se silanizan portaobjetos de vidrio según protocolos convencionales, usando trimetoxiaminopropilsilano. La carga superficial se reforzó activando, en primer lugar, los portaobjetos con fenileno-diisocianato seguid por el tratamiento con el dendrímero PAMAM de 4^a generación (poliamidoamina) (64 grupos amino) (véase, por ejemplo, <http://www.dendritech.com/pamam.html>). Seguidamente, los portaobjetos se acopla con un enlazador anfífilo de 18 átomos usando la química de fosforamidita. Los portaobjetos derivatizados se sometieron, entonces, a un método de síntesis de enmascaramiento para producir diferentes parches usando fosforamiditas triplicadoras (véase la Figura 4). De forma alternativa, las fosforamiditas se aplican para crear un gradiente continuo en cada dimensión del portaobjetos de vidrio. Las características que se pueden variar usando esta síntesis de enmascaramiento o gradiente incluyen, sin limitaciones, hidrofobia/hidrofília y carga positiva/negativa.

65

ES 2 319 634 T3

Ejemplo 4

I. Preparación de mezclas de proteínas a partir de células o a través de la traducción de proteínas de células o tejidos preparados de una biblioteca de ADNc

Las mezclas de proteínas se pueden dividir selectivamente en las técnicas de separación físicas o bioquímicas.

1. Preparación de combinaciones de proteínas de complejidad limitada usando cultivo celular o tejido

Las proteínas se aíslan de cultivos celulares o tejidos, de acuerdo con métodos bien conocidos por los expertos en la técnica. Las proteínas aisladas se purifican usando métodos bien conocidos por los expertos en la técnica (por ejemplo, TPAE, precipitación diferencial de proteínas (precipitación por sales, pH y polímeros iónicos), cristalización diferencial de proteínas y fraccionamiento a granel, electroforesis (PAGE, enfoque isoeléctrico, capilar) y cromatografía (inmunoadinidad, HPLC, LC)). Se recogen fracciones de columnas individuales que contienen mezclas de proteínas de complejidad limitada para su uso como antígenos.

2. Preparación de combinaciones de proteínas de complejidad limitada usando bibliotecas de expresión de ADNc (Figura 13)

a. Aislamiento de ARN

i. Aislamiento de ARN Total

Células o tejidos cultivados se homogeneizan en una solución desnaturalizante que contiene tiocianato de guanidina 4M. El homogeneizado se mezcla secuencialmente con acetato sódico 2M (pH 4), fenol y, por último, cloroformo/alcohol isoamílico o bromocloropropano. La mezcla resultante se centrifuga, dando una fase acuosa superior que contiene el ARNm total. Tras la precipitación de isopropanol, el granulado de ARN se disuelve en solución desnaturalizante (que contiene tiocianato de guanidina 4M), se precipita con isopropanol y se lava con etanol al 75%.

ii. Aislamiento de ARN Citoplasmático

Las células se lavan con solución salina fisiológica-tampón fosfato helada y se conserva en hielo para las subsecuentes manipulaciones. El granulado de células cosechadas se resuspende en una solución tampón de lisis que contiene el detergente no iónico Nonidet P-40. La lisis de las membranas plasmáticas se produce de forma casi inmediata. Se retiran los núcleos intactos con un breve ciclo de microcentrifugación, y se agrega dodecilsulfato sódico al sobrenadante citoplasmático para desnaturalizar la proteína. La proteína se digiere con proteasa y se retira por extracciones con fenol/cloroformo y cloroformo. Por precipitación con etanol se recupera el ARN citoplasmático.

b. Purificación de ARNm

El ARN mensajero se purifica a partir de la preparación de ARN total o citoplasmático usando procedimientos convencionales. El ARN Poli(A)⁺ se puede separar del ARN total por unión con oligo(dT) a la cola Poli(A) del ARNm. El ARN total se desnaturaliza para exponer las colas de Poli(A) (poliadeniladas). A continuación, el ARN que contiene Poli(A) se fija a perlas magnéticas recubiertas con oligo(dT) y se destila a partir del ARN total o citoplasmático por fuerzas magnéticas. La población de ARNm se puede enriquecer adicionalmente para obtener moléculas de longitud completa a través de la selección de especies de ARNm que contienen un casquete 5'.

c. Síntesis de ADNc

Se pueden usar diferentes tipos de cebadores para sintetizar bibliotecas de ADNc que contienen longitud completa o el extremo 5' a partir del ARNm aislado.

I. Cebador oligo(dT), que generará ADNc para todas las especies de ARNm (Figura 14)

En la Figura 14 se ofrece un ejemplo de la producción de una biblioteca de ADNc cebada con oligo(dT) adaptada.

II. Oligonucleótidos degenerados, específicos para el motivo funcional de proteínas; estos cebadores generarán un número limitado de genes pertenecientes a la misma familia de proteínas o de proteínas relacionadas funcionalmente (Figura 15)

Un ejemplo de la producción de una biblioteca de ADNc específica para el motivo de secuencia adaptada se ofrece en la Figura 15.

III. *Los oligonucleótidos específicos para genes producirán ADNc sólo para una especie particular de ARNm (Figura 16)*

Los oligonucleótidos utilizados para la producción de ADNc pueden contener secuencias adicionales, 1) secuencias específicas de etiquetado de proteínas para una purificación más sencilla de las proteínas recombinantes (6X HIS, Figura 7), 2) sitios de enzimas de restricción, 3) extremo 5' modificado para la purificación de ADNc o con fines de construcción de ADN (Figura 17).

La conversión de ARNm es ADNc de doble cadena para la inserción en un vector se lleva a cabo en dos partes. En primer lugar, se copia por transcriptasa inversa el ARNm intacto, hibridado con un cebador oligonucleótido, y los productos se aíslan por extracción con fenol y precipitación en etanol. El ARN en el híbrido ARN-ADN se retira con RNasa H a medida que la ADN polimerasa I de *E. coli* rellena las separaciones. Los fragmentos de la segunda cadena producidos de este modo se ligan mediante ADN ligasa de *E. coli*. Se completa la síntesis de la segunda cadena, se degrada el ARN residual y el ADNc se torna como con RNasa H, RNasa A, ADN polimerasa T4 y ADN ligasa de *E. coli*.

d. *Ligadura del adaptador*

Las moléculas adaptadoras pueden ligarse a los dos extremos del ADNc de extremos romos, o sólo a un extremo de ADNc. La ligadura del adaptador dirigido al sitio puede lograrse mediante el uso de oligonucleótidos 5' modificados (por ejemplo, biotinilados, aminados) durante la síntesis de ADNc, que impide la ligadura del adaptador al extremo 3' del ADNc. Las moléculas de ADNc resultantes contienen una biblioteca de ADNc de extremo 5' compuesta por una región 5' no traducida, el codón de inicio de la traducción AUG, que codifica una metionina, seguido de la región codificadora del gen o de los genes. Las moléculas de ADNc están flanqueadas por una secuencia conocida de ADN en sus extremos 5' y 3' (Figuras 14, 15 y 16).

e. *Amplificación de ADNc*

Se pueden sintetizar cebadores de PCR para las secuencias 5' y 3' conocidas o las secuencias internas conocidas, usándolos en la amplificación de la biblioteca completa o de subpoblaciones específicas de ADNc, usando un cebador de amplificación 5' o 3' extendido, en combinación con el cebador localizado en el sitio opuesto de las moléculas de ADNc (Figura 18).

f. *Diseño de cebador para la amplificación de subpoblaciones de genes*

Los cebadores de subpoblaciones contienen dos porciones (Figura 19). La parte 5' del cebador es complementaria para la secuencia de una secuencia conocida, que se extiende con su extremo 3' hacia una secuencia desconocida de ADNc. Dado que cada nucleótido en la parte ADNc de la biblioteca puede tener un residuo adenosina, citidina, guanosina o timidina, existen 4 posibilidades diferentes de nucleótidos para cada posición de nucleótido. Se pueden sintetizar cuatro cebadores de amplificación diferentes, en donde cada uno contiene la misma secuencia conocida y se extiende por un nucleótido hacia la zona de ADNc de la biblioteca. Los 4 cebadores difieren solamente en su nucleótido 3' extremo, que es A, C, G o T. Si se supone que cada nucleótido (A, C, G, T) está igualmente representado en un tramo de ADN, cada uno de los 4 cebadores de amplificación amplificará una cuarta parte de los genes totales representados en la biblioteca de ADNc. Extendiendo adicionalmente la secuencia del cebador de amplificación y aumentando el número de cebadores de amplificación, es posible reducir aún más la complejidad de los productos de amplificación. La extensión de la secuencia en 2 nucleótidos requiere la síntesis de 16 cebadores diferentes, reduciendo la complejidad en 16 veces; 3 nucleótidos requieren 64 cebadores diferentes, y la extensión de nucleótidos requiere n^4 cebadores diferentes.

g. *Amplificación por PCR*

La amplificación por PCR comprende mezclar ADN de molde, dos cebadores oligonucleótidos apropiados (cebadores de extremos 5' y 3' situados en las secuencias conocidas añadidas en orientación complementaria), Taq u otras ADN polimerasas termoestables, trifosfatos de desoxirribonucleósido (dNTPs), y una solución tampón. Los productos de PCR se analizan después de ciclar sobre geles de ADN o por el análisis en ABI 377, usando el software de análisis genescan. Estos métodos de análisis permiten determinar la complejidad del combinado de ADNc amplificado.

h. *Producción de una biblioteca de expresión de proteína*

Cada subpoblación amplificada de biblioteca de ADNc se clona 5' a 3' en un sistema de expresión de proteína bacteriano (*E. coli*, etc.) o eucariótico (Baculovirus, levadura, mamífero). Se introduce el gen con su propia señal de inicio de la traducción y una etiqueta 6x His en los 3 marcos. Por ejemplo: el ADNc se restringe con dos enzimas de restricción de corte diferentes (BglIII extremo 5' y NotI de extremo 3') y se clona en la orientación 5' a 3' en el vector de transferencia de Baculovirus pVL1393, bajo control directo del promotor polihedra.

ES 2 319 634 T3

i. Expresión de la proteína

Se co-transfectan ADN linealizado d Baculovirus y ADN recombinante del vector de transferencia en células susceptibles de insecto SF9 con fosfato cálcico. Para la co-transfección, se preparan 10 μg de ADN plasmídico purificado. Se prepara un stock recombinante inicial y las células SF9 se infectan para la producción de proteína recombinante.

j. Purificación de la proteína

Las proteínas recombinantes expresadas contienen una etiqueta de afinidad (el ejemplo es una etiqueta 6XHis). Se purifican sobre agarosa Ni-NTA. Se obtienen, habitualmente, alrededor de 1 a 2 mg de proteína de fusión recombinante 6xHis por litro de cultivo de células de insecto.

II. Generación de anticuerpos por inmunización de diferentes animales con mezclas de proteínas individuales

3. Preparación de Anticuerpos reactivos de captura de proteínas

Una preparación de proteína purificada, traducida de una combinación de ADNc, se inyecta por vía intramuscular, intradérmica o subcutánea, en presencia de adyuvante en un animal de la especie seleccionada (conejo). Se inician las inmunizaciones de recuerdo 4 a 8 semanas después de la primera inmunización, y se continúa a intervalos de 2-3 semanas. El antisuero policlonal se purifica usando métodos convencionales conocidos por los expertos en la técnica.

Los lotes de anticuerpos purificados se pueden usar directamente como reactivos de captura de proteínas sin modificación. En este caso, los lotes de anticuerpo de diferentes animales deben mantenerse separados (cada lote es un reactivo de captura).

III. Las proteínas de anticuerpo se aíslan y conjugan con secuencias de ácido nucleico correspondientes a la preparación original de antígeno, resultantes en reactivos de captura de anticuerpos

Generación de moléculas bifuncionales de captura/separación para la separación de la mezcla compleja de proteínas en una fase sólida.

El dominio C_H^2 glicosilado de los anticuerpos policlonales se conjuga con oligonucleótidos modificados en 5', usando métodos de conjugación convencionales. La molécula resultante tiene un resto de captura de proteína (anticuerpo) y un resto de ácido nucleico (oligonucleótido) (Figura 20).

Los lotes de anticuerpos tras la inmunización de un animal con una combinación de proteínas de complejidad reducida se conjugan con la secuencia de oligonucleótidos. Los anticuerpos producidos a partir de múltiples episodios de inmunización con diferentes combinaciones de proteínas se conjugan en un oligonucleótido con una secuencia diferente (Figura 20).

4. Captura de proteínas diana usando la funcionalidad de reactividad y separación por hibridación de oligonucleótidos

Se han desarrollado dos métodos diferentes para unir los oligonucleótidos a un soporte sólido: se les puede sintetizar *in situ*, o pre-sintetizar y fijar al soporte. En cualquier caso, es posible utilizar los oligonucleótidos unidos al soporte en una reacción de hibridación con oligonucleótidos en fase líquida para formar híbridos (dúplex); a continuación, se puede eliminar por lavado el exceso de oligonucleótidos.

El soporte puede adoptar la forma de partículas, por ejemplo, esferas de vidrio o perlas magnéticas. En este caso, las reacciones se podrían llevar a cabo en tubos, o en los pocillos de una placa de microtitulación. Los métodos para sintetizar oligonucleótidos y para fijar oligonucleótidos pre-sintetizados a estos materiales son conocidos (véase, por ejemplo, Stahl *et al.* (1988) *Nucleic Acids Research* 16(7):3025-3039).

a. Preparación de un soporte sólido funcionalizado con amina

Se sintetizan oligonucleótidos de una secuencia definida sobre un soporte de vidrio funcionalizado con amina. Se unió una función amina a localizaciones discretas del portaobjetos de vidrio, usando una solución de 700 μl de $\text{H}_2\text{N}(\text{CH}_2)_3\text{Si}(\text{OCH}_2\text{CH}_3)_3$ en 10 ml de etanol al 95% a temperatura ambiente durante 3 horas. El soporte tratado se lava una vez con metanol y, luego, una vez con éter etílico. El soporte se secó a temperatura ambiente y se coció luego a 110°C durante 15 horas. Se le lavó entonces con agua, metanol y agua, y se secó por último.

El portaobjetos de vidrio se hizo reaccionar durante 30 min a temperatura ambiente con 250 mg (1 milimol) de anhídrido ftálico en presencia de 2 ml de piridina anhidra y 61 mg de 4-dimetil-aminopiridina.

ES 2 319 634 T3

El producto se enjuagó con dicloruro de metileno, alcohol etílico y éter, y se secó. Los productos en el portaobjetos se hicieron reaccionar con 330 mg de diciclohexilcarbodiimida (DCC) durante 30 min a temperatura ambiente. Se decantó la solución y se sustituyó con una solución de 117 mg de 6-amino-1-hexanol en 2 ml de dicloruro de metileno, y se dejó a temperatura ambiente durante aproximadamente 8 horas.

5

b. Síntesis de oligonucleótidos en un soporte sólido

El soporte sólido funcionalizado con amina se preparó para la síntesis de oligonucleótidos por tratamiento con 400 mg de anhídrido succínico y 224 mg de 4-dimetil aminopiridina en 3 ml de piridina anhidra, durante 18 horas a temperatura ambiente. El soporte sólido fue tratado con 2 ml de DMF, que contuvo 3 milimoles (330 mg) de DCC y 3 milimoles (420 mg) de p-nitrofenol a temperatura ambiente durante la noche. Se lavó el portaobjetos con DMF, CH₃CN, CH₂Cl₂ y éter etílico. Se hizo reaccionar una solución de 2 milimoles (234 mg) de H₂N(CH₂)₆OH en 2 ml de DMF con el portaobjetos durante la noche. El producto de esta reacción fue un soporte -O(CH₂)₃NHCO(CH₂)₂CONH(CH₂)₅CH₂OH. El portaobjetos se lavó con DMF, CH₂CN, metanol y éter etílico.

15

El éster funcionalizado resultante de la preparación del soporte de vidrio se usó para la síntesis de una secuencia de oligonucleótidos. Cada residuo nucleósido se agregó en forma de fosforamidita, según el procedimiento de Caruthers *et al.* (Patente de EE.UU. No. 4.415.732).

20

5. Análisis de proteína de las proteínas capturadas y comparación con la muestra compleja de proteínas

Los lotes purificados de anticuerpo pueden 1) unirse directamente a una superficie sólida, e incubarlos con muestras de proteína, 2) ser incubados con las muestras y, subsiguientemente, unirse a un soporte sólido sin utilizar la molécula bifuncional de captura, 3) se puede usar la molécula bifuncional de captura para capturar su correspondiente proteína en una muestra y, subsiguientemente, separar las proteínas capturadas por hibridación específica de nucleótidos (Figura 21).

30

IV. Los reactivos de captura de oligonucleótidos antisentido se inmovilizan en localizaciones discretas y conocidas en una superficie sólida para crear un ensayo de captura de anticuerpo

Se sintetizan oligonucleótidos 5'-aminados usando la química de fosforamidatos y se unen a ésteres N-oxisuccinimida. Las secuencias de oligonucleótidos fijadas son complementarias a los oligonucleótidos de separación de las moléculas bifuncionales de anticuerpo (Figura 20). Se capturan proteínas a través de la hibridación de ácidos nucleicos de sus oligonucleótidos separados a la secuencia complementaria unida al oligonucleótido de la superficie sólida.

35

V. Los reactivos de captura de anticuerpos se agregan a la mezcla total de proteínas (etapa de reactividad). Seguidamente, se agrega la mezcla de reacción a la matriz de superficie sólida bajo condiciones que permiten la hibridación de oligonucleótidos (etapa de separación)

40

7. Captura y separación de reactivo bifuncional/proteína

45

Los anticuerpos bifuncionales se incuban con la muestra de proteína bajo condiciones que permiten que los anticuerpos se unan a su correspondiente antígeno. La molécula bifuncional del anticuerpo con la proteína capturada se agrega a la matriz de captura preparada del oligonucleótido. Bajo condiciones convencionales de hibridación de ADN, que no desnaturalizan la unión antígeno-anticuerpo, el anticuerpo bifuncional hibridará con su resto de ácido nucleico en el oligonucleótido complementario.

50

VI. La proteína de captura se identifica usando espectrometría de masa MALDI

8. Análisis de las proteínas de captura

55

Las proteínas unidas se analizarán usando métodos convencionales de análisis de proteínas tales como Espectrometría de Masa.

Dado que las modificaciones resultarán evidentes para los expertos en la técnica, se pretende que la presente invención esté limitada únicamente por el alcance de las reivindicaciones adjuntas.

60

65

REIVINDICACIONES

1. Una matriz de gradiente para capturar biomoléculas, que comprende:

5 una matriz bidimensional de compuestos de captura que comprenden restos X e Y; y

una superficie que contiene loci para presentar los compuestos de captura que contienen los restos X e Y, en donde:

10 la superficie comprende un soporte sólido con los compuestos sobre el mismo, en loci dispuestos en una matriz bidimensional;

la superficie comprende al menos 10 compuestos de captura diferentes, en donde en cada locus están situados compuestos diferentes;

15 los restos X se seleccionan independientemente para fijarse de forma covalente a las cadenas laterales de aminoácidos de las proteínas;

20 cada resto Y modula una o más de las propiedades de afinidad, estéricas y electrónicas de X, aumentando de este modo la selectividad de la unión por medio de X, de manera que X se une a menos proteínas cuando el resto de selectividad Y está presente que en su ausencia;

25 a lo largo de cada fila de loci (eje X) en la matriz, se altera una propiedad seleccionada de X en cada locus, que se selecciona entre hidrofobia, lipofilia, carga, tamaño o especificidad del compuesto de captura en cada locus, de manera predeterminada a lo largo de la fila para proporcionar un gradiente de la propiedad;

a lo largo de cada columna de loci (eje Y) en la matriz, se altera una propiedad de Y en cada locus, que se selecciona entre hidrofobia, carga, tamaño, especificidad, de manera predeterminada a lo largo del eje Y para proporcionar un gradiente de la propiedad;

30 para cada matriz, la propiedad en el eje X que se altera no es la misma que la propiedad que se altera en el eje Y, de manera que la matriz presenta un gradiente bidimensional de dos o más de las propiedades, y presenta loci con diferentes afinidades para las biomoléculas.

35 2. La matriz de gradiente de la reivindicación 1, en la que los compuestos a lo largo del eje X comprenden un grupo azobenceno, y se crea un gradiente de hidrofilia por el aumento de exposición a la luz.

3. La matriz de gradiente de la reivindicación 1, en la que los compuestos a lo largo del eje Y comprenden un compuesto azobenceno, y se crea un gradiente de hidrofilia por el aumento de exposición a la luz.

40 4. La matriz de gradiente de la reivindicación 1, en la que los compuestos a lo largo del eje X comprenden un grupo cargado, y se crea un gradiente de carga por el aumento de exposición a la corriente.

45 5. La matriz de gradiente de la reivindicación 1, en la que los compuestos a lo largo del eje Y comprenden un grupo cargado, y se crea un gradiente de carga por el aumento de exposición a la corriente.

6. La matriz de gradiente de la reivindicación 1, en la que la propiedad de los compuestos que se varía a lo largo del eje X es especificidad para un grupo NH₂, SH, SS u OH.

50 7. La matriz de gradiente de la reivindicación 1, en la que la propiedad de los compuestos que se varía a lo largo del eje Y es especificidad para un grupo NH₂, SH, SS u OH.

55 8. La matriz de gradiente de la reivindicación 1, en la que X es un α -halo éter, un grupo α -halo carbonilo, maleimido, un complejo metálico, un epóxido, un isotiocianato, o un anticuerpo contra péptidos/proteínas fosforilados o glicosilados.

9. La matriz de gradiente de la reivindicación 1, en la que X es -C(=O)O-Ph-pNO₂, -C(=O)O-C₆F₅, -C(=O)-O-(N-succinimidilo), -OCH₂-I, -OCH₂-Br, -OCH₂-Cl, -C(O)CH₂I, -C(O)CH₂Br o -C(O)CH₂Cl.

60 10. La matriz de gradiente de la reivindicación 1, en la que los loci o los restos X o Y en cada locus comprenden una etiqueta modificadora de la masa.

11. Un método para analizar biomoléculas, que comprende:

65 a) poner en contacto una composición que comprende una biomolécula con una matriz de cualquiera de las reivindicaciones 1 a 10 para formar complejos de biomoléculas con los restos X y/o Y en loci de la matriz; y

b) identificar o detectar las biomoléculas unidas.

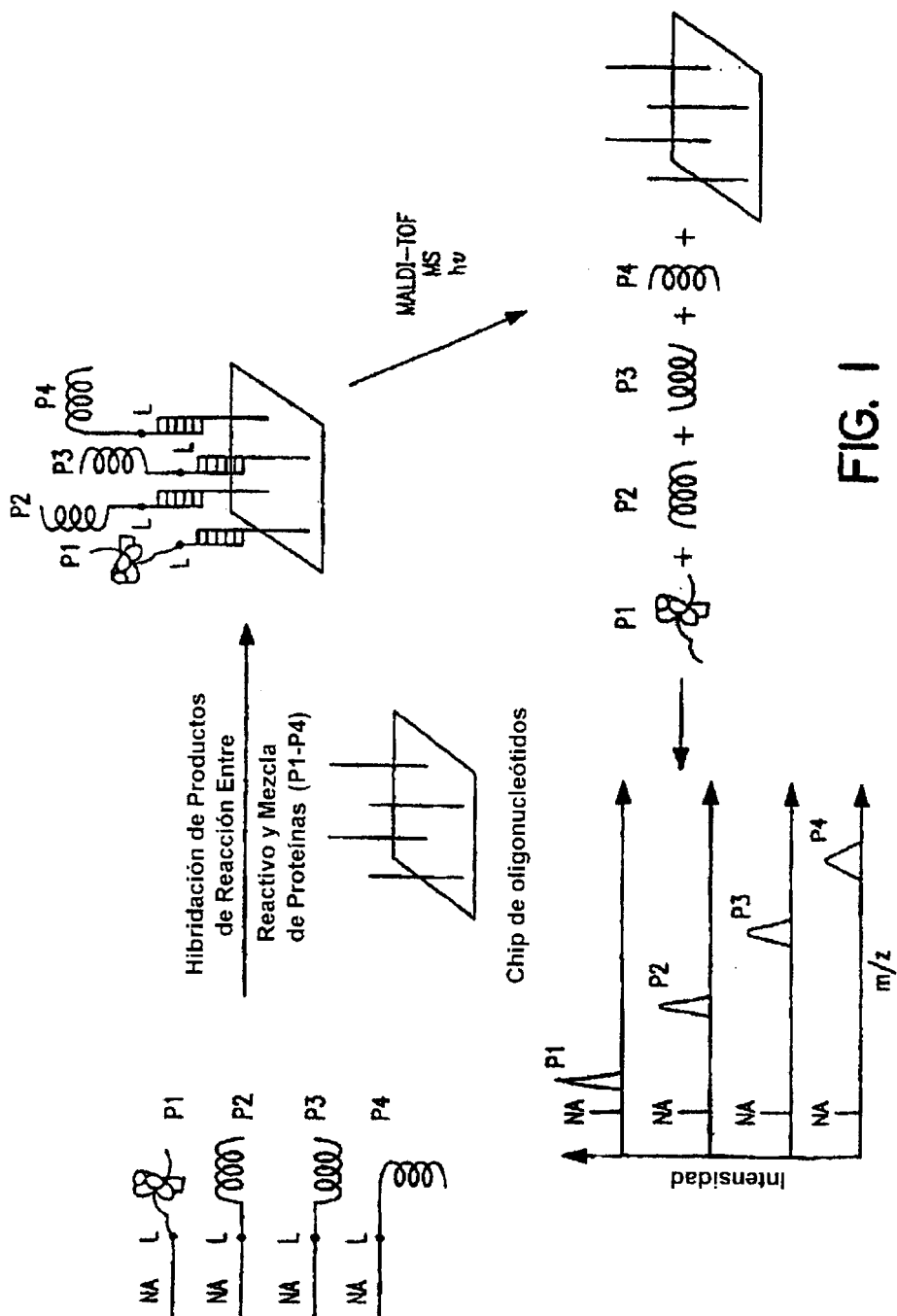


FIG. 1

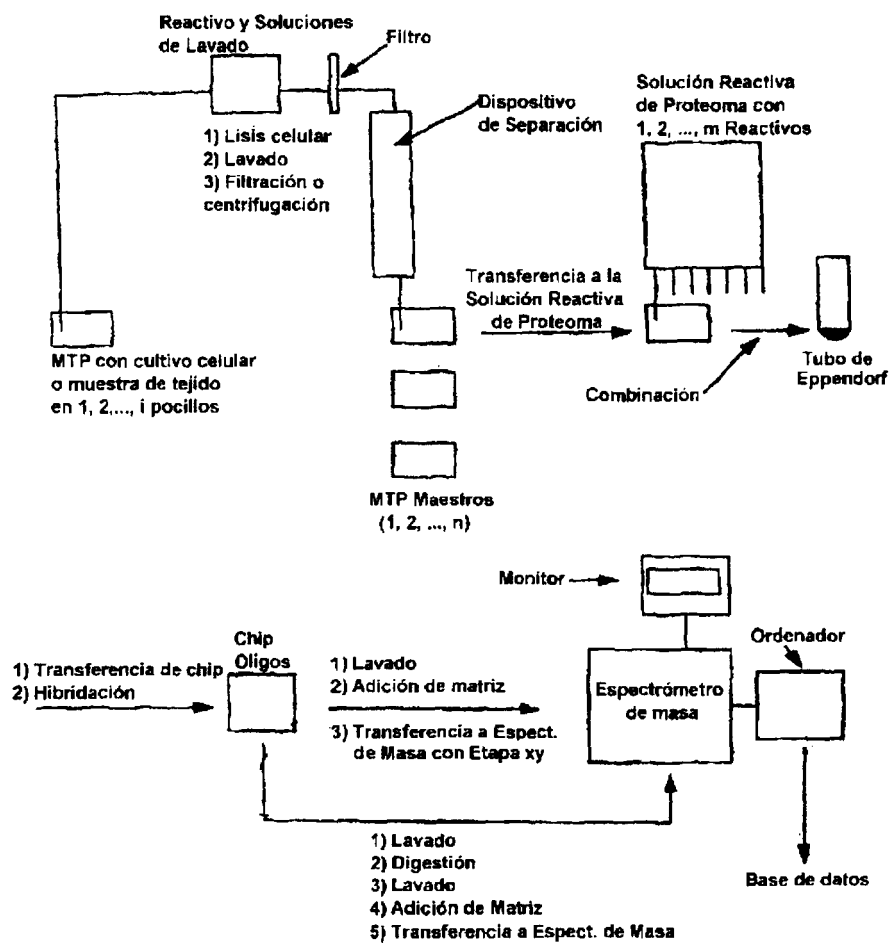


FIG. 2

A

#	Tipo Uniplex	Secuencia	Enlace
1	lineal	5'-(atg.gtt.agt.g)	no
2	lineal	5'-(cac.taa.cca.t)	no
3	lineal	5'-(atg.gtt.agt.g)	no
4	lineal	5'-(cac.taa.cca.t)	no
5	dend3	5'-(atg.gtt.agt.g)	3'
6	dend3	5'-(cac.taa.cca.t)	3'
7	dend2	5'-(cac.taa.cca.t)	5'
8	dend9	5'-(atg.gtt.agt.g)	3'
9	dend9	5'-(cac.taa.cca.t)	3'
10	dend9	5'-(cac.taa.cca.t)	5'
11	dend27	5'-(atg.gtt.agt.g)	3'
12	dend27	5'-(cac.taa.cca.t)	3'
13	lineal	5'-(gaa.gag.aaa.gag.aaa)	3'
14	lineal	5'-(gaa.gag.aaa.gag.aaa)	5'
15	lineal	5'-(ttt.ctt.ttt.ctt.ttc)	no
16	lineal	5'-(gaa.gaa.gag.aaa.gaa)	3'
17	lineal	5'-(gaa.gaa.gag.aaa.gaa)	5'
18	dend9	5'-(ttt.ttt.ctt.ttc.ctt)	3'
19	dend27	5'-(ttt.ttt.ctt.ttc.ctt)	3'
20	lineal	5'-(atg.gtt.agt.g)	no
21	lineal	5'-(cac.taa.cca.t)	no
22	dend2	5'-(atg.gtt.agt.g)	5'
23	dend2	5'-(cac.taa.cca.t)	5'
24	dend6	5'-(atg.gtt.agt.g)	3'
25	dend6	5'-(cac.taa.cca.t)	3'

B

#	de Tipo Duplex	Apareamiento más sencillo	T _m , °C	
Fig. 3A			3,5xTMA	6xSSC
Solución				
1+2	lineal	=	44,0	41,6
3+4	lineal concatenar	(=) ₃	77,0	76,5
6+1	dend3		46,0	
5+6	dend3		54,7	50,0
5+7	dend3 'starfish'		49,5	
8+9	dend9		62,5	56,4
8+10	dend9 'starfish'		62,1	
11+12	dend27		67,9	59,2
20+21	lineal concatenar	(=) ₆	81,6	81,5
22+23	dend2		51,2	47,2
24+25	dend6		80,4	54,5
Soporte Sólido				
dimensiones				
h.nm d.nm				
13+15	lineal		45,0	5 2
14+15	lineal		45,0	5 2
16+18	dend9 favorable		57,5	5 8
17+18	dend9 desfavorable		55,0	
18+19	dend27 favorable		58,0	5 10

A) Los compuestos nº 22 y nº 23 se sintetizaron usando fosforamiditas 3'-nucleótido y hexaetilenglicol fosforamidita (Cruachem), con subsiguiente síntesis del segundo oligonucleótido sobre el primero, usando fosforamiditas 5'-nucleótido

B) Soluciones tampón: TMA, cloruro de trimetilamonio; SSC: citrato sodico. Para los estudios de solución, se determinaron los volúmenes T_m de forma automática integrando las curvas (software Beckman DU 640). Las mediciones se llevaron a cabo de tres a cinco veces y las temperaturas de fusión estuvieron dentro de 1°C (para curvas no suavizadas) del promedio indicado en la tabla.

FIG. 3

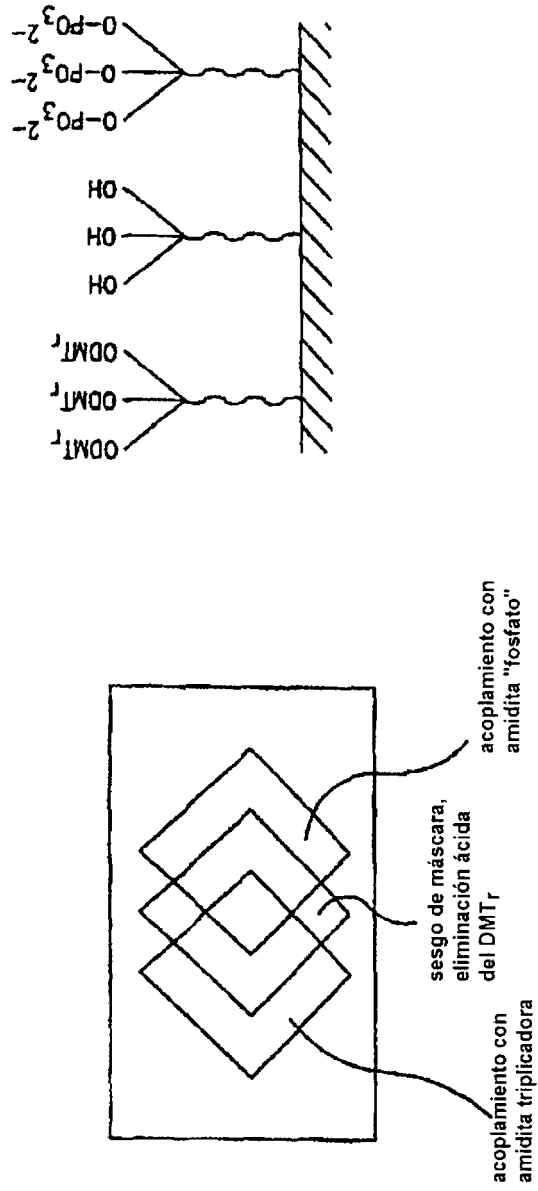


FIG. 4

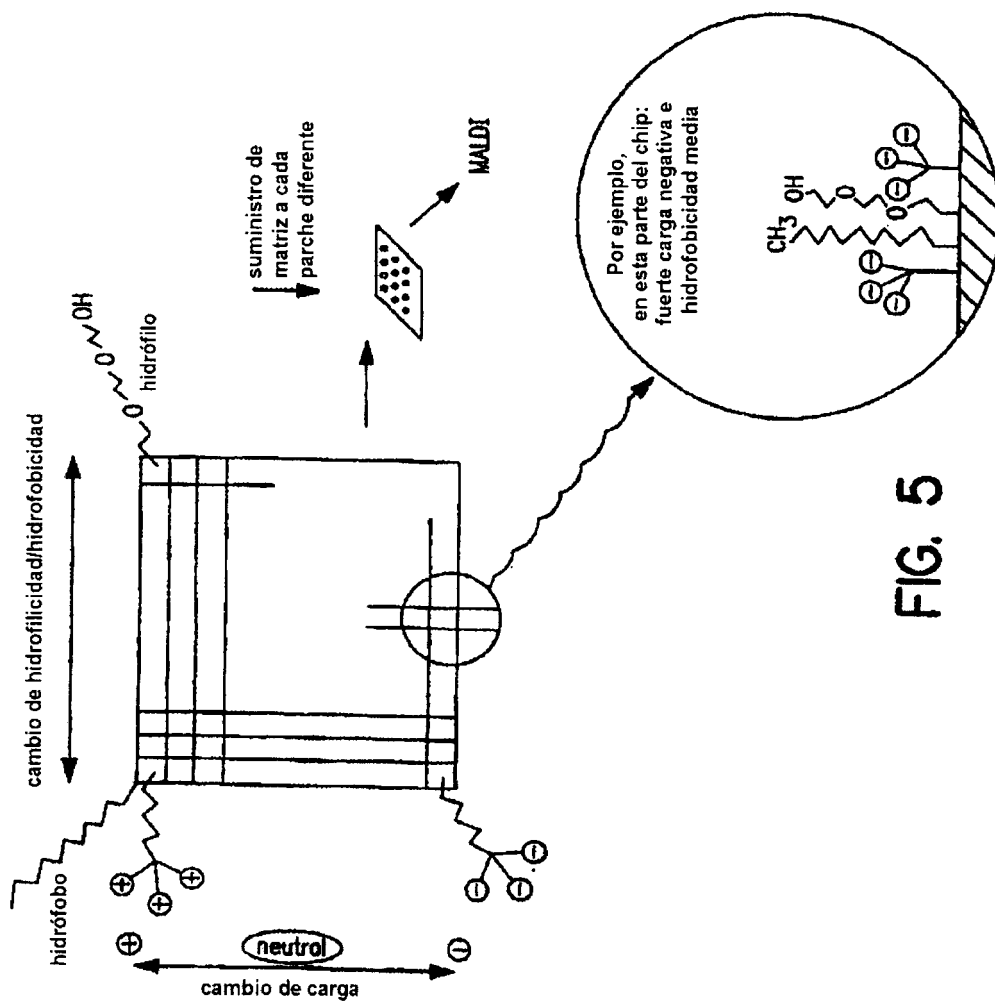


FIG. 5

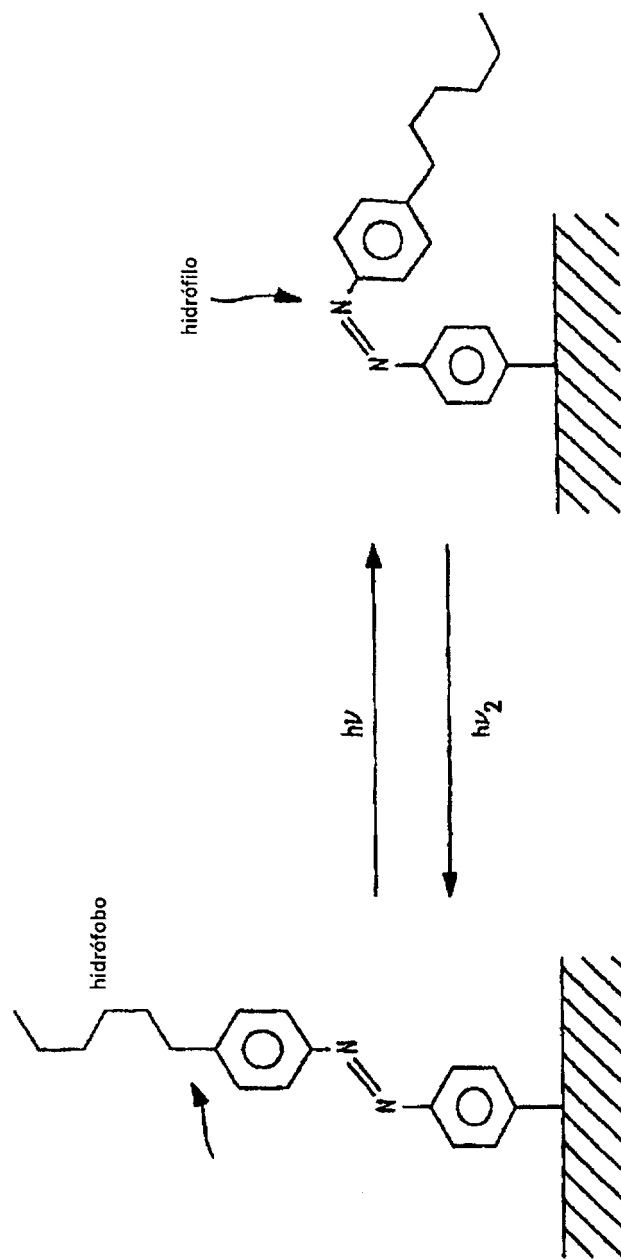


FIG. 6

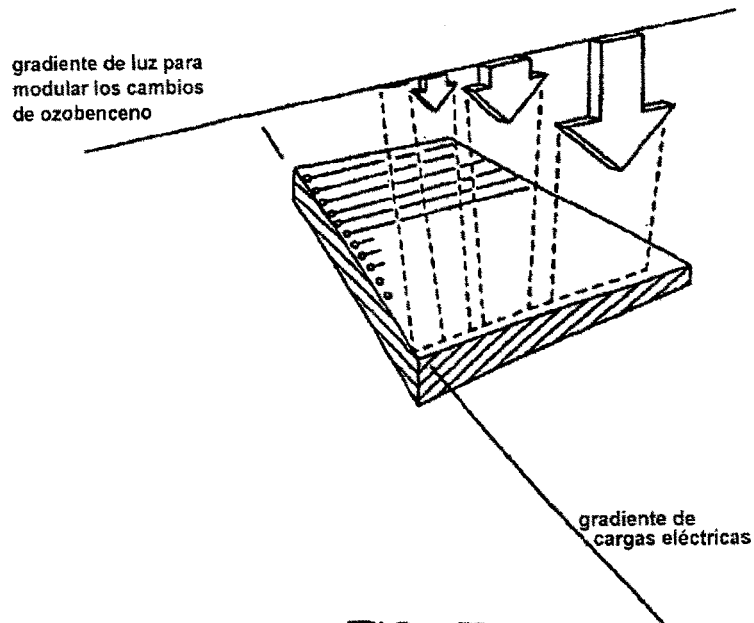


FIG. 7

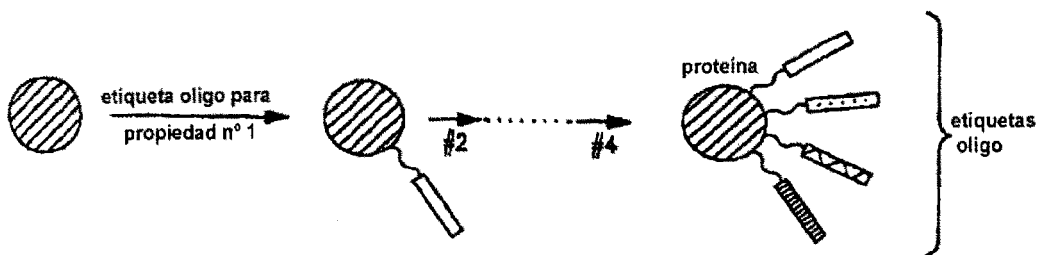


FIG. 8

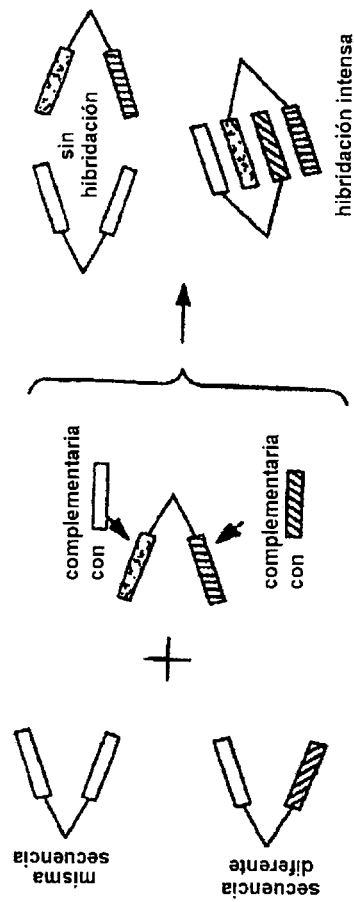


FIG. 9

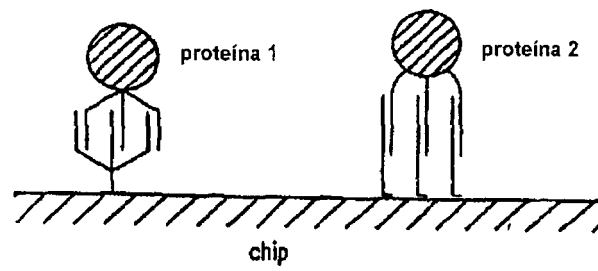


FIG. 10

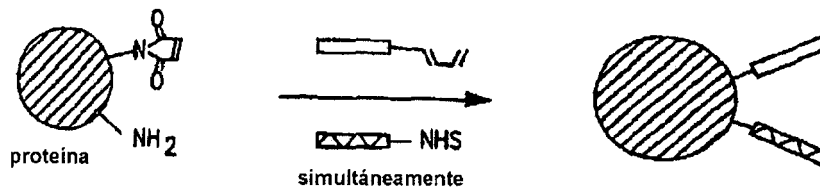


FIG. 11

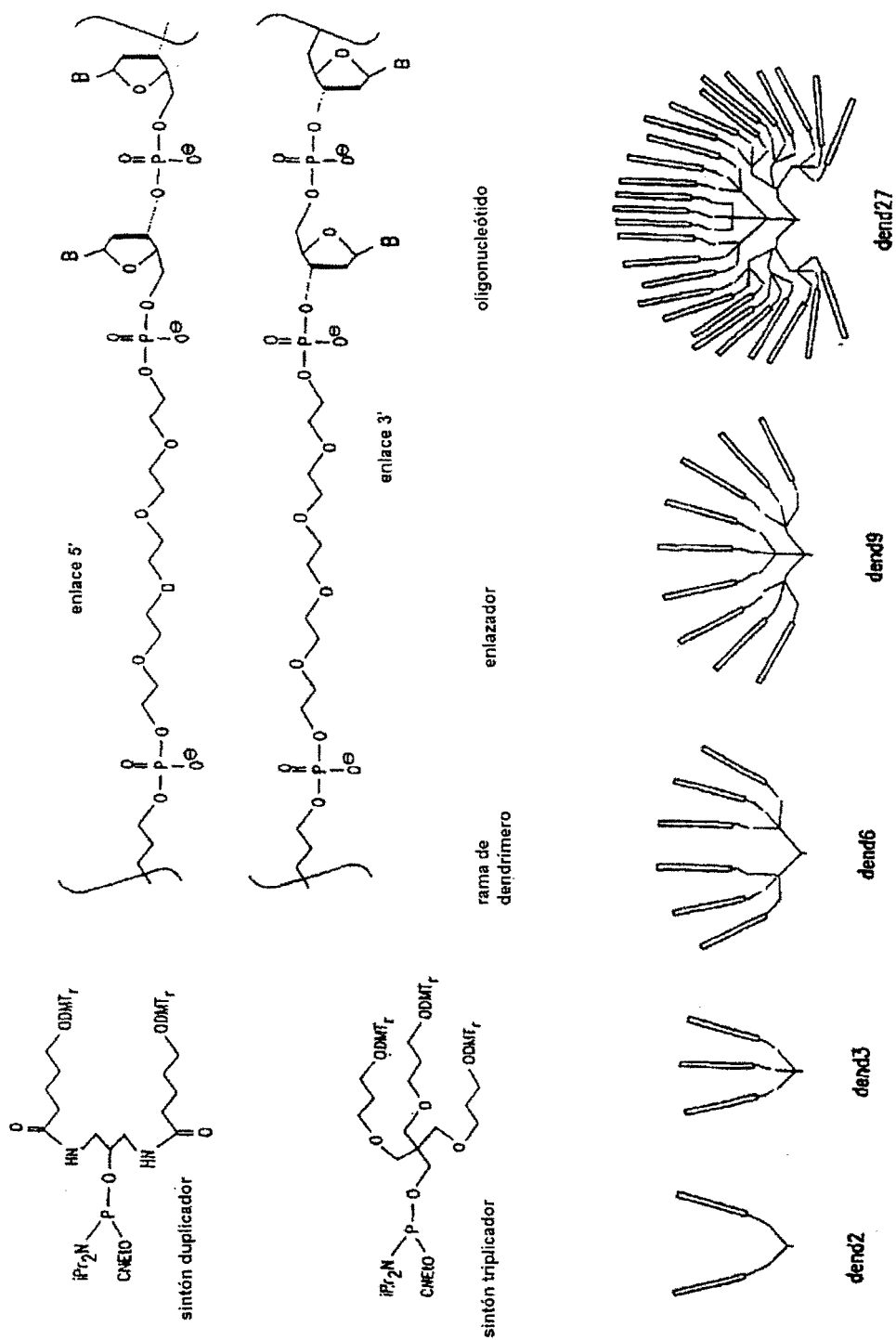


FIG. 12

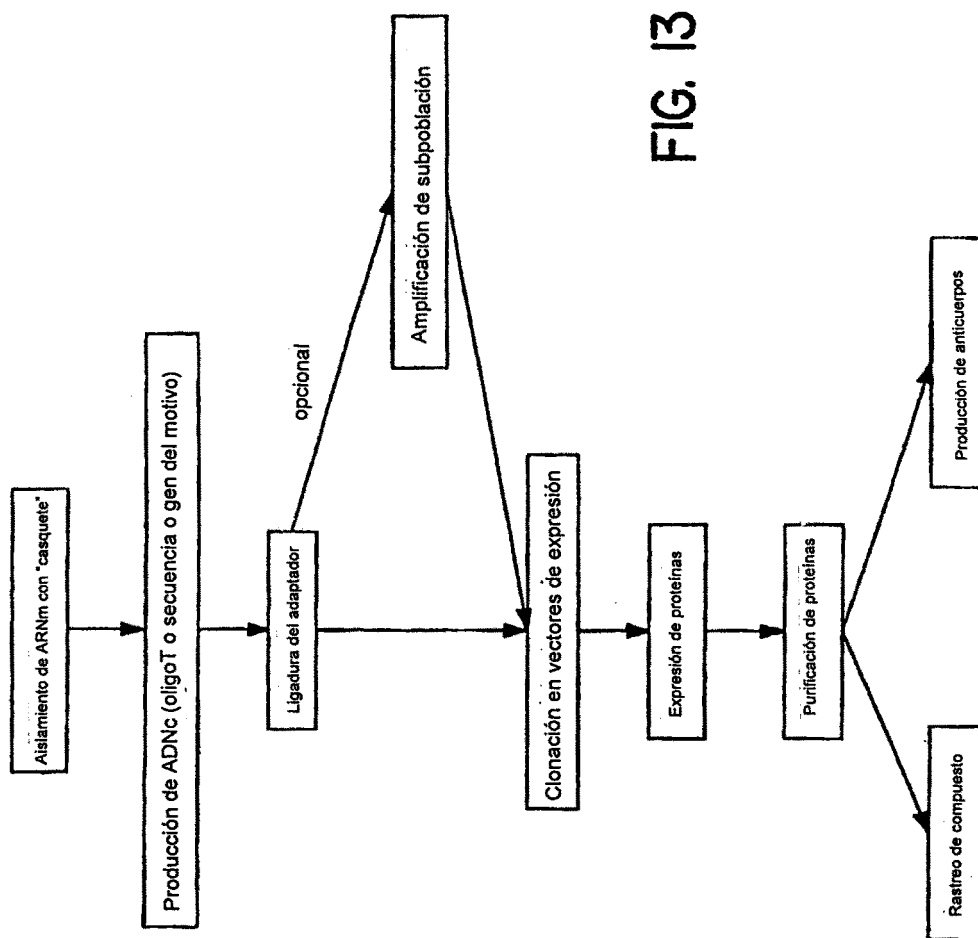


FIG. 13

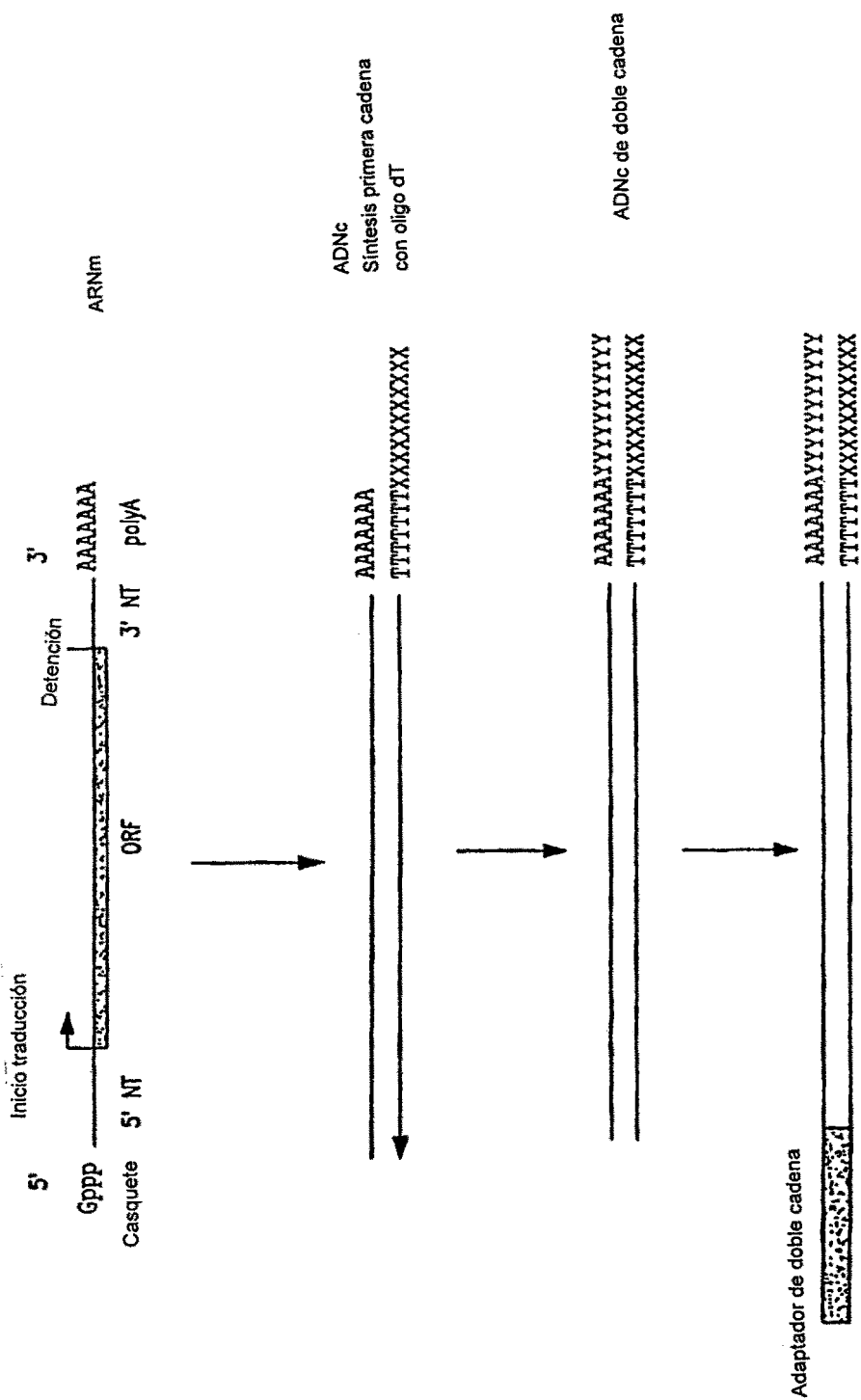


FIG. 14

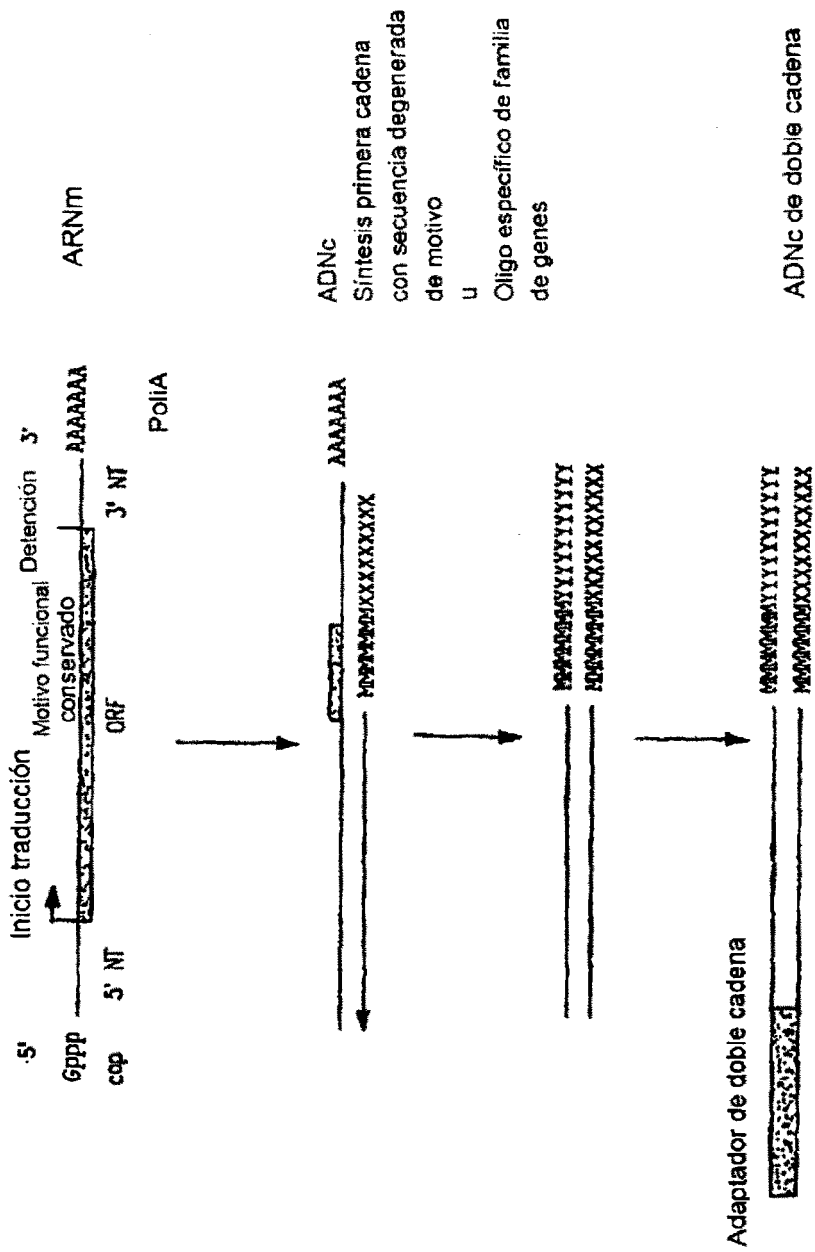


FIG. 15

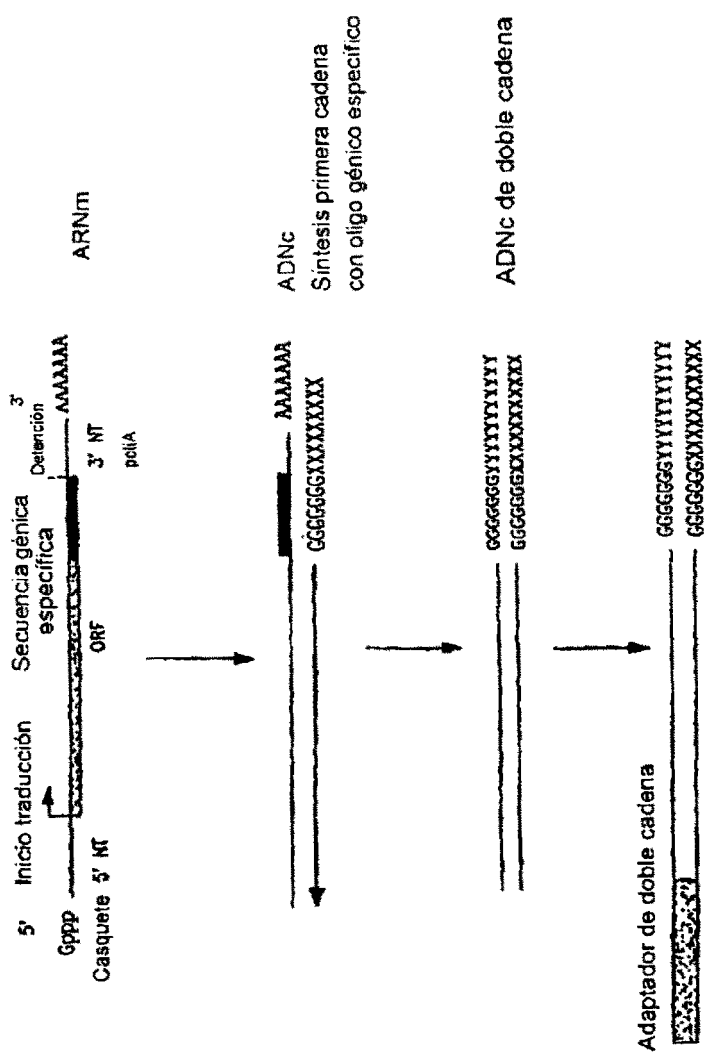


FIG. 16

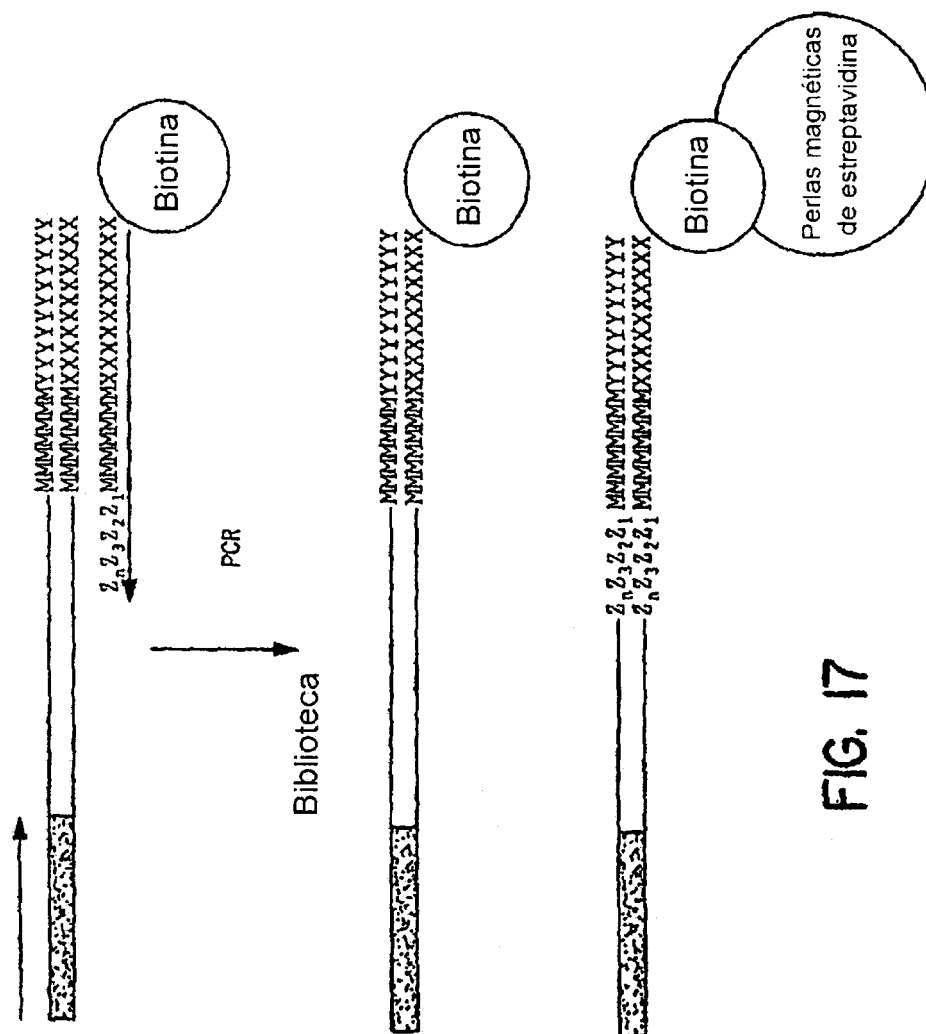


FIG. 17

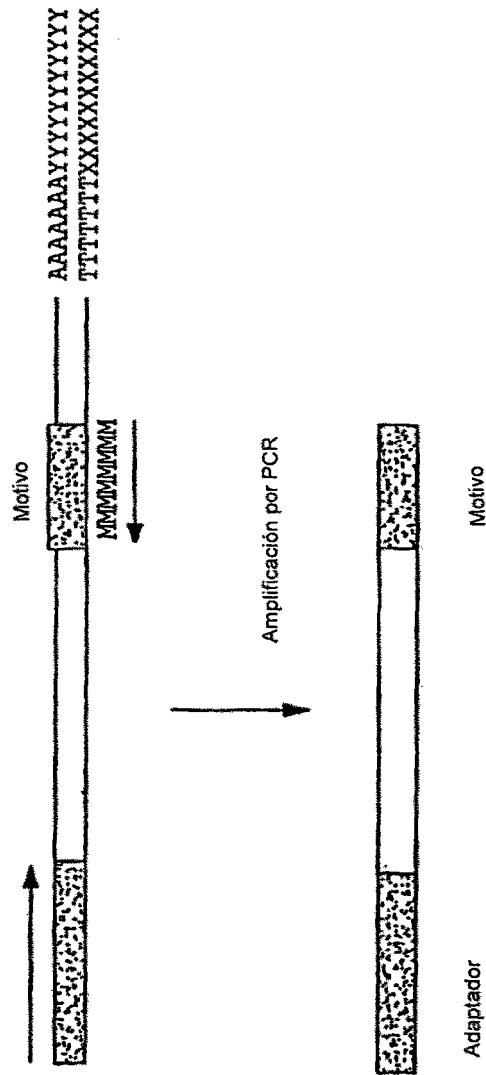


FIG. 18

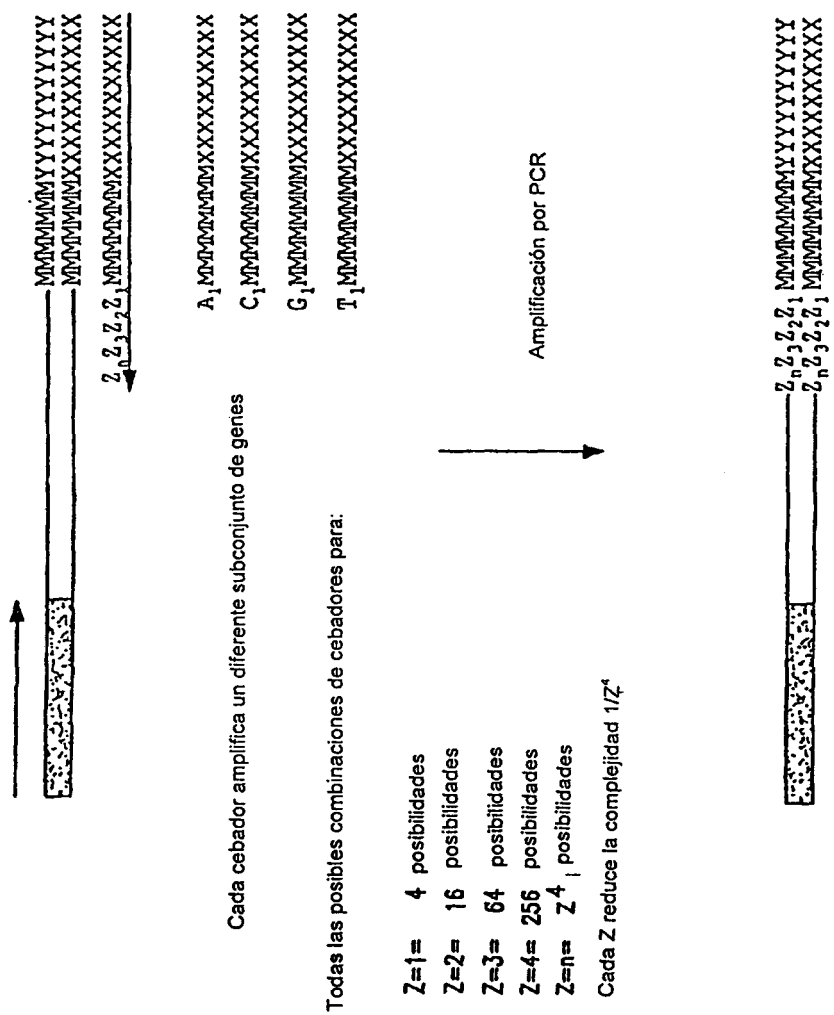
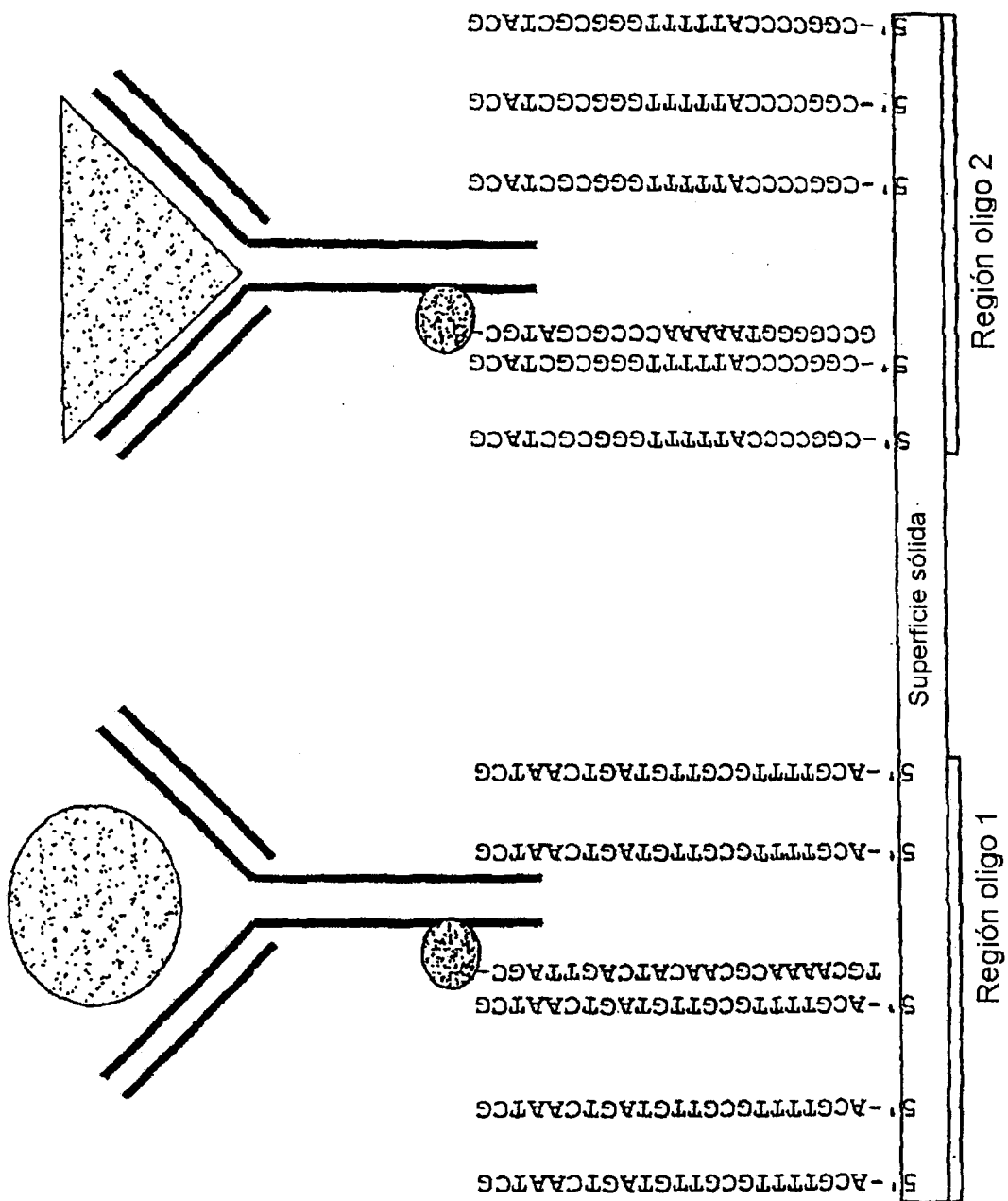


FIG. 19

FIG. 20



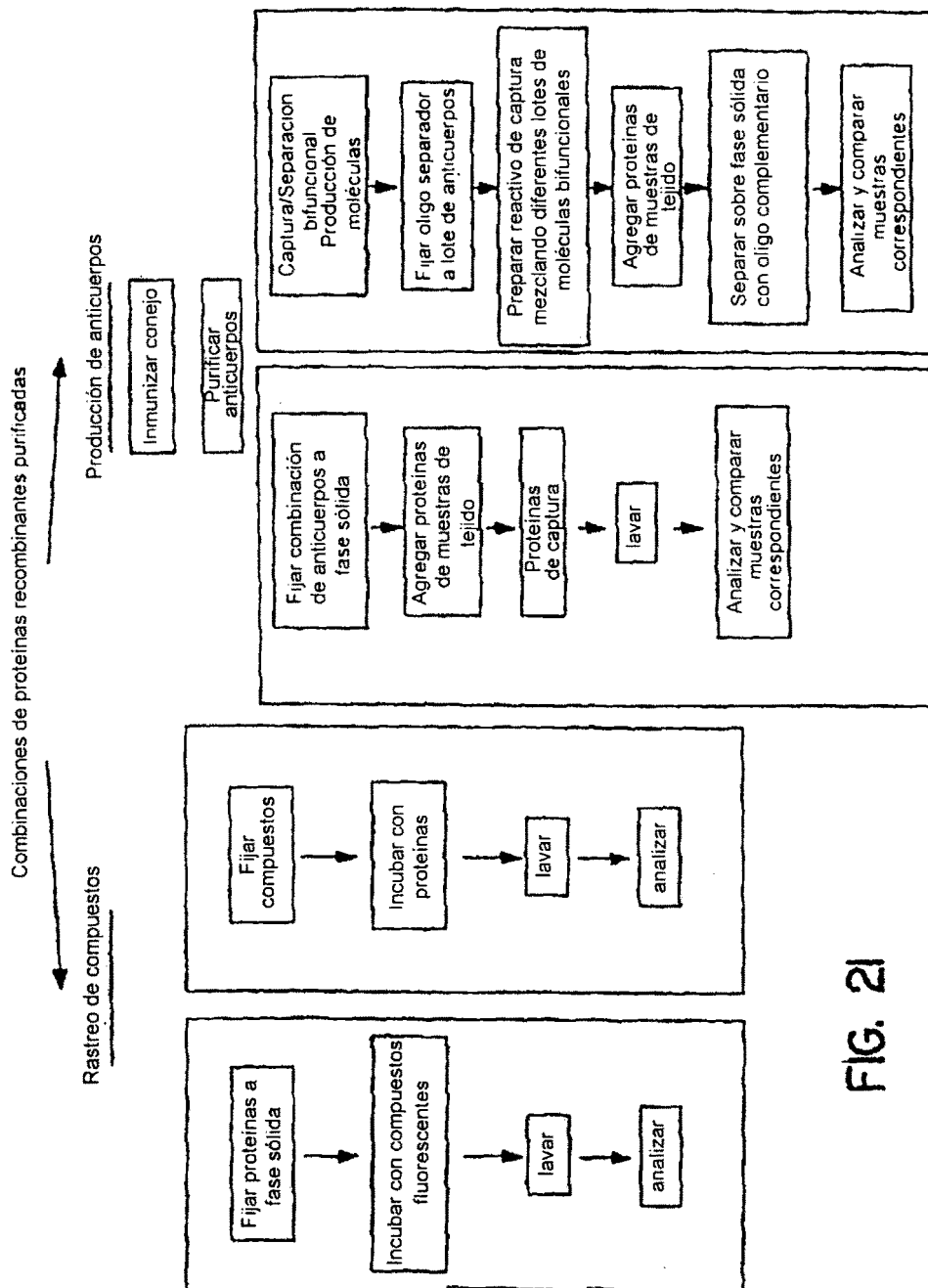


FIG. 21