



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 319 422**

51 Int. Cl.:
G01N 33/68 (2006.01)
C12Q 1/37 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **04719437 .8**
96 Fecha de presentación : **11.03.2004**
97 Número de publicación de la solicitud: **1604209**
97 Fecha de publicación de la solicitud: **14.12.2005**

54 Título: **Ensayo de detección.**

30 Prioridad: **12.03.2003 GB 0305656**

45 Fecha de publicación de la mención BOPI:
07.05.2009

45 Fecha de la publicación del folleto de la patente:
07.05.2009

73 Titular/es: **Bioinvent International AB.**
223 70 Lund, SE

72 Inventor/es: **Nilsson, Fredrik**

74 Agente: **Urizar Anasagasti, José Antonio**

ES 2 319 422 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

Ensayo de detección.

5 La presente invención se relaciona con un ensayo de detección.

Antecedentes de la invención

10 La genómica funcional es un campo de investigación cuyo propósito es entender lo que hace cada gen, cómo es regulado y cómo interactúan los diferentes genes y productos de genes. Un aspecto importante de la genómica funcional es entender la estructura y función de los productos de genes, como las proteínas, así como ser capaz de determinar dónde, cuándo y en qué grado son expresados los genes. El término patrones de expresión o análisis de expresión comprende generalmente tanto estudios de expresión del ARNm (análisis de transcripción) como análisis de proteína (análisis del proteoma o proteómica).

15 El análisis de transcripción se realiza típicamente utilizando ADN en un formato de micromatriz, permitiendo la detección paralela de miles o decenas de miles de moléculas de ARNm simultáneamente (e.g., utilizando micromatrices disponibles comercialmente de e.g., Affymatrix, EE.UU.). Típicamente estos micromatrices se utilizan para mapear la distribución de transcritos en diferentes tejidos o para estudiar las diferencias en los niveles de expresión del ARNm entre, e.g., individuos sanos y enfermos. Las aplicaciones en el desarrollo fármacos y en el descubrimiento de fármacos incluyen la identificación de una diana y la estratificación del paciente.

20 El patrón de expresión de proteína, o proteómica, es el análisis global del contenido de proteína en, por ejemplo, un tejido o una población celular.

25 La electroforesis bidimensional unida a la espectrometría de masas es una técnica bien establecida para el análisis de muestras de proteínas complejas con un poder de resolución lo suficientemente alto como para separar miles de proteínas. Las principales desventajas de esta técnica son la falta de un rango dinámico (las proteínas estructurales y las abundantes enzimas metabólicas tienden a enmascarar especies menos abundantes), su baja producción y una alta intensidad de trabajo.

30 La desorción/ionización por láser de superficie (SELDI) (Weinberger *et al.*, 2002, Journal of Chromatography B, 782, 307-316) es una técnica basada en el enriquecimiento selectivo de una subpoblación de proteínas en una superficie de afinidad (e.g., intercambio iónico, fase reversa, anticuerpos) seguida de un análisis espectrométrico de masas realizado por espectrometría de masas en tiempo de vuelo con desorción/ionización mediante láser asistida por matriz (MALDI-TOF), una técnica en la cual un coprecipitado de una matriz que absorbe luz UV y de biomoléculas es irradiado mediante pulsos láser de nanosegundos. La mayoría de la energía del láser es absorbida por la matriz, lo cual previene la fragmentación no deseada de la biomolécula. Las biomoléculas ionizadas son aceleradas en un campo eléctrico y separadas de acuerdo a su proporción masa-carga en un tubo de vuelo. No obstante, el poder de resolución del sistema es limitado debido a la restringida resolución de la espectrometría de masas MALDI-TOF para el análisis de proteínas grandes y la separación de proteínas por debajo del nivel óptimo que se alcanza mediante el tipo de extracción por pasos en fase sólida de la técnica de separación empleada.

35 Otra técnica alternativa, conocida como marcaje isotópico diferencial (ICAT) (Gygi *et al.*, 1999, *Nature Biotechnology*, 17(1), 994-9), utiliza un marcaje de biotina a una cisteína específica para comparar el patrón de expresión de la proteína en dos muestras diferentes. El marcaje permite la extracción de un péptido que contiene cisteína de mezclas de proteínas digeridas con tripsina, lo cual reduce la complejidad de los fragmentos del péptido a un nivel en que se puede realizar un análisis más fácilmente. Mediante el uso de dos marcajes diferentes con composiciones isotópicas diferentes, se pueden distinguir péptidos que se originan a partir de las dos muestras haciendo el análisis mediante espectrometría de masas y se puede obtener una estimación relativa de la cantidad. Sin embargo, las limitaciones de la técnica ICAT incluyen una reducción insuficiente de la complejidad de muestras altamente complejas, requiriendo una separación ulterior mediante cromatografía líquida, y el hecho de que no se detectan las proteínas que carecen de cisteína.

55 Todas las técnicas mencionadas anteriormente presentan una serie de limitaciones concernientes, por ejemplo, a la sensibilidad, velocidad, resolución y capacidad para ser aplicadas en diferentes tipos de proteínas, e.g., proteínas solubles y unidas a membranas.

60 La WO 02/086081 describe un método de identificación de una proteína que incluye la ruptura de la proteína con un agente proteolítico para producir fragmentos de péptidos, contactando los fragmentos con una matriz que comprende reactivos ligantes, detecta uniones y compara uniones a un set de referencia.

Otros métodos para analizar muestras de proteínas conocidos en la técnica anterior incluyen la captura de péptidos generados por la tripsina utilizando anticuerpos, cada uno de los cuales se une específicamente a un péptido conocido de una proteína conocida (Scrivener, E. *et al.*, 2003, *Proteomics* 3(2), 122-8; WO 02/25287). Los péptidos capturados son caracterizados después por espectrometría de masas MALDI-TOF. Un acercamiento similar está descrito por Nelson *et al.* (1996, *Anal Chem* 67, 1153-8) donde anticuerpos específicos capturan proteínas intactas y las proteínas capturadas son extraídas y analizadas por espectrometría de masas.

Ambos tipos de acercamientos presuponen la identidad de los componentes de la proteína a ser analizados y requieren la generación de moléculas ligantes para cada proteína individual. Por lo tanto, para diseñar una matriz que detecte y meda, e.g., 2000 proteínas, estas 2000 proteínas o péptidos deben ser aislados o sintetizados y después hay que generar 2000 anticuerpos específicos u otras moléculas ligantes. En contraste, la presente invención puede detectar un gran número de péptidos, como 10.000, los cuales pueden representar otras tantas proteínas, mediante la utilización de muchos menos, tal como sólo 200, ligantes diferentes.

Descripción de la invención

Por consiguiente, un primer aspecto de la presente invención proporciona un método para el análisis de una muestra heterogénea de péptidos, o fragmentos de proteínas o péptidos, comprendiendo el método,

(a) separar la mezcla heterogénea de péptidos, o fragmentos de proteínas o péptidos, en clases heterogéneas mediante la unión de los miembros de cada clase a una distancia espaciada en una posición definida sobre una matriz, en donde los miembros de cada clase tienen una configuración común a esa clase; y

(b) caracterizar los péptidos o fragmentos de proteínas o péptidos, en cada clase mediante la determinación de las masas de los péptidos, o fragmentos de proteínas o péptidos, en clases heterogéneas, y la determinación de la cantidad de péptidos, o fragmentos de proteínas o péptidos, de masas diferentes en clases heterogéneas.

La mezcla heterogénea de péptidos o de proteínas puede ser extraída a partir de una muestra celular o de tejido, o derivada de la fragmentación de una muestra heterogénea de péptidos o proteínas extraídas a partir de una muestra celular o de tejido, típicamente (pero no necesariamente) de origen humano. La muestra celular o de tejido puede ser derivada de un tejido normal o enfermo. La muestra celular o de tejido puede ser derivada de tejidos en varios estados de diferenciación o de actividad. Fuentes apropiadas adicionales de proteínas y péptidos incluyen las líneas celulares eucariotas y procariotas, materiales tisulares de ratones knockout y otros modelos animales así como plantas transgénicas y material de planta.

La muestra heterogénea puede ser procesada antes del análisis para extraer la cantidad de proteínas o péptidos en particular, como la albúmina y/o inmunoglobulinas en una muestra de suero, o para aumentar una muestra para una proteína o péptido en particular o grupo de proteínas o péptidos.

Cada clase heterogénea de péptidos o proteínas consta de todos los péptidos o proteínas en la muestra heterogénea que se unirá a una molécula ligante específica presente en la matriz. La molécula ligante se selecciona por su capacidad para unirse a la configuración, preferiblemente a una proteína o péptido particular, y por ello la molécula ligante puede unir tipos diferentes de proteínas y péptidos que contienen la misma configuración. Preferiblemente cada molécula ligante es específica para una configuración dada. Por lo tanto, una clase heterogénea de proteínas o péptidos se une a una molécula ligante dada en un método de la presente invención que comprende típicamente, como media, al menos dos, típicamente más de dos, p. ej. 10, 20, 50, 100, 200, 500, 1000 o más, tipos diferentes de proteínas o péptidos. Por "tipo diferente" incluimos el significado de la diferencia de las proteínas y péptidos en la secuencia aminoacídica, masas, modificaciones post-traduccionales y semejantes.

Por consiguiente, las proteínas y péptidos son clasificadas mediante la presente invención según su capacidad para ser capturadas y retenidas por una molécula ligante específica. Una clase heterogénea de péptidos o proteínas se unirá a una molécula ligante específica debido a la presencia de una configuración común de todos los miembros de una clase particular. La identidad de la unión a la configuración en cada clase de péptidos es, por lo tanto, una consecuencia de la especificidad de unión de la molécula ligante que define a cada clase.

La configuración puede ser una secuencia lineal o no lineal de aminoácidos tal como cuatro, cinco, seis, siete, ocho, nueve, diez o más aminoácidos. Una configuración lineal es formada a partir de aminoácidos adyacentes. Una configuración no lineal comprende aminoácidos no adyacentes en la secuencia pero que están generados en una proximidad cercana a los otros como resultado del pliegue tridimensional de la proteína o péptido.

Las moléculas ligantes en la matriz pueden ser específicas a las secuencias en puntos particulares dentro de una proteína o péptido, como las secuencias del C-terminal, el N-terminal, o en una posición definida relacionado a una característica interna, como una secuencia o un aminoácido modificado. Por ejemplo, todas las moléculas ligantes en la matriz pueden ser específicas para las secuencias del C-terminal, pero cada tipo de molécula ligante puede ser específica para una secuencia diferente de C-terminal que otros tipos de moléculas ligantes en la matriz.

Similarmente, las moléculas ligantes sobre la matriz pueden ser específicas a secuencias que contienen una mezcla de aminoácidos "constantes" y variables. Los aminoácidos constantes (como se define más adelante) pueden proporcionar una característica constante común a todas las uniones a las configuraciones mediante todas las moléculas ligantes en la matriz. Sin embargo, la identidad exacta de la unión a la configuración para cada tipo de molécula ligante en la matriz puede diferir según la inclusión, en cada configuración, de un set diferente de aminoácidos variables.

Generalmente la configuración de cada péptido o proteína contendrá tres, cuatro o cinco aminoácidos variables. Estos aminoácidos variables pueden ser identificados como parte de la configuración debido a su posición dentro del péptido o proteína (e.g., relacionado al C-terminal, el N-terminal, o a una característica interna) y/o por formar parte de una configuración mayor que también contiene aminoácidos "constantes".

Adicionalmente o alternativamente una característica de la configuración puede ser la presencia de un aminoácido modificado, como un aminoácido fosforilado o un aminoácido glicosilado. Preferiblemente, la configuración debe contener al menos un aminoácido no modificado. De más preferencia, todos los aminoácidos en la configuración son no modificados.

Fragmentación de la muestra

El método de la invención puede comprender el paso inicial de fragmentación de la muestra heterogénea de proteínas o péptidos para producir una muestra heterogénea de fragmentos de péptidos.

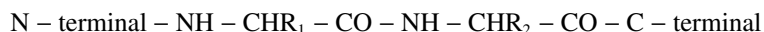
La fragmentación de una muestra heterogénea de proteínas o péptidos puede ser ventajosa debido a que puede incrementar el número de moléculas de péptidos que representan cada proteína o péptido originales. Por ejemplo, si una proteína es fragmentada en la muestra original, la unión a cualquiera de sus múltiples fragmentos puede ser utilizada como un marcador o como presencia y cantidad de tal proteína. En otras palabras, la fragmentación incrementa las probabilidades para cualquier proteína o péptido particular que estarán representadas en cualquier clase heterogénea dada. Esto significa que menos moléculas ligantes pueden ser utilizadas sin reducir la información que puede ser obtenida a partir de cada muestra analizada.

La fragmentación también permite la detección de proteínas de transmembrana, la cual, sin la fragmentación, no puede ser analizada.

TABLA 1

Enzima	Sitio Preferido
tripsina	R ₁ = Lys, Arg
quimiotripsina	R ₁ = Tyr, Phe, Leu, Ile, Val, Trp e His a un elevado pH
pepsina	R ₁ = Phe, Leu, muchos otros
trombina	R ₁ = Arg
papaína	R ₁ -Arg, Lys, Phe-X (región del CO del residuo próximo a Phe
Bromolaina	R ₁ = Lys, Ala, Tyr, Gly
Proteasa de <i>Staphylococcus aureus</i>	R ₁ = Glu, Asp
Factor Xa	R ₁ = Ile-Glu-Gly-Arg
termolisina	R ₂ = Tyr, Phe, Leu, Ile, Trp e His

En donde, R₁ y R₂ están definidos conforme a la fórmula siguiente:



El paso de fragmentación de la muestra heterogénea de proteínas, polipéptidos o péptidos puede ser alcanzada mediante cualquier método conocido en la técnica. Por ejemplo, puede ser utilizada la escisión química o enzimática. Numerosos métodos de escisión química o enzimática son conocidos en la técnica (i.e., dirigida por proteasas). Por ejemplo, las proteasas incluyen a la tripsina, quimiotripsina, pepsina, trombina, papaína, bromolaina, termolisina, subsilisina, Factor Xa, la proteasa de *Staphylococcus aureus*, y la carboxipeptidasa A. En una realización preferida, el método de fragmentación escindirá las proteínas, polipéptidos o péptidos en localizaciones definidas. La escisión enzimática es típicamente dirigida por secuencia, como es mostrado previamente en la Tabla 1. Los métodos de escisión química pueden también ser dirigida por secuencia, e.g., la fragmentación con bromuro de cianógeno, la cual escindirá una proteína o péptido en el sitio C-terminal de la metionina.

Así, por ejemplo, la escisión con tripsina es un medio de fragmentación dirigida por secuencia, ya que la escisión está dirigida por la presencia de residuos de arginina o lisina en una proteína, polipéptido o péptido, y por consiguiente produce la escisión de fragmentos que tienen, como residuo en su C-terminal, ya sea arginina o lisina. Un experto conocerá muchos otros métodos de fragmentación "dirigida", tales como los descritos en WO 02/25287.

Generalmente, la configuración en cada fragmento estará en la misma posición en cada fragmento respecto al sitio de escisión. De este modo, por ejemplo, donde son creados los fragmentos por el mecanismo de escisión dirigida por secuencia (ver debajo), entonces la configuración puede comprender uno o más aminoácidos adyacentes al sitio del terminal creado por la escisión, algunos de los cuales pueden ser constantes como resultado del mecanismo de escisión dirigida por secuencia.

De este modo, uno o más aminoácidos que forman la secuencia que dirige la escisión pueden ser retenidos en el fragmento. Por ejemplo, donde es utilizada la escisión con tripsina como método de fragmentación entonces, los fragmentos producidos tienen, en sus residuos terminales, bien arginina o bien lisina. De este modo la configuración puede incluir aminoácidos que forman parte del sitio de escisión.

Por tanto, la configuración en cada fragmento generado puede comprender uno o más, p. ej. dos, tres, cuatro o más aminoácidos constantes. Para los propósitos de la presente invención, un experto apreciará que el término “constante”, cuando es utilizado en el contexto de un aminoácido dentro de una configuración, incluye posiciones de aminoácidos en las cuales hay un bajo nivel de variabilidad, p. ej. 2, 3, 4, 5, 6, 7, 8, 9 ó 10 posibilidades diferentes. Son preferidos los números más bajos. Por ejemplo, la configuración en los fragmentos trípticos puede comprender el aminoácido del C-terminal, que es entonces un residuo constante ya sea de arginina o de lisina. En otras palabras, la identidad de un aminoácido “constante” no es tan aleatoria como la de otras posiciones “variables”.

De este modo, la configuración puede ser formada a partir de una mezcla de aminoácidos constantes y no constantes (i.e., variables). Generalmente la configuración contendrá tres, cuatro, o cinco aminoácidos variables, siendo los otros aminoácidos en la configuración (cualquiera que haya), constantes entre todos los fragmentos.

Matrices

El paso de separación de la mezcla heterogénea de proteínas, péptidos y/o fragmentos de éstos en clases heterogéneas es alcanzado en función de la presencia de una configuración mediante la unión de los miembros de cada clase a una distancia espaciada en una posición definida en la matriz.

Las matrices son bien conocidas *per se* en la técnica. Típicamente están formadas de una estructura lineal o bidimensional teniendo regiones (“puntos”) a una distancia espaciada (i.e., separadamente), teniendo cada una un área finita, formada de un soporte sólido en la superficie. Una matriz puede ser una estructura esférica en donde cada perla puede ser identificada mediante un código molecular o un código de color o identificada en un flujo continuo. El análisis puede ser también realizado secuencialmente en donde la muestra es pasada sobre una matriz de puntos adsorbiendo cada una las clases de moléculas a partir de la solución. El soporte sólido es típicamente de vidrio o de un polímero, siendo los polímeros más comúnmente utilizados los de celulosa, poliacrilamida, nylon, poliestireno, cloruro de polivinilo o de polipropileno. Los soportes sólidos pueden estar en forma de tubos, perlas, discos, chips de silicio, microplatos, membrana de fluoruro de polivinilideno (PVDF), membrana de nitrocelulosa, otras membranas porosas, membranas no porosas (e.g., plástico, polímero, vidrio acrílico, silicio, entre otras), una variedad de ejes poliméricos, o una variedad de pocillos de microtitulación, o cualquier otra superficie conveniente para la inmovilización de proteínas, polinucleótidos y otras moléculas convenientes y/o conducción de un inmunoensayo. Los procesos de unión son bien conocidos en la técnica y consisten generalmente del entrecruzamiento covalentemente de la unión o de la adsorción físicamente de una molécula de proteína, polinucleótido, o semejante al soporte sólido. Mediante la utilización de técnicas bien conocidas, como el marcaje por contacto o no contacto, enmascaramiento o la fotolitografía, pueden ser conocidas la posición de cada punto. Para revisiones ver a Jenkins, R.E., Pennington, S.R. (2001, *Proteomics*, 2,13-29) y Laf et al., (2002, *Drug Discov Today* 15; 7 (18 Suppl): S143-9).

Típicamente la matriz es una micromatriz. Por “micromatriz” incluimos el significado de una matriz de regiones que tienen una densidad de regiones discretas de al menos unos 100/cm², y preferiblemente de al menos unos 1000/cm². Las regiones de una micromatriz tienen dimensiones típicas, e.g., diámetros, en un rango aproximado de 10-250 µm, y se separan de otras regiones de la matriz por la misma distancia aproximadamente.

Típicamente los puntos sobre la matriz comprenden un número de tipos diferentes de moléculas ligantes (como se define más adelante), cada tipo siendo inmovilizado en un punto separado en las matrices. Así, utilizando un método que genera puntos con localizaciones definidas, es posible conocer la identidad y/o la afinidad de unión a cada punto en la matriz.

Preferiblemente, cada tipo de molécula ligante, y por lo tanto, cada punto, es capaz de unir específicamente a una configuración previamente definida y los tipos diferentes de moléculas ligantes tienen especificidades diferentes de unión. De este modo las proteínas, péptidos y/o fragmentos de éstos que se unen a un punto compartirán una configuración común. Contrariamente, las proteínas, péptidos y/o fragmentos de éstos sobre los diferentes puntos están separados en clases heterogéneas basados en la presencia de configuraciones diferentes.

De este modo, en donde la configuración sea una secuencia terminal, como la secuencia del C-terminal, la molécula ligante de un punto se unirá específicamente a las proteínas, péptidos y/o fragmentos de éstos que comprenden una primera secuencia dada del C-terminal, así como una molécula ligante en otro punto se unirá específicamente a proteínas, péptidos y/o fragmentos de éstos que comprenden una segunda secuencia dada del C-terminal, siendo diferentes la primera y segunda secuencias del C-terminal.

En una realización, todas las moléculas ligantes en la matriz son específicas para la configuración del C-terminal. En otra realización, todas las moléculas ligantes en la matriz son específicas para las configuraciones del N-terminal. En otra realización, todas las moléculas ligantes en la matriz son específicas para las configuraciones que no están conservadas posicionalmente.

Donde son fragmentadas las proteínas o péptidos antes del análisis, las configuraciones diana definidas pueden ser seleccionadas dependiendo del método de fragmentación utilizado. Por ejemplo, donde es utilizada la escisión con tripsina como método de fragmentación entonces, como se discutió previamente, los fragmentos producidos tienen, en su residuo C-terminal, ya sea arginina o lisina. De este modo, debe ser útil separar fragmentos basados, por ejemplo, en los primeros cuatro residuos del C-terminal. Dado que cada fragmento tendrá ya sea arginina o lisina como su residuo C-terminal, entonces la variabilidad será encontrada en las posiciones 2, 3, y 4 (respecto al residuo C-terminal que, en este contexto, está designado como posición 1). En este ejemplo, el nivel máximo de variabilidad desplegada por el tetrapéptido del C-terminal será de $2 \times 20 \times 20 \times 20 = 16.000$ posibles diferentes configuraciones. Utilizando el mismo esquema, si la configuración utilizada para clasificar los fragmentos tripticos está basada en, por ejemplo, sus primeros cinco residuos del C-terminal, entonces el nivel de variabilidad máximo desplegado será de $2 \times 20 \times 20 \times 20 \times 20 = 320.000$ posibles diferentes configuraciones.

Un experto apreciará que el número total de diferentes configuraciones terminales generados puede ser incrementado mediante el aumento del número de aminoácidos variables en cada configuración diana y disminuidas mediante el reemplazo de aminoácidos variables por aminoácidos constantes. Por otra parte, la cantidad de cada configuración en la muestra heterogénea de proteínas, péptidos o fragmentos de éstos puede ser por lo tanto incrementada mediante la reducción del tamaño de la configuración y la disminuidos mediante el incremento del tamaño de la configuración.

Así, puede usarse un método de fragmentación que utiliza un mecanismo de escisión dirigida por secuencia para generar fragmentos que tienen un aminoácido terminal definido o una secuencia terminal definida para reducir el número total de configuraciones terminales diferentes ante cualquier longitud de la configuración dada.

Una matriz conveniente para utilizar en un método previamente definido puede comprender un número de tipos diferentes de moléculas ligantes, cada tipo inmovilizada en una posición diferente y definida en la matriz, en donde cada tipo de molécula ligante es capaz de unirse específicamente a una configuración como se describió previamente y en donde los tipos diferentes de moléculas ligantes tienen especificidades de unión diferentes.

No es necesario que la matriz tenga tantos tipos diferentes de moléculas ligantes como posibles configuraciones diferentes. Esto es debido a que cada molécula ligante es específica solo para una configuración, no para una proteína particular (a diferencia de los métodos de la técnica anterior, como en WO 02/25287), y de esta manera diferentes proteínas múltiples, péptidos o fragmentos de éstos pueden ser unidos en un punto dado de la matriz. Por otra parte, donde la proteína o péptido de la muestra sea fragmentada antes del análisis, cada proteína o péptido en la muestra original puede generar entonces múltiples fragmentos. De este modo la matriz puede proporcionar un número conveniente de tipos diferentes de moléculas ligantes tal como de al menos un fragmento a partir de cada proteína o péptido en la muestra puede unirse específicamente a una molécula ligante.

De hecho, un experto apreciará que la muestra heterogénea de proteínas o péptidos puede ser caracterizada útilmente aún si no pueden ser representadas todas las proteínas o péptidos de la muestra no fragmentada. Idealmente, el número de tipos diferentes de moléculas ligantes proporcionadas en una matriz es conveniente para capturar al menos 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, 98, 99 o sustancialmente el 100% de los tipos de proteínas o péptidos en la muestra, o al menos un fragmento derivado del porcentaje anterior de los tipos de proteínas o péptidos en una muestra. El porcentaje como es utilizado aquí no se refiere al contenido total de proteína en masa, ya que una muestra puede comprender muchas proteínas diferentes pero una proteína particular puede predominar y, en este caso, la unión de la proteína predominante a la exclusión de todas las otras podría representar la captura de un alto porcentaje de la proteína de la muestra, aún brindaría poca o ninguna información proteómica. Más bien, el porcentaje es utilizado para reflejar la variedad de las diferentes especies proteináceas de la muestra, independientemente de la cantidad de cada especie. De este modo cada diferente tipo de proteína o péptido en la muestra no fragmentada representa "uno" y el porcentaje de captura de las proteínas o péptidos de la muestra puede ser determinado mediante el método de la presente invención por la suma de todos los fragmentos diferentes de proteínas o péptidos en la muestra no fragmentada como son determinados por los métodos conocidos en la técnica anterior como la electroforesis bidimensional unida a una espectrometría de masas, y multiplicando por cien.

Como en el ejemplo *in silico*, una degradación simulada por la tripsina de 10.000 secuencias de proteínas extraídas a partir de SwissProt resulta en 400.000 fragmentos de péptidos. La cantidad de péptidos que tiene cada tipo de la configuración posible del tetrapéptido con C-terminal varía entre un 0-10%. Una matriz conveniente puede ser formada mediante la elección de moléculas ligantes con una afinidad para unas cantidades de configuraciones convenientes, y así un número limitado de diferentes moléculas ligantes serán capaces de capturar un gran set de fragmentos diferentes. Por ejemplo, sólo 200 diferentes de dichas moléculas ligantes, cada una capturando como promedio 100 péptidos, capturarán 20.000 fragmentos a partir de la digestión con tripsina de una preparación de proteína hechas de una muestra de tejido. El análisis *in silico* del proteoma teórico que consta de todas las secuencias de proteínas humanas en SwissProt (aprox. 10.500 secuencias) indica que, si las configuraciones son seleccionadas aleatoriamente de todas las configuraciones posibles con una frecuencia teórica de aproximadamente 100 en el proteoma definido previamente, los péptidos capturados contendrían uno o más péptidos del 75% de todas esas proteínas. Una selección racional de las moléculas ligantes para evitar el solapamiento innecesario (mediante la captura de muchos péptidos de ciertas proteínas y ninguno de otras) incrementará más la cobertura.

Por consiguiente, la matriz puede tener al menos 10, 50, 100, 150, 200, 250, 300, 350, 400, 500, 600, 700, 800, 900, 1000 o más tipos diferentes de moléculas ligantes como se han definido previamente.

Cada punto sobre la matriz puede unirse como promedio a 2, 4, 6, 8, 10, 20, 40, 60, 80, 100, 200, 400, 600, 800, 900, 1000, 1500, 2000, o más tipos diferentes de proteínas, péptidos o fragmentos de éstos, cada uno teniendo la misma configuración. En este contexto, “tipos diferentes” de proteínas, péptidos o fragmentos de éstos se refiere a proteínas, péptidos o fragmentos de éstos que tienen al menos lo siguiente: secuencias diferentes; masas moleculares diferentes; y/o diferentes modificaciones post-translacionales.

Moléculas ligantes

Las moléculas ligantes pueden ser seleccionadas a partir de una librería, según sus capacidades de unión a una configuración dada, como se explica más adelante.

Al menos un tipo, más típicamente todos los tipos de moléculas ligantes pueden ser un anticuerpo o fragmentos o variantes de éstos.

De este modo, un fragmento puede contener uno o más de los dominios de variable pesada (V_H) o variable ligera (V_L). Por ejemplo, el término fragmento de anticuerpo incluye moléculas semejantes al Fab (Better *et al* (1988) *Science* 240, 1041); moléculas Fv (Skerra *et al* (1988) *Science* 240, 1038); moléculas Fv de simple cadena (ScFv) en donde los patrones de los dominios V_H y V_L están unidos vía un oligopéptido flexible (Bird *et al* (1988) *Science* 242, 423; Houston *et al* (1988) *Proc. Natl. Acad. Sci. EE.UU.*, 5879) y anticuerpos de dominio lineal (dAbs) que comprenden dominios V aislados (Ward *et al* (1989) *Nature* 341, 544).

El término “variante de anticuerpo” incluye cualquiera de los anticuerpos sintéticos, anticuerpos recombinantes o anticuerpos híbridos, tales como pero no limitados a, molécula de anticuerpo de simple cadena producida mediante el desarrollo de inmunoglobulina en la luz del fago y/o las regiones constantes y/o variables de la cadena pesada, u otras moléculas inmunointeractivas capaces de unirse a un antígeno en un formato de inmunoensayo que es conocido por los expertos en la técnica.

Una revisión general de las técnicas complejas en la síntesis de fragmentos de anticuerpos los cuales retienen sus sitios de unión específicos es encontrada en Winter & Milstein (1991) *Nature* 349, 293-299.

Adicionalmente o alternativamente al menos un tipo, más típicamente todos los tipos de moléculas ligantes es un aptámero.

Adicionalmente o alternativamente al menos un tipo, más típicamente todos los tipos de moléculas ligantes es un polinucleótido.

Selección de las moléculas ligantes

Las bibliotecas moleculares como las bibliotecas de anticuerpos (Clackson *et al*, 1991, *Nature* 352, 624-628; Marks *et al*, 1991, *J Mol Biol* 222(3): 581-97), bibliotecas de péptidos (Smith, 1985, *Science* 228(4705): 1315-7), bibliotecas expresadas en ADNc (Santi *et al* (2000) *J Mol Biol* 296(2): 497-508), bibliotecas de soportes diferentes a los marcos de anticuerpos como son los affibodies (Gunneriusson *et al*, 1999, *Appl Environ Microbiol* 65(9): 4134-40) o bibliotecas basadas en aptámeros (Kenan *et al*, 1999, *Methods Mol Biol* 118, 217-31) pueden ser utilizadas como fuente a partir de la cual las moléculas ligantes son seleccionadas para una configuración dada para utilizarse en los métodos de la invención.

Las bibliotecas moleculares pueden expresadas *in vivo* en células procariotas (Clackson *et al*, 1991, *op. cit.*; Marks *et al*, 1991, *op. cit.*) o en células eucariotas (Kieck *et al*, 1999, *Proc Natl Acad Sci EE.UU.*, 96(10): 5651-6) o bien expresarse *in vitro* sin la participación de células (Hanes & Pluckthun, 1997, *Proc Natl Acad Sci EE.UU.* 94(10):4937-42; He & Taussig, 1997, *Nucleic Acids Res* 25(24):5132-4; Nemoto *et al*, 1997, *FEBS Lett*, 414(2):405-8).

Cuando son utilizadas las bibliotecas basadas en proteínas con frecuencia los genes que codifican las moléculas ligantes potenciales de las bibliotecas son empaquetados en virus y la molécula ligante potencial es expresada en la superficie del virus (Clackson *et al*, 1991, *op. cit.*; Marks *et al*, 1991, *op. cit.*; Smith, 1985, *op. cit.*).

El más comúnmente utilizado hoy en día de estos sistema son los bacteriófagos filamentosos que expresan fragmentos de anticuerpos en sus superficies siendo expresados como una fusión en la proteína secundaria de la cubierta del bacteriófago (Clackson *et al*, 1991, *op. cit.*; Marks *et al*, 1991, *op. cit.*). Sin embargo, también han sido utilizados otros sistemas de expresión utilizando otros virus (EP 39578), bacterias (Gunneriusson *et al*, 1999, *op. cit.*; Daugherty *et al*, 1998, *Protein Eng* 11(9):825-32; Daugherty *et al*, 1999, *Protein Eng* 12(7):613-21), y levaduras (Shusta *et al*, 1999, *J Mol Biol* 292(5):949-56).

Además, recientemente, los sistemas presentados que utilizan uniones del producto del polipéptido que codifica su ARNm en los llamados sistemas de expresión de los ribosomas (Hanes & Pluckthun, 1997, *op. cit.*; He & Taussig, 1997, *op. cit.*; Nemoto *et al*, 1997, *op. cit.*), o alternativamente ha sido presentado la unión del producto del polipéptido a la codificación del ADN (ver patente EE.UU. No. 5.856.090 y WO 98/37186).

Cuando las moléculas ligantes potenciales son seleccionadas a partir de las bibliotecas son generalmente empleados uno o unos pocos selectores de péptidos que tienen configuraciones definidas. Los residuos de aminoácidos que mantienen la estructura, la disminución de la flexibilidad en el péptido o las cadenas laterales cargadas hidrofóbicas o polares permitiendo la interacción con las molécula ligante puede ser utilizada en el diseño de las configuraciones para los selectores de péptidos. Por ejemplo,

- (i) La prolina puede estabilizar la estructura del péptido cuando su cadena lateral está unida al carbono alfa así como al nitrógeno;
- (ii) La fenilalanina, tirosina y el triptófano tienen cadenas laterales aromáticas y son altamente hidrofóbicos, mientras que la leucina y la isoleucina presentan cadenas laterales alifáticas y son también hidrofóbicas;
- (iii) La lisina, arginina, y la histidina tienen cadenas laterales básicas y estarán cargadas positivamente a pH neutro, mientras que el aspartato y el glutamato tienen cadenas laterales ácidas y estarán cargados negativamente a pH neutro;
- (iv) La serina, treonina, y la tirosina tienen cadenas laterales que contienen grupos hidroxilos, los cuales pueden participar en los puentes de hidrógenos.

La selección de las moléculas ligantes puede involucrar típicamente el uso de las tecnologías y sistemas de matrices para analizar la unión a los puntos correspondientes a los tipos de moléculas ligantes.

Las moléculas ligantes potenciales, e.g., fragmentos de anticuerpos en una biblioteca, pueden ser clonadas y reveladas en un formato de la matriz. La posición del punto puede correlacionarse con la identidad del clon. Después, los selectores de péptidos que tienen configuraciones definidas deberían permitir la unión a la matriz. Para los puntos que resultaron contener las moléculas ligantes contra la configuración definida de un selector de péptido particular, el selector de péptido particular se une, y la unión brinda una señal que permite al usuario determinar la posición del punto y, de este modo la identidad del clon a partir del cual fue obtenida la molécula ligante positiva. Los falsos positivos (e.g., moléculas ligantes que se unen a regiones del selector de péptido diferente a la configuración) pueden ser evitados, mediante la medición de la capacidad de los positivos presuntivos de unirse a péptidos similares sin la configuración, en donde la unión a estos péptidos similares indica que el unidor presuntivo es un falso positivo.

Similarmente, las bibliotecas de péptidos de polinucleótidos potenciales que se unen a las moléculas ligantes pueden ser investigadas para la capacidad de unirse a los péptidos selectores que tienen configuraciones definidas (e.g., utilizando el chip Affymatrix disponible comercialmente).

Una vez aislado un número conveniente de moléculas ligantes, un experto puede fabricar una matriz.

Un método para realizar una biblioteca de moléculas ligantes puede comprender los siguientes pasos:

- (a) proporcionar, como primer componente, un péptido selector que comprende una configuración como la definida previamente;
- (b) proporcionar, como segundo componente, una fuente de moléculas ligantes candidatas, como una biblioteca molecular como está definido previamente;
- (c) combinación del primer y segundo componentes; y
- (d) identificación de las moléculas ligantes candidatas que son capaces de unirse específicamente a la configuración del péptido escogido en el primer componente.

Una biblioteca, típicamente una biblioteca en donde los componentes han sido preseleccionados mediante el método precedente puede comprender aproximadamente al menos 10, 50, 100, 150, 200, 250, 300, o más tipos diferentes de moléculas ligantes, cada tipo siendo capaz de unirse específicamente a una configuración como está definido previamente y los tipos diferentes teniendo especificidades de unión diferentes. Al menos una molécula ligante en la biblioteca, generalmente todas las moléculas ligantes de la biblioteca, pueden ser anticuerpos o fragmentos o variantes de éstos, como el Fv, scFv o Fab; aptámeros; y/o polinucleótidos.

Una biblioteca de moléculas ligantes como la definida previamente puede ser utilizada para producir una matriz como la descrito previamente.

Un método para producir una matriz conveniente para utilizarla en un método en concordancia con el primer aspecto de la invención puede comprender,

- (a) proporcionar una biblioteca de tipos diferentes de moléculas ligantes, cada tipo siendo capaz de unirse específicamente a una configuración como la definida previamente y los tipos diferentes teniendo especificidades de unión diferentes; e

- (b) inmovilizar las moléculas ligantes en una matriz tal que los tipos diferentes de moléculas ligantes estén inmovilizados en localizaciones definidas y discretas.

Los métodos de inmovilización de moléculas ligantes como los anticuerpos, aptámeros, polinucleótidos y semejantes en localizaciones definidas y discretas en una matriz se han discutido previamente, y en cualquier caso son bien conocidos en la técnica.

Un sistema para analizar una muestra heterogénea de proteínas y péptidos puede comprender una matriz como la descrita previamente y un portador de datos que comprende la información de la identidad y/o propiedades de unión y la posición de cada tipo diferente de molécula ligante en la matriz. El portador de datos puede ser un portador de datos electrónico, típicamente en forma de portador de datos legible por ordenador. La información puede correlacionar la posición (punto) en la matriz con la identidad de un clon de librería que aportó la molécula ligante al punto de la matriz, permitiendo así al usuario investigar más sobre las características de una molécula ligante producida mediante un clon dado. De forma adicional o alternativa, el portador de datos puede comprender información sobre las características de unión de una molécula ligante en una posición dada en la matriz.

Condiciones de detección

Habiendo proporcionado una matriz conveniente, es posible analizar una muestra en concordancia con el método de la invención. Con el fin de separar una muestra heterogénea de proteínas, péptidos y/o fragmentos de éstos en clases heterogéneas mediante la unión de cada componente de la clase a una posición definida espaciada en la matriz, teniendo cada clase heterogénea una configuración común a esa clase, es importante que las condiciones de unión sean convenientemente estrictas para evitar básicamente una unión no específica.

La formación de los complejos molécula ligante:configuración puede ser realizada en una variedad de condiciones. Las soluciones de reacciones que contienen un fragmento de péptido puede contener grados variables de sal o estar presentadas en niveles de pH variables. Además, la reacción de unión puede ser llevada a cabo a diferentes temperaturas. En general, las condiciones de pH variarán entre 2-10 (de más preferencia en torno a un pH 8), temperaturas de 0°C-100°C y condiciones de salinidad de 1 μ M a 5M (en el caso del NaCl).

Después del paso de combinación de la muestra heterogénea de proteínas, péptidos y/o fragmentos de éstos con la matriz en condiciones que permitan especificidad, la matriz es típicamente lavada para extraer las proteínas, péptidos o fragmentos de éstos no unidos. Las soluciones apropiadas para lavar pueden contener sales, como el cloruro de sodio, agentes tampones como el tampón fosfato, agentes caotrópicos como la urea y detergentes como el Tween-20. La concentración de estos componentes, así como el pH de la solución, puede ser optimizada para obtener unas condiciones de lavado estrictas adecuadas. Antes del análisis de la espectrometría de masas MALDI-TOF (ver más abajo) la matriz debería ser lavada con agua destilada para eliminar las sales, detergentes, polímeros u otros componentes que puedan interferir con el análisis.

Un experto puede adaptar la reacción de unión y las condiciones de lavado para alcanzar una condición apropiada para evitar las uniones no específicas aplicando una mezcla de proteínas, péptidos y/o fragmentos de éstos que tengan secuencias conocidas para una matriz y la determinando si las proteínas, péptidos y/o fragmentos de éstos se unen no específicamente (i.e a las manchas que tienen moléculas ligantes de un tipo específico para una configuración que no esté contenida en las proteínas, péptidos y/o fragmentos de éstos en la mezcla). Si se dan uniones no específicas, puede incrementarse la severidad de las condiciones utilizadas. Alternativamente, el usuario puede reemplazar la molécula ligante responsable de la baja especificidad por una molécula ligante de mayor especificidad.

Las constantes de afinidad son una medida de la interacción entre un ligando particular y su co-receptor. La "afinidad de unión" o la medición de la fortaleza de asociación entre una molécula ligante particular y la diana de su configuración generalmente se mide mediante las constantes de afinidad para las concentraciones en equilibrio de las configuraciones asociadas y disociadas de la molécula ligante y su diana. Preferiblemente la unión de una molécula ligante a su configuración debiera ocurrir a una afinidad de aprox. $K_D = 10^{-6}$ M o más para ser útil para la presente invención, siendo más preferiblemente mayor que 10^{-7} M, y más preferiblemente entre aprox. 10^{-8} M y aprox. 10^{-11} M. Los fragmentos de anticuerpos tendrán por lo general afinidades de unión en el rango de entre aprox. 10^{-7} M a aprox. 10^{-8} M.

Caracterización de las clases heterogéneas de uniones de proteínas, péptidos y/o fragmentos de éstos

Una vez separados en clases heterogéneas en la matriz, las proteínas, péptidos o fragmentos de éstos en cada clase pueden ser entonces más caracterizados mediante las técnicas analíticas conocidas en la técnica tales como la espectrometría de masas por desorción (e.g., la espectrometría de masas MALDI-TOF, ver Roepstorff, P, 2000, *EXS*, 86: 81-97), para proporcionar información en forma de espectrogramas de masa, en los cuales cada pico indicará la presencia, la masa y la cantidad relativa de un péptido específico.

Cuando se realiza la fragmentación de la muestra antes del análisis de la muestra, la identidad de la proteína o péptido a partir del cual se derivan los fragmentos capturados (i.e., "proteína parental") puede ser determinada mediante colisión inducida por la espectrometría de masas por disociación, la cual puede ser utilizada para obtener información estructural de un péptido.

También, si se conoce la especificidad de la molécula ligante y se utilizan condiciones suficientemente estrictas, uno puede saber que una proteína, péptido o fragmento de éstos capturados sobre un mancha dada comprenden una configuración dada. Por ejemplo, si la configuración son los primeros cuatro aminoácidos del C-terminal, entonces es posible deducir la secuencia del tetrapéptido del C-terminal de todas las proteínas, péptidos o fragmentos de éstos en una mancha dada.

La información del contenido de la configuración, en combinación con una determinación precisa de la masa obtenida mediante la espectrometría de masas, puede ser suficiente para enlazar la información contra una proteína, péptido o fragmentos de éstos generados mediante un análisis *in silico* de una base de datos de secuencias de proteínas, o en una digestión *in silico* de las secuencias presentes en ellas.

Por consiguiente, el paso de caracterización de las proteínas, péptidos o fragmentos de éstos en cada clase heterogénea comprende caracterizar las uniones de proteínas, péptidos o fragmentos de éstos en cada posición definida y diferente de la matriz, mediante la determinación de la masa de las proteínas, péptidos, o fragmentos de éstos. Generalmente esto se hace mediante espectrometría de masas por desorción. El paso de caracterización de los fragmentos en cada clase heterogénea puede comprender adicionalmente la determinación de la identidad de las proteínas o péptidos en la muestra heterogénea no fragmentada a partir de la cual se derivan los fragmentos detectados (i.e., "los parentales"). Esto se realiza típicamente mediante la colisión inducida por la espectrometría de masas. Los datos adquiridos de este modo pueden proporcionar información de la secuencia o pueden ser utilizados para investigar en las bases de datos de secuencias de proteínas para enlazar las secuencias.

La intensidad relativa de la señal obtenida a partir del péptido específico mediante espectrometría de masas es dependiente de la concentración, peso molecular, y las características de ionización del péptido. La calidad de la cuantificación puede ser mejorada mediante la adición de proteínas de referencia marcadas con isótopos (Goshe B G y Smith D S (2003) *Opinion Curr Biotech*, 14: 101-109).

La información concerniente a la cantidad de un fragmento y a la identidad de la proteína parental puede ser utilizada para cuantificar la proteína o péptido parentales en la muestra heterogénea no fragmentada.

Uno de los beneficios de la presente invención puede ser visto en los análisis de cada clase heterogénea de proteínas, péptidos o fragmentos de éstos. La presente invención proporciona un método en el cual cada clase heterogénea es analizada sin la necesidad de una separación adicional de los componentes de cada clase. De este modo la presente invención tiene ventajas sobre los métodos de la técnica precedente la cual utiliza pasos múltiples de separación por afinidad (como en WO 02/060377), dado que los métodos de la técnica precedente dependen de múltiples pasos de captura/extracción del péptido y de un complejo sistema de manejo de fluidos, los cuales son laboriosos y consumen tiempo. Por el contrario, la presente invención proporciona un método en un solo paso para el sub-fraccionamiento de las proteínas, péptidos, o fragmentos de éstos en clases heterogéneas diferentes seguido de una caracterización directa de cada clase, e.g., mediante espectrometría de masas.

Adicionalmente, la presente invención proporciona información cualitativa y cuantitativa acerca de cada clase heterogénea. Por ejemplo, puede ser determinado el peso molecular y la cantidad de cada especie dentro de cada clase. Esto es una mejora sobre la técnica precedente (e.g., WO 02/060377), la cual solo proporciona la determinación de la cantidad total de la proteína en cualquier mancha.

45 Aplicaciones

Una aplicación de la invención es en la comparación entre muestras diferentes. Un experto apreciará que los datos generados mediante un método en concordancia con la presente invención pueden ser extremadamente complejos e implicar varios miles de unidades de datos diferentes. Puede ser apropiado reunir, almacenar y analizar los datos generados por medios electrónicos. Por lo tanto, un portador de datos puede comprender información obtenida por un método acorde con el primer aspecto de la presente invención. Puede proporcionarse un sistema electrónico de procesamiento de datos, como una computadora, que comprenda un portador de datos que comprenda información obtenida por un método acorde con el primer aspecto de la invención y medios para la comparación de la información obtenida a partir del análisis de las diferentes muestras. En este contexto, un medio para comparar es típicamente un programa de ordenador diseñado para comparar los datos generados a partir del análisis de una variedad de muestras y resaltar diferencias entre las muestras, permitiendo con ello al usuario identificar fácilmente las proteínas y péptidos candidatos de interés.

Dichas comparaciones pueden incluir muestras de, e.g., tejido normal y tejido enfermo o de e.g., tejidos en varios niveles de diferenciación o de activación. La invención puede así ser utilizada para comparar rápidamente y eficientemente un set grande de muestras con el fin de investigar las diferencias en la composición de proteína o péptido. Dichas diferencias pueden usarse para identificar moléculas con potencial como dianas de fármacos.

Por tanto, un método para identificar las diferencias en la composición entre dos o más muestras heterogéneas de proteínas, polipéptidos o péptidos puede comprender analizar cada muestra mediante un método acorde con el primer aspecto de la presente invención, consecuentemente para identificar cualquier diferencia.

Una matriz o sistema como el descrito previamente puede utilizarse para analizar una o más muestras heterogéneas de proteínas, péptidos y/o fragmentos de éstos, utilizando los métodos descritos previamente. El uso puede ser para identificar una proteína relacionada a una enfermedad mediante el análisis de al menos una muestra, típicamente una muestra *ex vivo*, procedente de un individuo con la enfermedad y al menos una otra muestra, típicamente una muestra *ex vivo*, procedente de un individuo sin la enfermedad. Las enfermedades convenientes para el análisis incluyen enfermedades neurodegenerativas, cáncer, enfermedades infamatorias, enfermedades cardiovasculares y desórdenes metabólicos.

De este modo, un método para la identificación de una proteína, polipéptido o péptido relacionado a una enfermedad puede comprender la identificación de diferencias entre dos o más muestras mediante el método precedente, en donde al menos una de las muestras analizadas es procedente de un individuo con la enfermedad y la otra de las muestras analizadas es procedente de un individuo sin la enfermedad.

Además, una vez identificada la proteína o el péptido relacionado con la enfermedad, puede suministrarse un método para diagnosticar el estado de la enfermedad de un individuo consistente en analizar una muestra, típicamente una muestra *ex vivo* tomada del individuo, mediante un método según el primer aspecto de la presente invención, y determinar si los resultados corresponden con la proteína, polipéptido o péptido asociados a la enfermedad mediante el método descrito previamente.

Después de diagnosticar una dolencia o enfermedad a un individuo usando los métodos precedentes, dicho individuo puede ser caracterizado por tener la necesidad de un régimen de tratamiento adecuado a la enfermedad dada diagnosticada. Por tanto, un método de tratar a un individuo al que se le ha detectado la necesidad del mismo mediante un método de diagnóstico como el definido previamente, puede comprender administrar una cantidad efectiva de un agente farmacéutico apropiado para el estado de la enfermedad del individuo. Los médicos serán capaces de determinar la cantidad efectiva del agente farmacéutico según la edad, peso, género y condición del paciente.

Un agente farmacéutico puede ser utilizado en la fabricación de un medicamento para el tratamiento de un individuo identificado por tener la necesidad de este mediante un método diagnóstico como el definido previamente.

La invención será descrita ahora en más detalle con referencia a las siguientes figuras y ejemplos no limitantes en donde:

La figura 1 muestra una visión general esquemática de una realización de la presente aplicación.

Las figuras 2-14 muestran el espectro de masas generado mediante el análisis de los fragmentos de péptidos trípticos unidos a las moléculas ligantes seleccionadas por sus capacidades de unión a los C-terminal de los tetrapéptidos y hexapéptidos que tienen ya sea arginina o lisina como residuo en el C-terminal.

Ejemplo 1

Este ejemplo describe cómo una micromatriz puede ser producida y utilizada para detectar péptidos generados a partir de una mezcla heterogénea de proteínas. En este ejemplo, elegimos fragmentar las proteínas en péptidos mediante la digestión con tripsina y capturar las subclases de fragmentos de péptidos que utilizan moléculas Fv de cadena simple (scFv) con propiedades de unión dirigidas hacia el C-terminal de los péptidos.

Generación de las moléculas ligantes

Diseño de los péptidos selectores: Los péptidos sintéticos se usan como agentes receptores cuando se aíslan moléculas Fv de simple cadena (scFv) convenientes a partir de una biblioteca de expresión de fagos. Los péptidos son diseñados para capturar partículas de fagos que expresan scFv con afinidad por el C-terminal de un tetrapéptido en el cual el último aminoácido (i.e., C-terminal) era la arginina o la lisina. Se puede añadir un espaciador al lado del N-terminal de este tetrapéptido así como del N-terminal de la biotina. Las secuencias de aminoácidos son diseñadas para incluir aminoácidos que es muy probable que generen buenos epítopes, como los aminoácidos hidrofóbicos (fenilalanina, tirosina, triptófano, leucina e isoleucina) o aminoácidos cargados (aspartato, glutamato, asparagina, glutamina e histidina). La metionina es excluida debido a su tendencia a oxidarse, y la cisteína es excluida para evitar problemas con la dimerización debido a la formación de puentes disulfuros. Las secuencias de los tetrapéptidos también se decidieron en base a su frecuencia en proteínas que aparecen de forma natural. Ejemplos de secuencias convenientes son la biotina-SGSG-XXXX-COOH en donde XXXX puede ser, e.g., EDFR, EPER, HPDK, LPSR, LQSK, OJEADA, WDSR o YLDK.

Selección de los ligantes específicos a partir de una biblioteca de expresión de fagos

La selección de unidores específicos a partir de la biblioteca n-CoDeR puede ser realizada utilizando perlas magnéticas cubiertas con estreptavidina (Hawkins, R.E., Russel, S.J. and Winter, G. (1992) *J. Mol. Biol.*, 226, 889-896). La construcción y el manejo de la biblioteca de expresión de scFv del fago n-CoDeR está descrita en Söderlind *et al* (2000) *Nature Biotech.*, 18, 852-856.

Un volumen que contiene $1-2 \times 10^{13}$ UFC de las cepas de fagos de la biblioteca es mezclada con el péptido selector biotinilado (concentración final del péptido de aprox. 10^{-7} M). Se añade BSA a una concentración final del 3%, azida sódica a una concentración final del 0,02% y Tween 20 a una concentración final del 0,05%. Se incuba a temperatura ambiente con una agitación moderada durante 1 h. Se añaden las perlas magnéticas (prebloqueadas con albúmina) y se incuba durante 15 minutos a temperatura ambiente con una agitación moderada. Se concentra las perlas con el imán y se saca el sobrenadante. Se lava las perlas con 3x1 mL de BSA al 3%, Tween 20 al 0,05%, azida sódica al 0,02% en PBS, seguido de 3x1 mL de Tween 20 al 0,05% en PBS y finalmente 3x1 mL de PBS. Se extraen los fagos unidos mediante la adición de 400 μ L de una solución stock de tripsina (1 mg/mL, Boehringer-Mannheim). Se incuba durante 30 minutos a temperatura ambiente. Se transfiere el adsorbente a un tubo nuevo y se añaden 40 μ L de una solución stock de un inhibidor de la tripsina y la aprotinina (2 mg/mL). Se determina la cantidad de fagos en el adsorbente (mediante la medición de la cantidad de UFC después de la infección con *E. coli*).

Las nuevas cepas del scFv del fago son producidas a partir del adsorbente mediante la infección del crecimiento logarítmico de *E. coli* con los fagos extraídos. Se añade ampicilina para eliminar las bacterias no infectadas. Las bacterias infectadas son amplificadas durante 3 h aproximadamente, seguido de la infección con los fagos de ayuda y la inducción de IPTG por la producción de los scFv expresados por los fagos. El ciclo de selección descrito previamente es repetido dos veces, pero a una concentración de antígeno de 10^{-8} M para la segunda ronda y 10^{-9} para la tercera. El adsorbente final resultante es almacenado a 4°C.

20 *Ensayo primario de las moléculas ligantes*

El proceso de selección puede generar decenas de miles de clones de fagos, incluyendo unidores inespecíficos y unidores específicos de calidad diferentes. También, no todos los clones darán scFv funcionales. Los grupos de fagos extraídos de la tercera selección se utilizan para infectar *E. coli* y es aislado el ADN de los plásmidos (fagomido). El ADN específico del fago es eliminado mediante la digestión con enzimas de restricción y el material re-ligado es transformado dentro de *E. coli*. Los clones transformados que expresan, i.e., el scFv, se seleccionan usando ampicilina. Para identificar los clones que generarán las mejores moléculas ligantes para la aplicación dada, se emplea un procedimiento de ensayo a doble escala. El ensayo primario está diseñado para evaluar las propiedades de unión de un gran número de scFv expresados (típicamente 10.000) contra un ligando previsto y un no ligando previsto, y diferenciará scFv con una interacción específica *versus* interacción inespecífica con el péptido seleccionado así como para proveer una medida aproximada de la calidad relativa entre los unidores específicos.

El ensayo primario es típicamente realizado utilizando sistemas automatizados de alta productividad para la recogida de clones, la expresión y la evaluación.

Típicamente 10.000 colonias son colectadas por un colector de colonias Qbot (Genetix; Hampshire, UK) y transferidas a placas de 384 pocillos para el crecimiento individual durante toda la noche. Se transfieren 5 μ L de la suspensión bacteriana (replicado) a placas de expresión para el crecimiento y expresión en un sistema automatizado (Thermo CRS; Burlington, Ontario, Canada).

En el sistema ELISA (Thermo CRS), las placas a evaluar son precubiertas con estreptavidina (0,1 μ g/pocillo), incubadas toda la noche y lavadas. Las placas son entonces cubiertas con los péptidos biotinilados (1pmol/pocillo), incubadas durante 1 hora (o durante toda la noche a +4°C), lavadas y bloqueadas (tampón de bloqueo: gelatina al 0,45% en 1xPBS con Tween al 0,05%).

Los sobrenadantes se añaden entonces desde las placas de expresión (10 μ L) a las placas a evaluar y se incuban durante 1 hora, seguido de un paso de lavado.

Los anticuerpos secundarios (anticuerpos anti-his del ratón conjugado con HRP) son entonces añadidos e incubados durante 1 hora, seguido de un paso de lavado.

El sustrato (Pierce Supersignal ELISA Pico) es añadido seguido de 10 minutos de incubación antes de leerse en un modo luminiscente.

Se reúnen y reevalúan (confirmación del pico) los activos (clones de más de 10 veces el radio de la señal de ELISA entre los péptidos dianas y no dianas).

La especificidad de los clones es típicamente realizada en un ensayo secundario en donde se evalúa un set mayor de péptidos. Las picos seleccionados con una alta especificidad son después secuenciados para obtener clones de pico únicos. Hasta 96 picos son secuenciadas por colonia RCP y la secuenciación del ciclo de terminación con colorante, utilizando el Analizador de ADN ABI PRIS M 3100 (Applied Biosystems, Warrington, UK).

Secuenciación

Los clones identificados como ligantes específicos durante el ensayo son analizados mediante secuenciación del ADN para identificar los clones únicos.

El gen que codifica para el scFv es secuenciado en concordancia con el método de terminación de cadena o dideoxi utilizando un ADN amplificado por RCP como plantilla, cebadores hechos a medida y el kit Big Dye Terminator RR (Applied Biosystems, EE.UU.). Los fragmentos terminados son separados y analizados utilizando un Analizador Genético 3100 (Applied Biosystems).

Caracterización de los ligandos

Un método para determinar si los scFv capturarán realmente un número apropiado de tipos de péptidos a partir de la muestra digerida con tripsina es la extracción por inmovinoafinidad unida a un análisis de espectrometría de masas.

Una muestra que contiene proteínas de plasma es reducida (e.g., con mercaptoetanolamina), alquilizada (e.g., con yodoacetamida), y digerida con tripsina (20 µg de tripsina/mg de proteína de plasma, incubación 6h a 37°C).

Los scFv 6 x His marcados pueden ser capturados en una pequeña columna (ZipTip™, Millipore), antes de ser modificados con iones Ni²⁺ (protocolo TN229, Millipore, EE.UU.). En principio, la inmovilización de los scFv que selecciona los péptidos a partir de las proteínas de interés hidrolizadas con tripsina es realizada mediante ciclos consecutivos de aspiración-dispensión de una solución de scFv (10-50 µg/ml en un tampón neutral o ligeramente básico, ≈10 µl) en el ZipTip™ modificado con Ni. Después de la extraer las moléculas de scFv no unidas, los antígenos son capturados dentro de las columnas de afinidad en un método similar al descrito previamente (e.g., mediante ciclos consecutivos de aspiración-dispensión a partir de ≈ 10 µl de tripsina digerida, previamente diluida a una concentración de 2-3 mg de proteína/ml en PBS). Después de capturar a los antígenos, la columna es lavada repetidamente para remover a los péptidos no unidos. Este paso de lavado puede ser realizado con PBS, si es requerido un lavado más riguroso, con una solución que contiene una alta concentración de sal (e.g., cloruro de sodio), agentes desnaturalizantes (por ejemplo, guanidina o la urea) o un detergente, como el Tween 20. Los péptidos capturados son extraídos en un medio de extracción de ≈ 1 µl (e.g., 5% de ácido acético o 50% de acetonitrilo + 0,1% de ácido trifluoroacético (TFA)) directamente en una placa diana MALDI-TOF (desorción/ionización por láser asistido por matriz-tiempo de vuelo). La solución de la matriz (e.g., ácido alfa-ciano-4-hidroxicinámico saturado en 1% de TFA, 75% de acetonitrilo) es entonces añadida en la superficie de cada una de las manchas de muestra y puesta a secar. Alternativamente, el compuesto de la matriz puede ser disuelto directamente en una solución utilizada para la extracción de los péptidos a partir de la columna de inmuoextracción.

Las muestras preparadas así son entonces analizadas mediante espectrometría de masas MALDI-TOF.

Generación de las matrices de afinidad

Los scFv seleccionados 6 x His marcados son expresados en *E. coli*, dializados y purificados en una columna Ni-NTA. Después de la extracción, los scFv son concentrados hasta 1-3 mg/ml en PBS. Luego, los scFv con selectividades diferentes son marcados (utilizando cualquiera de las tecnologías actuales existentes para el marcaje de proteínas, por ejemplo un marcaje por contacto o por no contacto) en un soporte conveniente (e.g., láminas de vidrio derivadas o el fondo del pocillo de una placa de microtitulación). Los scFv pueden ser inmovilizados ya sea covalentemente (e.g., a través de grupos amino, aldehído o epoxi reactivos) sobre la superficie del soporte o no covalentemente (por ejemplo, adsorción pasiva en superficies modificadas de poliestireno o nitrocelulosa: para una revisión, ver Jenkins R.E y Penninton, S.R. (2001) *Proteomics*, 1, 13-29). Además, es posible la inmovilización orientada de los scFv, ya sea por vía de una lámina de vidrio modificada con níquel capaz de unir el 6 x His marcado, o mediante una unión covalente a las láminas de vidrio modificadas con maleimida, uniéndose covalentemente a los Cys marcados, introducidos previamente en la estructura del scFv. La alta producción de la micromatriz puede ser explotada mediante el marcaje de 1000-2000 scFv diferentes en la misma lámina para un análisis simultáneo de varios antígenos a partir de la misma muestra. En este ejemplo, 200-300 moléculas ligantes diferentes pueden ser suficientes, cada una marcada por duplicado o triplicado, dando un número total de 400-1000. Las matrices pueden ser almacenadas a 4°C por varias semanas.

Análisis de la muestra compleja

Preparación de la muestra: La muestra a ser analizada, e.g., plasma, puede ser directamente digerida con tripsina después de transferirla o diluirla en un tampón conveniente (e.g., 50 mM de bicarbonato de sodio, pH 7.0). Alternativamente, la muestra puede ser prefraccionada para aumentar las proteínas de interés o para eliminar determinados componentes como la albúmina e inmunoglobulinas (Anderson NL, Anderson NG. (2002) *Mol Cell Proteomics*, 1(11): 845-67) para incrementar el límite de detección. Las proteínas de la muestra pueden ser reducidas y carboximetiladas para evitar puentes disulfuros entre los péptidos que contienen cisteína.

Aplicación de la muestra: Se aplicó 10-200 µl de la muestra digerida con tripsina sobre la micromatriz impresa y se incubó durante 2 horas, utilizando ya sea una cámara de incubación (Arrayit Hibbridization Casete, TeleChem International Inc, EE.UU.) o un instrumento de procesamiento de muestra automatizado (e.g., Protein Array Workstation, Perkin-Elmer, EE.UU.). Se lava la micromatriz repetidamente con e.g., 50 mM de tampón fosfato, pH 7.0, 0,1% de Tween, y 100 mM de cloruro de sodio. Para condiciones de lavado más rigurosas, pueden añadirse diferentes detergentes o sales a varias concentraciones.

Detección: La matriz que absorbe la UV (ácido alfa-ciano-4-hidroxicinámico, saturado en 1% de TFA, 75% acetonitrilo) se añade a la matriz (100-500 nL/mancha). La matriz es montada en una placa diana MALDI-TOF (Borrebaeck CAK, Ekström, Malmberg Hager AC, Nilsson J, Laurell T y Marko-Varga G (2001) *Biotechniques* 30, 1126-1132) y son adquiridos los espectros de masas de cada mancha usando un espectrómetro de masas en modo reflector.

Ejemplo 2

Este ejemplo describe cómo una matriz de afinidad puede ser producida y utilizada para detectar péptidos generados a partir de una mezcla heterogénea de proteínas. En este ejemplo, escogemos fragmentar las proteínas en péptidos mediante la digestión de tripsina y capturar las subclases de los fragmentos de péptidos utilizando moléculas Fv de simple cadena (scFv) con propiedades de unión dirigidas al C-terminal de los péptidos.

Generación de las moléculas ligantes

Diseño de los péptidos selectores: Se usaron péptidos sintéticos como agentes receptores al aislar moléculas Fv de simple cadena (scFv) convenientes a partir de una biblioteca de expresión de fagos. Los péptidos fueron diseñados para capturar partículas de fagos que expresan scFv con afinidad por el C-terminal de un tetra o hexapéptido en el cual el último aminoácido era arginina o lisina. Se añadió un espaciador en la región N-terminal de este péptido así como una biotina N-terminal. Las secuencias de aminoácidos fueron diseñadas para incluir aminoácidos que fuera muy probable que generaran buenos epítopes, como los aminoácidos hidrofóbicos (fenilalanina, tirosina, triptófano, leucina e isoleucina) o aminoácidos cargados (aspartato, glutamato, asparagina, glutamina e histidina). La metionina fue excluida debido a su tendencia a oxidar, y la cisteína fue excluida para evitar problemas con la dimerización debido a la formación de puentes disulfuros. Las secuencias de los péptidos también se deciden según su frecuencia en proteínas que se desarrollan naturalmente. Los péptidos utilizados como selectores y competidores en este ejemplo se describen en la Tabla 2.

TABLA 2

Péptidos utilizados durante la selección

Nombre	Secuencia
FN1	Biotina-SGSG-EDFR (-COOH)
FN2	Biotina-SGSG-EPER (-COOH)
FN3	Biotina-SGSG-EPFR (-COOH)
FN4	Biotina-SGSG-HPDK (-COOH)
FN5	Biotina-SGSG-LPSR (-COOH)
FN6	Biotina-SGSG-LQSK (-COOH)
FN7	Biotina-SGSG-PEEK (-COOH)
FN8	Biotina-SGSG-TGEK (-COOH)
FN9	Biotina-SGSG-WDSR (-COOH)
FN10	Biotina-SGSG-YLDK (-COOH) <u>u</u>
FN11	SGSG-ASAK (-COOH)
FN12	SGSG-ASAR (-COOH)
FN13	Biotina-SGSG-LYEIAR (-COOH)
FN14	Biotina-SGSG-DFAEDK (-COOH)
FN15	Biotina-SGSG-LTEFAK (-COOH)
FN16	Biotina-SGSG-TEEQLK (-COOH)
FN17	Biotina-SGSG-SSAYSR (-COOH)

Selección de los unidores específicos a partir de una biblioteca de expresión de fagos

La selección de unidores específicos fue realizada a partir de la biblioteca n-CoDeR utilizando perlas magnéticas cubiertas con estreptavidina (Hawkins, R.E., Russel, S.J. e Invierno, G. (1992) *J. Mol. Biol.*, 226, 889-896). La construcción y el manejo de la biblioteca n-CoDeR que muestra fago de scFv se describe en Söderlind *et al* (2000) *Nature Biotech.* 18, 852-856. Fueron realizadas tres rondas consecutivas de selección; Selección 1. La biblioteca de fagos n-CoDeR (Lib 2000) fue primero preseleccionada contra un péptido biotinilado irrelevante (biotina-GIVKYLYEDEG, 10^{-7} M). El péptido fue capturado en las perlas magnéticas cubiertas con estreptavidina y las perlas fueron extraídas por centrifugación. Esta preselección extrae los unidores contra la estreptavidina, biotina y el ligador SGSG.

ES 2 319 422 T3

Las cepas de los fagos preseleccionados (una biblioteca equivalente por cada grupo de péptido) fueron seleccionadas contra cuatro grupos de péptidos (5×10^{-8} M de cada péptido). La composición de los grupos se muestra en la Tabla 3. Los péptidos competidores FN11 (10^{-6} M) y FN12 (10^{-6} M) fueron añadidos a los grupos R y a los grupos K, respectivamente.

TABLA 3

Grupos de péptidos dianas utilizados en la selección 1

Tetra-Grupo R	Tetra- Grupo K	Tetra- Grupo R	Tetra- Grupo K
FN1	FN4	FN13	FN14
FN2	FN6	FN17	FN15
FN3	FN7		FN16
FN5	FN8		
FN9	FN10		

Los péptidos fueron capturados con las perlas magnéticas cubiertas con estreptavidina y fueron extraídos los fagos no específicos mediante lavado (las perlas fueron concentradas utilizando un imán). Los fagos que se unen a las perlas fueron extraídos utilizando tripsina y fueron amplificados los grupos de fagos extraídos en *E. coli* HB101F⁺. Las cepas de fagos amplificadas a partir de la selección 1 fueron preseleccionadas contra un péptido irrelevante como está descrito previamente. Las cepas de fagos preseleccionados fueron entonces utilizadas para los unidores seleccionados a los péptidos individuales biotinilados (2×10^{-8} M de cada péptido). Fueron realizadas 15 selecciones separadas. Esta vez ambos péptidos competidores, FN11 y FN12 (2×10^{-7} M de cada uno), fueron añadidos a todas las selecciones.

Los péptidos fueron capturados en las perlas magnéticas cubiertas con estreptavidina y los fagos no específicos fueron extraídos mediante lavado. Los fagos unidos a las perlas fueron extraídos utilizando ácido. Los grupos de fagos extraídos no fueron amplificados sino que se usaron directamente en la selección 3. La selección 3 fue realizada como una selección en fase sólida en placas ELISA de 96 pocillos. Los grupos de fagos extraídos a partir de la selección 2 fueron primero preseleccionados contra la estreptavidina (0,5 μ g/pocillo, 8 pocillos por selección) y después contra la avidina (0,5 μ g/pocillo, 8 pocillos por selección).

La cepas de fagos preseleccionados fueron utilizadas para seleccionar fagos contra los péptidos dianas cargados sobre avidina (10 pmol de péptido/pocillo, 8 pocillos por selección). Ambos péptidos competidores (2×10^{-7} M de cada uno) fueron añadidos a todas las selecciones. Los fagos no específicos fueron extraídos mediante lavado y los fagos unidos a los pocillos fueron extraídos utilizando tripsina.

La calidad de los grupos de fagos fue evaluada a partir de la selección 3 en una ELISA de fagos. Los grupos de fagos extraídos fueron amplificados en *E. coli* HB101F⁺ y fueron probadas diluciones seriadas de los grupos amplificados contra un péptido diana y un péptido no diana. Para identificar los clones que generarán las mejores moléculas ligantes para la aplicación dada, se empleó un procedimiento de ensayo a doble escala. El ensayo principal es diseñado para evaluar las propiedades de unión de un gran número de scFv expresados (típicamente 10.000) contra un ligando previsto y un no ligando previsto, y diferenciará entre el scFv con interacción específica *versus* no específica con el péptido escogido y proporcionará una medición aproximada de la calidad relativa entre los unidores específicos. En base al ELISA de fagos, se identificaron las selecciones que mostraron resultados de enriquecimiento de los unidores específicos. Los grupos de fagos extraídos de la selección 3 fueron utilizados para infectar *E. coli* HB101F⁺ y fue aislado el ADN del fago. El ADN-específico del fago fue eliminado mediante digestión con enzimas de restricción y el material re-ligado fue transformado en *E. coli* TOP10 químicamente competente. Los transformantes, i.e., clones que expresan el scFv, fueron seleccionados en una placa LA que contenía ampicilina.

Los clones individuales bacterianos fueron escogidos y el scFv fue expresado en LB en placas de 384 pocillos para un ensayo ulterior con un ELISA de luminiscencia (ELISA lum). Fueron escogidas 1920 colonias para cada diana excepto para el FN9 (768 colonias) y el FN15 (1008 colonias).

El ensayo del ELISA lum fue realizado en un formato de 384 pocillos. Cada scFv fue ensayado contra un péptido diana y un péptido no diana. Los péptidos biotinilados (1pmol/pocillo) fueron cargados con estreptavidina (0,1 μ g/pocillo) y detectados utilizando un anticuerpo anti-His conjugado a HRP.

Todas las selecciones de hexapéptidos (FN13-FN17) y tres de las selecciones de tetrapéptidos (FN1, FN3, FN9) mostraron la presencia de unidores específicos al scFv en el ensayo primario por robot.

Los clones identificados como unidores específicos durante el ensayo fueron analizados por secuenciación del ADN para identificar los clones únicos.

Los genes que codifican el scFv fueron secuenciados según el método de terminación de cadena o dideoxi utilizando como plantilla un ADN amplificado por RCP, cebadores hechos a medida y el equipo Big Dye Terminador RR (Applied Biosystems, EE.UU.). Los fragmentos terminados fueron separados y analizados utilizando un Analizador Genético (Applied Biosystems).

Para determinar qué scFv capturará un número y tipo conveniente de péptidos a partir de una muestra digerida con tripsina, los scFv fueron unidos a un medio de cromatografía (Poros AL, Applied biosystems) y empaquetados en puntas cargadas de gel para generar columnas de poca afinidad.

Las muestras fueron reducidas con mercaptoetanolamina, alquilizadas con yodoacetamida, y digeridas con tripsina (PBS pH 7,4; 20µg de tripsina/mg de proteína, 6 h de incubación a 37°C). Las columnas de afinidad fueron utilizadas para capturar los péptidos a partir del homogenizado de hígado de ratón digerido con tripsina, los péptidos capturados fueron extraídos y analizados mediante una espectrometría de masas por desorción/ionización mediante láser asistida por matriz.

Generación de una matriz de columnas de afinidad

Los 14 scFv fueron seleccionados en base a su capacidad para capturar subgrupos diferentes de péptidos a partir de proteínas tripsinadas de hígado de ratón. La reacción de acoplamiento del scFv al medio de cromatografía POROS-AL (Applied Biosystems, Foster City, EE.UU.), fue realizada en concordancia con las instrucciones del fabricante. La suspensión fue empaquetada en puntas cargadas de gel (Invitrogen) para generar columnas de afinidad con una longitud de lecho de aproximadamente 2 cm.

Análisis de muestra complejas

El homogenizado de hígado de ratón fue alquilado y fragmentado como se indica arriba y diluido 2 veces en PBS pH 7,4. Las columnas de afinidad fueron lavadas con 2 x 10 µl de 5% de ácido acético y equilibradas con 2 x 10 µl de PBS pH 7,4. Fueron cargados 10 µl de la muestra en la columna seguido de un lavado con 2 x 10 µl de PBS pH 7,4. La columna fue extraída con una diana Massprep MALDI (Micromass, UK) con 7 µl de 5% de ácido acético. El extracto fue puesto a secar y el pocillo diana fue lavado dos veces con 0,1% de ácido trifluoroacético. Finalmente fue añadido 1 µl de 0,5 mg/ml de ácido α-ciano-4-hidroxi-cinámico en 75% de acetonitrilo/1% de ácido trifluoroacético. Las muestras fueron analizadas utilizando en espectrómetro de masas Micromass M @Idi Reflectron.

Resultados

Las figuras 2-15 muestran los espectros de masas generados. Cada espectro contiene aproximadamente de 20-100 picos diferentes con, una señal que tiene una señal/ruido por encima de 3, casi todos los picos correspondiendo a un único péptido. Unos pocos picos pueden ser detectados en todos los espectros, correspondiendo éstos a péptidos que se unen inespecíficamente al material Poros. El número total de péptidos que pueden ser detectados utilizando esta matriz está bien por encima de 500.

Referencias citadas en la descripción

Este listado de referencias citadas por el solicitante tiene como único fin la conveniencia del lector. No forma parte del documento de la Patente Europea. Aunque se ha puesto gran cuidado en la compilación de las referencias, no pueden excluirse errores u omisiones y la OEP rechaza cualquier responsabilidad en este sentido.

Documentos de patentes citados en la descripción

- WO 02086081 A
- WO 0225287 A
- EP 39578 A
- US 5856090 A
- WO 9837186 A
- WO 02060377 A

Literatura no relacionada con patentes citada en la descripción

- **WEINBERGER** *et al.* *Journal of Chromatography B*, 2002, vol. 782, 307-316
- 5 • **GYGI** *et al.* *Nature Biotechnology*, 1999, vol. 17 (1), 994-9
- **SCRIVENER**, E. *et al.* *Proteomics*, 2003, vol. 3 (2), 122-8
- **NELSON** *et al.* *Anal Chem*, 1995, vol. 67, 1153-8
- 10 • **JENKINS**, R.E.; **PENNINGTON**, S.R. *Proteomics*, 2001, vol. 2, 13-29
- **LAL** *et al.* *Drug Discov Today*, 2002, vol. 15 (18), S143-9
- 15 • **BETTER** *et al.* *Science*, 1988, vol. 240, 1041
- **SKERRA** *et al.* *Science*, 1988, vol. 240, 1038
- **BIRD** *et al.* *Science*, 1988, vol. 242, 423
- 20 • **HUSTON** *et al.* *Proc. Natl. Acad. Sci. USA*, 1988, vol. 85, 5879
- **WARD** *et al.* *Nature*, 1989, vol. 341, 544
- 25 • **WINTER**; **MILSTEIN**. *Nature*, 1991, vol. 349, 293-299
- **CLACKSON** *et al.* *Nature*, 1991, vol. 352, 624-628
- **MARKS** *et al.* *J Mol Biol*, 1991, vol. 222 (3), 581-97
- 30 • **SMITH**. *Science*, 1985, vol. 228 (4705), 1315-7
- **SANTI** *et al.* *J Mol Biol*, 2000, vol. 296 (2), 497-508
- 35 • **GUNNERIUSSON** *et al.* *Appl Environ Microbiol*, 1999, vol. 65 (9), 4134-40
- **KENAN** *et al.* *Methods Mol Biol*, 1999, vol. 118, 217-31
- **KIEKE** *et al.* *Proc Natl Acad Sci USA*, 1999, vol. 96 (10), 5651-6
- 40 • **HANES**; **PLUCKTHUN**. *Proc Natl Acad Sci USA*, 1997, vol. 94 (10), 4937-42
- **HE**; **TAUSSIG**. *Nucleic Acids Res*, 1997, vol. 25 (24), 5132-4
- 45 • **NEMOTO** *et al.* *FEBS Lett*, 1997, vol. 414 (2), 405-8
- **DAUGHERTY** *et al.* *Protein Eng*, 1998, vol. 11 (9), 825-32
- **DAUGHERTY** *et al.* *Protein Eng*, 1999, vol. 12 (7), 613-21
- 50 • **SHUSTA** *et al.* *J Mol Biol*, 1999, vol. 292 (5), 949-56
- **GOSHE** BG; **SMITH** DS. *Curr Opinion Biotech*, 2003, vol. 14, 101-109
- 55 • **HAWKINS**, R.E.; **WINTER**, G. *J. Mol. Biol.*, 1992, vol. 226, 889-896
- **SÖDERLIND** *et al.* *Nature Biotech*, 2000, vol. 18, 852-856
- 60 • **ANDERSON** NL; **ANDERSON** NG. *Mol Cell Proteomics*, 2002, vol. 1 (11), 845-67
- **BORREBAECK** CAK; **EKSTRÖM** S; **MALMBORG** HAGER AC; **NILSSON** J; **LAURELL** T; **MARKO-VARGA** G. *Biotechniques*, 2001, vol. 30, 1126-1132
- 65 • **HAWKINS**, R.E.; **RUSSEL**, S.J.; **WINTER**, G. *J. Mol. Biol.*, 1992, vol. 226, 889-896

REIVINDICACIONES

1. Método para analizar una muestra heterogénea de péptidos, o fragmentos de péptidos o proteínas, comprendiendo el método:
 - (a) separar la muestra heterogénea de péptidos, o fragmentos de péptidos o proteínas, en clases heterogéneas mediante la unión de componentes de cada clase en una posición definida espaciada en una matriz, en donde los componentes de cada clase tienen una configuración común a esa clase; y
 - (b) caracterizar los péptidos, o fragmentos de péptidos o proteínas, en cada clase determinando la masa de los péptidos, o fragmentos de péptidos o proteínas, en las clases heterogéneas, y determinando la cantidad de péptidos, o fragmentos de péptidos o proteínas, de diferente masa en las clases heterogéneas.
2. Método según la Reivindicación 1 en donde la muestra heterogénea de péptidos, o fragmentos de péptidos o proteínas, es un extracto del contenido total de proteínas de una célula o tipo de tejido.
3. Método según las Reivindicaciones 1 ó 2, en donde antes de realizar el paso (a) de la Reivindicación 1, la muestra heterogénea de fragmentos se forma fragmentando una mezcla heterogénea de proteínas o péptidos.
4. Método según la Reivindicación 3 en donde la fragmentación se realiza por escisión química o enzimática.
5. Método según las Reivindicaciones 3 ó 4 en donde la fragmentación se realiza utilizando un mecanismo de escisión dirigida por secuencia.
6. Método según una cualquiera de las Reivindicaciones 3 a 5 en donde la fragmentación se realiza mediante la digestión de la muestra heterogénea de proteínas o péptidos con tripsina.
7. Método según cualquier reivindicación precedente en donde la configuración en cada péptido, o fragmento de péptido o proteína, está en la misma posición en cada péptido, o fragmento de péptido o proteína, respecto al C-terminal, el N-terminal, o una característica interna.
8. Método según cualquier reivindicación precedente en donde la muestra es una muestra heterogénea de fragmentos de péptidos o proteínas y la configuración en cada fragmentos está en la misma posición en cada fragmento, respecto a su sitio de escisión.
9. Método según cualquier reivindicación precedente en donde la configuración en cada péptido, o fragmento de péptido o proteína, tiene tres, cuatro, cinco, seis, o más aminoácidos de longitud.
10. Método según cualquier reivindicación precedente en donde la configuración contiene tres, cuatro, o cinco aminoácidos variables, siendo los otros aminoácidos constantes en la configuración entre todos los péptidos, o fragmentos de péptidos o proteínas.
11. Método según cualquier reivindicación precedente en donde la configuración está en el C-terminal.
12. Método según cualquier reivindicación precedente en donde la configuración está en el N-terminal.
13. Método según cualquiera de las reivindicaciones precedentes en donde la matriz comprende una serie de tipos diferentes de moléculas ligantes, inmovilizado cada tipo en una posición definida espaciada en la matriz, en donde cada tipo de molécula ligante es capaz de unirse específicamente a una configuración definida como la definida en cualquiera de las reivindicaciones precedentes y en donde los diferentes tipos de moléculas ligantes tienen especificidades de unión diferentes.
14. Método según la Reivindicación 13 en donde el número de tipos diferentes de moléculas ligantes suministradas en la matriz es conveniente para capturar al menos el 10% de los péptidos en la muestra no fragmentada o, donde la muestra es una muestra heterogénea de fragmentos de proteínas o péptidos, al menos un fragmento de al menos el 10% de las proteínas o péptidos en la muestra no fragmentada.
15. Método según la Reivindicación 13 en donde el número de tipos diferentes de moléculas ligantes suministradas en la matriz es conveniente para capturar al menos el 50% de los péptidos en la muestra no fragmentada o, donde la muestra es una muestra heterogénea de fragmentos de péptidos o proteínas, al menos un fragmento de al menos el 50% de las proteínas o péptidos en la muestra no fragmentada.
16. Método según la Reivindicación 13 en donde el número de tipos diferentes de moléculas ligantes suministradas en la matriz es conveniente para capturar significativamente el 100% de los péptidos en la muestra no fragmentada o, donde la muestra es una muestra heterogénea de fragmentos de péptidos o proteínas, al menos un fragmento de sustancialmente el 100% de las proteínas o péptidos en la muestra no fragmentada.

ES 2 319 422 T3

17. Método según una cualquiera de las reivindicaciones 13 a 16, en donde la matriz dispone de al menos unos 10, 50, 100, 200, 250, 300, o más tipos diferentes de moléculas ligantes.

18. Método según una cualquiera de las reivindicaciones 13 a 17, en donde al menos un tipo de las moléculas ligantes es un anticuerpo o un fragmento o variante de éste, como el Fv, el scFv, o el Fab.

19. Método según una cualquiera de las reivindicaciones 13 a 18 en donde al menos uno de los tipos de moléculas ligantes es un aptámero.

20. Método según una cualquiera de las reivindicaciones 13 a 19 en donde al menos uno de los tipos de moléculas ligantes es un polinucleótido.

21. Método según una cualquiera de las reivindicaciones previas, en donde el paso (b) de la Reivindicación 1 comprende caracterizar la unión de péptidos, o fragmentos de péptidos o proteínas, en las posiciones discretas y definidas en la matriz.

22. Método según una cualquiera de las reivindicaciones precedentes en donde el paso (b) de la Reivindicación 1 comprende caracterizar los péptidos, o fragmentos de péptidos o proteínas, en las clases heterogéneas mediante espectrometría de masas de desorción o espectrometría de masas de colisión inducida por disociación.

23. Método según cualquier reivindicación precedente en donde la información derivada del paso (b) de la Reivindicación 1 es utilizada para determinar la identidad de la proteína o péptido parental en la muestra heterogénea no fraccionada de la cual se deriva un fragmento de péptido o proteína detectado.

24. Método según cualquier reivindicación precedente en donde la información derivada del paso (b) de la Reivindicación 1 es utilizada para determinar la cantidad de la proteína o péptido parental en la muestra heterogénea no fraccionada de la cual se deriva un fragmento de péptido o proteína detectado.

25. Método para identificar diferencias en la composición entre dos o más muestras heterogéneas de péptidos, o fragmentos de péptidos o proteínas, fraccionados o no fraccionados, que comprende analizar cada muestra mediante el método según una cualquiera de las reivindicaciones precedentes y comparar los resultados para identificar, con ello, cualquier diferencia.

26. Método para identificar una proteína o péptido relacionado con una enfermedad que comprende la identificación de diferencias entre dos o más muestras mediante el método de la Reivindicación 25, en donde al menos una de las muestras analizadas procede de un individuo con la enfermedad y otra de las muestras analizadas procede de un individuo sin la enfermedad.

27. Método para analizar una muestra heterogénea de péptidos, fragmentos de péptidos o proteínas, comprendiendo el método,

(a) proporcionar, como primer componente, un péptido selector que comprende una configuración definida en una cualquiera de las reivindicaciones 7 a 12;

(b) proporcionar, como segundo componente, una fuente de moléculas ligantes candidatas;

(c) combinar el primer y el segundo componente;

(d) identificar las moléculas ligantes candidatas que son capaces de unirse específicamente a la configuración del péptido escogido en el primer componente;

(e) inmovilizar las moléculas ligantes identificadas en el paso (d) en una matriz de forma que los diferentes tipos de moléculas ligantes sean inmovilizadas en posiciones definidas y discretas;

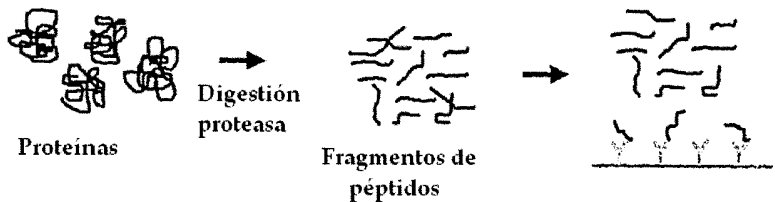
(f) proporcionar una muestra heterogénea de péptidos, o fragmentos de péptidos o proteínas, muestra que comprende péptidos, o fragmentos de péptidos o proteínas, que tienen cada uno una configuración a la que se une una molécula ligante inmovilizada en el paso (e);

(g) separar la muestra heterogénea de péptidos, o fragmentos de péptidos o proteínas, en clases heterogéneas mediante la unión de los componentes de cada clase a las moléculas ligantes inmovilizadas en el paso (e); y

(h) caracterizar los péptidos, fragmentos de péptidos o proteínas, en cada clase mediante la determinación de la masa de los péptidos, o fragmentos de péptidos o proteínas, en clases heterogéneas, y determinar la cantidad de péptidos, o fragmentos de péptidos o proteínas, de diferente masa en clases heterogéneas.

FIG. 1

Captura de afinidad



Espectrometría de masas de desorción

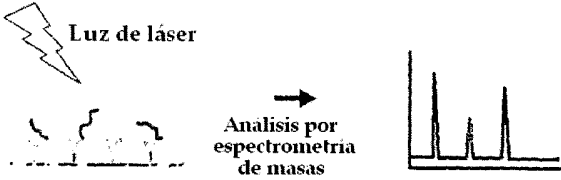
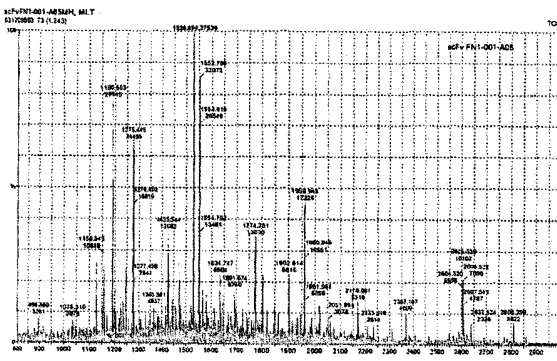


Fig. 2

Fig. 3



ES 2 319 422 T3

Fig. 10

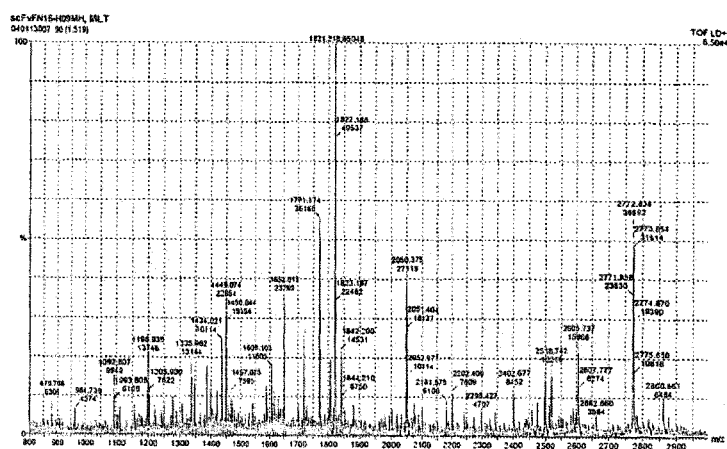


Fig. 11

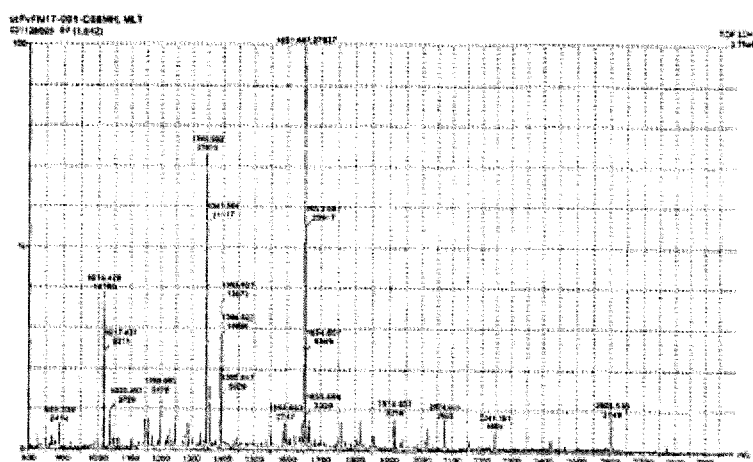
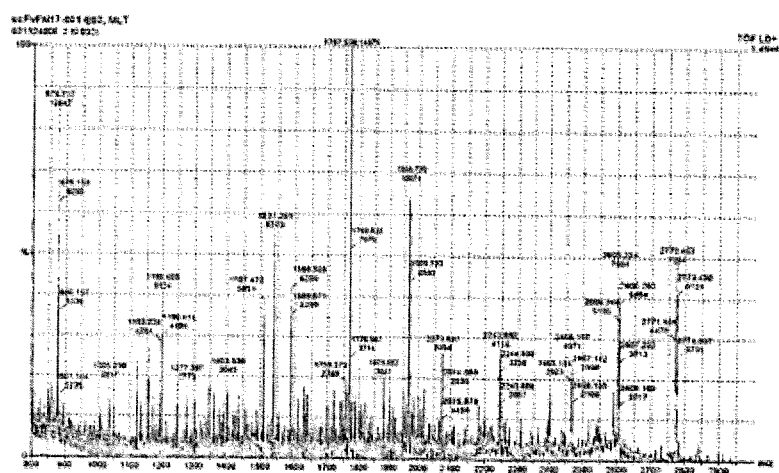


Fig. 12



(4/4)

[illegible]

6CP-BITM7-001-A12M04_MLT
031120501 140 (2.435)

TOP LD= 7.18e3

Mass spectrum plot showing relative intensity (0 to 100) versus mass-to-charge ratio (m/z) from 0 to 300. The base peak is at m/z 134.754. Other significant peaks are labeled with their m/z values.

m/z	Relative Intensity (approx)
67.617	15
87.067	10
89.580	5
92.586	5
94.443	5
104.1	5
103.0	5
115.439	15
116.439	10
118.617	15
124.7	10
128.7	10
147.854	10
150.7	10
156.437	10
160.7	10
172.743	15
176.8	10
179.815	10
180.8	10
191.818	10
199.9	10
201.648	10
204.812	10
212.1	10
219.987	10
241.100	10
250.0	10
256.188	10
257.1	10
260.233	10
262.133	10
272.250	10
274.205	10
275.0	10
286.002	10
287.1	10
290.812	10
292.0	10
294.714	10
295.0	10
296.0	10
297.0	10
298.0	10
299.0	10
300.0	10

cF₂BTHM12-261 E16MN, MLT
03117603 52 (0.85) 145.47636151

TOF LDI 3.14e4

Mass spectrum showing relative intensity (0 to 100) versus m/z (0 to 2500). The base peak is at m/z 145.47636151. Other labeled peaks include:

m/z	Relative Intensity (approx)
877.037	10
878.041	10
1122.188	10
1183.071	10
1318.708	10
145.47636151	100
154.482	10
160.000	10
177.722	10
181.2	10
182.713	10
183.0	10
185.0	10
186.0	10
187.0	10
188.0	10
189.0	10
190.0	10
191.0	10
192.0	10
193.0	10
194.0	10
195.0	10
196.0	10
197.0	10
198.0	10
199.0	10
200.0	10
201.0	10
202.0	10
203.0	10
204.0	10
205.0	10
206.0	10
207.0	10
208.0	10
209.0	10
210.0	10
211.0	10
212.0	10
213.0	10
214.0	10
215.0	10
216.0	10
217.0	10
218.0	10
219.0	10
220.0	10
221.0	10
222.0	10
223.0	10
224.0	10
225.0	10
226.0	10
227.0	10
228.0	10
229.0	10
230.0	10
231.0	10
232.0	10
233.0	10
234.0	10
235.0	10
236.0	10
237.0	10
238.0	10
239.0	10
240.0	10
241.0	10
242.0	10
243.0	10
244.0	10
245.0	10
246.0	10
247.0	10
248.0	10
249.0	10
250.0	10
251.0	10
252.0	10
253.0	10
254.0	10
255.0	10
256.0	10
257.0	10
258.0	10
259.0	10
260.0	10
261.0	10
262.0	10
263.0	10
264.0	10
265.0	10
266.0	10
267.0	10
268.0	10
269.0	10
270.0	10
271.0	10
272.0	10
273.0	10
274.0	10
275.0	10
276.0	10
277.0	10
278.0	10
279.0	10
280.0	10
281.0	10
282.0	10
283.0	10
284.0	10
285.0	10
286.0	10
287.0	10
288.0	10
289.0	10
290.0	10
291.0	10
292.0	10
293.0	10
294.0	10
295.0	10
296.0	10
297.0	10
298.0	10
299.0	10
300.0	10