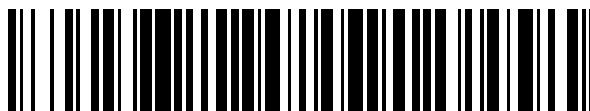


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 383 191**

51 Int. Cl.:
C12N 15/10 (2006.01)
C40B 30/04 (2006.01)
C40B 40/02 (2006.01)
C40B 40/08 (2006.01)
C40B 50/06 (2006.01)
G01N 33/569 (2006.01)
G01N 33/68 (2006.01)
A61K 47/48 (2006.01)
C07K 1/107 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **09708496 .6**
96 Fecha de presentación: **04.02.2009**
97 Número de publicación de la solicitud: **2257624**
97 Fecha de publicación de la solicitud: **08.12.2010**

54 Título: **Métodos y composiciones**

30 Prioridad:
05.02.2008 GB 0802079
08.10.2008 GB 0818399

45 Fecha de publicación de la mención BOPI:
19.06.2012

45 Fecha de la publicación del folleto de la patente:
19.06.2012

73 Titular/es:
Medical Research Council
2nd Floor, David Phillips Building
Polaris House,
North Star Avenue
Swindon, SN2 1FL , GB

72 Inventor/es:
WINTER, Gregory y
HEINIS, Christian

74 Agente/Representante:
de Elzaburu Márquez, Alberto

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).

ES 2 383 191 T3

DESCRIPCIÓN

Métodos y composiciones

CAMPO DE LA INVENCION

- La invención se refiere a la modificación y al constreñimiento de polipéptidos, en particular a polipéptidos
 5 codificados genéticamente en complejos con un ácido nucleico que los codifica tal como en el contexto de la exposición en fagos.

ANTECEDENTES DE LA INVENCION

- La generación de las moléculas con afinidad y especificidad elevadas por las dianas biológicas es un problema
 10 central en la química, en la biología y en las ciencias farmacéuticas. En particular, los ligandos de fijación son importantes para la creación de fármacos que son capaces de interferir con los procesos biológicos. La creación de ligandos que se fijan a un ligando diana elegido suele implicar un proceso de generación de una colección de posibles moléculas de fijación y el ensayo de las propiedades de fijación de dichas moléculas.

- Mientras que las técnicas biológicas de selección *in vitro* se utilizaron eficazmente para aislar grandes estructuras biopoliméricas, tales como anticuerpos, hasta la fecha eran menos practicables para el aislamiento de fármacos
 15 que fueran moléculas pequeñas. Las técnicas biológicas de selección *in vitro* se limitan por lo general a polímeros biológicos tales como polipéptidos, ARN o ADN. Los biopolímeros pequeños como, por ejemplo, los péptidos, también se pueden fijar a dianas biológicas, pero pueden carecer de flexibilidad conformacional y pueden ser propensos a la degradación proteolítica en los líquidos corporales. Además está la debilidad de la afinidad de unión de los péptidos lineales pequeños. Se conocen diferentes estrategias de circularización para constreñir las
 20 genotecas de péptidos pequeños codificados genéticamente. Se sabe que los repertorios de péptidos expuestos en fagos, por ejemplo, se circularizan mediante la oxidación de dos restos de cisteína flanqueantes. Se sabe que las genotecas de péptidos cíclicos codificados por ARNm se generan al formar un enlace entre la amina aminoterminal y un resto de lisina del péptido con un reactivo de entrecruzamiento químico. Esta estrategia se utilizó para aislar macrociclos insensibles a la óxido-reducción que se fijan a la proteína de señalización G $\pm$ i1
 25 (Millward, S. W. et al., *ACS Chem. Biol.*, 2007). También se conocen diferentes estrategias para uso en la incorporación de bloques de construcción no naturales en genotecas de polipéptidos codificados genéticamente para expandir la diversidad de las genotecas o para insertar propiedades que no pueden proporcionar los aminoácidos naturales. Sin embargo, las estrategias permitieron sólo la adición de un número limitado de añadidos orgánicos pequeños a los polipéptidos lineales codificados genéticamente. Frankel, A. et al., por ejemplo,
 30 habían incorporado aminoácidos no naturales a polipéptidos naturales que se codificaron mediante exposición de ARNm (Frankel, A et al., *Chem. Biol.*, 2003). Jespers L. et al habían conectado químicamente una molécula indicadora fluorescente a un lazo hipervariable de un repertorio de anticuerpos expuestos en fagos, y seleccionaron del repertorio los que se fijaban a antígenos (Jespers, L., et al., *Prot. Eng.*, 2004). Dwyer, M. A. et al habían juntado péptidos sintéticos a un repertorio de péptidos expuestos en fagos mediante la ligación química nativa para la
 35 generación de una colección de inhibidores de proteasas que contenían un aminoácido no natural (Dwyer, M. A. et al., *Chemistry & Biology*, 2000). También se han conectado moléculas orgánicas pequeñas a repertorios combinatorios de péptidos codificados por ARNm. El equipo de investigación de Roberts, R. W. había pegado un resto de penicilina a una posición fija de una genoteca de péptidos de exposición de ARNm para seleccionar los inhibidores de la proteína 2a de fijación a la penicilina en *Staphylococcus aureus* (Li, S y Roberts, W. R., *Chem &*
 40 *Biol.*, 2003).

- Se han propuesto varias metodologías para aplicar la selección *in vitro* a colecciones combinatorias de compuestos que tienen arquitecturas moleculares más diversas (p. ej., moléculas ramificadas) y que están
 45 formadas por bloques de construcción no naturales. A diferencia de los métodos biológicos de selección *in vitro*, estas metodologías utilizan estrategias químicas para pegar etiquetas de ADN a las moléculas orgánicas pequeñas. Brenner S y Lerner R. A. habían propuesto un proceso de síntesis combinatoria paralela para codificar cada miembro de una gran quimioteca con secuencias nucleotídicas únicas en perlas (Brenner, S. y Lerner, R. A., *PNAS*, 1992). Después de que la entidad química se fije a la diana, el código genético se descodifica por secuenciación de la etiqueta nucleotídica. Liu D. R. y colaboradores habían conjugado una pequeña colección de moléculas orgánicas a oligonucleótidos de ADN y realizaron selecciones por afinidad con antígenos diferentes
 50 (Doyon, J. B. et al., *JACS*, 2003). Neri D. y colaboradores habían generado grandes repertorios de parejas moleculares mediante el autoensamblaje de subquimiotecas más pequeñas codificadas por ADN mediante la hibridación de dos hebras de ADN (Melkko, S. et al., *Nature Biotechnol.*, 2004). La metodología se utilizó con éxito para la maduración de la afinidad de los ligandos que son moléculas pequeñas. Halpin D. R. y Harris P. B. desarrollaron una estrategia para la evolución *in vitro* de quimiotecas combinatorias que implica la amplificación de
 55 determinados compuestos para realizar varios ciclos de selección (Halpin, D. R. y Harbury, P. B., *PLOS Biology*,

2004). Woiwode T. F. et al pegaron colecciones de compuestos sintéticos a las proteínas de la cubierta de las partículas de bacteriófago de tal forma que la identidad de la estructura química está especificada en el genoma del fago (Woiwode, T. F., *Chem. & Biol.*, 2003). Todas estas estrategias que emplean entidades químicas especificadas por ADN han resultado ser eficientes en experimentos modelo y algunos incluso han producido

5 nuevas moléculas pequeñas aglutinantes. Sin embargo, quedó muy claro que la codificación de genotecas grandes de compuestos y la amplificación de los compuestos seleccionados es mucho más exigente que los procedimientos equivalentes en los sistemas biológicos de selección.

Jespers et al (2004, «Protein engineering design and selection», volumen 17, n.º 10, páginas 709-713) describen la selección de biosensores ópticos a partir de colecciones de anticuerpos quimiosintéticos. Este documento trata de

10 la unión de una molécula indicadora fluorescente a través del lazo hipervariable de un repertorio de anticuerpos expuestos en fagos. En particular, este documento describe la conexión de una molécula indicadora fluorescente en un lazo hipervariable (región determinante de la complementariedad o CDR) de un repertorio de anticuerpos sintéticos. La molécula indicadora fluorescente está conectada mediante un enlace covalente único a un resto de

15 cisteína introducido artificialmente en el lazo hipervariable. Se realiza la unión del uno al otro. Los restos cisteína en las partículas de fagos se redujeron con DTT y se retiró el exceso del agente reductor mediante precipitación convencional con polietilenglicol (PEG) como se conoce bien en la técnica.

Dwyer et al describen la exposición en fagos biosintéticos, que describe una nueva herramienta de ingeniería de proteínas que combina la diversidad química y genética. Dwyer et al (*Chem. Biol.* 2000, volumen 7, n.º. 4, páginas 263-274) describen la ligación química de un péptido sintético que tiene un aminoácido no natural a una colección

20 de péptidos sintéticos que comprende los principales restos estructurales de una proteína de interés. La motivación para realizar esto fue para generar un amplio abanico de secuencias de proteasas, y cada una contiene un segmento constante que incorpora un aminoácido no natural. El péptido sintético que comprende el aminoácido no natural se pegó simplemente mediante ligación química nativa, lo que da lugar al acoplamiento de los dos fragmentos peptídicos juntos. No se describe ningún compuesto conector. No se describe la adhesión de ninguna

25 molécula pequeña. No se consiguió ninguna constricción ni restricción conformacional del péptido resultante. No se describe ninguna formación de enlaces covalentes entre determinados restos y la cadena polipeptídica.

Diferentes equipos de investigación han sujetado previamente polipéptidos con restos de cisteína a una estructura molecular sintética (Kemp, D. S y McNamara, P. E., *J. Org. Chem.* 1985; Timmerman, P. et al., *ChemBioChem*, 2005). Meloen y colaboradores habían utilizado tris(bromometil)benceno y moléculas relacionadas para la ciclación

30 cuantitativa y rápida de varios lazos peptídicos sobre armazones sintéticos para la imitación estructural de las superficies proteicas (Timmerman, P. et al., *ChemBioChem*, 2005). Los métodos para generar compuestos farmacológicos candidatos en donde dichos compuestos se generan mediante la conexión de polipéptidos que contienen cisteína a un armazón molecular como, por ejemplo, tris(bromometil)benceno, se describen en las patentes internacionales WO 2004/077062 y WO 2006/078161.

Los métodos dados a conocer en las patentes internacionales WO 2004/077062 y WO 2006/078161 se basan en

35 tomar muestras de compuestos individuales, por ejemplo, en un procedimiento de escrutinio. El escrutinio de cada compuesto o de conjuntos pequeños de compuestos es tedioso y puede ser caro si se analizan grandes cantidades de compuestos. El número de compuestos que se puede ensayar con los ensayos de escrutinio no suelen exceder unos pocos miles. Además, las condiciones de reacción descritas en la patente internacional WO 2004/077062

40 para sujetar un péptido que contiene cisteína a un armazón que contiene halometilo como, por ejemplo, tris(bromometil)benceno, no son adecuadas para modificar un péptido que contiene cisteína codificado genéticamente.

La patente internacional WO 2004/077062 describe un método para seleccionar un compuesto farmacológico candidato. En particular, este documento describe diferentes armazones moleculares que comprenden un primer y

45 segundo grupo reactivo, y que pone en contacto dicho armazón con otra molécula para formar al menos dos enlaces entre el armazón y la otra molécula en una reacción de acoplamiento. Este método adolece de muchas restricciones. Primero, se basa en el uso de péptidos sintéticos y en reacciones químicas *in vitro* en recipientes independientes. Por este motivo resulta muy trabajoso. No existe la posibilidad de automatizar ni de aplicar el método al escrutinio de muchas variantes peptídicas sin producir manualmente cada variante mediante la

50 realización de numerosas reacciones independientes en paralelo. No se menciona la diversidad codificada genéticamente en este documento, y ciertamente no se menciona la aplicación a genotecas de fagos codificadas genéticamente. Realmente, las condiciones de reacción descritas en este documento dan a entender que sería difícil o imposible realizar las reacciones descritas sobre partículas de fagos.

La patente internacional WO 2006/078161 describe compuestos de fijación, compuestos inmunógenos y

55 peptidomiméticos. Este documento describe la síntesis artificial de diferentes colecciones de péptidos tomados de las proteínas existentes. Luego, estos péptidos se combinan con un péptido sintético constante al que se han

introducido algunos cambios de aminoácidos para producir colecciones combinatorias. La introducción de esta diversidad a través de enlace químico con cada péptido que presenta diferentes cambios de aminoácido proporciona un incremento de la probabilidad de encontrar la actividad de fijación deseada. La figura 7 de este documento muestra una representación esquemática de la síntesis de distintas construcciones peptídicas con lazo.

- 5 No hay ninguna descripción de genotecas peptídicas codificadas genéticamente en este documento. No hay ninguna descripción del uso de las técnicas de exposición en fagos en este documento. Este documento describe un proceso que se considera que es incompatible con la exposición en fagos. Por ejemplo, la química establecida en este documento probablemente dé lugar a la molécula de conexión que reacciona con la cubierta de los fagos. Existe el riesgo de que pudieran entrecruzarse partículas de fagos. Es probable que se inactiven las partículas de los
- 10 fagos (p. ej., perderían su infectividad) si se sometieran a la química descrita. Este documento se centra en la manipulación de varios péptidos sintéticos en reacciones independientes de conjugación química.

- Millward et al (2007 *Chemical Biology*, volumen 2, n.º 9, páginas 625-634) describen el diseño de péptidos cíclicos que fijan a las superficies proteicas con una afinidad similar a la de los anticuerpos. Este documento describe la ciclación de diferentes péptidos producidos por una genoteca codificada genéticamente. Los polipéptidos se ciclan
- 15 a través de la reacción de un entrecruzador químico con la amina del extremo amino y una amina de una lisina del polipéptido. En este documento, la genoteca codificada genéticamente es una genoteca de exposición de ARNm. Este documento no describe la adhesión de ningún compuesto conector a los polipéptidos resultantes. Este documento trata de la producción de péptidos ciclados insensibles a la oxidación-reducción. La química descrita en este documento es la ciclación a través de la reacción de un entrecruzador químico con la amina del extremo
- 20 amino y una amina de una lisina del polipéptido. La reacción de ciclación se realiza en tampón fosfato a 50 mM a pH 8 por la adición de DSG (1 mg/ml en DMF). A lo sumo, este documento describe la unión de dos partes de una cadena polipeptídica mediante un resto de entrecruzamiento para proporcionar un péptido cíclico.

- La patente de los EE.UU. n.º US 2003/0235852 describe genotecas de exposición de péptidos y ácidos nucleicos que contienen péptidos con restos de aminoácidos no naturales, y métodos para fabricar éstos mediante agentes
- 25 modificadores de péptidos. En otras palabras, este documento describe genotecas de polipéptidos codificados genéticamente que contienen o bien un aminoácido no natural o bien un aminoácido en donde un bloque de construcción no natural (p. ej., penicilina) se pega postraduccionalmente mediante una reacción química. Este documento se centra en los métodos conocidos para asociar un péptido traducido con el ácido nucleico que lo codifica. El otro problema tratado por este documento es cómo incorporar los aminoácidos no naturales a ese
- 30 péptido. Esto se consigue principalmente mediante el uso de ARNt supresores para incorporar aminoácidos no naturales en respuesta a codones ámbar/ocreo/ópalo como se conoce bien en la técnica. En otras realizaciones menores, los aminoácidos no naturales se crean postraduccionalmente mediante el tratamiento del péptido traducido con un «agente modificador de péptidos». Este reactivo típicamente tiene por objeto alterar un resto de aminoácido existente para convertirlo en un resto de aminoácido no natural, o de si no, convertirlo en
- 35 funcionalmente reactivo o receptivo a la adhesión de otro resto químico. Específicamente, este documento da a conocer la conjugación postraduccionale de un resto de cisteína en el polipéptido de interés al antibiótico ²-lactámico ácido 6-bromoacetilpenicilámico. Esto da lugar a la conjugación de este análogo de penicilina con el polipéptido de interés a través de un solo enlace con la cadena lateral del resto de cisteína. No se describe que la molécula que se va a ligar al polipéptido forme varios enlaces covalentes. No se describe ningún constreñimiento
- 40 conformacional del polipéptido. Ningún lazo peptídico ni ninguna otra estructura ternaria compleja se forma mediante los métodos descritos en este documento: es simplemente un modo de unir un único grupo molecular adicional a un polipéptido a través de un único enlace. La química de conjugación convencional se utiliza en este documento para introducir las modificaciones en los polipéptidos.

COMPENDIO DE LA INVENCION

- 45 La presente invención permite ventajosamente la combinación de la diversidad genéticamente codificada, en particular de genotecas de polipéptidos codificados genéticamente, con una modificación química y un constreñimiento conformacional.

- Además, las técnicas descritas en la presente memoria dan a conocer la conexión de un compuesto conector a una molécula polipeptídica mediante al menos tres enlaces covalentes. Esto proporciona la ventaja del
- 50 constreñimiento conformacional del polipéptido, en particular el constreñimiento conformacional de al menos dos segmentos del polipéptido, uno respecto a otro. Por el contrario, el uso de un conector que forma sólo dos enlaces covalentes constreñirá sólo un único segmento del polipéptido.

- Las ventajas de la invención se deducen de estas características técnicas, por ejemplo, debido a su construcción con tres enlaces, las moléculas conjugadas de la invención tienen dos o más lazos peptídicos que son capaces de
- 55 interactuar con una diana. Con varios lazos de fijación se puede conseguir mayor afinidad de fijación que con las moléculas que sólo tienen un único lazo peptídico.

Además, la superficie de interacción de una molécula de la invención con dos o más lazos de fijación para la interacción con una diana es mayor que la de una molécula con un único lazo peptídico con una diana. La mayor superficie de unión puede proporcionar una mejora de la afinidad de fijación, y/o también puede proporcionar una mejora de la especificidad.

- 5 Por lo tanto, en un aspecto, la invención da a conocer un complejo que comprende una partícula de fago, comprendiendo dicha partícula de fago

(i) un polipéptido;

(ii) un ácido nucleico que codifica el polipéptido de (i) y que se expone sobre la superficie del fago;

(iii) un compuesto conector unido a dicho polipéptido

- 10 en donde dicho compuesto conector está unido al polipéptido mediante al menos tres enlaces covalentes independientes.

Los enlaces covalentes son enlaces covalentes independientes adecuados en el sentido de que cada uno es un enlace distinto entre el compuesto conector y una parte del polipéptido. Por ejemplo, un único puente entre el polipéptido y el compuesto conector cuyo único puente está hecho de tres enlaces covalentes (p. ej., conector - x -

- 15 y - polipéptido, donde «-» representa un enlace covalente) no se consideraría que comprende al menos tres enlaces covalentes independientes porque los tres enlaces no son tres puentes separados ni tres conexiones desde el compuesto conector al polipéptido diana. El principio subyacente es que el compuesto conector/núcleo molecular y el polipéptido estén juntos mediante al menos tres enlaces covalentes de puente independientes.

- 20 Convenientemente, cada uno de los al menos tres enlaces covalentes se formado con un resto de aminoácido distinto del polipéptido. En otras palabras, cada resto de aminoácido es convenientemente un resto de aminoácido individual o diferente: se puede formar más de un enlace con una única especie o tipo de resto de aminoácido, p. ej., cada uno de los dos enlaces puede estar formado por restos cisteína, pero convenientemente los dos restos de cisteína serán restos cisteína independientes.

La parte compuesto conector-polipéptido del complejo descrito más arriba se denomina a veces un «conjugado».

- 25 En algunas realizaciones, el conjugado (a saber, un resto de polipéptido-compuesto conector que corresponde al comprendido por el complejo de la invención) se puede sintetizar por separado. En esta realización, el conjugado puede no estar formando un complejo con un ácido nucleico. Esto se explica más detalladamente a continuación.

- 30 Convenientemente, «codificar» tiene su significado natural en la técnica, a saber, codificar en el sentido del código universal de tripletes para convertir una secuencia de nucleótidos en una secuencia polipeptídica. En la técnica anterior, «codificar» podría haberse utilizado en el sentido de «etiquetar» o «desconvolucionar», p. ej., cuando se utiliza una única secuencia de nucleótidos para etiquetar un resto y que ese conocimiento de la secuencia de nucleótidos sea capaz de «descodificar», a saber, decir al usuario qué resto etiquetado estaba presente, incluso sin que tenga ninguna relación biológica con su estructura. Sin embargo, en la presente invención, «codificar» y «descodificar» se utilizan de la manera natural tradicional para referirse a codificar en el sentido de la traducción
- 35 desde la secuencia de nucleótidos a la secuencia de aminoácidos.

Convenientemente, el compuesto conector comprende una molécula orgánica. Convenientemente, el compuesto conector comprende una molécula orgánica pequeña.

Convenientemente, los enlaces covalentes se forman entre el compuesto conector y los restos de aminoácidos del polipéptido.

- 40 Convenientemente, dicho polipéptido comprende un resto de cisteína, y convenientemente al menos uno de dichos tres enlaces covalentes independientes que unen dicho compuesto conector al polipéptido comprende un enlace con dicho resto de cisteína.

Convenientemente, el compuesto conector tiene una simetría molecular que corresponde al número de enlaces covalentes mediante los cuales está unido al polipéptido.

- 45 Convenientemente, el compuesto conector posee una simetría molecular ternaria y el compuesto conector está unido al polipéptido mediante tres enlaces covalentes.

Convenientemente, el compuesto conector comprende un grupo químico estructuralmente rígido.

Convenientemente el compuesto conector comprende tris-(bromometil)benceno (TBMB).

El ácido nucleico tiene su significado usual en la técnica y puede comprender ADN, ARN o cualquier otro ácido nucleico adecuado. El ácido nucleico puede comprender oligonucleótido(s) o genoma(s) de fagos o cualquier otro ejemplo adecuado de ácidos nucleicos conocidos por el experto en la técnica.

Dicho ácido nucleico está comprendido por dicha partícula de fago.

- 5 En otro aspecto, la invención se refiere a una genoteca de polipéptidos codificados genéticamente que comprende al menos dos complejos diferentes como los descritos más arriba.

En otro aspecto, la invención se refiere a un método para construir un complejo, comprendiendo dicho método

- (i) proporcionar una partícula fágica que comprende un ácido nucleico, y un polipéptido expresado por el ácido nucleico y expuesto sobre la partícula del fago
- 10 (ii) proporcionar un compuesto conector
- (iii) unir dicho compuesto conector a dicho polipéptido mediante la formación de al menos tres enlaces covalentes entre dicho compuesto conector y el polipéptido.

- 15 Convenientemente, se reducen los grupos reactivos del polipéptido expuestos sobre la partícula del fago y, convenientemente, la partícula de fago que expone el polipéptido que comprende grupos reactivos reducidos se purifica mediante filtración antes de la etapa (iii). Convenientemente, cuando los grupos reactivos comprenden cisteína, están reducidos; en esta realización, la purificación es purificación para retirar el agente reductor, por ejemplo mediante filtración.

- 20 Convenientemente, después de la etapa de purificación por filtración, el polipéptido se mantiene en el estado reducido para formar un enlace con el compuesto conector mediante la incubación en un tampón desgasificado y en presencia de un quelante.

Convenientemente, la etapa (iii) comprende la incubación del polipéptido y del compuesto conector juntos a 30 °C a pH 8 en tampón acuoso que comprende acetonitrilo.

Convenientemente, el compuesto conector comprende tris-(bromometil)benceno (TBMB).

Convenientemente, el tris-(bromometil)benceno está presente a 10 µM.

- 25 Convenientemente, el tris-(bromometil)benceno está presente a 10 µM, el quelante es ácido etilendiaminotetraacético (EDTA), el acetonitrilo está presente al 20% y la etapa de incubación (iii) se realiza durante 1 hora.

- 30 Convenientemente, dicho método comprende otra etapa (iv) de escisión de uno o más enlaces de la cadena polipeptídica. Esto tiene la ventaja de modificar la cadena polipeptídica. Por ejemplo, esto puede tener el beneficio de producir varios polipéptidos unidos a un solo compuesto conector, p. ej., cuando la escisión tiene lugar en la cadena polipeptídica entre enlaces entre el polipéptido y el compuesto conector. Convenientemente, dicha etapa de escisión comprende poner en contacto dicho polipéptido con una proteasa.

En otro aspecto, la invención se refiere a un complejo obtenido mediante un método como el descrito más arriba.

- 35 En otro aspecto, la invención se refiere a un método para identificar un complejo de acuerdo con una cualquiera de las reivindicaciones precedentes que es capaz de unirse a un ligando, comprendiendo el método

- (i) proporcionar un complejo como el descrito más arriba
- (ii) poner en contacto dicho complejo con el ligando, y
- (iii) seleccionar los complejos que se fijan a dicho ligando.

- 40 Tal método de selección se puede realizar en cualquier formato adecuado. Convenientemente, el ligando está inmovilizado. A continuación, el complejo se pone en contacto con el ligando inmovilizado. Luego se lavan el o los complejos que no se unieron. De este modo, los complejos que se fijan al ligando inmovilizado se enriquecen o se seleccionan. En una realización, es posible que los complejos se puedan recuperar mediante la liberación del ligando, a saber, la liberación o la elución del resto de ligando-complejo. Sin embargo, convenientemente, los complejos se recuperan mediante elución (separación) del ligando inmovilizado. En esta realización, los complejos
- 45 eluidos ya no están unidos al ligando en la etapa de elución.

Los complejos, o el o los polipéptidos de dichos complejos, o los conjugados polipéptido-compuesto conector de dichos complejos, pueden ser útiles en otros contextos. Por ejemplo, pueden ser útiles como una base para el diseño de fármacos tales como fármacos pequeños, o pueden ser útiles como CDR o como restos de fijación (p. ej., para etiquetar o detectar su o sus homólogos de unión) u otras aplicaciones donde se puede explotar el conocimiento íntimo de la interacción.

En otro aspecto, la invención se refiere a un método como el descrito más arriba que además comprende determinar la secuencia del ácido nucleico de dicho complejo.

En otro aspecto, la invención se refiere a un método como el descrito más arriba que además comprende la etapa de fabricar una cantidad del complejo aislado con capacidad para fijarse a dicho ligando.

- 10 En otro aspecto, la invención se refiere a un método como el descrito más arriba que además comprende la etapa de fabricar una cantidad del resto polipéptido-compuesto conector comprendido por el complejo aislado que es capaz de fijarse a dicho ligando. En esta realización, el resto polipéptido-compuesto conector puede sintetizarse ventajosamente en ausencia de ácido nucleico.

- 15 En otro aspecto, la invención se refiere a un método como el descrito más arriba que además comprende la etapa de fabricar una cantidad de un polipéptido aislado o identificado mediante un método de la invención, comprendiendo dicha fabricación la unión del compuesto conector al polipéptido, en donde dicho polipéptido se expresa recombinantemente, o se sintetiza químicamente. En otra realización, la invención se refiere a un método como el descrito más arriba que además comprende la etapa de fabricar una cantidad de un polipéptido aislado o identificado mediante un método de la invención, comprendiendo dicha fabricación unir un compuesto conector al polipéptido, en donde el compuesto conector puede ser diferente del compuesto conector unido durante el aislamiento o la identificación del polipéptido, siempre y cuando dicho compuesto conector se una a dicho polipéptido mediante al menos tres enlaces covalentes, y en donde dicho polipéptido se expresa recombinantemente o se sintetiza químicamente.

DESCRIPCIÓN DETALLADA DE LA INVENCION

- 25 La invención trae nuevas características y le acompañan ventajas que se pueden explicar con más detalle en conexión con la generación de moléculas genéticamente codificadas con una estructura central. En particular, la invención da a conocer una restricción conformacional que no se logra mediante las técnicas conocidas de ciclación de péptidos. Además, el interconector en los sistemas conocidos, tales como los de Roberts (*ibid*), no tiene el carácter de un núcleo central/compuesto conector de la presente invención. En los sistemas conocidos, se utilizó el interconector simplemente para reemplazar un enlace disulfuro para generar un péptido cíclico insensible a la óxido-reducción. No se menciona ni sugiere el concepto de un núcleo central con varias adiciones tales como un complejo polipéptido-compuesto conector unido por tres enlaces covalentes como se da a conocer en la presente memoria. De hecho, debe notarse que en la presente invención el polipéptido está conectado a la estructura central mediante al menos tres enlaces covalentes, que proporcionan una diferencia estructural clave comparada con los sistemas conocidos. La conexión de una estructura central (compuesto conector) a un polipéptido codificado genéticamente vía tres o más enlaces es una reacción compleja que no se ha mostrado antes.

- Además, la conexión de un polipéptido a un compuesto conector a través de al menos tres enlaces covalentes podría producir varios productos diferentes. Esto podría ocasionar dificultades en el proceso de selección y en el procedimiento descodificador. Sin embargo, de acuerdo con la presente invención, se da a conocer una solución que utiliza un compuesto conector con tres grupos reactivos y, preferentemente una simetría rotacional ternaria, cuya combinación tiene la ventaja de producir un sólo producto. Por supuesto, el lector experto apreciará que en determinadas circunstancias oscuras, un compuesto conector con una simetría rotacional ternaria puede producir varios productos, muy especialmente en el ejemplo de una molécula tetrahédrica con tres grupos reactivos idénticos; esto también tiene una simetría rotacional ternaria, pero produciría dos estereoisómeros. Sin embargo, para facilitar la comprensión, se reconoce que son posibles tales excepciones teóricas para la formación de un único producto, produciendo los compuestos conectores adecuados con una simetría rotacional ternaria un único producto de acuerdo con la presente invención; en las raras circunstancias observadas más arriba, entonces se elige el polipéptido adecuado para evitar la formación de moléculas tetrahédricas y, por consiguiente, mantener la formación de tan sólo un único producto.

El compuesto conector utilizado en los métodos y composiciones descritos en la presente memoria es diferente de los interconectores bivalentes conocidos (p. ej., como se utilizó por Millward et al., *ibid*) en el requerimiento clave de que un compuesto conector de la invención tiene al menos tres grupos reactivos que son capaces de formar al menos tres enlaces covalentes con el polipéptido diana. Esta peculiaridad produce numerosos beneficios técnicos

para la invención. Primero, mediante la formación de enlaces entre el compuesto conector y el polipéptido a través de al menos tres enlaces covalentes se crean al menos dos lazos de polipéptido. Estos lazos se forman entre el primer y el segundo enlace, y entre el segundo y el tercer enlace del compuesto conector con el polipéptido. El conector conocido descrito por Millward et al sólo puede conectar dos grupos funcionales de un péptido, y no puede formar dos o más lazos peptídicos constreñidos.

Ventajas del enlace entre compuesto conector y polipéptido

Hay una serie de propiedades que diferencian las moléculas de la invención que tienen 3 o más enlaces a un compuesto conector desde otras moléculas tales como las que sólo contienen 2 enlaces. Algunas de ellas se explican a continuación.

10 Primero, se apreciará que las moléculas con dos enlaces a un compuesto conector están constreñidas al conectar entre sí los extremos flexibles de un péptido lineal. También es el caso de las moléculas de acuerdo con la presente invención con tres o más enlaces covalentes a un compuesto conector. Sin embargo, la conformación de las moléculas de acuerdo con la presente invención con 3 o más enlaces covalentes a un compuesto conector está constreñida por dos efectos adicionales que no se aplican a una molécula con solo dos enlaces:

- 15 (i) El polipéptido unido al compuesto conector a través de al menos tres enlaces covalentes comprenderá al menos dos lazos polipeptídicos constreñidos.
- (ii) Los lazos polipeptídicos son capaces de interactuar entre sí a través de interacciones no covalentes para generar una constricción adicional, y
- (iii) Cada uno de los lazos ocupa un espacio que no puede ser ocupado por el otro o los otros lazos, lo que
- 20 adicionalmente restringe su flexibilidad conformacional.

Para ilustrar estos puntos se pueden imaginar las posibles sendas que puede tomar un polipéptido anclado a los puntos A y C a un compuesto conector. La introducción de un punto de anclaje B, entre los puntos A y C, y al mismo compuesto conector, limitará además las posibles sendas tomadas por el polipéptido y, por este motivo, su entropía conformacional. Como la fijación del péptido a un ligando requiere la pérdida de entropía conformacional

25 (y siempre y cuando el péptido pueda adoptar una conformación que es complementaria a un ligando), se espera que la afinidad de unión entre el péptido ABC constreñido en el punto B intermedio y el ligando sea mayor que la del péptido AC. Por lo tanto, con la utilización de las moléculas constreñidas de la presente invención se consiguen mayores afinidades de unión de lo que era posible en la técnica anterior.

Además de estos puntos clave, a continuación se presentan más ventajas de tres o más enlaces covalentes entre

30 el polipéptido y el compuesto conector.

Las moléculas de la invención pueden fijarse a una diana a través de la interacción de dos o más lazos peptídicos constreñidos conformacionalmente. Cuantos más lazos de unión, más elevadas se pueden obtener las afinidades y especificidades. Se produce un efecto paralelo con los anticuerpos: se unen mejor cuando varias CDR interactúan con la diana. Las moléculas de acuerdo con la presente invención proporcionan por lo tanto

35 ventajosamente este beneficio técnico de varios lazos para la interacción, beneficio que está ausente en las moléculas con menos de tres enlaces.

Además de la provisión actual de un segundo (o posterior) lazo peptídico, es importante observar que tal lazo también ofrece la ventaja de constreñir conformacionalmente el otro o los otros lazos. Esto puede ocurrir a través de la ocupación de algunos de los pocos espacios tridimensionales que pueden entonces no estar ocupados por el

40 otro o los otros lazos. Alternativamente, esto puede ocurrir a través de interacciones no covalentes entre varios lazos.

De estas ventajas también se puede observar que los ligandos más estructurados se fijan generalmente con mayor afinidad (se pierde menos entropía durante la unión) y especificidad.

Como se explicó en la presente memoria, la invención también da a conocer la producción de estructuras peptídicas con lazos en las cuales cada uno de dos (o más) lazos tiene una propiedad diferente. Tales estructuras se denominan de «especificidad dual» para reflejar el hecho de que una única entidad molecular tiene especificidades duales atribuibles a dos partes diferentes (lazos) de la misma estructura global. La ventaja de tales realizaciones es que cada lazo se puede seleccionar o construir para que se fije a una diana diferente (tal como un «antígeno»). Esto representa otra particularidad distintiva de un sistema de tres enlaces de acuerdo con la

50 presente invención.

Además de estos efectos, las moléculas de la invención también proporcionan la posibilidad de formar un sándwich con un único antígeno (u otra entidad) entre dos segmentos de la cadena polipeptídica. Esta posibilidad está por supuesto ausente de las construcciones polipeptídicas con menos de dos lazos. Por supuesto, la disposición particular adoptada puede depender de la geometría de la construcción particular a utilizar, pero la invención lo hace posible, a diferencia de lo que ocurre con los métodos de la técnica anterior.

Por supuesto, en la explicación anterior se han mencionado los lazos generados de acuerdo con la presente invención. En algunas realizaciones, los lazos se escinden entonces. Es importante observar que incluso en tales realizaciones se forman los lazos, es simplemente que la molécula con lazo se trata como un intermediario que luego se procesa adicionalmente mediante la escisión de los lazos para producir una estructura peptídica lineal con varias fijaciones. En estas realizaciones, debido a que el péptido se conecta inicialmente al compuesto conector a través de tres o más enlaces covalentes, las moléculas quedarán decoradas con tres o más restos peptídicos después de la escisión del polipéptido. Tales moléculas pueden formar más interacciones con dianas, y se esperan afinidades de unión/especificidades más elevadas, que es una ventaja más del sistema de tres enlaces de la invención.

Es una ventaja que las moléculas de la invención que tienen dos o más lazos polipeptídicos puedan formar más interacciones con un ligando diana y, por consiguiente, pueden tener mayor afinidad y/o especificidad que las moléculas polipeptídicas con sólo un lazo. Por ejemplo, puede ser deseable refinar el segundo lazo para que mejore su afinidad, lo que claramente no es posible para las moléculas con un único lazo.

Es una ventaja que, en los complejos de la invención, el compuesto conector tenga al menos dos lazos polipeptídicos cercanos en el espacio. Estos dos o más lazos pueden interactuar simultáneamente con diferentes epítomos sobre el mismo ligando diana.

Además, el beneficio de tener dos o más lazos se puede explotar para la fabricación de moléculas de «especificidad dual», en donde un lazo tiene, o se seleccionado por ello, una propiedad particular o afinidad de unión, y el otro lazo tiene una propiedad o afinidad diferente. Estas moléculas se denominan «biespecíficas» o de «especificidad dual». Hay varios tipos posibles. Por ejemplo,

(a) biespecíficas hechas seleccionando el lazo 1 y luego el lazo 2 (o más)

(b) dos macrociclos bicíclicos unidos

(c) un macrociclo bicíclico más péptido o fármaco.

Para (a), esto podría típicamente realizarse haciendo/seleccionando una alícuota de una genoteca contra un primer antígeno, y otra alícuota contra un segundo antígeno. Los lazos seleccionados entonces podrían combinarse por pares, por ejemplo, mediante técnicas estándares tales como recombinar los segmentos de ácidos nucleicos que codifican los dos lazos para proporcionar una nueva genoteca de diferentes combinaciones del primer y segundo lazos. Las moléculas combinadas por pares (p. ej., fago) luego se pueden detectar y/o seleccionar por fijarse a ambos antígenos secuencialmente. De este modo se pueden fabricar biespecíficos capaces de fijarse a dos antígenos independientes. Naturalmente, este método puede aumentarse con otras etapas opcionales, e igualmente la afinidad de unión para cada antígeno podría mejorarse mediante la mutación de cada lazo, que puede ser por mutación dirigida o incluso aleatoria.

En una variación de esta técnica se podría seleccionar una alícuota de una colección para fijación a un primer antígeno. El lazo más importante para la fijación podría identificarse, por ejemplo, mediante la inspección de «secuencias consenso» entre las seleccionadas como aglutinantes, y el otro lazo se podría aleatorizar y seleccionar contra el segundo antígeno.

La mayoría de los biespecíficos adecuados de este tipo tales como los descritos en (a) se fabricarían en los fagos.

Las moléculas variantes observadas en (b) y (c) más arriba podrían igualmente realizarse como un fago (de una manera similar a la de más arriba). Alternativamente, lo más probable es que las dos entidades unidas se podrían seleccionar por separado y luego fusionarlas en la etapa de la síntesis química, lo que podría simplificar su selección/construcción.

Otra característica ventajosa significativa más de los aspectos de la invención es que, además del compuesto conector de la invención que sirve para conectar los segmentos del polipéptido mediante enlaces covalentes con los restos de aminoácidos en la base de cada lazo peptídico, el compuesto conector también se emplea en otras interacciones no covalentes (tales como enlaces iónicos, interacciones hidrófobas, puentes de hidrógeno, interacciones de Van der Waals) con otros elementos de la cadena polipeptídica, tales como otros restos de

aminoácidos. Por el contrario, el conector bivalente de Millward et al. es lineal y muy flexible (propilo) y su única función es la de conectar dos extremos de un polipéptido para crear un péptido cíclico insensible a la óxido-reducción. No sería esperable que el conector de Millward et al. estableciera interacciones no covalentes significativas con el lazo polipeptídico, al ser pequeño y muy flexible, y desde luego no hay pruebas de que lo haga.

- 5 Esta ventaja de la invención se ilustra adicionalmente en la sección de ejemplos, junto con las pruebas de las interacciones no covalentes ventajosas. Así pues, convenientemente, el polipéptido se une al compuesto conector mediante una o más interacciones no covalentes, además del o de los enlaces covalentes explicados en la presente memoria. Esto tiene la ventaja adicional de proporcionar otro nivel de constreñimiento estructural al complejo/conjugado de la invención.
- 10 Es una ventaja de la invención que una molécula con varios lazos peptídicos esté por lo general más estructurada que un polipéptido con un único lazo peptídico. Las moléculas muy estructuradas tienen a ser más específicas. Igualmente, las moléculas bien estructuradas tienen por lo general mejor afinidad de unión. Además, una molécula con varios lazos peptídicos es capaz de formar más interacciones con un ligando diana que con un polipéptido con un único lazo peptídico.
- 15 Es otro beneficio de los aspectos de la invención que el compuesto conector de la invención también impone constreñimientos conformacionales que se derivan de su estructura química. Por ejemplo, algunos grupos químicos se sabe que son inflexibles, que impiden la rotación, que proporcionan impedimentos o restricciones estéricos, que presentan una estructura rígida o si no, que proporcionan el armazón o constreñimiento al complejo. Así pues, convenientemente, el compuesto conector de la invención comprende un grupo de armazón tal como un
- 20 grupo de armazón rígido. La función de este grupo de armazón es proporcionar la estructura o constreñimiento molecular al complejo de la invención. En conexión con un compuesto conector preferente de la invención, el tris-(bromometil)benceno (TBMB), esta peculiaridad se ilustraría con referencia a la estructura plana del grupo benceno del TBMB. Este grupo bencénico es rígido debido a su carácter plano, y así es capaz de servir de grupo de armazón del compuesto conector, en particular un grupo de armazón rígido.
- 25 Así pues, en una realización más preferente de la invención, el compuesto conector proporciona constreñimientos conformacionales al polipéptido impuestos por los al menos tres enlaces covalentes, proporciona estructura adicional a través de los enlaces no covalentes entre el compuesto conector y el polipéptido, y además el compuesto conector de la invención también impone constreñimientos conformacionales por la naturaleza de su propia estructura química que sirve de armazón rígido. Por ejemplo, la estructura plana del grupo bencénico
- 30 cuando el compuesto conector comprende el mismo, tal como cuando el compuesto conector es tris-(bromometil)benceno (TBMB).

COMPUESTO CONECTOR

El compuesto conector se denomina a veces el «núcleo molecular». Convenientemente, el compuesto conector posee simetría molecular. Convenientemente, el compuesto conector posee tres grupos reactivos y posee simetría ternaria. Esto tiene la ventaja de producir solo un único producto de reacción. Si el compuesto conector no es una molécula simétrica, entonces se pueden generar numerosos productos de reacción. Esto puede conllevar complicaciones o requerir que el isómero deseado quede separado de los otros productos de reacción. Al usar un compuesto conector que tiene la simetría adecuada, tales problemas se mejoran ventajosamente.

- 40 Es una ventaja de la invención que los polipéptidos producidos tengan una mayor complejidad que los péptidos cíclicos de la técnica anterior. Por ejemplo, los polipéptidos producidos de acuerdo con la presente invención podrían poseer más de dos lazos para la interacción con otras entidades químicas. Además, los polipéptidos producidos de acuerdo con la presente invención disfrutan de un mayor nivel de constreñimiento que los polipéptidos basados en la técnica anterior. Estos dos efectos juntos crean la ventaja adicional de que varios lazos (o «ciclos») del polipéptido quedan retenidos muy cerca unos de otros a través de sus enlaces con el compuesto
- 45 conector común. Esto proporciona un mayor nivel de constreñimiento sobre la conformación de los polipéptidos.

- Típicamente, los polipéptidos cíclicos de la técnica anterior se unen mediante varios restos de cisteína tales como dos restos de cisteína de un puente entre dos partes del péptido y mediante lo cual se forma un polipéptido cíclico. Sin embargo, tales moléculas son sensibles a la óxido-reducción. El método de Millward et al. está directamente enfocado a la producción de péptidos cíclicos que son insensibles a la óxido-reducción. En este sentido, el método
- 50 de Millward et al. se aparta de la técnica anterior y da a conocer el uso de las cisteínas como grupos reactivos para la modificación de polipéptidos. Por el contrario, de acuerdo con la presente invención, las cisteínas son los grupos reactivos preferentes.

Cuando hay tres o más grupos reactivos para al menos tres enlaces covalentes independientes con el compuesto conector, dichos grupos reactivos no necesitan ser cada uno una cisteína. Por ejemplo, los tres grupos reactivos

pueden comprender una cisteína y otros dos grupos reactivos adecuados, que por ejemplo podrían comprender lisina, selenocisteína u otro(s). Más convenientemente, los tres grupos reactivos son cisteína.

Los métodos de la técnica anterior solo han conducido a la producción de polipéptidos con un único lazo. De acuerdo con la presente invención, se pueden producir al menos dos lazos o incluso más mediante la sujeción de
5 diferentes puntos del polipéptido al compuesto conector.

El método de la presente invención implica un mínimo de tres enlaces con el polipéptido. Esto tiene la ventaja de un mayor constreñimiento molecular. En esto consiste la ventaja adicional de la presentación de varios lazos polipeptídicos para la interacción con otros restos.

En las técnicas conocidas se ha introducido o unido al polipéptido, tal como un polipéptido codificado
10 genéticamente, un mejor agente de interconexión. Por el contrario, la presente invención da a conocer un compuesto conector para la coordinación múltiple de diferentes partes del mismo polipéptido.

Convenientemente, el compuesto conector puede ser una molécula pequeña. Convenientemente, el compuesto conector es una molécula orgánica pequeña.

Convenientemente, el compuesto conector puede ser monómeros naturales tales como nucleósidos, azúcares o
15 esteroides, o puede basarse en ellos. Convenientemente, el compuesto conector puede comprender un polímero pequeño de tales entidades, tal como un dímero o un trímero.

Convenientemente, el compuesto conector es un compuesto de toxicidad conocida, convenientemente de baja toxicidad. Ejemplos de compuestos adecuados incluyen colesterol, nucleótidos, esteroides o fármacos existentes tal como tamazapán.

20 Convenientemente, el compuesto conector puede ser una macromolécula. Convenientemente, el compuesto conector es una macromolécula compuesta de aminoácidos, nucleótidos o glúcidos.

Convenientemente, el compuesto conector comprende grupos reactivos que son capaces de reaccionar con uno o más grupos funcionales del polipéptido diana para formar enlaces covalentes.

El compuesto conector puede comprender grupos químicos como aminas, tioles, alcoholes, cetonas, aldehídos,
25 nitrilos, ácidos carboxílicos, ésteres, alquenos, alquinos, azidas, anhídridos, succinimidas, maleimidas, haluros de alquilo y haluros de acilo.

Convenientemente, el compuesto conector puede comprender o puede consistir en tris(bromometil)benceno o un derivado del mismo.

Convenientemente, el compuesto conector tiene una simetría rotacional ternaria de tal forma que al hacer
30 reaccionar tres grupos funcionales del polipéptido diana con el compuesto conector se genera un único isómero del producto.

En algunas realizaciones, el compuesto conector puede tener geometría tetrahédrica de tal forma que la reacción de cuatro grupos funcionales del polipéptido codificado con el compuesto conector genera no más de dos isómeros del producto.

35 Un compuesto conector adecuado es el 1,3,5-tris(bromometil)benceno (TBMB).

Un compuesto conector adecuado es el 2,4,6-tris(bromometil)mesitileno. Es similar al 1,3,5-tris(bromometil)benceno, pero contiene otros tres grupos metilo unidos al anillo bencénico. Esto tiene la ventaja de que los grupos metilo adicionales pueden formar otros contactos con el polipéptido y así añadir otros constreñimientos estructurales.

40 El compuesto conector de la presente invención se selecciona entre o bien una molécula pequeña o bien una estructura macromolecular. Dicho compuesto conector está formado por componentes orgánicos, inorgánicos u orgánicos e inorgánicos.

En una realización preferente, el compuesto conector es una molécula orgánica pequeña como por ejemplo, un
45 alcano lineal. Más convenientemente, el compuesto conector es un alcano ramificado, un alcano cíclico, un alcano policíclico, un aromato, un alcano heterocíclico o un aromato heterocíclico, lo cual ofrece la ventaja de ser menos flexible (esto es, más rígido).

En otra realización, el compuesto conector se selecciona entre una estructura macromolecular como por ejemplo

un polipéptido, un polinucleótido o un polisacárido.

El compuesto conector de la invención contiene grupos químicos que permiten que los grupos funcionales del polipéptido de la librería codificada de la invención formen enlaces covalentes con el compuesto conector. Dichos grupos químicos se seleccionan entre una amplia abanico de grupos funcionales, que incluyen aminas, tioles, alcoholes, cetonas, aldehídos, nitrilos, ácidos carboxílicos, ésteres, alquenos, alquinos, anhídridos, succinimidas, maleimidas, azidas, haluros de alquilo y haluros de acilo.

En una realización, el compuesto conector de la invención es tris(bromometil)benceno o un derivado del mismo.

POLIPÉPTIDO

Los grupos funcionales de los polipéptidos codificados son convenientemente proporcionados por las cadenas laterales de los aminoácidos naturales y no naturales. Los grupos funcionales de los polipéptidos codificados se seleccionan convenientemente entre grupos tiol, grupos amino, grupos carboxilo, grupos guanidinio, grupos fenólicos o grupos hidroxilo. Los grupos funcionales de los polipéptidos codificados se podrían seleccionar convenientemente entre grupos azida, cetocarboxilo, alquino, vinilo o haluro de arilo. Los grupos funcionales de los polipéptidos codificados con los que unirse a un compuesto conector pueden convenientemente ser el extremo amino o carboxilo del polipéptido.

En algunas realizaciones, cada uno de los grupos funcionales del polipéptido con los que se une al compuesto conector son del mismo tipo. Por ejemplo, cada grupo funcional puede ser un resto de cisteína.

En algunas realizaciones, los grupos funcionales con los que se une al compuesto conector pueden comprender dos o más tipos diferentes, o pueden comprender tres o más tipos diferentes. Por ejemplo, los grupos funcionales pueden comprender dos restos de cisteína y un resto de lisina, o pueden comprender un resto de cisteína, un resto de lisina y una amina del extremo amino.

En algunas realizaciones, aminoácidos alternativos, tales como los aminoácidos naturales, pueden resultar adecuados para modificar químicamente los polipéptidos, tal como los péptidos expuestos en fagos de la invención.

La cisteína es el aminoácido más adecuado porque tiene la ventaja de que su reactividad es la más diferente entre los aminoácidos. Los grupos reactivos que se pueden usar en el compuesto conector para que reaccionen con los grupos tiol de las cisteínas son haluros de alquilo (o también denominados halogenoalcanos o haloalcanos). Ejemplos son el bromometilbenceno (el grupo reactivo de ejemplo en el TBMB) o la yodoacetamida. Las maleimidas son otros grupos reactivos que se usan para conjugar selectivamente los compuestos a las cisteínas de las proteínas. Ejemplos de maleimidas que se pueden usar como compuestos conectores en la invención incluyen: tris-(2-maleimidoetil)amina, tris-(2-maleimidoetil)benceno, tris-(maleimido)benceno. La selenocisteína también es un aminoácido natural que tiene una reactividad similar a la cisteína y que se puede usar para las mismas reacciones. Así pues, siempre que se mencione la cisteína, es típicamente aceptable su sustitución por selenocisteína a menos que el contexto sugiera otra cosa. La cisteína se usa con más preferencia.

Las lisinas (y las aminas primarias del extremo amino de los péptidos) también son grupos funcionales adecuados para modificar los péptidos de los fagos mediante la unión con un compuesto conector. Sin embargo, en las proteínas de los fagos abundan más que las cisteínas, y el riesgo de que las partículas de fago se queden entrecruzadas o pierdan su infectividad es elevado. No obstante, encontramos que las lisinas son especialmente útiles en las reacciones intramoleculares (por ejemplo, cuando un compuesto conector ya está unido al péptido en el fago) para formar un segundo o sucesivos enlaces con el compuesto conector. En este caso, el compuesto conector reacciona preferentemente con las lisinas del péptido expuesto (en particular con las lisinas que estén muy cerca). Los grupos funcionales que reaccionan selectivamente con las aminas primarias son succinimidas, aldehídos o haluros de alquilo. Con respecto a los haluros de alquilo, el lector sabrá que existen haluros de alquilo de diferente reactividad. In el grupo bromometilo que hemos usado en una serie de ejemplos acompañantes, los electrones del anillo bencénico son capaces de estabilizar el estado de transición catiónico. Este haluro de alquilo particular es por tanto 100-1000 veces más reactivo que los haluros de alquilo que no están conectados a un grupo bencénico. Los ejemplos de succinimidas para uso como compuesto conector incluyen tris(succinimidil-aminotriacetato), ácido 1,3,5-benciltriacético. Los ejemplos de aldehídos para uso como compuesto conector incluyen el triformilmetano. Los ejemplos de haluros de alquilo para uso como compuesto conector incluyen 1,3,5-tris(bromometil)-2,4,6-trimetilbenceno, 1,3,5-tris(bromometil)benceno, 1,3,5-tris(bromometil)-2,4,6-trietilbenceno.

En algunas realizaciones, los conectores moleculares o las modificaciones se pueden añadir a (o crear en) los grupos funcionales de los polipéptidos codificados antes de que se unan al compuesto conector, en donde dichos conectores o modificaciones son capaces de reaccionar con el compuesto conector.

Los aminoácidos con grupos funcionales para unir a un compuesto conector pueden estar localizados en cualquier posición adecuada dentro del polipéptido codificado. Para influir en la creación de lazos o de estructuras particulares, el experto en la técnica sabe cómo variar las posiciones de los aminoácidos que tienen los grupos funcionales, por ejemplo mediante manipulación del ácido nucleico que codifica el polipéptido para mutar el

5 polipéptido producido.

Cada uno de los aminoácidos del polipéptido codificado puede ser una diana para la mutagénesis (por ejemplo, mutagénesis de variación restringida) de acuerdo con las necesidades del experto o el propósito con el que se aplica la invención. Claramente se requiere que el polipéptido de interés presente al menos tres grupos funcionales para formar enlaces con el compuesto conector. Los aminoácidos que no sean los requeridos para formar enlaces con el compuesto conector pueden variar con libertad de acuerdo con las necesidades del operador y se denominan «aminoácidos variables». Dichos aminoácidos variables del polipéptido codificado (por ejemplo, uno o varios miembros de la colección de polipéptidos) pueden estar aleatorizados, parcialmente aleatorizados o constantes.

10

El polipéptido diana comprende un segmento de fijación al compuesto conector. Esta es la región a la que se une el compuesto conector. Convenientemente, a este segmento de fijación se aplica el comentario con respecto a los grupos funcionales del polipéptido. Convenientemente, el segmento de fijación al compuesto conector que tiene el polipéptido diana comprende de 1 a 20 restos de aminoácidos. Convenientemente, el segmento de fijación al compuesto conector que tiene el polipéptido diana comprende menos de 10 aminoácidos. Esto tiene la ventaja de imponer más constricciones conformacionales al segmento del polipéptido cuando está unido al compuesto conector.

20

El polipéptido diana comprende convenientemente la secuencia $AC(X)_6C(X)_6CG$, en la que X se refiere a un aminoácido natural al azar, A es alanina, C es cisteína y G es glicina.

El polipéptido diana comprende convenientemente la secuencia $(X)_lY(X)_mY(X)_nY(X)_o$, en donde Y representa un aminoácido con un grupo funcional, X representa un aminoácido al azar, m y n son números entre 1 y 20 que definen la longitud de los segmentos polipeptídicos interpuestos, y l y o son números entre 0 y 20 que definen la longitud de los segmentos polipeptídicos flanqueantes.

25

En algunas realizaciones, el complejo de la invención puede comprender un polipéptido con la secuencia $AC(X)_6C(X)_6CG$. En una realización, un miembro o complejo de una genoteca de la invención puede comprender un compuesto conector mesitileno y un polipéptido con la secuencia $AC(X)_6C(X)_6CG$, en donde el polipéptido se sujeta a los grupos metilo exocíclicos del compuesto conector mediante los restos de cisteína del polipéptido al formar tres enlaces tioéter con ellos, y en donde X se refiere a un aminoácido (convenientemente un aminoácido natural), A es alanina, C es cisteína y G es glicina.

30

Convenientemente, el polipéptido diana comprende un inhibidor de la calicreína plasmática humana y el polipéptido comprende una o más de las secuencias de aminoácidos

35 **GCSDFRNCPEALCG, ACSDFRNCPLWSGTCG, ACSTERYCPIEFPCG, ACAPWRTACYEDLMWCG, ACGTGEGRVNVWTPCG** o una secuencia relacionada.

Con secuencia relacionada se quiere decir una secuencia de aminoácidos que tiene al menos una identidad del 50%, convenientemente al menos una identidad del 60%, convenientemente al menos una identidad del 70%, convenientemente al menos una identidad del 80%, convenientemente al menos una identidad del 90%, convenientemente al menos una identidad del 95%, convenientemente al menos una identidad del 98%, convenientemente al menos una identidad del 99%. La identidad se juzga convenientemente a lo largo de un segmento contiguo de al menos 10 aminoácidos, convenientemente al menos 12 aminoácidos, convenientemente al menos 14 aminoácidos, convenientemente al menos 16 aminoácidos, convenientemente al menos 17 aminoácidos o la longitud completa de la secuencia de referencia.

40

Convenientemente, el polipéptido diana comprende un inhibidor de la catepsina G humana y el polipéptido comprende una o más de las secuencias de aminoácidos

45 **ACEYDGLWCGWDPVCG, ACIFDLGFCHNDWWNCG, ACLRAQEDCYDRGFPG** o una secuencia relacionada.

Convenientemente, el polipéptido diana comprende un inhibidor del activador humano del plasminógeno del tipo urocinasa y el polipéptido comprende una o más de las secuencias de aminoácidos

50 **ACNSRFSGCQIDLLMCG, ACSRYEVDGRGRGACG** o una secuencia relacionada.

Convenientemente, el polipéptido diana está incluido en una colección de polipéptidos que contiene al menos 10^5 miembros, más convenientemente al menos 10^9 miembros. La invención también se refiere a tales colecciones.

GRUPOS REACTIVOS DEL POLIPÉPTIDO

El compuesto conector de la invención puede estar fijado al polipéptido a través de grupos funcionales o reactivos del polipéptido. Están formados típicamente por las cadenas laterales de aminoácidos determinados que se encuentran en el polímero polipeptídico. Tales grupos reactivos pueden ser una cadena lateral de cisteína, una

5 cadena lateral de lisina o un grupo amina del extremo amino o cualquier otro grupo reactivo adecuado.

Convenientemente, al menos un grupo funcional es un grupo cisteína. Grupos tales como las aminas de las lisinas o del extremo amino no suelen ser suficientemente reactivos para formar un enlace con el compuesto conector por sí mismos dentro de un intervalo de tiempo conveniente. Sin embargo, una vez que el compuesto conector ha sido atraído o enlazado a al menos una cisteína, entonces la cinética ordinaria de la reacción da a entender que los

10 enlaces de lisina o amina pueden formarse rápida y establemente a partir de ese momento. Por este motivo, convenientemente, al menos uno de los grupos funcionales es un grupo cisteína.

Si fueran deseables los grupos reactivos en el polipéptido diferentes de los grupos de cisteína/lisina/amina, entonces se podría elegir un compuesto conector diferente para acoplarse con los grupos reactivos funcionales determinados de elección en el polipéptido diana.

15 Convenientemente, los grupos cisteína, lisina o amina se utilizan como grupos funcionales o reactivos en el polipéptido de interés.

Convenientemente, se forman al menos tres enlaces covalentes entre el compuesto conector y el polipéptido de interés.

En algunas realizaciones se pueden formar cuatro enlaces o incluso más entre el compuesto conector y el

20 polipéptido de interés. Sin embargo, si se utilizan más de cuatro enlaces, entonces típicamente las mezclas de producto formadas se vuelven cada vez más complejas y pueden obstaculizar los posteriores usos o aplicaciones. Por este motivo se prefieren tres enlaces o cuatro enlaces entre el compuesto conector y el polipéptido de interés.

En cualquier realización se prefiere que el compuesto conector tenga simetría molecular. Lo más preferente son los compuestos conectores que tienen tres grupos funcionales o reactivos. Lo más preferente son los compuestos

25 conectores que tienen una simetría molecular ternaria.

Los grupos funcionales de los polipéptidos codificados genéticamente de la invención son capaces de formar enlaces covalentes con el núcleo molecular/compuesto conector. Los grupos funcionales son grupos específicos de átomos dentro de aminoácidos tanto naturales como no naturales. Preferentemente, los grupos funcionales con una reactividad química distintiva se utilizan para conectar el polipéptido al compuesto conector para formar el

30 complejo de la invención. El uso de dichos grupos funcionales distintivos permite que el núcleo molecular/compuesto conector forme enlace exclusivamente con los grupos funcionales designados del polipéptido, pero no con otros grupos químicos de o bien los distintos elementos del polipéptido, o bien el ácido nucleico, o bien otros componentes del complejo.

Los grupos funcionales adecuados de aminoácidos naturales son el grupo tiol de la cisteína, el grupo amino de la lisina, el grupo de carboxilo del aspartato o del glutamato, el grupo guanidinio de la arginina, el grupo fenol de la tirosina o el grupo hidroxilo de la serina. Los aminoácidos no naturales pueden proporcionar un gran abanico de grupos funcionales, entre ellos, una azida, un cetocarbonilo, un alquino, un vinilo o un grupo haluro de arilo. El grupo amino y el carboxilo de los extremos del polipéptido también pueden servir de grupos funcionales para formar enlaces covalentes con un compuesto conector/núcleo molecular.

35 Los grupos funcionales adecuados de aminoácidos naturales son el grupo tiol de la cisteína, el grupo amino de la lisina, el grupo de carboxilo del aspartato o del glutamato, el grupo guanidinio de la arginina, el grupo fenol de la tirosina o el grupo hidroxilo de la serina. Los aminoácidos no naturales pueden proporcionar un gran abanico de grupos funcionales, entre ellos, una azida, un cetocarbonilo, un alquino, un vinilo o un grupo haluro de arilo. El grupo amino y el carboxilo de los extremos del polipéptido también pueden servir de grupos funcionales para formar enlaces covalentes con un compuesto conector/núcleo molecular.

Los polipéptidos codificados de la invención contienen convenientemente al menos tres grupos funcionales. Dichos polipéptidos también pueden contener cuatro o más grupos funcionales. Cuantos más grupos funcionales se utilizan, más diversidad de segmentos se pueden sujetar al núcleo molecular/compuesto conector. Sin embargo, no se recomienda la unión de un número excesivo de grupos funcionales a un núcleo molecular/compuesto conector porque puede conducir a un número inmanejable de isómeros de productos. Convenientemente se utilizan

45 tres, cuatro o cinco enlaces covalentes a un compuesto conector; lo más convenientemente tres o cuatro enlaces covalentes; lo más convenientemente tres enlaces covalentes.

En una realización preferente se generan polipéptidos codificados con tres grupos funcionales. La reacción de dichos polipéptidos con un núcleo molecular/compuesto conector que tiene una simetría rotacional ternaria genera un único isómero del producto. La generación de un único isómero del producto es favorable por varias razones.

50 Los ácidos nucleicos (a veces denominados los «códigos genéticos») de las genotecas compuestas codifican sólo las secuencias primarias del polipéptido, pero no la isomería de las moléculas que se forman al hacer reaccionar el polipéptido codificado con el núcleo molecular. Sólo en el caso de que se forme un único isómero del producto se definirá claramente la asignación del ácido nucleico al isómero del producto. Si se forman varios isómeros del

producto, el ácido nucleico no puede dar información sobre la naturaleza del isómero del producto que se aisló en un procedimiento de escrutinio o de selección. La formación de un único isómero del producto también es ventajosa si se sintetiza un miembro específico de una genoteca de la invención. En este caso, la reacción química del polipéptido con el compuesto conector produce un único isómero del producto en vez de una mezcla de isómeros.

En otra realización de la invención, se generan polipéptidos codificados con cuatro grupos funcionales. La reacción de dichos polipéptidos con un núcleo molecular/compuesto conector que tiene una simetría tetrahédrica genera dos isómeros del producto. Incluso aunque los dos diferentes isómeros del producto estén codificados por uno y el mismo ácido nucleico («código genético»), la naturaleza isomérica del isómero aislado se puede determinar mediante la síntesis química de ambos isómeros, separando los dos isómeros y comprobando si ambos isómeros se fijan a un ligando diana.

En una realización de la invención, al menos uno de los grupos funcionales de los polipéptidos es ortogonal a los demás grupos funcionales. El uso de los grupos funcionales ortogonales permite dirigir dichos grupos funcionales ortogonales a sitios específicos del núcleo molecular. Las estrategias de conexión que implican grupos funcionales ortogonales se puede utilizar para limitar el número de isómeros del producto formados. En otras palabras, al elegir grupos funcionales diferentes o distintos para uno o más de los al menos tres enlaces con los elegidos por el que queda de los al menos tres enlaces, se puede conseguir de manera útil un orden determinado de fijación o dirección de los grupos funcionales específicos del polipéptido a posiciones específicas del compuesto conector.

En otra realización, los grupos funcionales del polipéptido codificado de la invención se hacen reaccionar con conectores moleculares, en donde dichos conectores son capaces de reaccionar con un armazón molecular/compuesto conector, de modo que el conector se interpondrá entre el compuesto conector y el polipéptido en el estado fijado final.

Los aminoácidos adecuados de los miembros de las genotecas combinatorias de entidades químicas codificadas genéticamente se pueden reemplazar por cualquier aminoácido natural o no natural. Quedan excluidos de estos aminoácidos intercambiables los que albergan grupos funcionales con los que entrecruzar los polipéptidos a un núcleo molecular. Un grupo de aminoácidos adyacentes que puede variar se define como un segmento polipeptídico. El tamaño de un único segmento polipeptídico oscila de 1 a 20 aminoácidos. Los segmentos polipeptídicos tienen secuencias aleatorias, secuencias constantes o secuencias con aminoácidos constantes y aleatorios. Los aminoácidos con grupos funcionales se localizan o bien en posiciones definidas o bien en posiciones aleatorias del polipéptido codificado de la invención.

En una realización, los segmentos de polipéptido que están fijados por dos aminoácidos que albergan grupos funcionales para fijarse a un núcleo molecular/compuesto conector son secuencias de aminoácidos cortas de 10 o menos aminoácidos. La reacción de dichas secuencias de polipéptidos codificados con un núcleo molecular genera miembros de la genoteca con un constreñimiento conformacional elevado. Los ligandos con constreñimiento conformacional son generalmente más específicos y tienen mayor afinidad de unión. El constreñimiento conformacional también puede proteger los ligandos de la degradación proteolítica, por ejemplo en los líquidos corporales.

En una realización, un polipéptido codificado con tres grupos funcionales tiene la secuencia $(X)_l Y(X)_m Y(X)_n Y(X)_o$, en donde Y representa un aminoácido con un grupo funcional, X representa un aminoácido aleatorio, m y n son números entre 1 y 20 que definen la longitud de los segmentos polipeptídicos interpuestos, y l y o son números entre 0 y 20 que definen la longitud de los segmentos polipeptídicos flanqueantes.

En una realización preferente, una genoteca de polipéptidos codificados de la invención tiene la secuencia $AC(X)_6 C(X)_6 CG$, en donde A representa alanina, C representa cisteína, X representa un aminoácido natural al azar y G representa glicina.

Se pueden utilizar alternativas a conjugaciones a través de tioles para unir el compuesto conector al péptido mediante interacciones covalentes. Alternativamente se pueden utilizar estas técnicas para modificar o unir otros restos (tales como moléculas pequeñas de interés que son distintas al compuesto conector) al polipéptido después de que se haya seleccionado o aislado según la presente invención; en esta realización, la unión no necesita ser entonces claramente covalente y puede abarcar una unión no covalente. Estos métodos se puede utilizar en vez de (o en combinación con) los métodos en los que intervienen tioles mediante la producción de fagos que exponen proteínas y péptidos que llevan aminoácidos no naturales con el requisito de grupos funcionales químicos, en combinación con moléculas pequeñas que llevan el grupo funcional complementario, o mediante la incorporación de los aminoácidos no naturales en un polipéptido sintetizado recombinantemente o químicamente cuando la molécula se fabrica después de la fase de selección/aislamiento.

- Los aminoácidos no naturales incorporados en los péptidos y proteínas expuestos en fagos pueden incluir 1) un grupo funcional cetónico (como el encontrado en la *para*- o *meta*-acetilfenilalanina) que se puede hacer reaccionar específicamente con hidrazinas, hidroxilaminas y sus derivados («Addition of the keto functional group to the genetic code of *Escherichia coli*». Wang L, Zhang Z, Brock A, Schultz P. G. *Proc. Natl. Acad. Sci. USA*. 2003 Jan 7; 100 (1): 56-61; *Bioorg. Med. Chem. Lett.* 15 de octubre de 2006; 16 (20): 5356-9. «Genetic introduction of a diketone-containing amino acid into proteins». Zeng H, Xie J, Schultz PG), 2) las azidas (como las encontradas en la *p*-azido-fenilalanina) que se pueden hacer reaccionar con alquinos mediante la «química clic» catalizada por cobre o cicloadiciones (3+2) promovidas por tensión para formar los correspondientes triazoles («Addition of *p*-azido-L-phenylalanine to the genetic code of *Escherichia coli*». Chin J. W., Santoro S. W., Martin A. B., King D. S., Wang L., Schultz P. G.. *J. Am. Chem. Soc.* 7 de agosto de 2002; 124 (31): 9026-7; «Adding amino acids with novel reactivity to the genetic code of *Saccharomyces cerevisiae*». Deiters A, Cropp T. A., Mukherji M., Chin J. W., Anderson J. C., Schultz P. G. *J. Am. Chem. Soc.* 1 de octubre de 2003; 125 (39): 11782-3), o azidas que se pueden hacer reaccionar con fosfinas de arilo, mediante una ligación de Staudinger («Selective Staudinger modification of proteins containing *p*-azidophenylalanine». Tsao M. L., Tian F., Schultz P. G. «*ChemBiochem.*», diciembre de 2005; 6 (12): 2147-9), para formar las correspondientes amidas, 4) alquinos que se pueden hacer reaccionar con azidas para formar el correspondiente triazol («In vivo incorporation of an alkyne into proteins in *Escherichia coli*», Deiters A, Schultz P. G. *Bioorg. Med. Chem. Lett.* 1 de marzo de 2005; 15(5): 1521-4), 5) ácidos borónicos (boronatos) que se pueden hacer reaccionar específicamente con compuestos que contienen más de un grupo hidroxilo espaciado adecuadamente o que se someten a acoplamiento mediado por paladio con compuestos halogenados (*Angew Chem. Int. Ed. Engl.* 2008; 47 (43): 8220-3). «A genetically encoded boronate-containing amino acid», Brustad E., Bushey M. L., Lee J. W., Groff D, Liu W, Schultz P. G.), 6) aminoácidos quelantes de metales, incluidos los que llevan bipyridilos, que se pueden coordinar específicamente a un ion metálico (*Angew Chem. Int. Ed. Engl.* 2007; 46(48): 9239-42. «A genetically encoded bidentate, metal-binding amino acid». Xie J, Liu W, Schultz P. G.).
- 25 Los aminoácidos no naturales se pueden incorporar en las proteínas o en los péptidos expuestos en fagos mediante la transformación de *E. coli* con plásmidos o combinaciones de plásmidos que llevan: 1) la aminoacil-ARNt sintetasa ortogonal y el ARNt que dirige la incorporación del aminoácido no natural en respuesta a un codón, 2) El ADN del fago o el plásmido fagémido alterado para contener el codón seleccionado en el sitio de la incorporación de los aminoácidos no naturales (*Proc. Natl. Acad. Sci. USA*, 18 de noviembre de 2008; 105(46): 17688-93. «Protein evolution with an expanded genetic code». Liu C. C., Mack A. V., Tsao M. L., Mills J. H., Lee H. S., Choe H, Farzan M, Schultz P. G., Smider V. V.; «A phage display system with unnatural amino acids». Tian F, Tsao M. L., Schultz P. G. «*J. Am. Chem. Soc.*», 15 de diciembre de 2004; 126 (49): 15962-3). La aminoacil-ARNt sintetasa ortogonal y el ARNt se pueden obtener de la pareja de tirosilos de *Methanococcus janashii* o una sintetasa («Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*», Chin J. W., Martin A. B., King D. S., Wang L., Schultz P. G. *Proc. Natl. Acad. Sci. USA*, 20 de agosto de 2002; 99 (17): 11020-4) y la pareja del ARNt que incorpora pirrolisina de forma natural («Multistep engineering of pyrrolysyl-tARN synthetase to genetically encode N(epsilon)-o-azidobenzoyloxycarbonyllysine for site-specific protein modification». Yanagisawa T, Ishii R, Fukunaga R, Kobayashi T, Sakamoto K, Yokoyama, S. *Chem. Biol.* 24 de noviembre de 2008; 15 (11): 1187-97; «Genetically encoding N(epsilon)-acetyllysine in recombinant proteins». Neumann H, Peak-Chew S. Y., Chin J. W., *Nat. Chem. Biol.*, abril de 2008; 4(4): 232-4. *Epub*, 17 de febrero de 2008). El codón para la incorporación puede ser el codón ámbar (UAG), otro codón de parada (UGA, o UAA), alternativamente puede ser un codón de cuatro bases. La aminoacil-ARNt sintetasa y el ARNt se pueden sintetizar desde los vectores existentes, entre ellos, la serie de vectores pBK, vectores pSUP («Efficient incorporation of unnatural amino acids into proteins in *Escherichia coli*», Ryu Y, Schultz P. G. *Nat. Methods*, abril de 2006; 3(4): 263-5), y vectores pDULE (*Nat. Methods.*, mayo de 2005, 2(5): 377-84., «Photo-cross-linking interacting proteins with a genetically encoded benzophenone». Farrell I. S., Toroney R., Hazen J. L., Mehl R. A., Chin J. W.). La cepa de *E. coli* utilizada expresará el pelo F' (generalmente mediante un operón *tra*). Cuando se utiliza la supresión ámbar, la cepa de *E. coli* no contendrá un gen del ARNt supresor ámbar activo. El aminoácido se añadirá a los medios de crecimiento, preferentemente a una concentración final de 1 mM o mayor. La eficacia de la incorporación de aminoácidos se puede realzar mediante el uso de una construcción de expresión con un sitio ortogonal de unión del ribosoma y traduciendo el gen con ribo-X («Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion». Wang K, Neumann H, Peak-Chew S. Y., Chin J. W. *Nat. Biotechnol.* julio de 2007; 25(7): 770-7). Esto puede permitir que el aminoácido no natural que proporciona numerosos sitios de la unión del ligando se incorpore con eficacia en varios sitios.

55 PURIFICACIÓN DE FAGOS

Se puede utilizar cualquier medio adecuado para la purificación del fago. Se pueden aplicar técnicas estándares en la presente invención. Por ejemplo, se pueden purificar fagos mediante la filtración o mediante la precipitación, tal como la precipitación con PEG; se pueden sintetizar partículas de fagos y purificarlas mediante la precipitación con

polietilenglicol (PEG) como se describió previamente.

En caso de que se necesite más ayuda, se hace referencia a Jespers et al. (*Protein Engineering Design and Selection*, 2004, 17(10): 709-713. «Selection of optical biosensors from chemisynthetic antibody libraries»). Convenientemente, se pueden purificar los fagos como se da a conocer en dicha memoria. El texto de esta publicación está específicamente incorporado en la presente memoria por referencia para el método de purificación de fagos; en particular, se hace referencia al apartado de los materiales y métodos que comienza a partir de la columna derecha de la página 709 de Jespers et al.

Además, se puede purificar el fago como publicaron Marks et al. *J. Mol. Biol.*, vol. 222, pág. 581-597, que se incorpora específicamente en la presente memoria por referencia, para la descripción particular de cómo se realiza la purificación/producción de fagos.

En el caso de que se necesite más ayuda, se puede reducir el fago y purificarlo como sigue. Aproximadamente, 5×10^{12} partículas de fagos se hacen reaccionar con ditiotreitól (DTT) a 1 mM durante 30 minutos a temperatura ambiente, luego se precipitan con PEG. Después de enjuagar con agua, se resuspende el sedimento en 1 ml de tampón de reacción (tampón de fosfato a 10 mM, EDTA a 1 mM, pH 7,8). A continuación, los fagos se hacen reaccionar opcionalmente con 50 μ l de N-[(2-yodoacetoxi)etil]-N-metilamino-7-nitrobenz-2-oxa-1,3-diazol (NBDIA) a 1,6 mM (Molecular Probes) durante 2 horas a temperatura ambiente, o más convenientemente se hace reaccionar con el compuesto conector como se describe en la presente memoria. La reacción se termina mediante la precipitación con PEG de las partículas de fagos.

Otro modo más para producir/purificar el fago es como dan a conocer Schreier y Cortese («A fast and simple method for sequencing DNA cloned in the single-stranded bacteriophage M13». *Journal of Molecular Biology*, 129 (1): 169-72, 1979). Esta publicación trata del procedimiento de secuenciación de ADN de terminación de la cadena de Sanger et al. (1977) que requiere ADN monocatenario como plantilla. El ADN del fago M13 existe como una única hebra y, por consiguiente, cada secuencia de ADN clonada en M13 se puede obtener fácilmente de este modo. Schreier y Cortese muestran que, mediante la simple centrifugación de las partículas de fagos y la extracción con fenol, se puede preparar el ADN monocatenario de M13 suficientemente puro como para utilizarlo como plantilla para determinar la secuencia. La publicación de Schreier y Cortese se incorpora específicamente en la presente memoria mediante referencia para el método de purificación del fago. Para despejar cualquier duda, la extracción con fenol no se realiza para formar complejos de acuerdo con la presente invención, ya que es para el propósito de la purificación del ácido nucleico. Por lo tanto, la etapa de fenol de Schreier y Cortese se omite convenientemente. El método de Schreier y Cortese se sigue sólo hasta el punto de partículas de fagos purificadas.

Así pues, hay miles de técnicas bien conocidas en la técnica para la purificación de fagos. En el contexto de la presente invención, tal purificación se utiliza para retirar el agente reductor utilizado para la reducción de los grupos funcionales del polipéptido de interés para que formen enlace con el compuesto conector, particularmente cuando tal fijación es a través de los restos de cisteína.

Opcionalmente, las técnicas especialmente ventajosas para la purificación de fagos se pueden adoptar como se trató en el apartado de la química de la reacción que viene a continuación. Se debe advertir expresamente que estas técnicas no se consideran esenciales para la invención, pero pueden representar métodos especialmente útiles o incluso el mejor modo de fabricar las partículas de fagos de la invención. Sin embargo, hay que prestar especial atención para evitar la reoxidación de los grupos reactivos/funcionales reducidos del fago en la etapa de la retirada del reductor antes de pegar el compuesto conector, y después, en principio, se puede utilizar cualquier técnica para llevar a cabo esto. Las técnicas de filtración descritas son particularmente eficaces, pero también más complicadas que las técnicas estándares por lo que el operador elegirá la técnica mejor adaptada a su utilización particular de la invención. Lo más convenientemente, se emplea la técnica de filtración.

45 QUÍMICA DE LA REACCIÓN

Además de dar a conocer más conocimientos en conexión con los conjugados polipéptido-compuesto conector fijados por tres enlaces y las partículas de fagos de la invención, los inventores también han obtenido un conjunto preciso de condiciones químicas que se pueden desplegar para conseguir la conexión química al mismo tiempo que se mantiene la integridad de la porción genéticamente codificada del producto. Las tecnologías de la técnica anterior para la modificación de los polipéptidos han implicado una química rigurosa y reacciones independientes de modificación del polipéptido. En cambio, la presente invención da a conocer nuevas condiciones químicas para la modificación de los polipéptidos al mismo tiempo que conserva ventajosamente el funcionamiento o la integridad del elemento codificado genéticamente del producto. Específicamente, cuando el elemento codificado genéticamente es un polipéptido expuesto sobre la superficie de un fago que lo codifica, la química ventajosamente

no compromete la integridad biológica del fago. Se describe en la presente memoria que hay una estrecha ventana de condiciones para las cuales estas reacciones químicas se pueden mejorar o facilitar. En particular, como se explicará con más detalle más adelante, los solventes y las temperaturas utilizadas son importantes para una reacción eficaz. Es más, la modificación de la concentración de los reactivos utilizados también puede favorecer la fijación correcta, mientras que mejora o elimina los entrecruzamientos o el daño de los restos polipeptídicos que se han de modificar.

En particular, se describe que se requiere la reducción de las cisteínas en el polipéptido diana para que la reacción sea más eficaz. Claramente, para realizar la unión deseada, se debe retirar el agente reductor con el que se reducen químicamente las cisteínas. Una técnica conocida es utilizar el ditioneitol (DTT) para la reducción de las cisteínas, y precipitar las partículas, tales como las partículas de fagos, en una reacción de precipitación para retirar el reductor. Tales reacciones de precipitación implican típicamente polietilenglicol (PEG) al 20% junto con NaCl a 2,5 M que conduce a la precipitación de las partículas de fagos. Sin embargo, los inventores describen que, en algunos experimentos, estas condiciones estándares específicas no condujeron a una reacción eficaz de los restos de cisteína del polipéptido con el compuesto conector, muy probablemente debido a la reoxidación de una proporción de restos cisteína que se habían reducido. Esto no se pudo predecir a partir de los conocimientos de la técnica anterior. Se debería advertir que esta técnica estándar todavía podría encontrar una aplicación en la invención, en particular cuando el experto en la técnica está alerta dado que se ha descrito la necesidad de estar vigilante para valorar/evitar la reoxidación. Sin embargo, los inventores han tratado este problema críptico de cómo retirar el agente reductor al mismo tiempo que se mantienen las cisteínas en su estado reducido. Como se describirá con más detalle a continuación, las soluciones para extraer las partículas en condiciones químicas favorables se encuentran en un abanico de estrategias que incluyen el uso de tris-carboxietil-fosfina, de tampón desgasificado, el uso de quelantes en la mezcla de reacción, y la filtración.

Se deben elegir cuidadosamente las condiciones de reacción, p. ej., para unir el compuesto conector al polipéptido diana. La elección de las condiciones puede variar según la aplicación para la cual se ha montado la invención. Se requiere un cuidado particular cuando el polipéptido diana está comprendido por una partícula de fago. Se proporciona ayuda a lo largo de la especificación y el apartado de ejemplos.

Las condiciones de reacción como la temperatura de reacción, la concentración del compuesto conector, el solvente y/o el pH, se deben elegir para permitir que los grupos funcionales del polipéptido diana reaccionen con eficacia con el compuesto conector, pero dejen el ácido nucleico que codifica el polipéptido en una condición que permita descodificar (p. ej., para secuenciar) y/o propagar las moléculas aisladas (p. ej., por PCR o mediante la propagación de fagos o por cualquier otra técnica adecuada). Además, las condiciones de reacción deben dejar la proteína de la cubierta del fago en una condición que le permita propagar el fago.

Los grupos tiol de un polipéptido codificado en fagos se puede reducir con un agente reductor antes de unir el compuesto conector. En tales realizaciones, en particular en las realizaciones de exposición en fagos, o en particular cuando el agente reductor es TCEP, el exceso de agente reductor se retira convenientemente por filtración, p. ej., la filtración del fago. Esto es especialmente ventajoso, ya que los presentes inventores describen por primera vez que las técnicas convencionales para retirar los reductores, tales como la precipitación con PEG/NaCl, puede a veces conducir a la reacción subóptima con el compuesto conector, probablemente debido a la reoxidación de los grupos laterales funcionales reducidos del polipéptido diana. Por lo tanto, se trata de una ventaja de las realizaciones en las que el polipéptido diana se prepara mediante reducción y posterior purificación (se retira el agente reductor) por filtración, con la que se consigue la mejor conservación de los grupos funcionales reducidos (y por lo tanto reactivos) del polipéptido.

En la presente invención se aplican las condiciones de reacción que, por una parte, permiten unir eficazmente el polipéptido codificado a un compuesto conector y, por otra parte, dejar el ácido nucleico con la adición (y las proteínas de la cubierta de los fagos) en una condición que permite su propagación o descodificación. Dichas condiciones de reacción son, por ejemplo, la temperatura de reacción, la concentración del compuesto conector, la composición del solventes o el pH.

En una realización de la presente invención, los grupos tiol de los restos cisteína se utilizan como grupos funcionales para conectar polipéptidos a un núcleo molecular. Para algunas reacciones químicas, hay que reducir los grupos tiol de los polipéptidos. Los grupos tiol de los polipéptidos expuestos en fagos se reducen eficazmente mediante la adición de un reductor como, por ejemplo, tris(carboxietil)fosfina (TCEP). Ya que un exceso de reductor puede interferir con la reacción de unión, se retira eficazmente mediante la filtración del fago.

La reoxidación de los grupos tiol después de retirar el TCEP se impide convenientemente con la desgasificación del tampón de reacción.

La reoxidación de los grupos tiol también se impide convenientemente mediante la formación de complejos de iones metálicos por quelación, por ejemplo, la quelación con el ácido etilendiaminotetraacético (EDTA).

Lo más convenientemente, la reoxidación de los grupos tiol se impide o inhibe mediante la quelación y el uso de tampones desgasificados.

- 5 En una realización de la presente invención, la unión del polipéptido al compuesto conector se lleva a cabo haciendo reaccionar los grupos reactivos del polipéptido, tales como los grupos tiol de un polipéptido codificado en fago, con el compuesto conector durante una hora.

Convenientemente, se hacen reaccionar a 30 °C.

- 10 Convenientemente, se hacen reaccionar con el compuesto conector (tal como tris(bromometil)benceno) a una concentración de 10 µM.

Convenientemente, la reacción es en tampón acuoso.

Convenientemente, la reacción es a pH 8.

Convenientemente, el tampón de reacción contiene acetonitrilo. Convenientemente el tampón de reacción contiene acetonitrilo al 20%.

- 15 Lo más convenientemente, la reacción consta de dos o más de las condiciones anteriores. Convenientemente, la reacción consta de tres o más de las condiciones anteriores. Convenientemente, la reacción consta de cuatro o más de las condiciones anteriores. Convenientemente, la reacción consta de cinco o más de las condiciones anteriores. Convenientemente, la reacción consta de seis o más de las condiciones anteriores. Convenientemente la reacción consta de cada una de las condiciones anteriores.

- 20 Estas condiciones de reacción están optimizadas para que los grupos tiol de un polipéptido reaccionen cuantitativamente con los grupos reactivos del tris(bromometil)benceno. En las mismas condiciones de reacción, aproximadamente el 20% de las partículas de fago permanecen infectantes para llevar el código genético al interior de las células bacterianas para la propagación y descodificación.

- 25 En una realización, el compuesto conector, tal como TBMB, se puede unir al polipéptido diana, tal como un polipéptido codificado en fago, mediante la reacción (incubación) de los grupos tiol del polipéptido durante una hora a 30 °C con TBMB (a saber, tris(bromometil)benceno) a una concentración de 10 µM en tampón acuoso a pH 8 que contiene acetonitrilo al 20%.

- 30 Los inventores también describen el efecto de la concentración del compuesto conector en la reacción sobre la infectividad de los fagos. En particular, la invención convenientemente disminuye al mínimo la concentración del compuesto conector utilizado en la reacción. En otras palabras, será mejor cuanto menor sea la concentración del compuesto conector utilizado en el tiempo de reacción con el polipéptido del fago, siempre y cuando se pegue al polipéptido de los fagos una cantidad suficiente del compuesto conector. Las ventajas de disminuir al mínimo el compuesto conector presente de este modo es que se conserva mejor la infectividad de los fagos después de la conjugación del compuesto conector. Por ejemplo, cuando el compuesto conector es TBMB, cualquier concentración del compuesto conector mayor de 100 µM puede comprometer la infectividad. Por lo tanto, convenientemente, cuando el compuesto conector es el TBMB, entonces, convenientemente, la concentración del TBMB presente durante el tiempo de formación de enlace con el polipéptido es de menos de 100 µM. Lo más convenientemente, la concentración es como se describe en el apartado de los ejemplos.

MODIFICACIÓN DESPUÉS DE LA UNIÓN

- 40 En algunas realizaciones, el complejo entre el compuesto conector y el polipéptido se puede modificar en cuanto se termina la unión.

En algunas realizaciones, los elementos del polipéptido de la invención se escinden proteolíticamente una vez que se sujetan a un núcleo molecular/compuesto conector. La escisión genera ligandos que tienen fragmentos peptídicos independientes inmovilizados en un compuesto conector/núcleo molecular.

- 45 Por ejemplo, uno o varios de los enlaces amida del polipéptido se pueden escindir proteolíticamente después de la inmovilización del polipéptido al núcleo molecular. Esto presenta la ventaja de crear polipéptidos cortos, cada uno unido al compuesto conector mediante al menos un enlace covalente, pero que puede presentar diferentes estructuras moleculares que se conservan en un complejo que comprende el ácido nucleico que codifica el polipéptido progenitor. La escisión de los polipéptidos está catalizada por cualquier medio adecuado conocido en la

técnica, tal como una hidrólisis controlada o, más convenientemente, mediante escisión enzimática por una proteasa adecuada. La proteasa puede ser cualquier proteasa adecuada, pero es preferentemente una proteasa con una secuencia o motivo de reconocimiento específico del polipéptido. Esto conduce ventajosamente a la producción de productos de escisión polipeptídicos más definidos y/o más predecibles. De hecho, en esta
 5 realización, las secuencias de reconocimiento de las proteasas se pueden añadir sistemáticamente del polipéptido diana o se pueden retirar de él, por ejemplo, mediante la manipulación del o de los ácidos nucleicos que lo codifican. Esto proporciona ventajosamente un mayor grado de control y permite una mayor diversidad para ser sintetizados en las moléculas expuestas de acuerdo con la presente invención. Lo más convenientemente, el polipéptido comprende al menos un sitio de reconocimiento de la proteasa. Convenientemente, cada dicho sitio de
 10 escisión está comprendido dentro de la secuencia o secuencias de aminoácidos entre grupos funcionales del polipéptido utilizado para la unión covalente al compuesto conector. Convenientemente, cada dicho sitio de reconocimiento está comprendido dentro de la secuencia o secuencias de aminoácidos entre grupos funcionales del polipéptido utilizado para la unión covalente al compuesto conector.

Los lazos polipeptídicos se escinden convenientemente con una proteasa que reconoce y procesa los polipéptidos en la posición de aminoácidos específicos, tales como la tripsina (arginina o lisina en posición P1) o la termolisina (cadenas laterales alifáticas en la posición P1). La enzima se utiliza a una concentración que permite procesar con eficacia los lazos peptídicos de la molécula expuesta, pero respeta las partículas de fagos. Las condiciones óptimas pueden variar según la longitud de los lazos de los polipéptidos y la proteasa utilizada. La tripsina, por ejemplo, se utiliza típicamente a 200 nM en tampón de TBS-Ca (Tris-HCl a 25 mM, NaCl a 137 mM y CaCl₂ a 1
 20 mM, pH 7,4) durante 10 minutos a 10 °C. En Kristensen, P, y Winter, G («Proteolytic selection for protein folding using filamentous bacteriophages». *Fold. Des.* 1998; 3(5): 321-8) se describe un amplio abanico de proteasas que son adecuadas para modificar los polipéptidos expuestos, pero que respetan el fago. El procesamiento enzimático del péptido en fagos puede ser una «proteólisis parcial» porque no se puede descartar que se escinda un pequeño número de proteínas de la cubierta de los fagos. Así pues, en la optimización de las condiciones se elige
 25 convenientemente el mejor equilibrio entre la máxima escisión posible de la diana y el máximo respeto de las partículas de los fagos.

Convenientemente, el polipéptido diana comprende al menos un sitio de escisión proteolítica. Convenientemente, el polipéptido diana comprende al menos dos sitios de escisión proteolítica. Convenientemente, el polipéptido diana comprende al menos tres sitios de escisión proteolítica.

30 En cada realización de proteólisis, convenientemente, el primero de tales sitios de proteasa está puesto distal con respecto al primer enlace covalente entre el polipéptido diana y el compuesto conector. Esto tiene la ventaja de que el compuesto conector se conserva en el complejo porque si el polipéptido diana se escinde antes del primer enlace covalente, entonces el complejo polipéptido-compuesto conector se separará del ácido nucleico que codifica el polipéptido diana, lo que es indeseable para la mayoría de aplicaciones de la invención.

35 El uso de lazos pequeños (pequeño significa, p. ej., 6 restos de aminoácidos o menos) puede comprender la capacidad de algunas proteasas para escindir dentro de los lazos. En este caso, puede ser deseable seleccionar lazos más largos que son probablemente más accesibles a la proteasa. Además, después de la escisión de los lazos mediante endoproteasa, puede ser deseable volver a cortar los lazos además con otras endoproteasas, o de hecho mediante exoproteasas, tales como carboxipeptidasas o aminopeptidasas.

40 Cuando el polipéptido diana comprende más de un sitio para tal proteasa, convenientemente cada uno de los sitios está puesto entre dos enlaces covalentes establecidos entre el polipéptido diana y el compuesto conector. Pueden ponerse varios sitios de escisión entre los enlaces si es necesario.

En las realizaciones de escisión, convenientemente, el polipéptido progenitor entero se considerará en la valoración de si está unido o no al compuesto conector mediante al menos tres enlaces covalentes. Más conveniente, el
 45 polipéptido diana se considerará que es el polipéptido intacto (sin escindir) cuando se evalúe si se une o no al compuesto conector mediante al menos tres enlaces covalentes. Tales polipéptidos sin escindir serán típicamente bicíclicos.

SÍNTESIS

Se debe advertir que una vez que el polipéptido de interés está aislado o identificado de acuerdo con la presente
 50 invención, entonces su síntesis posterior se puede simplificar siempre que sea posible. Por ejemplo, la secuencia del polipéptido de interés se puede determinar, y se puede fabricar sintéticamente, mediante técnicas estándares tras la reacción con un compuesto conector *in vitro*. Cuando se realiza, se puede utilizar la química estándar porque ya no hay ninguna necesidad de conservar la funcionalidad ni la integridad de la partícula portadora genéticamente codificada. Esto permite la preparación rápida a gran escala de material soluble para otros

experimentos posteriores o la validación. En este sentido, la preparación a gran escala de los candidatos o protagonistas identificados por los métodos de la presente invención podría llevarse a cabo mediante la química convencional, tal como la descrita en Meloen y Timberman.

- 5 Así pues, la invención también se refiere a la fabricación de polipéptidos o conjugados seleccionados como se ha propuesto en la presente memoria, en donde la fabricación comprende otras etapas funcionales como se explica a continuación. Lo más convenientemente, estas etapas se llevan a cabo sobre el conjugado/polipéptido del producto final fabricado mediante síntesis química, en vez de en el fago.

Opcionalmente, los restos de aminoácidos en el polipéptido de interés pueden contener sustituciones cuando se fabrica un conjugado o un complejo, p. ej., después de la etapa de aislamiento/identificación inicial.

- 10 Para ilustrar las modificaciones o adiciones que se describen, es útil considerar el ejemplo de selección de un polipéptido que reacciona con un receptor. Puede ser deseable extender el péptido en su extremo amino o en su extremo carboxilo. Esto puede ser útil, por ejemplo, al fabricar un péptido macrocíclico que se une a una diana, con una cola tal como una cola lineal que se une a una diana, por ejemplo, un péptido que penetra en una célula tal como los derivados de tal como VP22, HIV-Tat, una proteína homeótica de *Drosophila* (Antennapedia) o
- 15 proteínas diseñadas químicamente tal como la poliarginina u otro péptido, p. ej., como se describe en (Chen y Harrison, *Biochemical Society Transactions* (2007), volumen 35, parte 4, pág. 821, «Cell-penetrating peptides in drug development: enabling intracellular targets»). Esto tendría la ventaja de asistir o permitir que un macrociclo que se había seleccionado contra una diana determinada, tal como una diana intracelular, entre en una célula.

- Para extender el péptido se puede simplemente extender químicamente por su extremo amino o por su extremo carboxilo con el uso de química estándar de fase sólida o de fase líquida. La química estándar de proteínas se puede utilizar para introducir un extremo amino o carboxilo activable. Alternativamente, se pueden realizar adiciones de fragmentos mediante condensación o ligación química nativa, p. ej., como se describe en (Dawson P. E., Muir T. W., Clark-Lewis I., Kent, S. B. H. 1994 «Synthesis of proteins by native chemical ligation», *Science*, 266: 776-779), o mediante enzimas, por ejemplo, utilizando la subtiligasa como se describe en («Subtiligase: a tool
- 25 for semisynthesis of proteins», Chang T. K., Jackson D. Y., Burnier J. P., Wells J. A., *Proc. Natl. Acad. Sci. USA*, 20 de diciembre de 1994; 91 (26): 12544-8 o en *Biorganic & Medicinal Chemistry Letters*, volumen 18, número 22, 15 de noviembre de 2008, páginas 6000-6003, «Tags for labeling protein N-termini with subtiligase for proteomics», Hikari A. I., Yoshihara, Sami Mahrus y James A. Wells).

- Alternativamente, los péptidos se pueden extender o modificar mediante otra conjugación a través de enlaces disulfuro. Tiene la ventaja adicional de permitir que el primer y el segundo péptido se disocien el uno del otro dentro del entorno reductor de la célula. En este caso, el compuesto conector (p. ej., TBMB) se podría añadir durante la síntesis química del primer péptido para que reaccione con los tres grupos cisteína; entonces se podría añadir otra cisteína al extremo amino del primer péptido, de manera que esta cisteína sólo reaccionó con una cisteína libre del segundo péptido.
- 30

- 35 Se aplican técnicas similares del mismo modo a la síntesis o conjugación de dos macrociclos bicíclicos.

Además, se puede llevar a término la adición de otros fármacos del mismo modo, utilizando la química apropiada, conjugándolos al extremo amino o carboxilo, o a través de las cadenas laterales. Convenientemente, la conjugación se realiza de tal manera que no bloquea la actividad de ninguna entidad.

- 40 Así pues, la invención se refiere además a un método como el descrito más arriba que comprende adicionalmente la etapa de extender el polipéptido en uno o más de los extremos amino o carboxilo del polipéptido.

Así pues, la invención además se refiere a un método como el descrito más arriba que comprende adicionalmente la etapa de conjugar a otro polipéptido dicho complejo o dicho conjugado entre polipéptido y compuesto conector.

Así pues, la invención además se refiere a un método como el descrito más arriba en el que dicha conjugación se realiza mediante

- 45 (i) la adición de otra cisteína al polipéptido después de unirse covalentemente al compuesto conector, y
- (ii) la conjugación de dicho polipéptido a dicho otro polipéptido mediante la formación de enlace disulfuro con dicha cisteína adicional.

DIVERSIDAD CODIFICADA GENÉTICAMENTE

- Los polipéptidos de interés están codificados genéticamente de una manera conveniente. Esto ofrece la ventaja de
- 50 favorecer la diversidad junto con la facilidad de manipulación. Convenientemente, los polipéptidos de interés están

codificados genéticamente como una genoteca de exposición en fagos.

Por lo tanto, el complejo de la invención comprende una partícula de fago. El ácido nucleico está comprendido en el genoma del fago. En estas realizaciones, el polipéptido está comprendido en la cubierta del fago.

- 5 En algunas realizaciones, la invención se puede utilizar para producir una genoteca combinatoria codificada genéticamente de polipéptidos que se generan traduciendo una serie de ácidos nucleicos en los correspondientes polipéptidos y formando un enlace entre las moléculas de dicho compuesto conector y dichos polipéptidos.

La genoteca combinatoria codificada genéticamente de los polipéptidos se puede generar mediante la exposición en fagos.

- 10 Los polipéptidos se exponen en fagos de acuerdo con técnicas establecidas, tales como las que se describen más adelante. Lo más convenientemente, tal exposición se lleva a cabo mediante la fusión del polipéptido diana de interés con un gen modificado genéticamente que permite la exposición externa del polipéptido de interés; convenientemente, dicho gen modificado genéticamente comprende la modificación genética de un gen 9 (p9 o gen IX), gen 8 (gen VIII), gen 7 (p7 o gen VII), gen 6 (p6 o gen VI) o gen 3 (p3 o gen III) del fago. Estas proteínas ofrecen la ventaja de que contienen pocas o ninguna cisteína que pueda reaccionar con los compuestos
- 15 conectores, tal como TBMB, y producir productos secundarios. Para p6, es ventajoso mutar la cisteína 84 en serina. Las cisteínas de p7 y p9 están muy probablemente ocultas y, por lo tanto, pueden no ser necesario mutarlas para retirarlas. p8 ofrece la ventaja de que no contiene ningún resto cisteína. Así pues, más convenientemente, dicho gen modificado genéticamente comprende la modificación genética de un gen 8 (gen VIII), gen 6 (gen VI) o gen 3 (gen III) del fago.
- 20 Lo más convenientemente, tal exposición se lleva a cabo mediante la fusión del polipéptido diana de interés a una proteína del gen 3 modificada genéticamente para que carezca de restos de cisteína en los dominios 1 y 2. Esta fusión se puede llevar a cabo mediante cualquier método adecuado conocido en la técnica, tal como mediante la manipulación del ácido nucleico que codifica la proteína del gen III del fago para cambiar los codones que codifican la cisteína por codones que codifican otros aminoácidos, y mediante la inserción en fase de una
- 25 secuencia de ácido nucleico que codifica el polipéptido diana en la secuencia codificante del gen III, por lo que se expone como una proteína de fusión con la del gen III en el exterior de la partícula de fago.

- Es un beneficio de la invención que el gen o los genes modificados genéticamente resultantes permitan que el fago sea infectante, a saber, capaz de infectar y propagarse. Incluso cuando el gen modificado genéticamente es un gen diferente al gen 3, (tal como el gen 6 o el gen 8), puede ser deseable aún modificar genéticamente el gen 3 para
- 30 retirar uno o más de los restos de cisteína (tal como todos los restos de cisteína).

Mediante exposición en fagos se han generado repertorios de polipéptidos codificados que comprenden hasta 10^{13} miembros diferentes. El número de ligandos diferentes que se puede generar de acuerdo con la invención supera claramente el número de moléculas diferentes que se ensayan generalmente en los escrutinios convencionales.

- 35 Cuando los fagos se producen en una célula bacteriana, el polipéptido se expresa como una fusión de la proteína de la cubierta. Al ensamblarse una partícula de fago, el polipéptido queda expuesto en la superficie del fago. Al poner en contacto un repertorio de fagos con un antígeno inmovilizado, algunos fagos permanecen fijados al antígeno, mientras que otros se eliminan con los lavados. El fago se puede eluir y propagar. Se puede secuenciar el ADN que codifica el polipéptido del fago seleccionado. La exposición en fagos se puede utilizar para codificar más de 10^{10} polipéptidos diferentes. Un aspecto favorable de la exposición en fagos es que el código genético, un ADN
- 40 monocatenario, está empaquetado en una cubierta. La cubierta puede proteger el ADN de la reacción con el núcleo molecular.

- En otra realización preferente, la genoteca de polipéptidos de la invención se expone en fagos como una fusión con la proteína del gen 3. Cada partícula de fago tiene unas 3 a 5 copias de la proteína de la cubierta de dicho fago. Como resultado de la exposición de varias copias del polipéptido modificado, los ligandos con afinidades en el
- 45 orden micromolar (fijadores débiles) también se pueden aislar en selecciones de fagos. Alternativamente, los fagémidos se utilizan para reducir el número de polipéptidos por fago para evitar los efectos de aidez y seleccionar los ligandos con las afinidades más altas.

- En otra realización preferente, el fago con las proteínas modificadas de la cubierta se utiliza para codificar las genotecas de polipéptidos de la invención. En particular se utilizan los fagos que carecen o que tienen un número
- 50 reducido de un tipo específico de aminoácido en las proteínas de la cubierta. La utilización de dichas proteínas de la cubierta puede ser ventajoso cuando el núcleo molecular es capaz de reaccionar con dicho tipo específico de aminoácidos. Este es explícitamente el caso cuando los grupos funcionales del polipéptido expuesto para entrecruzar un núcleo molecular son aminoácidos naturales, y el mismo tipo de aminoácido natural está presente

en una región expuesta de la superficie en la cubierta de los fagos. La utilización de dichos fagos con proteínas modificadas de la cubierta puede impedir el entrecruzamiento de las partículas de fagos mediante la reacción de los aminoácidos de varios fagos con el mismo núcleo molecular. Además, la utilización de dicho fago es capaz de reducir el entrecruzamiento de ambas cadenas laterales de aminoácidos, de los grupos funcionales del polipéptido, y de una proteína de la cubierta del fago con el mismo núcleo molecular.

Aún en otra realización preferente, el fago con una proteína del gen 3 que carece de los restos cisteína de los puentes disulfuro C7-C36, C46-C53, C188-C201 del dominio 1 y el 2 se utilizan para exponer genotecas de polipéptidos de la invención. Schmidt F. X. y colaboradores (Kather, I et al., *J. Mol. Biol.*, 2005) generaron un fago con mutaciones en dichas posiciones (C7C, C36I, C46I, C53V, C188V, C201A) y 14 mutaciones adicionales en la proteína del gen 3 para compensar la reducción de la estabilidad térmica (T13I, N15G, R29W, N39K, G55A, T56I, 160V, T101I, Q129H, N138G, L198P, F199L, S207L, D209Y). Los fagos sin grupos tiol en dicha proteína minoritaria de la cubierta son adecuados si son restos de cisteína uno o más de los aminoácidos funcionales con los que se entrecruzan el polipéptido y un núcleo molecular. La retirada de los restos de cisteína en la proteína de la cubierta del fago impide que interfieran con dicha reacción de formación de enlaces entre el polipéptido y el compuesto conector.

Este fago de ejemplo para la aplicación en la invención se describe a continuación con más detalle.

El fago sin puentes disulfuro de F. X. Schmid (dominios D1-D2) comprende el fago fd procedente del vector fCKCBS (Krebber, C., 1997, *J. Mol. Biol.*). El vector fCKCBS se basa en un vector del fago fd que procede de la American Type Culture Collection (ATCC: 15669-B2).

- 20 La secuencia de aminoácidos de los dominios 1 y 2 de la p3 de fago fd de tipo silvestre es pública, por ejemplo, en la base de datos PubMed. Para facilidad de referencia, una secuencia de ejemplo es:

AETVESCLAKPHTENSFTNVWKDDKTLDRYANYEGCLWNATGVVVCTGDETQCYGTWVPIGLA
IPENEGGGSEG GGSEGGGSEGGGKTPPEYGDTPIPGYIYNPLDGTYPGTEQNPNPNPSLEES
QPLNTFMFQNNRFRNRQGALT VYTGVTQGTD PVKTYQYTPVSSKAMYDAYWNGKFRDCAF
HSGFNEDPFVCEYQGQSSDLPQPPVNAPSG

F. X. Schmid y colaboradores han estabilizado evolutivamente la p3 de este fago (Martin, A. y Schmid, F. X., 2003, *J. Mol. Biol.*) con la mutación de 4 aminoácidos. En el siguiente trabajo, F. X. Schmid y colaboradores habían mutado 6 cisteínas para eliminar los 3 puentes disulfuro e insertaron otras mutaciones para compensar la pérdida de estabilidad (Kather, I. y Schmid F. X., 2005, *J. Mol. Biol.*). En varios ciclos evolutivos habían generado los clones 19, 20, 21 y 23, que todos tienen una p3 sin puentes disulfuro con una estabilidad térmica diferente.

Como se explica con más detalle en el apartado de los ejemplos, el mutante 21 («clon 21») se puede fabricar como está descrito, o simplemente obtenerlo de F. X. Schmid y colaboradores. De acuerdo con la publicación de F. X. Schmid, este clon contiene las mutaciones siguientes en los dominios 1 y 2: C7S, T13I, N15G, R29W, C36I, N39K, C46I, C53V, G55A, T101I, Q129H, C188V, F199L, C201A, D209Y. Además encontramos las mutaciones siguientes en los dominios 1 y 2 cuando secuenciamos el clon y lo comparamos con el fago fd de tipo silvestre: P11S y P198L. Sin el deseo de comprometernos con la teoría, se supone que estas mutaciones ya estaban presentes en el fago del vector fCKCBS.

- 35 Los dominios D1 y D2 del clon 21 tienen la siguiente secuencia de aminoácidos:

AETVESSLAKSHIEGSFTNVWKDDKTLDWYANYEGILWKATGVVVITGDETQVYATWVPIGLA
IPENEGGGSEG GGSEGGGSEGGGKTPPEYGDTPIPGYIYNPLDGTYPGTEQNPNPNPSLEES
HPLNTFMFQNNRFRNRQGALT VYTGVTQGTD PVKTYQYTPVSSKAMYDAYWNGKFRDVAF
HSGFNEDLLVAEYQGQSSYLPQPPVNAPSG

Se puede generar una genoteca mediante los métodos de acuerdo con la invención.

La invención se puede aplicar al escrutinio de moléculas o entidades que se fijan a un complejo de la invención (o que influyen en su fijación). Un ejemplo de un complejo de la invención es un polipéptido diana con un compuesto conector unido a él. El experto en la técnica podrá adoptar cualquier formato de escrutinio convencional. El formato concreto utilizado dependerá de los objetivos del operador. Por ejemplo, si se desea un escrutinio de alto rendimiento, entonces será primordial un funcionamiento fácil y un recambio rápido. Típicamente, las técnicas tales como las rondas de criba de fagos, la exposición de ARNm y similares se pueden aplicar en la presente invención igual a como se aplican en la técnica. Los beneficios clave de la invención son la formación de tres

enlaces covalentes del compuesto conector con el polipéptido de interés y el formato determinado en el que se realiza el escrutinio de los complejos resultantes (o el uso de estos complejos como moduladores candidatos de otras interacciones o en otros escrutinios) es una materia de elección para la persona que pone en funcionamiento la invención.

- 5 En una realización se puede realizar el escrutinio al poner en contacto una genoteca generada mediante los métodos de acuerdo con la invención con un ligando diana y con el aislamiento de uno o más miembros de la genoteca que se fijan a dicho ligando.

En otra realización, cada uno de los miembros de dicha genoteca se ponen en contacto con un ligando diana en un escrutinio y se identifican miembros de dicha genoteca que se fijan a dicho ligando diana.

- 10 En otra realización, los miembros de dicha genoteca se ponen en contacto simultáneamente con un ligando diana y se seleccionan los miembros de dicha genoteca que se fijan a dicho ligando diana.

El o los ligandos diana pueden ser un péptido, una proteína, un polisacárido, un lípido, un ADN o un ARN.

El ligando diana puede ser un receptor, un ligando del receptor, una enzima, una hormona o una citocina.

- 15 El ligando diana puede ser una proteína procariótica, una proteína eucariótica o una proteína de arqueas. Más específicamente, el ligando diana puede ser una proteína de mamífero o una proteína de insectos o una proteína bacteriana o una proteína fúngica o una proteína vírica.

El ligando diana puede ser una enzima, tal como una proteasa. Más específicamente, el ligando diana puede ser una elastasa, la calicreina plasmática, la cathepsina G o el activador del plasminógeno de tipo urocinasa.

- 20 Se debe observar que la invención también engloba a miembro(s) de genotecas aislado(s) en un escrutinio de acuerdo con la invención. Convenientemente, el o los métodos de escrutinio de la invención comprenden adicionalmente la etapa de: fabricar una cantidad de ligando aislado que es capaz de fijarse al complejo de la invención. Cuando se realiza el escrutinio en el formato opuesto (esto es, cuando el o los complejos de la invención se identifican en virtud de su capacidad para fijarse a un ligando proporcionado), convenientemente, el o los métodos de escrutinio de la invención comprenden además la etapa de: fabricar una cantidad del complejo de la invención aislado que es capaz de unirse a dicho ligando.

- 25 La invención también se refiere a miembros de genotecas que son, o que son capaces de ser, aislados mediante un escrutinio de acuerdo con la presente invención, en donde dicho miembro se genera o fabrica posteriormente sin el uso adicional del ácido nucleico que codificó dicho polipéptido cuando era parte del complejo de la invención. Por ejemplo, los métodos adecuados de la invención comprenden además otra etapa de fabricación de una cantidad de un polipéptido aislado o identificado mediante un método de la invención por unión del compuesto conector al polipéptido, en donde dicho polipéptido se expresa recombinantemente o se sintetiza químicamente. Por ejemplo, cuando el polipéptido se sintetiza recombinantemente en esta realización, el ácido nucleico que lo codifica originalmente como parte de un complejo de la invención puede que ya no esté directamente presente, pero puede haber estado presente en una etapa intermedia, p. ej., de amplificación por PCR o de clonación del ácido nucleico original del complejo, lo que conduce a la producción de una plantilla de ácido nucleico a partir de la cual se puede sintetizar el polipéptido en esta etapa adicional.

MÁS VENTAJAS

- 30 Es una ventaja de la invención que los propios complejos sean capaces de propagarse. Por lo tanto, los complejos o genotecas de la invención pueden -cultivarse-seleccionarse-(iterativamente si se desea)-enriquecerse. Esto contrasta con los métodos de la técnica anterior que necesitan ser desconvolucionados después de una única ronda de selección.

La presente invención permite ventajosamente la construcción y el escrutinio de genotecas muy grandes.

- 40 Convenientemente, el compuesto conector puede ser flexible o rígido, más convenientemente el compuesto conector es rígido. Esto presenta la ventaja de un constreñimiento molecular mayor sobre la molécula del producto.

- 45 En algunas realizaciones, el compuesto conector no sólo constriñe la molécula al sujetarla por tres o más enlaces, sino también por actuar como un armazón. Los aminoácidos del péptido son capaces de interactuar con el armazón y formar una estructura compacta. Este fenómeno también se puede encontrar en anticuerpos, en donde los aminoácidos de los CDR interactúan con los aminoácidos del armazón. Así pues, la invención da a conocer esta peculiaridad ventajosa por primera vez en los polipéptidos conjugados, tales como los comprendidos en las

partículas de fagos.

En algunas realizaciones se utilizan los compuestos conectores con una geometría simétrica. Esto presenta la ventaja de que se produce un solo producto en vez de mezclas de productos (p. ej, isómeros).

- 5 Las reacciones sintéticas se han establecido para conectar un compuesto conector con un péptido mediante al menos tres enlaces covalentes. Las condiciones de reacción química de la técnica anterior no pueden aplicarse fácilmente a los péptidos codificados genéticamente. Describimos un conjunto de condiciones específicas que encuentran aplicación en la conjugación, toda vez que se conserva ventajosamente la infectividad.

Es una ventaja de la invención que se obtenga una lectura directa, en particular para una combinación de péptido + entidad química (a saber, péptido + compuesto conector).

- 10 Es una ventaja de la invención la creación de una genoteca de entidades químicas sintéticas que se pueda propagar. En otras palabras, los métodos de la técnica anterior han creado genotecas de entidades químicas en modos en los que no es posible amplificar ni leer la molécula pequeña de interés de esta manera, a saber, incluso aunque los ácidos nucleicos hayan estado presentes en las genotecas de entidades químicas de la técnica anterior, no se ha podido hacer crecer ni propagar, sino que sólo se ha permitido la hibridación u otras técnicas.
- 15 Se deducen ventajas similares de las técnicas descritas en la presente memoria tales como las condiciones químicas utilizadas para juntar el compuesto conector al polipéptido a la vez que se conserva ventajosamente la infectividad del complejo cuando el complejo comprende una partícula de fago.

MÁS APLICACIONES

- 20 La invención también da a conocer un método para generar una genoteca combinatoria de entidades químicas codificadas genéticamente que comprende polipéptidos sujetos a un núcleo molecular, comprendiendo el método: (a) la generación de una genoteca codificada genéticamente de polipéptidos que comprenden grupos funcionales capaces de formar enlaces covalentes con un núcleo molecular; (b) la conexión química de dicha genoteca con dicho núcleo mediante al menos tres enlaces covalentes.

- 25 En un amplio aspecto, la invención se refiere a un polipéptido que comprende un compuesto conector unido a dicho polipéptido, en donde el compuesto conector está unido al polipéptido mediante al menos tres enlaces covalentes distintos. En particular, la invención se refiere a tales polipéptidos que se obtienen, o son obtenibles, mediante los métodos de la presente invención.

La invención también se puede aplicar al diseño y/o la selección de imitadores de péptidos o imitadores de moléculas pequeñas para uso como fármacos o como dianas de fármacos.

- 30 La invención también da a conocer métodos para generar genotecas combinatorias de entidades químicas codificadas genéticamente y para aislar ligandos de las mismas.

- La invención se puede aplicar a la identificación de dianas relevantes a partir de secuenciación del ADN e identificación de secuencias consenso en los péptidos de las dianas recuperadas y luego sintetizar los péptidos. Por ejemplo, con este análisis se puede diseñar un péptido consenso, pudiendo tener dicho péptido consenso una
- 35 secuencia de aminoácidos no necesariamente idéntica a ninguna de las dianas recuperadas en la fase de escrutinio, y este péptido consenso se puede entonces sintetizar de acuerdo con la presente invención.

El complejo comprende una partícula de fago.

- Así pues, se da a conocer un método para generar genotecas combinatorias de entidades químicas codificadas genéticamente, en donde dichas genotecas comprenden polipéptidos sujetos a un núcleo molecular mediante al
- 40 menos tres enlaces covalentes. También se dan a conocer las genotecas generadas con dicho método. Además, se da a conocer un método para poner en contacto dichas genotecas con un ligando diana y aislar los miembros que se fijan a dicho ligando, pues son miembros de genotecas generadas con dicho método.

- En marcado contraste con los métodos conocidos de las patentes internacionales WO 2004/077062 y WO 2006/078161, la presente invención da a conocer métodos para generar y ensayar genotecas grandes de
- 45 complejos. De acuerdo con las patentes internacionales WO 2004/077062 y WO 2004/077062, se dan a conocer métodos reputados para proporcionar y realizar el escrutinio de cientos o miles de compuestos. La presente invención da a conocer métodos para codificar genéticamente genotecas compuestas. Esto permite generar y ensayar millones, miles de millones, o más, compuestos diferentes.

A diferencia de los métodos conocidos de las patentes internacionales WO 2004/077062 y WO 2006/078161, la

presente invención da a conocer métodos para ensayar grandes genotecas compuestas en un único compartimento de reacción utilizando principios de selección *in vitro*. A diferencia de los compuestos generados de acuerdo con las patentes internacionales WO 2004/077062 y WO 2006/078161, los complejos de la presente invención comprenden un ácido nucleico que permite la identificación de los complejos aislados; convenientemente, dicho ácido nucleico codifica el polipéptido del complejo.

La presente invención da a conocer condiciones de reacción tales como la concentración del compuesto conector, el tiempo de reacción, la temperatura de reacción, y similares, que respetan el ácido nucleico del complejo como, por ejemplo, una partícula de fago y, en particular, respetan la infectividad de la partícula de fago. En otras palabras, la química presentada en la presente memoria conserva la función del ácido nucleico del complejo y conserva la función biológica del complejo. En el ejemplo del complejo que comprende una partícula de fago, la química presentada en la presente memoria refuerza ventajosamente la conservación de la funcionalidad de la partícula de fago, y hace posible, o más posible, que se pueda utilizar en la propagación del ácido nucleico después de la formación del complejo.

La presente invención comprende también genotecas combinatorias de compuestos codificadas genéticamente que se generan con los métodos descritos.

También se conciben polipéptidos tricíclicos unidos a un compuesto conector. Se pueden crear, por ejemplo, uniendo los extremos amino y carboxilo de un polipéptido bicíclico unido a un compuesto conector de acuerdo con la presente invención. De este modo, los extremos amino y carboxilo unidos crean un tercer lazo, haciendo un polipéptido tricíclico. Esta realización no se realiza convenientemente en el fago, sino que se realiza convenientemente en un conjugado de la invención entre el polipéptido y el compuesto conector. La unión de los extremos amino y carboxilo es un asunto cotidiano de la química de péptidos. En el caso de que se necesite alguna ayuda, se puede activar el extremo carboxilo y/o se pueden extender los extremos amino y carboxilo, por ejemplo, para añadir una cisteína a cada extremo, y luego unirlos mediante un puente disulfuro. Alternativamente, la unión se puede llevar a cabo mediante el uso de una región conectora incorporada en los extremos amino o carboxilo. Alternativamente, los extremos amino y carboxilo se pueden unir mediante un enlace peptídico convencional. Alternativamente, se puede emplear cualquier otro medio adecuado para unir los extremos amino y carboxilo, por ejemplo, la N-C-ciclación se podría realizar mediante técnicas estándares, por ejemplo, como se describe en Linde et al., *Peptide Science*, 90, 671-682 (2008). «Structure-activity relationship and metabolic stability studies of backbone cyclization and N-methylation of melanocortin peptides», o como en Hess et al. *J. Med. Chem.* 51, 1026-1034 (2008) «Backbone cyclic peptidomimetic melanocortin-4 receptor agonist as a novel orally administered drug lead for treating obesity». Una ventaja de tales moléculas tricíclicas es evitar la degradación proteolítica de los extremos libres, en particular mediante la acción exoproteásica. Otra ventaja de un polipéptido tricíclico de esta naturaleza es que el tercer lazo se puede utilizar para funciones aplicables de forma general, tales como la fijación de SAB, la entrada en las células o los efectos de transporte, el etiquetado o cualquier otro uso. Se observará que este tercer lazo típicamente no estará disponible para la selección (porque no se produce en el fago, sino sólo en el conjugado entre el polipéptido y el compuesto conector) y, por lo tanto, su uso para otras funciones biológicas aún deja ventajosamente ambos lazos 1 y 2 para la selección/creación de la especificidad. Así pues, la invención también se refiere a tales polipéptidos tricíclicos y a su fabricación y usos.

La presente invención da a conocer más métodos para poner en contacto las genotecas compuestas codificadas genéticamente con un ligando diana y para identificar los ligandos que se fijan a dicho ligando diana. Las genotecas compuestas codificadas genéticamente se ensayan mediante procedimientos de selección o de escrutinio.

En un procedimiento de escrutinio se ensaya cada miembro de la genoteca. Por ejemplo, varias copias de un miembro determinado de la genoteca se incuban, por ejemplo, con un ligando diana. El ligando diana se inmoviliza antes o después de ponerlo en contacto con los miembros de la genoteca y los miembros sin unir se retiran por lavado. Los ligandos unidos se detectan, por ejemplo, en un ensayo inmunoensayo de adsorción (ELISA). La secuencia de aminoácidos de los miembros de la genoteca que se fijan al ligando diana se determina mediante la secuenciación del código genético.

En un procedimiento de selección, varios miembros de la genoteca compuesta codificada se ponen en contacto con un ligando diana. El ligando diana se inmoviliza antes o después de ponerlo en contacto los miembros de la genoteca, y los miembros sin unir se retiran por lavado. Se secuencian el código genético de los ligandos unidos. Los ligandos seleccionados se propagan alternativamente para realizar más rondas de selección.

En una realización de la invención, las genotecas compuestas están codificadas por exposición en fagos y se realizan las selecciones mediante rondas de criba de fagos.

El ligando diana de la presente invención puede ser una proteína, un ADN, un ARN o un polisacárido. La proteína puede ser un receptor, una enzima, una hormona, una citocina o una proteína vírica. Un posible ligando diana proteico es una proteasa, en donde dicha proteasa puede ser elastasa, calicreína plasmática, catepsina G o activador del plasminógeno de tipo urocinasa.

- 5 La presente invención comprende también miembros de las genotecas combinatorias de entidades químicas codificadas que se aíslan con los métodos de la invención. Dichos miembros se pueden sintetizar con o sin el código genético adjunto. En una realización preferente, dichos miembros que carecen del ácido nucleico se utilizan como fármacos o precursores de fármacos.

- 10 Con un método de la presente invención se aislaron varios miembros de las genotecas combinatorias de entidades químicas codificadas que son capaces de fijarse al ligando diana. Dichos miembros están compuestos por un núcleo de mesitileno y por un polipéptido con la secuencia $AC(X)_6C(X)_6CG$, en donde el polipéptido está sujeto por los grupos metilo exocíclicos del núcleo a través de los restos de cisteína que forman tres enlaces tioéter, y en donde X se refiere a un aminoácido natural, A es alanina, C es cisteína y G es glicina. La porción peptídica de dichos miembros se puede expresar recombinantemente o se puede sintetizar químicamente.

- 15 Los inhibidores de la calicreína plasmática humana se pueden aislar con los métodos de la invención a partir de las genotecas combinatorias de entidades químicas codificadas de la invención. Dichos inhibidores tienen o bien las secuencias polipeptídicas

GCSDRFRNCPADEALCG, ACSDRFRNCPLWSGTCG,

ACSTERRYCPIEIFPCG, ACAPWRTACYEDLMWCG, ACGTGEGRVNWTPCG, o bien secuencias relacionadas, en donde

- 20 los grupos tiol de las cisteínas están unidos a los núcleos mesitileno.

Los inhibidores de la catepsina G humana se pueden aislar con los métodos de la invención a partir de las genotecas combinatorias de entidades químicas codificadas de la invención. Dichos inhibidores tienen o bien las secuencias polipeptídicas **ACEYGLWCGWDPPVCG, ACIFDLGFCHNDWWNCG, ACLRAQEDCVYDRGFCG**, o bien secuencias relacionadas, en donde los grupos tiol de las cisteínas están unidos a los núcleos mesitileno.

- 25 Los inhibidores del activador del plasminógeno humano de tipo urocinasa se pueden aislar con los métodos de la invención a partir de las genotecas combinatorias de entidades químicas codificadas de la invención. Dichos inhibidores tienen o bien las secuencias polipeptídicas **ACNSRFGSCQIDLLMCG, ACSRYEVDRCGRGSACG**, o bien secuencias relacionadas, en donde los grupos tiol de las cisteínas están unidos a los núcleos de mesitileno.

DIANAS BIOLÓGICAS

- 30 Es importante crear y ensayar tantas moléculas como sea posible, ya que la probabilidad de identificar un ligando con las propiedades deseadas aumenta cuando se comprueban más moléculas. De igual forma, en general, los ligandos con mayor afinidad se obtienen cuando se ensayan los mayores repertorios de moléculas.

- Típicamente, los investigadores evalúan las moléculas mediante metodologías de selección o de escrutinio. El escrutinio es un procedimiento mediante el cual a los compuestos se les ensaya por separado su capacidad para 35 modificar una diana. Los procedimientos de escrutinio son versátiles y permiten el ensayo de repertorios de moléculas que tienen estructuras muy distintas. No obstante, el escrutinio mediante ensayos por separado puede llevar mucho tiempo y el número de moléculas únicas a las que se puede comprobar la fijación a una diana específica no suele exceder por lo general 10^6 entidades químicas. Por el contrario, los métodos de selección suelen permitir el muestreo de un número mucho mayor de moléculas diferentes. Así pues, los métodos de 40 selección se usan más convenientemente en las aplicaciones de la invención. En los procedimientos de selección, las moléculas se ensayan en un único recipiente de reacción y las que tienen propiedades favorables (por ejemplo, fijación) se separan físicamente de las moléculas inactivas. Las estrategias de selección que están disponibles permiten generar y ensayar simultáneamente más de 10^{13} compuestos distintos. Los ejemplos de técnicas poderosas de selección por afinidad son la exposición en fagos, la exposición en ribosomas, la exposición de 45 ARNm, la exposición en levaduras, la exposición en bacterias y los métodos de aptámeros de ADN/ARN. Estos métodos de selección biológica *in vitro* tienen en común que los repertorios de ligandos están codificados por ADN o ARN. Permiten la propagación y la identificación por secuenciación de los ligandos seleccionados. La tecnología de exposición en fagos se ha usado por ejemplo para el aislamiento de anticuerpos con afinidad de unión muy alta por virtualmente cualquier diana.

50 APLICACIÓN INDUSTRIAL

La presente invención es aplicable al descubrimiento de moléculas que son útiles en los campos de la biología, la biotecnología y las ciencias farmacéuticas. En particular, la presente invención se refiere a los métodos para la generación de fármacos o precursores de fármacos.

La presente invención comprende métodos para la generación de genotecas combinatorias de entidades químicas codificadas genéticamente y métodos para el aislamiento de miembros de dichas genotecas. Además, la invención comprende genotecas generadas con dichos métodos y miembros de las genotecas aisladas con dichos métodos.

La presente invención da a conocer un método para la generación de genotecas combinatorias de entidades químicas codificadas genéticamente, en donde los miembros de dichas genotecas comprenden un núcleo molecular central y una gran diversidad de elementos que se añaden a dicho núcleo. Dichas genotecas codificadas genéticamente se generan en dos etapas: en la primera etapa se generan las genotecas de polipéptidos codificadas genéticamente que comprenden grupos funcionales capaces de formar enlaces covalentes con un núcleo molecular. En la segunda etapa, las genotecas de polipéptidos codificadas genéticamente se entrecruzan químicamente con dicho núcleo molecular al menos mediante tres enlaces covalentes.

Las moléculas de las genotecas combinatorias de entidades químicas codificadas genéticamente de la presente invención tienen una estructura central que se expande mediante diferentes añadidos. A diferencia del estado actual de las genotecas combinatorias de entidades químicas codificadas genéticamente que se generan por métodos biológicos, las genotecas de la presente invención proporcionan moléculas con arquitecturas ramificadas, no lineales. Las moléculas con tales estructuras ramificadas son adecuadas para la fijación a los ligandos diana, ya que son capaces de fijarse a la diana mediante la interacción de los diferentes añadidos que apuntan hacia afuera del núcleo central.

A diferencia de las estructuras poliméricas lineales, los complejos de la presente invención tienen menos flexibilidad conformacional. En solución, adoptan solo un conjunto limitado de conformaciones. Como consecuencia, la fijación de dichos complejos o polipéptidos a un ligando diana no está asociada a una pérdida espectacular de entropía y se pueden obtener afinidades de unión elevadas.

Los polipéptidos de los complejos/genotecas de la invención están codificados genéticamente. Esto permite que se puedan aplicar para su producción los poderosos métodos biológicos de codificación, como por ejemplo exposición en fagos, exposición en ribosomas, exposición en levaduras, exposición en bacterias o exposición de ARNm, lo que permite generar genotecas de ligandos que contienen millones, miles de millones, o más, de miembros diferentes.

Se puede variar la secuencia de los añadidos a los polipéptidos de los miembros de las genotecas combinatorias de entidades químicas codificadas genéticamente. Las excepciones son los aminoácidos que portan grupos funcionales para entrecruzar el polipéptido con un núcleo molecular, lo que se explica con más detalle en la presente memoria. Los añadidos a los polipéptidos pueden comprender una diversidad combinatoria muy grande. Es importante que los repertorios combinatorios tengan una gran representación, ya que la probabilidad de aislar fijadores de gran afinidad por los ligandos diana aumenta con el tamaño de la genoteca.

A diferencia de los biopolímeros lineales como polipéptidos y aptámeros ADN o ARN, los complejos y miembros de las genotecas combinatorias de entidades químicas codificadas genéticamente de la invención no forman estructuras ternarias complejas. Los complejos y miembros de las genotecas combinatorias de entidades químicas codificadas genéticamente de la invención disfrutan de un riesgo muy reducido de inactivación por culpa del desplegamiento irreversible. La formación de agregados debido al desplegamiento es, por tanto, improbable, ventajosamente.

Las genotecas codificadas genéticamente de la invención convenientemente comprenden al menos 10^5 miembros distintos. Preferentemente, dichas genotecas comprenden millones o miles de millones, o más, de miembros distintos. El tamaño de dichas genotecas está determinado por los métodos que se usan para conectar el ácido nucleico que codifica un polipéptido con dicho polipéptido. En una realización preferente de la invención, se usan los métodos biológicos para generar repertorios de polipéptidos codificados genéticamente. El número de miembros diferentes de polipéptidos que se conectan al código polinucleotídico codificante puede exceder 10^{13} , según los métodos usados.

La invención se describe ahora mediante ejemplos en los que se hace referencia a las siguientes figuras:

BREVE DESCRIPCIÓN DE LOS DIBUJOS

En la figura 1 se muestra la generación de genotecas combinatorias de entidades químicas pequeñas codificadas genéticamente. Un péptido codificado en fago con tres restos de cisteína se sujeta al compuesto trifuncional tris-(bromometil)benceno en una reacción de sustitución nucleófila. Las entidades químicas resultantes se pueden modificar adicionalmente mediante reacciones enzimáticas.

En la figura 2 se muestra la valoración de las condiciones de reacción para unir los péptidos expuestos en fagos

con el tris-(bromometil)benceno (TBMB). A) Masa molecular de la proteína de fusión **GCGSGCGSGCG-D1-D2** antes y después de la reacción con TBMB a 10 μ M en NH_4HCO_3 a 20 mM, EDTA a 5 mM, pH 8, y ACN al 20% a 30 °C durante 1 hora, determinada por espectrometría de masas. La diferencia de masa de la proteína de fusión con el péptido que reaccionó y que no reaccionó corresponde a la masa del mesitileno, el núcleo molecular pequeño. C) 5 Títulos (unidades de transducción) de los fagos reducidos y tratados con diferentes concentraciones de TBMB en NH_4HCO_3 a 20 mM, EDTA a 5 mM, pH 8, y ACN al 20% a 30 °C durante 1 hora. Se muestran los títulos de los fagos de fdg3p0ss21 (negro) y de la genoteca 1 (blanco).

En la figura 3 se muestra el diseño de la genoteca en fagos y las secuencias de los clones seleccionados. A) 10 Secuencia de aminoácidos de las proteínas de fusión con péptidos expresadas por los clones de la genoteca 1. El péptido líder se elimina al secretar la proteína gracias a una proteasa de *E. coli* que deja el péptido con una alanina aminoterminal, dos secuencias de 6 aminoácidos aleatorios flanqueadas por tres cisteínas y un conector de Gly-Gly-Ser-Gly que conecta el péptido con la proteína del gen 3. B y C) Secuencias de aminoácidos de los clones seleccionados con la calicreína plasmática (B) y la cathepsina G (C) humanas. Se indica la actividad inhibidora de 15 las proteínas de fusión péptido-D1-D2 modificadas con TBMB. Las similitudes de las secuencias se resaltan con colores.

En la figura 4 se muestra la maduración de la afinidad de los inhibidores de la calicreína plasmática humana. A) 20 Diseño de las genotecas 2, 3 y 4. En cada genoteca, uno de los lazos peptídicos tiene la secuencia de un motivo consenso identificado en las primeras selecciones y el otro contiene seis aminoácidos al azar. B) Secuencias de aminoácidos de los clones seleccionados con la calicreína plasmática humana. Todos los clones proceden de la genoteca 2. Se indica la actividad inhibidora de las proteínas de fusión péptido-D1-D2 modificadas con TBMB. Los 25 colores resaltan las similitudes de las secuencias del segundo lazo de fijación.

En la figura 5 se muestra la inhibición de la calicreína plasmática humana por los péptidos sintéticos modificados con TBMB. La actividad inhibidora se expresa como la fracción de actividad (índice de inhibición/índice de no inhibición) a diferentes concentraciones de inhibidor.

25 En la figura 6 se muestra una estructura en solución representativa por RMN del péptido PK15 modificado con TBMB mostrado como una estructura en «salchicha». Los lazos peptídicos se muestran en azul (lazo 1) y verde (lazo 2). Los átomos de carbono $\pm$ de los aminoácidos del lazo peptídico y del extremo se muestran como esferas.

En la figura 7 se muestra la reacción química del compuesto trifuncional TBMB con péptidos que contienen una o dos cisteínas. A) Posible mecanismo de reacción del TBMB con una proteína de fusión al péptido que contiene dos 30 restos de cisteína. B) Espectro de masas de una proteína de fusión al péptido con dos cisteínas antes y después de la reacción con el TBMB. C) Posible mecanismo de reacción del TBMB con una proteína de fusión al péptido que contiene un resto de cisteína. D) Espectro de masas de una proteína de fusión al péptido con una cisteína antes y después de la reacción con el TBMB.

En la figura 8 se muestra la inhibición de la activación por contacto en el plasma humano mediante la aprotinina y 35 el péptido PK15 modificado con TBMB. Efecto de la aprotinina (A) y del péptido PK15 modificado con TBMB (B) sobre la generación de trombina desencadenada por la actina. Ambos inhibidores ocasionan una prolongación del tiempo de latencia dependiente de la dosis en comparación con la muestra de control. (C) La suma de las actividades del factor Xlla y de la calicreína plasmática se midió con el sustrato colorimétrico H-D-Pro-Phe-Arg-pNA en el plasma humano de tres donantes diferentes tratados con diferentes concentraciones del inhibidor. Se 40 inició la activación por contacto añadiendo caolín. Se indican los valores medios y las desviaciones estándares.

En la figura 9 se muestra (a) la generación de genotecas combinatorias de entidades químicas pequeñas codificadas en fagos. Un péptido codificado por fago con tres restos de cisteína se sujetó al compuesto trifuncional 45 tris-(bromometil)benceno mediante una reacción de sustitución nucleófila. Las entidades químicas resultantes se podrían modificar adicionalmente a través de reacciones enzimáticas. (b) Estructura química de un inhibidor macrocíclico de la calicreína plasmática aislado mediante la exposición en fagos (PK15).

En la figura 10 muestra la valoración de las condiciones de reacción para unir péptidos expuestos en fagos al tris-(bromometil)benceno (TBMB). (a) La masa molecular de la proteína de fusión **GCGSGCGSGCG-D1-D2** antes y 50 después de la reacción con TBMB a 10 μ M en NH_4HCO_3 a 20 mM, EDTA a 5 mM, pH 8, y ACN al 20% a 30 °C durante 1 hora se determinó mediante espectrometría de masas. La diferencia de masa de la proteína de fusión al péptido que reaccionaron y sin reaccionar corresponde a la masa del núcleo del mesitileno, un núcleo molecular pequeño. (b) Las titulaciones (unidades de transducción) de los fagos que se redujeron y se trataron con diferentes concentraciones de TBMB en NH_4HCO_3 a 20 mM, EDTA a 5 mM, pH 8, y ACN al 20% a 30 °C durante 1 hora. Se muestran los títulos de fagos de fdg3p0ss21 (negro) y de la genoteca 1 (blanco).

- En la figura 11 se muestra el diseño de las genotecas en fagos y las secuencias de los clones seleccionados. (a) La secuencia de aminoácidos de proteínas de fusión con péptidos expresadas por clones de la genoteca 1. La secuencia líder se elimina durante la secreción de la proteína mediante una proteasa de *E. coli* que deja un péptido con una alanina en el extremo amino, dos secuencias aleatorias de 6 aminoácidos flanqueadas por tres cisteínas y un conector de Gly-Gly-Ser-Gly que conecta el péptido a la proteína del gen 3. (b y c) Secuencias de aminoácidos de los clones seleccionados con la calicreína plasmática (b) y la catepsina G (c) humanas. Se indican las actividades inhibitoras de las proteínas de fusión péptido-D1-D2 modificadas con TBMB. Se resaltan las similitudes de secuencia con colores.
- 5
- En la figura 12 se muestra la maduración de la afinidad de los inhibidores de la calicreína plasmática humana. (a) Diseño de las genotecas 2, 3 y 4. En cada genoteca, uno de los lazos peptídicos tiene la secuencia de un motivo consenso identificado en las primeras selecciones y el otro contiene seis aminoácidos aleatorios. (b) Secuencias de aminoácidos de los clones seleccionados con calicreína plasmática humana. Todos los clones proceden de la genoteca 2. Se indican las actividades inhibitoras de las proteínas de fusión péptido-D1-D2 modificadas con TBMB. Los colores resaltan las similitudes de secuencia en el segundo lazo de fijación.
- 10
- En la figura 13 se muestra la inhibición de la calicreína plasmática humana mediante péptidos sintéticos modificados con TBMB. La actividad inhibitora se expresa como la fracción de actividad (velocidad inhibida/velocidad no inhibida) en concentraciones diferentes de inhibidor. Los clones PK2, PK4, PK6 y PK13 se aislaron en selecciones de fagos con la genoteca 1. PK15 procede de la genoteca 2 y es un inhibidor de afinidad madurada.
- 15
- En la figura 14 se muestra la estructura en solución por NMR del péptido PK15 modificado con TBMB. Los lazos del péptido se muestran en amarillo (lazo 1) y naranja (lazo 2). El núcleo de mesitileno, los tres restos de cisteína y la alanina (extremo amino) y glicina (extremo carboxilo) terminales se muestran en gris. Los átomos del esqueleto del péptido se representan como salchichas y las cadenas laterales de los aminoácidos se muestran como varillas.
- 20
- En la figura 15 se muestra la reacción química del compuesto trifuncional TBMB con péptidos que contienen uno o dos restos de cisteína. (a) unidos de reacción del TBMB con una proteína de fusión a péptido que contiene dos restos de cisteína. (b) La espectrometría de masas de una proteína de fusión a péptido con dos cisteínas antes y después de la reacción con TBMB. (c) Posible mecanismo de reacción del TBMB con una proteína de fusión a péptido que contiene una cisteína y un resto de lisina. (d) La espectrometría de masas de una proteína de fusión a péptido con una cisteína y un resto de lisina antes y después de la reacción con TBMB.
- 25
- En la figura 16 se muestra la supresión de la activación del factor XII en el plasma humano a través de la inhibición de la calicreína plasmática con la aprotonina o con el péptido PK15 modificado con TBMB. La vía de coagulación intrínseca en el plasma humano de tres donantes diferentes se inició mediante la adición de caolín. La superficie cargada negativamente del caolín activa cantidades pequeñas del factor XII. La precalicreína se convierte en calicreína mediante el factor XII activado (XIIa), y la calicreína ejerce una retroalimentación positiva para activar más factor XII. La actividad del factor XIIa se midió con el sustrato colorimétrico H-D-Pro-Phe-Arg-pNA. Se indican los valores medios y las desviaciones estándares de la actividad del factor XIIa.
- 30
- 35

Los siguientes ejemplos pretenden tener una naturaleza ilustrativa y no pretenden limitar el alcance de las reivindicaciones anexas.

Ejemplos

40 Visión general

En estos ejemplos mostramos la fabricación de genotecas combinatorias de entidades químicas codificadas en fagos.

- El descubrimiento de las moléculas sintéticas con una elevada afinidad y especificidad por dianas biológicas es un problema central en el descubrimiento de fármacos. Mientras que recientemente ha sido posible aislar grandes estructuras moleculares como anticuerpos u aptámeros a virtualmente cualquier diana mediante técnicas de selección *in vitro*, la generación de aglutinantes orgánicos pequeños con gran afinidad seguía siendo un gran reto. En esta invención describimos una estrategia para aislar estructuras de moléculas pequeñas que están formadas por un núcleo de molécula orgánica (compuesto conector) que está decorado con restos peptídicos (p. ej., polipéptido(s)). Por comodidad en estos ejemplos, se utilizó la tecnología de exposición en fagos para codificar la fracción de péptido de las moléculas pequeñas que permiten generar y seleccionar repertorios combinatorios de entidades químicas muy grandes. Se eligieron las condiciones de reacción para sujetar selectivamente un péptido de 17 aminoácidos mediante tres enlaces tioéter a una molécula de benceno, pero que respeta las proteínas de la cubierta de las partículas de fagos. Se obtuvieron aglutinantes muy específicos con afinidades del orden
- 45
- 50

submicromolar contra las dos serina proteasas humanas calicreína plasmática y catepsina G. Un inhibidor de afinidad madurada por la calicreína plasmática humana con una K_i aparente de 1,5 nM suprimió con eficacia la activación por contacto en el plasma humano.

Antecedentes de los ejemplos

- 5 Se necesitan moléculas con gran afinidad y especificidad por dianas biológicas para desarrollar tratamientos eficientes y selectivos contra un amplio abanico de enfermedades. El procedimiento de encontrar una nueva molécula orgánica pequeña contra una diana elegida suele implicar el escrutinio de alto rendimiento, en donde se ensayan grandes genotecas de entidades químicas por su capacidad para modificar la diana. Sin embargo, el
- 10 procedimiento tarda mucho tiempo y es costoso, y el número de moléculas únicas que se pueden ensayar contra una diana específica por lo general no supera un millón de entidades químicas. Los escrutinios a menudo sólo proporcionan precursores, que luego requieren más mejora mediante métodos empíricos o mediante el diseño químico. Los métodos más poderosos para generar moléculas de fijación son las técnicas biológicas de selección *in vitro*, como exposición en fagos, exposición en ribosomas, exposición de ARNm o técnicas de aptámeros de ARN/ADN. Permiten la generación rápida de grandes repertorios combinatorios (10^9 - 10^{13}) de polipéptidos, ARN o
- 15 ADN, y el aislamiento posterior de aglutinantes de gran afinidad. Sin embargo, la restricción de tales métodos a estructuras biopoliméricas grandes como anticuerpos o aptámeros imposibilita su utilización para el descubrimiento de moléculas pequeñas.

- Para aplicar la selección *in vitro* a genotecas combinatorias de entidades químicas se han propuesto diferentes metodologías para asociar moléculas orgánicas con una etiqueta que especifica su estructura. La mayoría de las
- 20 estrategias propusieron el uso de etiquetas de ADN para identificar las moléculas orgánicas pequeñas después de la selección por afinidad. Se ha propuesto un procedimiento de síntesis combinatoria paralela para codificar cada uno de los miembros de una gran genoteca de entidades químicas con secuencias de nucleótidos únicas sobre perlas (Brenner, S., y Lerner, R. A., *PNAS*, 1992). Después de que se fije la entidad de química a la diana, el código genético se descodifica secuenciando la etiqueta nucleotídica. Se ha conjugado una colección pequeña de
 - 25 moléculas orgánicas a oligonucleótidos de ADN y se han realizado selecciones de afinidad con diferentes antígenos (Doyon, J. B et al., *JACS*, 2003). Neri D. y colaboradores habían generado repertorios grandes de parejas de moléculas mediante el autoensamblaje de subgenotecas más pequeñas de entidades químicas codificadas por ADN mediante la hibridación de dos hebras de ADN (Melkko, S et al., *Nature Biotechnol.*, 2004). La metodología se utilizó exitosamente para la maduración de la afinidad de ligandos de moléculas pequeñas. Halpin
 - 30 D. R. y Harris P. B. desarrollaron una estrategia para la evolución *in vitro* de genotecas combinatorias de entidades químicas que implica la amplificación de los compuestos seleccionados para realizar varias rondas de selección (Halpin, D. R y Harbury, P. B., *PLOS Biology*, 2004). Woiwode T. F. et al. unieron genotecas de compuestos sintéticos a las proteínas de la cubierta de partículas de bacteriófagos de tal forma que la identidad de la estructura química está codificada en el genoma del fago (Woiwode, T. F., *Chem & Biol.*, 2003). Todas estas estrategias que
 - 35 emplean entidades químicas codificadas por ADN que se ha demostrado que han resultado ser eficientes en experimentos con modelos y algunas incluso han producido nuevas moléculas pequeñas aglutinantes. Sin embargo, se hizo evidente que codificar genotecas grandes de compuestos y la amplificación de los compuestos seleccionados exige mucho más que los procedimientos equivalentes en los sistemas biológicos de selección.

- En esta invención damos a conocer una estrategia para codificar estructuras híbridas de moléculas pequeñas-péptidos que están formadas por varios fragmentos polipeptídicos sujetos a una molécula orgánica pequeña central. La porción peptídica está codificada por partículas de fago que permiten generar y seleccionar
- 40 diversidades grandes y complejas. Imaginamos los siguientes procedimientos de reacción para conectar fragmentos peptídicos a una molécula pequeña (esquemáticamente representado en la figura 1). Una estructura química equipada con grupos reactivos se incubaba con un péptido expuesto en fago. Los aminoácidos específicos del péptido (p. ej., cisteínas) reaccionan con grupos funcionales de la molécula pequeña para formar enlaces covalentes, en donde un primer enlace acelera los enlaces posteriores. A continuación, las moléculas resultantes se someten a selecciones por afinidad. Alternativamente, determinados enlaces peptídicos de la estructura multicíclica se escinden enzimáticamente para obtener estructuras químicas pequeñas decoradas con entidades peptídicas diferentes. La adhesión de los repertorios de polipéptidos expuestos en fagos a estructuras moleculares
 - 45 pequeñas no es trivial, ya que la reacción necesita ser específica y selectiva para producir un único producto. De igual forma, los reactivos no deben deteriorar convenientemente la partícula de fago. Además, la conexión a moléculas pequeñas a través de varios sitios a un péptido añade un nivel adicional de complejidad, ya que podrían generarse fácilmente mezclas de los productos o se podrían entrecruzar las partículas de fagos. De hecho, no se conoce ningún ejemplo en la técnica donde se conectara una molécula pequeña a un polipéptido expuesto sobre
 - 50 fago a través de más de un enlace.

Materiales y métodos

Unión química de las proteínas de fusión péptido-D12 a un almacén químico

Se expresaron en *E. coli* los dominios D1-D2 de g3p (que comprende los restos de aminoácidos 2 a 217 de la fdg3p madura) con y sin el péptido ***ACGSGCGSGCG*** fusionado al extremo amino. El vector de expresión basado en pUC119 con una secuencia líder y el gen de D1-D2 con una etiqueta de seis histidinas en el extremo carboxilo (denominado aquí pUC119H6D12) fue proporcionado amablemente por Phil Holliger del Laboratorio de Biología Molecular (LMB) en Cambridge. Un plásmido para la expresión de D1-D2 con el péptido en el extremo amino se clonó mediante amplificación por PCR a partir del gen de D1-D2 con los cebadores pepd12ba (que codifica la secuencia peptídica) y d12fo, y la ligación en el pUC119H6D12 digerido con *Sfi*//*Not*I. El gen para la expresión de D1-D2 sin puentes disulfuro con un total de 20 mutaciones de aminoácidos fue amablemente proporcionado por Insa Kather y Franz Xaver Schmid de la Universidad de Bayreuth. El gen se amplificó por PCR a partir del vector fdg3p0ss21 con alguna de las parejas de cebadores d120ssba/d120ssfo, pepd120ssba/d120ssfo, P2cd120ssba/d120ssfo o P1cd120ssba/d120ssfo y se ligó en *Sfi*//*Not*I de pUC119H6D12 para que el D-D2 sin puentes disulfuro se exprese con y sin los péptidos ***ACGSGCGSGCG***, ***AGSGCGSGCG*** o ***AGSGKSGSGCG*** fusionados en el extremo amino. Las 6 proteínas se expresaron en las células TG1 de *E. coli* a 30 °C durante 8 horas y se purificó la fracción periplásmica por etapas mediante cromatografía de afinidad de Ni y filtración en gel en una columna Superdex 75 en NH₄HCO₃ a 20 mM, pH 7,4. Los grupos sulfhidrilo oxidados se redujeron mediante la incubación de la proteína (1 a 10 µM) con TCEP a 1 mM en NH₄HCO₃ a 20 mM, pH 8 a 42 °C durante 1 hora. Se retiró el agente reductor en un filtro Vivaspin 20 que tenía un MWCO de 10 000 (Vivascience, Stonehouse, UK) utilizando el tampón de NH₄HCO₃ a 20 mM y EDTA a 5 mM a pH 8. Los grupos tiol de las proteínas se hicieron reaccionar mediante incubación con TBMB a 10 µM en tampón de reacción (NH₄HCO₃ a 20 mM, EDTA a 5 mM, pH 8, y ACN al 20%) a 30 °C durante 1 hora. Para retirar el TBMB que no reaccionó y concentrar, se filtró la proteína con un microcon YM-30 (Millipore, Bedford, MA). Se determinó la masa molecular de las proteínas (5 a 20 µM) mediante desnaturalización en 4 volúmenes de MeOH al 50% y ácido fórmico al 1%, y análisis en un espectrómetro de masas de tiempo de vuelo con ionización por electropulverización (Micromass, Mildford, MA, USA). Se obtuvieron las masas moleculares desconvolucionando varios espectros de masas de proteínas cargadas mediante la versión 4.1 de MassLynx. El rendimiento de la reacción de modificación química en presencia del fago se evaluó mediante la adición de fago purificado en PEG a una concentración final de 10¹⁰ u.t. para la proteína antes de la reducción con TCEP. Se retiró el fago mediante filtración en gel con una columna de PD-10 (Amersham Pharmacia, Uppsala, Suecia) después de la reacción con TBMB.

Creación de una genoteca de péptidos en fagos

Los genes que codifican un péptido semialeatorio con la secuencia Ala-Cys-(Xaa)₆-Cys-(Xaa)₆-Cys, el conector Gly-Gly-Ser-Gly y los dos dominios sin puentes disulfuro D1 y D2 se clonaron en la orientación correcta en el vector fágico fd0D12 para obtener la genoteca 1 de fagos. El vector fd0D12, que carece de los genes de los dominios D1 y D2 del gen 3 y que tiene un segundo sitio de restricción *Sfi*I se creó previamente mediante la amplificación por PCR del plásmido completo de fdg3p0ss21 (Kather, I et al., *J. Mol. Biol.*, 2005) con los cebadores ecoG3pNba y pelbsfiecfo. Los genes que codifican el repertorio de péptidos y los dos dominios del gen 3 se crearon por etapas en dos PCR consecutivas. Primero, se amplificaron por PCR los genes de D1 y D2 con los dos cebadores prepcr y sfi2fo con el vector fdg3p0ss21 como plantilla. Segundo, el ADN que codifica los péptidos aleatorios se añadió en una PCR con los cebadores sficx6ba y sfi2fo. La ligación de 33 y 9 µg del plásmido fd0D12 digerido con *Sfi*I y el producto de PCR produjo 4,4 x 10⁹ colonias sobre 12 placas de 20 x 20 cm de 2YT con cloranfenicol (30 µg/ml). Se recuperaron las colonias de las placas con el medio 2YT, se complementó con glicerol al 15% y se almacenaron a -80 °C. Se diluyeron las reservas a una DO₆₀₀ = 0,1 en cultivos de un litro de 2YT/cloranfenicol (30 µg/ml) y los fagos se expresaron a 30 °C durante una noche (12 a 16 horas).

Unión química entre un péptido expuesto en fago y una molécula pequeña

Típicamente, 10¹¹-10¹² u.t. del fago purificado en PEG se redujeron en 20 ml de NH₄HCO₃ a 20 mM, pH 8, con TCEP a 1 mM a 42 °C durante 1 hora. Los fagos se centrifugaron a 4000 rpm en un filtro Vivaspin 20 (MWCO de 10 000) para reducir el volumen a 1 ml y se lavaron dos veces con 10 ml de tampón de reacción enfriado en hielo (NH₄HCO₃ a 20 mM y EDTA a 5 mM, pH 8). El volumen del fago reducido se ajustó a 32 ml con el tampón de reacción y se añadieron 8 ml de TBMB a 50 µM en ACN para obtener una concentración final de 10 µM. La reacción se incubó a 30 °C durante 1 hora antes de que el TBMB sin reaccionar se retirase mediante precipitación del fago con 1/5 de volumen de PEG al 20% y NaCl a 2,5 M en hielo, y centrifugación a 4000 rpm durante 30 minutos.

Selección de fagos con calicreína plasmática y catepsina G humanas

La calicreína plasmática de humano (activada con el factor XIIa) se compró a Innovative Research (Southfield, MI, EE.UU.) y se biotiniló a una concentración de 1,2 µM con un exceso molar de 5 veces de Sulfo-NHS-LC-biotina

(Pierce, Rockford, IL, EE.UU.) en PBS, pH 7,4 y DMSO al 5% a TA durante 1 hora. La proteína biotinilada se purificó en una columna PD-10 con un tampón de NaAc a 50 mM, pH 5,5, y NaCl a 200 mM. Más fácil de obtener, la catepsina G humana biotinilada se compró a Lee Biosolutions (St. Louis, MI, EE.UU.). Se incubaron los antígenos biotinilados (de 5 a 20 µg) con 50 µl de perlas magnéticas de estreptavidina (Dynal, M-280 de Invitrogen, Paisley, Reino Unido) durante 20 minutos a 4 °C. Las perlas revestidas de antígeno se lavaron dos veces con tampón de lavado (Tris-HCl a 10 mM, pH 7,4, NaCl a 150 mM, MgCl₂ a 10 mM y CaCl₂ a 1 mM) y se bloquearon en 0,5 ml de tampón de lavado que contenía SAB al 1% y Tween 20 al 0,1% durante 30 minutos. El fago modificado químicamente (típicamente 10¹⁰-10¹¹ u.t. disueltos en 2 ml de tampón de lavado) se bloqueó mediante la adición de 1 ml de tampón de lavado que contenía SAB al 3% y Tween 20 al 0,3%. Se pipetearon 3 ml de fagos bloqueados en 0,5 ml de perlas magnéticas bloqueadas y se incubaron en una rueda giratoria a TA. Las perlas se lavaron 8 veces con tampón de lavado que contenía Tween 20 al 0,1% y dos veces con tampón de lavado antes de la incubación con 100 µl de glicina a 50 µM, pH 2,2, durante 5 minutos. Los fagos eluidos se transfirieron a 50 µl de Tris-HCl a 1 M, pH 8, para la neutralización, se incubaron con 50 ml de células TG1 a una DO₆₀₀ = 0,4 durante 90 minutos a 37 °C, y las células se sembraron en placas grandes de 2YT/cloranfenicol. Se realizaron dos rondas adicionales de criba con los mismos procedimientos. En la segunda ronda de selección se utilizaron perlas magnéticas revestidas de neutravidina para impedir el enriquecimiento de péptidos específicos de estreptavidina. Las perlas de neutravidina se prepararon haciendo reaccionar 0,8 mg de neutravidina (Pierce, Rockford, IL, EE.UU.) con 0,5 ml de perlas magnéticas activadas con tosilo (Dynal, M-280 de Invitrogen, Paisley, Reino Unido) según las instrucciones del proveedor.

20 *Procedimiento de escrutinio para identificar inhibidores de las proteasas*

El ADN plasmídico de los clones seleccionados después de la segunda y tercera rondas del biocribado se amplificó por PCR en un único tubo con los cebadores 21seqba y flagfo, y se clonaron en el vector pUC119H6D12 en los sitios *SfiI* y *NotI* para la expresión periplasmática de los péptidos fusionados con los dominios D1 y D2 sin puentes disulfuro con una etiqueta FLAG y una etiqueta de seis histidinas en el extremo carboxilo. Los plásmidos ligados se electroporaron en las células TG1 y se sembraron en placas de 2YT/ampicilina (100 µg/ml). Los clones que expresan la proteína recombinante se identificaron como sigue: los cultivos de 1 ml de 2YT/ampicilina (100 µg/ml) en placas de 96 pocillos profundos se inoculaban con células de colonias diferentes y se incubaron a 37 °C. La expresión de la proteína se indujo con IPTG a 1 mM cuando los cultivos estaban turbios y las placas se agitaron a 300 rpm a 30 °C durante una noche. Las células se sedimentaron por centrifugación a 3500 rpm durante 30 minutos, se lisaron con tampón de lavado que contenía lisozima a 1 mg/ml, y se centrifugaron a 3500 rpm para sedimentar los desechos celulares. Los sobrenadantes se transfirieron a placas de 96 pocillos Polysorp (Nunc, Roskilde, Dinamarca) para la adsorción inespecífica. Los pocillos se enjuagaron dos veces con tampón de lavado que contenía Tween 20 al 0,1% y se bloquearon con tampón de lavado que contenía SAB al 1% y Tween 20 al 0,1% durante 1 hora. El anticuerpo anti-FLAG M2 conjugado a peroxidasa (Sigma-Aldrich, St. Louis, MO, EE.UU.) se diluyó 1:5000 y se bloqueó en tampón de lavado con SAB al 1% y Tween 20 al 0,1%, y se añadió a las placas durante 1 hora. Los pocillos se lavaron (5 veces con tampón de lavado que contenía Tween 20 al 0,1% y una vez sin el detergente) y se detectó la peroxidasa unida con la solución del sustrato TMB (eBiosciences, San Diego, EE.UU.). Se secuenció el ADN plasmídico de los clones que expresaban la proteína (Geneservice, Cambridge, Reino Unido). Los clones seleccionados se expresaron en una escala de 800 ml y se purificaron por cromatografía de afinidad de Ni y filtración en gel como se describe más arriba. Los péptidos se modificaron químicamente con el procedimiento descrito más arriba y la concentración de los productos se determinó con la medición de la absorción óptica a 280 nm. Se midió la CI₅₀ mediante la inoculación de distintas concentraciones de las proteínas de fusión con el péptido modificado (diluciones a la mitad) con la calicreína (0,1 nM) o con la catepsina G (20 nM) de plasma humano, y se determinó la actividad residual en Tris-HCl a 10 mM, pH 7,4, NaCl a 150 mM, MgCl₂ a 10 mM, CaCl₂ a 1 mM, SAB al 0,1% y Triton X-100 al 0,01%. La actividad de la calicreína plasmática humana se midió con el sustrato fluorógeno Z-Phe-Arg-AMC (Bachem, Bubendorf, Suiza) a una concentración de 100 µM en un lector de placas de fluorescencia Spectramax Gemini (excitación a 355 nm, lectura de emisión a 460 nm; Molecular Devices, Sunnyvale, CA, EE.UU.). La actividad de la catepsina G humana se midió con el sustrato colorimétrico N-Suc-Ala-Ala-Phe-Pro-pNA (Bachem, Bubendorf, Suiza) a una concentración de 1 mM con el lector de placas de absorción Sptmax (lectura a 410 nm; Molecular Devices, Sunnyvale, CA, EE.UU.).

Selección de fagos de inhibidores de la calicreína plasmática humana por maduración de la afinidad

Se crearon tres genotecas de péptidos en fagos esencialmente como la genoteca 1 (véase más arriba), pero usando los cebadores degenerados sficx6abc (genoteca 2), sficx6abb (genoteca 3) y sficx6aba (genoteca 4) en lugar del sficx6ba. La electroporación de las reacciones de ligación en las células TG1 produjo 9,5 x 10⁸ (genoteca 2), 1,1 x 10⁹ (genoteca 3) y 1,2 x 10⁹ (genoteca 4) transformantes. Los fagos de cada genoteca se amplificaron en cultivos de 1 l, se purificaron, se reunieron y se hicieron reaccionar con TBMB. Se realizaron tres rondas de criba esencialmente como en las selecciones descritas más arriba, pero con la calicreína plasmática humana biotinilada a la concentración más baja (1 mM en la 1.^a y 2.^a rondas, y 200 pM en la 3.^a ronda).

Síntesis química de péptidos bicíclicos

Los péptidos con una amina libre en el extremo amino y una amida en el extremo carboxilo se sintetizaron químicamente a una escala de 25 mg con química de fase sólida (JPT Peptide Technologies, Berlín, Alemania). Los péptidos brutos en 1 ml de NH_4HCO_3 al 60%, pH 8, y ACN al 30% (1 mM) se hicieron reaccionar con TBMB (1,2 mM) durante 1 hora a TA. El producto de reacción se purificó por cromatografía líquida de alta resolución (HPLC) en fase inversa con una columna C18 y gradiente de elución con una fase móvil compuesta por una solución de ACN y ácido trifluoroacético (TFA) al 0,1% a una velocidad de flujo de 2 ml/min. Los péptidos purificados se liofilizaron y se disolvieron en DMSO o un tampón de Tris-HCl a 50 mM, pH 7,8, y NaCl a 150 mM para medir la actividad.

10 *Medición de la actividad y de la especificidad de los inhibidores de la calicreína plasmática humana*

Se determinó la actividad inhibidora (CI_{50}) mediante la medición de la actividad residual de la enzima tras la incubación (30 min a TA) con diferentes concentraciones del inhibidor (que oscilan típicamente de 10 μM a 0,5 pM). La actividad de la calicreína plasmática humana (0,1 nM) y del factor XIa (0,8 nM; Innovative Research, Southfield, MI, EE.UU.) se midió con Z-Phe-Arg-AMC (100 μM), y la actividad de la trombina humana (2 nM; Innovative Research, Southfield, MI, EE.UU.) con Boc-Phe-Ser-Arg-AMC (100 μM) en Tris-HCl a 10 mM, pH 7,4, NaCl a 150 mM, MgCl_2 a 10 mM, CaCl_2 a 1 mM, SAB al 0,1%, Triton X-100 al 0,01% y DMSO al 5%. La calicreína recombinante plasmática de ratón de R&D Systems (Minneapolis, MN, EE.UU.) con un péptido señal se activó proteolíticamente con termolisina a 0,5 mg/ml a 37 °C durante 1 hora. La actividad de la calicreína plasmática de ratón (3 mM) se midió con Z-Phe-Arg-AMC (100 μM) en Tris-HCl a 50 mM, pH 7,4, NaCl a 250 mM, CaCl_2 a 10 mM, SAB al 0,1%, Triton X-100 al 0,01% y DMSO al 5%. El inhibidor hidrolizado en un lazo de fijación se generó por incubación del péptido PK15 modificado con TBMB con la calicreína plasmática humana a una relación molar de 5:1 durante 24 horas a 37 °C y posterior inactivación por calor de la enzima a 60 °C (30 min). Los valores de K_i aparente se calcularon de acuerdo con la ecuación de Cheng y Prusoff (Cheng, Y. y Prusoff, W. H. *Biochem. Pharmacol.* 1973).

25 *Medición de la activación por contacto en el plasma humano*

El plasma humano normal de donantes individuales se adquirió en 3H Biomedical (Uppsala, Suecia). El plasma se centrifugó a 1500xg a 20 °C durante 15 minutos para obtener plasma con pocas plaquetas (PPP). Las alícuotas del PPP se guardaron en tubos de propileno a 80 °C. Se prepararon muestras de 60 μl de PPP que contenían 5, 50, 500 o 5000 nM de aprotinina (Roche, Mannheim, Alemania) o del péptido PK15 modificado con TBMB. El tiempo de activación de la trombina se midió a 37 °C con la adición de 20 μl de actina FS diluida 1:10 (Dade Behring, Marburg, Alemania) y 20 μl de tampón HEPES a 20 mM, pH 7,4, CaCl_2 a 100 mM, SAB a 50 mg/ml y Z-Gly-Gly-Arg-AMC a 1 mM a la muestra de plasma, y siguiendo la intensidad de la fluorescencia con un lector de placas de fluorescencia (excitación a 355 nm, lectura de emisión a 460 nm; PHERAStar, Labtech, Offenburg, Alemania). La activación del factor XIIa y de la calicreína plasmática humana se midió como se indica a continuación. Se añadieron 2 μg de caolín a las muestras de plasma, se mezcló bien y se incubó durante 20 min a 37 °C. Se diluyeron las muestras 250 veces en Tris-HCl a 50 mM, pH 7,8 y NaCl a 150 mM. La actividad similar a la de la calicreína plasmática se midió con el sustrato cromógeno H-D-Pro-Phe-Arg-pNA (100 μM ; Bachem, Bubendorf, Suiza) con un lector de placas de absorción (absorción a 450 nm; Molecular Devices, Sunnyvale, CA, EE.UU.).

Determinación de la estructura del péptido PK15 modificado con TBMB

Se disolvió 1 mg del péptido PK15 modificado con TBMB en 550 μl de Tris-HCl deuterado a 10 mM, pH 6,6, NaCl a 150 mM, MgCl_2 a 10 mM y CaCl_2 a 1 mM para obtener una concentración de inhibidor de 1 mM. El espectro de los inhibidores se leyó a 800 MHz (Bruker Avance con criosonda TCI). Las asignaciones de espectro se basaron en los espectros TOCSY (espectroscopia de correlación total) y NOESY (espectroscopia del efecto nuclear de Overhauser). Las restricciones de distancias eran del espectro NOESY. Se calcularon 50 conformómeros estructurales. El programa PyMOL se usó para analizar y visualizar la estructura de los modelos moleculares.

EJEMPLO 1: Fabricación de un complejo

En este ejemplo mostramos la unión a moléculas pequeñas de los péptidos expuestos en fagos. El polipéptido de este ejemplo es un péptido expuesto en fago. El ácido nucleico está comprendido en la partícula de fago. El compuesto conector de este ejemplo es una molécula pequeña (TBMB en este ejemplo).

Usamos la molécula orgánica pequeña tris-(bromometil)benceno (TBMB) como armazón para anclar los péptidos que contienen tres restos de cisteína (Kemp, D. S. y McNamara, P. E., *J. Org. Chem.* 1985; figura 1B). Los halogenuros de alcano conjugados a un armazón aromático reaccionan específicamente con los grupos tiol de las cisteínas en solventes acuosos a temperatura ambiente (Stefanova, H. I., *Biochemistry*, 1993). Meloen y

- colaboradores habían utilizado previamente armazones sintéticos sustituidos con bromometilo para la inmovilización de los péptidos con varias cisteínas (Timmerman, P. et al., *ChemBioChem*, 2005). Las condiciones suaves que se necesitan para la reacción de sustitución son las ventajosas para respetar la funcionalidad del fago (Olofsson, L. et al., *J. of Molecular Recognition*, 1998). Elegimos las cisteínas como puntos de anclaje porque su
- 5 cadena lateral tiene la actividad más distintiva entre los 20 aminoácidos naturales. De igual forma, los restos de cisteína son poco frecuentes en las proteínas de la cubierta del fago (8 cisteínas en pIII, una cisteína en pVI, pVII y pIX; Petrenko V. A. y Smith, G. P., *Phage Display in Biotechnology and Drug Discovery*, 2005). La simetría rotacional ternaria de la molécula de TBMB garantiza la formación de un único isómero espacial y estructural tras la reacción con tres cisteínas de un péptido.
- 10 Las condiciones de reacción para la modificación de un péptido en un fago se elaboran a continuación. Como con las técnicas disponibles resulta difícil detectar sobre el fago el péptido modificado químicamente, expresamos el péptido **^NGCGSGCGSGCG^C** como una fusión aminoterminal con los dos dominios solubles D1 y D2 de la proteína pIII minoritaria de la cubierta del fago y analizamos la masa molecular de la proteína antes y después de la
- 15 reacción con el TBMB mediante espectrometría de masas. Fallaron los intentos por conectar selectivamente las tres cisteínas del péptido con el armazón a la vez que se respetaban los tres puentes disulfuro de los dominios D1 y D2 de pIII (C7-C36, C46-C53, C188-C201). Esto nos llevó a aprovecharnos de la proteína del gen 3 sin puentes disulfuro recientemente desarrollada por Schmidt F. X. y colaboradores (Kather, I. et al. *J. Mol. Biol.* 2005). El
- 20 péptido fusionado al dominio aminoterminal de la proteína pIII sin cisteínas se redujo con tris(carboxietil)fosfina (TCEP). Como se comprobó que el reductor hacía reacción con los grupos bromometilo del armazón de TBMB, se retiró antes de añadir el TBMB a la proteína. La reoxidación de los grupos tiol después de retirar el TCEP se podría evitar con la desgasificación del tampón de reacción y la formación de complejos de iones metálicos con EDTA 5 mM. La reacción de los grupos tiol con el TBMB a distintas concentraciones y el análisis de espectrometría de
- 25 masas del producto reveló que una concentración de TBMB de 10 µM es suficiente para la modificación cuantitativa del péptido a 30 °C en una hora. Se formaba predominantemente un producto con la masa molecular esperada (Δ masa esperada = 114 Da; figura 2A). Cuando el D1-D2 sin puentes disulfuro y sin fusionar a péptido se incubó con el TBMB, la masa no cambió, lo que indicaba que no se habían producido reacciones inespecíficas con otros aminoácidos. La adición de partículas de fago a las reacciones (10^{10} u.t. por mililitro) reveló que la alta densidad de proteínas de la cubierta del fago en el recipiente no entorpecía la reacción del polipéptido con el
- 30 TBMB. Inesperadamente, encontramos que la reacción del TBMB con los péptidos que contenían solo dos restos de cisteína (**^NAGSGCGSGCG^C-D1-D2**) producía un producto con una masa molecular que era coherente con la reacción del grupo bromometilo restante con la amina primaria del extremo amino (figuras 7A y 7B). De igual forma, la reacción del TBMB con un péptido que tiene una cisteína y una lisina (**^NAGSGKGS^CCG^C-D1-D2**) producía una masa molecular que era esperable cuando las aminas primarias de la lisina y del extremo amino reaccionaban con los dos grupos bromometilo restantes (figuras 7C y 7D).
- 35 A continuación comprobamos si el fago modificado con TBMB todavía tenía capacidad de infectar las bacterias. Encontramos que, a medida que subía la concentración del TBMB en la reacción, menos fago permanecía infectante (figura 2B). En las condiciones de reacción que permiten la modificación cuantitativa del péptido (TBMB a 10 µM, 30 °C, reacción durante 1 hora), el número de fagos infectantes caía a la quinta parte.

EJEMPLO 2: Escrutinio

- 40 Este ejemplo muestra la selección por afinidad de inhibidores de la calicreína plasmática de la catepsina G humanas.
- La viabilidad de seleccionar estructuras de híbridos de moléculas pequeñas y péptidos codificadas por fagos se aplicó al ensayo que utiliza los dos antígenos humanos calicreína plasmática y catepsina G. Se creó una genoteca de exposición de péptidos en fagos sobre la proteína minoritaria pIII de la cubierta con una complejidad de $4,4 \times$
- 45 10^9 variantes. Los péptidos se diseñaron para que tuvieran dos secuencias de seis aminoácidos aleatorios flanqueados por tres cisteínas (Cys-(Xaa)₆-Cys-(Xaa)₆-Cys; figura 3A). Se añadió una alanina al extremo amino del péptido para asegurar un procesamiento correcto de la secuencia señal. Se colocó el conector Gly-Gly-Ser-Gly entre la tercera cisteína y la proteína del gen 3. Como el fago con la proteína del gen 3 sin puentes disulfuro tenía aproximadamente la infectividad 100 veces menor en comparación con el fago de tipo silvestre, se produjo gran
- 50 cantidad de partículas de fagos. Un cultivo de 1 litro incubado durante una noche a 30 °C produjo típicamente de 10^{11} a 10^{12} partículas infectantes. Aproximadamente 10^{12} partículas infectantes de fago purificadas se modificaron químicamente con el armazón de TBMB y se incubaron con cada una de las dos proteasas biotiniladas. Los fagos fijados se capturaron en perlas magnéticas de estreptavidina y se sometieron a dos ciclos de selección más. El incremento de la cifra de fagos capturados en la segunda o en la tercera ronda de selección indicó que se
- 55 enriquecieron en aglutinantes específicos. La secuenciación de los péptidos reveló diferentes secuencias consenso bien en uno o incluso ambos lazos (figuras 38 y 3C). Se amplificó por PCR el ADN de los péptidos seleccionados

de la población y se insertó en un nuevo plásmido para la expresión periplasmática de los péptidos como proteína de fusión con D1-D2. Las proteínas de fusión con péptidos que mostraron o bien similitudes de secuencia con otros clones seleccionados o bien que se encontraron en varias copias, se expresaron, se purificaron, se modificaron químicamente y se les ensayó su actividad inhibidora. Los mejores inhibidores de la calicreína plasmática y de la cathepsina G tenían una CI_{50} de 400 nM (PK2 y PK4) y 100 nM (CG2 y CG4), respectivamente, cuando se ensayaron como una fusión con D1-D2.

Ejemplo 3: Escrutinio

En este ejemplo se describe la maduración de la afinidad de los inhibidores de la calicreína plasmática humana. La comparación de la secuencia de aminoácidos de los clones seleccionados contra la calicreína plasmática humana reveló que diferentes grupos de clones tenían una gran similitud de secuencia principalmente en uno de los posibles lazos de fijación. Suponemos que las moléculas bicíclicas interaccionan predominantemente con el lazo de fijación conservado, mientras que los lazos con diferentes composiciones de aminoácidos no habían evolucionado para la interacción óptima con la proteasa. Por consiguiente, se crearon nuevas genotecas de fagos con péptidos que tenían tanto un lazo con una secuencia de una de las tres regiones consenso encontradas en la selección con la calicreína plasmática, como un lazo con seis aminoácidos aleatorios (figura 4A). El cribado de fagos con una presión selectiva más elevada mediante el uso de concentraciones bajas (de 1 nM a 200 pM) de antígeno produjeron clones que tenían una secuencia consenso en el segundo lazo de interacción (figura 4B). Los ensayos de inhibición revelaron que la CI_{50} del mejor inhibidor (PK15) se mejoró aproximadamente en un factor de 20 (20 nM) cuando se ensayó como una fusión con D1-D2.

20 Ejemplo 4: Caracterización de los complejos

Se investigó la actividad y la especificidad de los inhibidores sintetizados químicamente.

Los péptidos sintéticos de los cuatro mejores inhibidores de la calicreína plasmática humana que se aislaron en la primera selección (PK2, PK4, PK6, PK13) y el mejor inhibidor de la selección por maduración de la afinidad (PK15) se produjeron mediante síntesis en fase sólida. Se diseñaron los péptidos para que tuvieran una alanina con un grupo amino libre en el extremo amino y una glicina amidada en el extremo carboxilo para representar exactamente la carga y el entorno químico de los péptidos expuestos en los fagos. Los péptidos sintéticos que reaccionan con el TBMB se vio que tenían una CI_{50} unas 10 veces menor que los correspondientes péptidos de fusión con D1-D2 que reaccionaron con TBMB (tabla A; figura 5). La menor afinidad de los péptidos como una fusión con D1-D2 puede tener su origen en la fijación intramolecular del péptido a los dominios de la proteína del gen 3 y, por lo tanto, en una concentración aparente de inhibidor más baja. La K_i aparente del péptido PK15 modificado con TBMB se calculó mediante la ecuación de Chang y Prusoff y se encontró que era 1,5 nM (Cheng. Y. y Prusoff, W. H., *Biochem. Pharmacol.*, 1973). La IC_{50} de los péptidos lineales sin constreñir es al menos 250 veces mayor que la de los péptidos modificados con TBMB (tabla A):

Tabla A

Clon	Secuencia de aminoácidos	Masa (Da)		CI_{50} (nM)	
		Péptido lineal	Péptido bicíclico	Péptido lineal	Péptido bicíclico
PK2	H-ACSDRFRNCPLWSGTCG-NH ₂	1871,2	1985,3	> 10 000	28,6
PK4	H-ACSTERRYCPIEIFPCG-NH ₂	1942,9	2055,9	7181	33
PK6	H-ACAPWRTACYEDLMWCG-NH ₂	1974,8	2088,7	5707	21,2
PK13	H-ACGTGEGRCRVNWTGCG-NH ₂	1764,8	1879,1	> 10 000	39,1
PK15	H-ACSDRFRNCPADEALCG-NH ₂	1825	1939,4	> 10 000	1,7

El análisis por espectrometría de masas del inhibidor incubado con calicreína plasmática humana mostró un descenso de la masa de 18 Da, lo que sugiere que se hidrolizó un enlace peptídico en uno de los lazos del inhibidor. Sin embargo, la actividad inhibidora (CI_{50}) del inhibidor tratado con calicreína era tan buena como la del péptido intacto PK15 modificado con TBMB bicíclico.

La especificidad de los cinco inhibidores se ensayó con la medición de la actividad inhibidora frente a la calicreína plasmática de ratón (identidad de secuencia del 79%) o las serina proteasas humanas homólogas factor XIa (que comparten la identidad de secuencia más elevada con la calicreína plasmática humana dentro de las serina proteasas humanas; 63%) y trombina (identidad de secuencia del 36%). Ni la calicreína plasmática de ratón ni ninguna de las proteasas homólogas del suero humano se inhibieron a la concentración más elevada ensayada (10 μ M).

Ejemplo 5: Uso de las entidades identificadas en los métodos de la invención

En este ejemplo se muestra la inhibición de la activación por contacto en el plasma humano mediante un inhibidor de la calicreína plasmática humana.

10 La calicreína plasmática humana desempeña una función clave en los primeros acontecimientos de la activación por contacto. La capacidad del péptido PK15 modificado con TBMB para inhibir la activación por contacto se ensayó con la medición de la prolongación del tiempo de activación de la trombina en el plasma humano en presencia de diferentes concentraciones del inhibidor. La trombina es la última enzima que se activa en la cascada de activación de la vía de la coagulación sanguínea. A una concentración del inhibidor de 50 nM, el péptido PK15
15 modificado con TBMB retrasó la formación de trombina, mientras que la aprotinina, un inhibidor proteico de 6 kDa de la calicreína plasmática humana, no produjo ningún efecto (figuras 8A y 7B). A una concentración elevada del inhibidor de 5 μ M, la demora de la activación de la trombina fue más prolongada con la aprotinina que con la molécula pequeña inhibidora. La aprotinina es un inhibidor de amplio espectro y puede inhibir otras proteasas en la vía intrínseca cuando se utilizaba a una concentración elevada. En un ensayo diferente, comprobamos si el péptido
20 PK15 modificado con TBMB es capaz de suprimir la activación del factor XIIa y de la calicreína plasmática en el plasma humano de tres donantes diferentes. La activación de dos proteasas podría suprimirse esencialmente a 5 μ M del péptido PK15 modificado con TBMB. Estimamos que se necesita una concentración unas 30 veces más elevada de aprotinina para obtener un mismo efecto de inhibición (figura 8C).

Ejemplo 6: determinación de la estructura del péptido PK15 modificado con TBMB.

25 Se determinó la conformación del péptido PK15 modificado con TBMB mediante espectroscopia 2D de 1 H-RMN en solución acuosa a pH 6,6. Se consiguieron asignaciones de desplazamiento químico mediante los métodos estándares. El análisis de los espectros NOESY probó que la conformación del esqueleto estaba bien definida. Son notables los desplazamientos químicos de los tres protones del anillo central de benceno que se podían resolver como resultado de las diferencias en su entorno espacial. Las estructuras medias en disolución se calcularon
30 utilizando limitaciones de la distancia procedentes de la NOESY (figura 6).

Resumen de los ejemplos 1 a 6

Hemos demostrado que la invención con referencia a la tecnología de exposición en fagos codifica la fracción peptídica de estructuras moleculares pequeñas no naturales (a saber, complejos de acuerdo con la presente invención). La codificación genética permite fácilmente generar, seleccionar y amplificar repertorios combinatorios
35 muy grandes. Una dificultad importante de este método fue inmovilizar los repertorios de péptidos codificados por fagos al núcleo de moléculas pequeñas. Desarrollamos una estrategia de síntesis cómoda y establecimos unas condiciones de reacción óptimas en una serie de experimentos. Se han tenido que escoger cuidadosamente la concentración de los reactivos, la composición del solvente y la temperatura de reacción para unir específicamente los péptidos lineales del fago con las moléculas pequeñas al mismo tiempo que se respetan las partículas de
40 fagos. Se utilizó un fago específico con proteínas del gen 3 sin puentes disulfuro para ayudar a impedir la generación de mezclas de productos a lo largo de la reacción de la molécula pequeña con los restos de cisteína de la cubierta del fago.

Hemos elegido la calicreína plasmática y la cathepsina G humanas como dianas para ensayar la eficacia de las técnicas de selección *in vitro* de la invención. Se aislaron frente a ambas dianas las moléculas cuya afinidad está
45 en el intervalo nanomolar inferior y se confirmó que la estrategia de selección y el diseño de moléculas propuestos son capaces de producir unos aglutinantes con gran afinidad. Cuando evaluamos la especificidad de los inhibidores de la calicreína plasmática humana, encontramos que no se inhibieron ni la calicreína plasmática de ratón, ni las proteasas plasmáticas humanas homólogas como el factor XIa o la trombina. Este hallazgo fue atractivo porque no resulta trivial la generación de los inhibidores de masa molecular pequeña específicos de la
50 calicreína plasmática humana (Young, W. B. et al., *Biorganic and Medicinal Chemistry Letters*, 2006) y de otras serina proteasas humanas (revisado en Abbenante, G, y Fairlie, D. P., *Medicinal Chemistry*, 2005 y Turk, B., *Nature Rev. Drug Discovery*, 2006). El acceso de las estructuras de moléculas pequeñas a la síntesis química permite reemplazar determinados aminoácidos por bloques de construcción no naturales y, por lo tanto, mejorar adicionalmente la afinidad de los inhibidores.

- La determinación de la estructura de uno de los inhibidores de la calicreína plasmática en solución mediante RMN sugirió que la molécula tiene el esqueleto en conformación definida. Como se anticipó, el anillo de benceno hidrófobo forma el núcleo de la molécula. Sin embargo, ninguna de las cadenas laterales de aminoácidos densamente empaquetadas con el anillo de benceno para esta combinación de compuesto conector y un único polipéptido particular. Se pueden utilizar ventajosamente armazones alternativos con estructuras químicas que ofrecen más posibilidades de interaccionar con el esqueleto peptídico o con las cadenas laterales de los aminoácidos para obtener un empaquetamiento más denso de la fracción de péptido si se desea. Los átomos de hidrógeno del armazón de 1,3,5-tris-(bromometil)-benceno en las posiciones del anillo 2, 4 y 6 podrían, por ejemplo, reemplazarse por tres sustituyentes químicos idénticos.
- En las selecciones demostradas en la presente memoria utilizamos un diseño de molécula en el cual se inmoviliza un péptido a través de tres enlaces con un armazón de molécula pequeña para obtener una estructura peptídica bicíclica. Por supuesto, en las realizaciones de selección/escrutinio también se puede utilizar la creación de arquitecturas moleculares alternativas en las que los lazos peptídicos los escinden las proteasas antes de la selección para obtener moléculas pequeñas con restos peptídicos independientes. De hecho, en este trabajo se generaron las estructuras con dos restos peptídicos independientes cuando se escindió el inhibidor PK15 modificado por TBMB mediante la calicreína plasmática humana al incubarlo con la enzima. La molécula digerida individualmente se encontró que tenía una actividad inhibidora que era tan buena como la de la forma no hidrolizada. La escisión de los lazos peptídicos también ofrece la posibilidad de unir otras estructuras químicas a los extremos nacientes amino o carboxilo mediante otras reacciones químicas.
- Hemos valorado el potencial terapéutico del inhibidor evolucionado de la calicreína plasmática humana mediante la comprobación de su capacidad para inhibir la activación por contacto en el plasma humano. En la cirugía cardíaca que implica una circulación extracorporeal, el contacto de la sangre con la superficie artificial de la máquina de circulación extracorporeal y la intubación activa varias vías de proteasas plasmáticas. Pueden sucederse complicaciones graves, entre ellas el síndrome de respuesta inflamatoria sistémica (SIRS, por su nombre en inglés), un estado inflamatorio pancorporal que puede afectar al funcionamiento del corazón y del pulmón en los pacientes (Miller, B. E. et al., *J. Of Cardiothoracic and Vascular Anesthesia*, 1997). La calicreína plasmática desempeña una función clave en los primeros acontecimientos de la activación por contacto y en la amplificación de otras vías de proteasas, tales como el sistema fibrinolítico y el sistema del complemento. Una estrategia habitual para suprimir la activación por contacto durante la intervención cardíaca es bloquear la actividad de la calicreína plasmática con la aprotinina, un inhibidor de proteasas de amplio espectro de 6 kDa procedente de tejido de pulmón bovino. El inhibidor se fija a la calicreína plasmática con una K_i de 30 nM y, por consiguiente, interrumpe la vía de coagulación intrínseca mediante la supresión de la activación del factor XII. Además, la inhibición de la calicreína plasmática disminuye la conversión del plasminógeno en plasmina y, por consiguiente, reduce la fibrinólisis y la hemorragia que le acompaña. La aprotinina es también un inhibidor directo de la plasmina (K_i H 3 nM). Se cree que la inhibición directa de la plasmina es el mecanismo principal por el que los efectos antifibrinolíticos conducen a la reducción de la pérdida de sangre y a la disminución de la necesidad de transfusión. El fármaco también tiene efectos adversos como la anafilaxis y la toxicidad renal (revisado por Mahdy A. M. y Webster N. R., *Br. J. Anaesth.*, 2004). Se ha desarrollado recientemente un inhibidor alternativo de la calicreína plasmática basado en el armazón del dominio Kunitz humano (6 kDa) (Markland, W. et al, *Biochemistry*, 1996). El fármaco tiene una afinidad (K_i = 30 pM) y especificidad significativamente más elevadas por la calicreína plasmática y se espera que sea menos inmunógeno debido a su marco estructural humano. Actualmente se está comprobando en ensayos clínicos de fase 2 (Dyax Corp., www.dyax.com). Incluso aunque nuestro inhibidor precursor recién desarrollado tenga una afinidad casi 50 veces menor por la calicreína plasmática humana que el producto inhibidor basado en el dominio Kunitz, se comprobó que suprimía de manera eficaz la activación por contacto *ex vivo*. Su menor tamaño (2 kDa) permite no sólo que se sintetice químicamente con facilidad, sino que también disminuye ventajosamente al mínimo el riesgo de una reacción inmunógena y hace al compuesto un inhibidor novedoso y atractivo para su desarrollo o su uso en las circulaciones extracorporales.

Ejemplo 7: interacciones no covalentes

- El compuesto conector de la invención proporciona la ventaja adicional de influir/establecer o imponer constreñimientos conformacionales en el polipéptido diana por medio de enlaces no covalentes formados entre el compuesto conector y el polipéptido diana. Estos se proporcionan ventajosamente además de los enlaces covalentes entre el compuesto conector y el polipéptido diana.

- Se debe observar que tales formaciones de enlaces y constreñimientos no se dan a conocer en los métodos de la técnica anterior, tal como el entrecruzamiento. Primero, los entrecruzadores son típicamente demasiado pequeños y/o demasiado flexibles para contribuir al constreñimiento conformacional. Segundo, en el ejemplo específico de entrecruzamiento conocido tratado más arriba (p. ej., Roberts, patente de los EE.UU. n.º US 2003/0235852A1), el conector bivalente es pequeño (propilo) y muy flexible al haberse diseñado a propósito, y no hay pruebas de que

produzca ninguna interacción no covalente ni imponga ningún constreñimiento conformacional más allá de juntar dos restos dentro del polipéptido. En cualquier caso, este entrecruzador de la técnica anterior es sólo bivalente.

En este ejemplo demostramos que se pueden formar enlaces no covalentes ventajosos entre el compuesto conector y el polipéptido diana de la invención.

- 5 La estructura de un inhibidor de la calicreína plasmática humana generado con el método de la invención (véanse los ejemplos anteriores) se resolvió por RMN. En la estructura propuesta, varios átomos de carbono del polipéptido se encuentran muy próximos ($< 4 \text{ \AA}$) a los átomos de carbono del compuesto conector. Esto sugiere que las interacciones no covalentes están presentes, en este ejemplo son interacciones hidrófobas, entre el núcleo y el polipéptido de la invención.

- 10 Estas interacciones son:

Ser3 CB	-	Rng C26	3,62 Å
Ser3 CB	-	Rng C2	4,0 Å
Ser3 CB	-	Rng CMe2	3,63 Å
Cys2 CB	-	Rng CMe2	2,56 Å
Cys9 CB	-	Rng C29	3,13 Å
Cys9 CB	-	Rng C9	3,32 Å
Pro10 CG	-	Rng CMe9	3,8 Å
Pro10 CD	-	Rng CMe9	3,13 Å
Cys16 CB	-	Rng C16	3,43 Å
Cys16 CB	-	Rng C26	3,79 Å

Además, las interacciones hidrógeno-hidrógeno entre los átomos de hidrógeno del polipéptido y los átomos de hidrógeno del compuesto conector se detectaron mediante espectroscopia ^1H -RMN NOESY.

Por lo tanto, se demuestra la existencia de varias clases de interacción no covalente entre el compuesto conector y el polipéptido diana de la invención. Éstas proporcionan ventajosamente más constreñimiento conformacional a los

- 15 polipéptidos de la invención.

Ejemplo 8: Genotecas combinatorias de entidades químicas codificadas en fagos

Visión general

- La tecnología de exposición en fagos se ha demostrado previamente que es eficaz para fabricar anticuerpos terapéuticos a partir de genotecas combinatorias, pero es difícil de aplicar para fabricar moléculas pequeñas farmacológicas. Aquí describimos una estrategia con fagos para seleccionar imitadores de compuestos macrocíclicos producidos por las péptido sintetas que no están en ribosomas. Los repertorios de péptidos se diseñaron con tres restos de cisteína reactivos, separados entre sí por restos de aminoácidos aleatorios, y se fusionaron a la proteína del gen 3 de los fagos. La conjugación con un compuesto conector (en este ejemplo tris-(bromometil)benceno) a través de las cisteínas reactivas generó repertorios de conjugados peptídicos con dos
- 20 lazos peptídicos anclados a un núcleo de mesitileno. La reiteración de la selección por afinidad produjo varios inhibidores enzimáticos; después de mutagénesis adicionales y selección, aislamos un inhibidor puntero (PK15) ($K_i = 1,5 \text{ mM}$) específico de la calicreína plasmática humana que interrumpió eficazmente la vía de coagulación intrínseca en el plasma humano ensayado *ex vivo*. Por consiguiente, demostró que esta estrategia proporciona medios poderosos para generar y detectar tales imitadores de macrociclos.

30 Antecedentes

- El descubrimiento de nuevos ligandos de dianas receptoras, enzimáticas o nucleotídicas representa la primera etapa del desarrollo de fármacos terapéuticos. Para los fármacos basados en ligandos orgánicos pequeños, el escrutinio de alto rendimiento (HTS, por su nombre en inglés) ha demostrado ser una estrategia popular; se sintetizan (o se compran) grandes genotecas de compuestos y a cada compuesto se le ensaya la unión a las
- 35 dianas. Con el uso de robots se puede detectar de 10^5 a 10^6 compuestos al día, pero las dianas relevantes suelen

- requerir más química para mejorar su afinidad de unión y su especificidad de diana^{1,2}. Para los fármacos basados en ácidos nucleicos, péptidos o proteínas, los métodos de selección biológica ofrecen una estrategia alternativa. Estos métodos (tales como la exposición en fagos, exposición en ribosomas, exposición de ARNm o tecnologías de aptámeros de ARN/ADN) se basan en (a) crear una genoteca heterogénea en la que el fenotipo (fijación a la diana) de cada miembro de la genoteca está conectado con su genotipo (el ADN o ARN codificante), y (b) un ciclo reiterativo en el que los miembros de la genoteca se seleccionan por la fijación a la diana y luego se amplifican (mediante la replicación en una célula hospedadora, o la copia *in vitro* del ácido nucleico codificado). En cada ronda de selección, los aglutinantes se enriquecen de este modo frente a los no aglutinantes. Se pueden realizar el escrutinio eficaz de genotecas muy grandes (de 10^9 a 10^{13} miembros) con unas pocas rondas de selección y los mejores casos se pueden refinar mediante mutación y posterior selección³. El método es muy poderoso y se ha utilizado para crear anticuerpos terapéuticos tales como Humira^{TM 4,5}. Se han intentado desarrollar varias veces métodos de selección para aislar ligandos orgánicos pequeños. El ADN típicamente se utiliza como una etiqueta que con facilidad se puede sintetizar, secuenciar, amplificar y/o hibridar. Por ejemplo, se pueden conjugar moléculas pequeñas a un etiqueta única de ADN⁶ (o bacteriófago⁷), y los conjugados se mezclan juntos para crear una genoteca de moléculas pequeñas etiquetadas. Después de seleccionar la genoteca contra la diana, las moléculas pequeñas «relevantes» se pueden identificar mediante la secuencia de sus etiquetas (amplificadas). Alternativamente, las etiquetas de ADN se pueden introducir durante la síntesis de las colecciones combinatorias de entidades químicas. Por ejemplo, las moléculas pequeñas y una correspondiente etiqueta se sintetizan en paralelo sobre la misma perla⁸, o la hibridación de la etiqueta se utiliza para gobernar la vía de la síntesis⁹ química. De tales colecciones, la vía (y por lo tanto estructura) sintética de las moléculas relevantes seleccionadas se pueden deducir de la secuencia de la etiqueta. A pesar de su ingeniosidad, estos métodos padecen una desventaja común; la molécula pequeña está conectada a la etiqueta de ADN sólo durante la primera ronda de selección, lo que lo imposibilita la reiteración de las rondas (y limitan la aplicación a colecciones pequeñas). Por lo tanto, la técnica anterior presenta numerosas dificultades.
- En este ejemplo demostramos que la invención se puede utilizar para modificar químicamente los péptidos sobre fago durante el procedimiento de selección^{10,11} para crear imitaciones de compuestos monocíclicos peptídicos. Recientemente se han descrito métodos para sujetar los péptidos por las cadenas laterales reactivas (p. ej., cisteínas) a los grupos funcionales de un armazón¹² orgánico, y, por consiguiente, generar conjugados de péptidos policíclicos que comprenden un núcleo orgánico decorado con lazos peptídicos. Como las estructuras son reminiscentes de los fármacos macrocíclicos peptídicos, exploramos la posibilidad de crear y seleccionar genotecas de tales conjugados en fagos filamentosos (figura 9a). Mientras que los macrociclos peptídicos se hacen normalmente *in vivo* mediante péptido sintetas^{13,14} no ribosómicas, nuestra estrategia utiliza la síntesis ribosómica *in vivo* y luego la conjugación química *ex vivo*.

Resultados

35 Conjugación de un armazón orgánico a los péptidos expuestos en fagos

Utilizamos el compuesto orgánico pequeño tris-(bromometil)benceno (TBMB) como armazón (compuesto conector) al que anclar los péptidos que contienen tres restos de cisteína^{12,15} (figura 9a). La reacción se produce en disolventes acuosos a temperatura ambiente, y la simetría de rotación ternaria de la molécula de TBMB asegura la formación de un isómero espacial y estructural único.

- Primero elaboramos las condiciones de reacción para conjugar el péptido ^NGCGSGCGSGCG^C fusionado a los dominios D1-D2 solubles de la pIII del fago, y se analiza la masa molecular de los productos mediante espectrometría de masas. Sin embargo, fuimos incapaces de conjugar selectivamente los tres restos de cisteína del péptido con el TBMB, a la vez que se respetaban los puentes disulfuro entre D1 y D2 (C7-C36, C46-C53, C188-C201). Esto nos invitó a aprovecharnos de una proteína del gen 3 sin puentes disulfuro desarrollada recientemente por Schmidt F. X. y colaboradores¹⁶. La proteína de fusión péptido-D1-D2 (sin puentes disulfuro) se redujo con tris-(carboxietil)fosfina (TCEP), se retiró el TCEP y se añadió el TBMB. Fue suficiente una concentración de TBMB de 10 μ M para la reacción cuantitativa con la proteína de fusión del péptido a 30 °C en una hora, lo que dio predominantemente un producto con la masa molecular esperada (Δ masa esperada = 114 Da; figura 10a). No se detectó ningún producto con la proteína D1-D2 (sin puentes disulfuro). Inesperadamente, encontramos que la reacción del TBMB con las fusiones péptido-D1-D2 (sin puentes disulfuro) que contienen sólo dos restos de cisteína (^NAGSGCGSGCG^C-D1-D2) produjeron un producto con una masa molecular coherente con la reacción de ambas cisteínas y el grupo amino $\pm$ del extremo amino del péptido (figuras 15A y 15b). De igual forma, la reacción del TBMB con las fusiones péptido-D1-D2 (sin puentes disulfuro) que tienen una cisteína y una lisina (^NAGSGKGS^CGCG^C-D1-D2) produjeron una masa molecular coherente con la reacción de la cisteína, el grupo amino $\pm$ del extremo amino y el grupo amino μ de la lisina (figuras 15c y 15d). Al haber identificado las condiciones adecuadas, hacemos reaccionar el TBMB con el fago (con p3 sin puentes disulfuro) que lleva el péptido

^NGCGSGCGSGCG^C. Esto conduce a una pequeña pérdida (de 5 veces) de infectividad del fago (figura 10b).

Creación de una genoteca de péptidos policíclicos y selección por afinidad

Diseñamos una genoteca de péptidos que contenían dos secuencias de seis aminoácidos aleatorios flanqueadas por tres cisteínas [Cys-(Xaa)₆-Cys-(Xaa)₆-Cys; figura 11a] para exposición en fagos (con p3 sin puentes disulfuro).

- 5 Se añadió un resto de alanina al extremo amino del péptido para garantizar que la secuencia señal se procesa correctamente. Se colocó un conector Gly-Gly-Ser-Gly entre la tercera cisteína y la proteína del gen 3. Como el fago (con p3 sin puentes disulfuro) tenía una infectividad 100 veces menor en comparación con el tipo silvestre, se produjo una gran cantidad de partículas de fagos a partir de la genoteca (se estimaban $4,4 \times 10^9$ variantes). Un litro de cultivo incubado durante una noche a 30 °C producía típicamente de 10^{11} a 10^{12} partículas infectantes.
- 10 Comprobamos la capacidad de unión e inhibición de las proteasas humanas cathepsina G y calicreína plasmática que tenía la genoteca de péptidos policíclicos. Unas 10^{12} partículas infectantes de fagos purificadas se modificaron químicamente con TBMB y luego se incubaron con las proteínas diana biotiniladas. Después de la captura sobre perlas magnéticas con avidina o estreptavidina, los fagos enriquecidos se trataron en dos rondas más de selección, donde cada ronda comprendía amplificación (mediante la infección bacteriana), conjugación química y
- 15 captura con las dianas biotiniladas. El título del fago se incrementó después de la segunda y tercera rondas, lo que sugería un enriquecimiento en los aglutinantes específicos. El ADN que codifica los péptidos se amplificó por PCR a partir de la población seleccionada de fagos en la tercera ronda, y se volvió a clonar para expresarlo en el periplasma como proteínas de fusión péptido-D1-D2 (D1-D2 sin puentes disulfuro) y para secuenciarlo. Esto reveló secuencias consenso en uno o ambos lazos peptídicos (figuras 11b y 11c) y varios de ellos se expresaron,
- 20 purificaron, conjugaron con TBMB y se les ensayó su actividad inhibidora de proteasas. Los mejores inhibidores de la calicreína plasmática y de la cathepsina G tenían una CI_{50} de 400 nM (PK2 y PK4) y 100 nM (CG2 y CG4), respectivamente, cuando se ensayaron como una fusión con D1-D2. Como realizamos el escrutinio de los clones seleccionados de fagos por su inhibición (en lugar de fijación), no podemos afirmar si también se seleccionaron las moléculas que se fijaban a las proteasas, pero que no las inhibían. Sin embargo, el hallazgo de que una gran
- 25 mayoría de los clones verificados después de la selección de fagos mostró actividad inhibidora sugiere que se seleccionaron predominantemente los inhibidores.

Maduración de la afinidad de los inhibidores de la calicreína plasmática humana

La mayoría de las secuencias que se unían a la calicreína revelaron secuencias consenso en uno u otro de los lazos del péptido. Se crearon tres nuevas genotecas con una de las tres regiones consenso en un lazo y seis aminoácidos aleatorios en el otro lazo (figura 12a). Se mezclaron las genotecas y se cribaron los fagos en condiciones rigurosas (de 1 nM a 200 pM de calicreína biotinilada). La secuencia aleatoria convergió a un nuevo consenso, y produjo clones con secuencias consenso en ambos lazos (figura 12b). Los ensayos de inhibición revelaron que la CI_{50} del mejor inhibidor (PK15) era de 20 nM cuando se ensayaba como una fusión con D1-D2.

Actividad y especificidad de los inhibidores sintetizados químicamente

- 35 Los péptidos sintéticos que corresponden a los cuatro inhibidores de la calicreína de la selección primaria (PK2, PK4, PK6 y PK13) y al mejor inhibidor de la selección por maduración de la afinidad (PK15) se sintetizaron mediante síntesis química en fase sólida. Los péptidos tenían un resto de alanina en el extremo amino y una glicina amidada en el extremo carboxilo para representar la carga y el entorno químico de los péptidos expuestos en fagos. Los péptidos sintéticos conjugados con TBMB inhibían la actividad de la calicreína con al menos 250
- 40 veces más potencia que los péptidos sin conjugar (tabla 1).

Tabla 1: Inhibidores peptídicos sintetizados químicamente. Se muestra la secuencia de aminoácidos de cinco inhibidores (17-meros) de la calicreína plasmática. La secuencia de los péptidos sintéticos procede de los clones PK2, PK4, PK6 y PK13 (aislados en selecciones de fagos con la genoteca 1) y del clon PK15 (un clon por maduración de la afinidad aislado de la genoteca 2). Se indica la masa molecular y la actividad inhibitoria antes y

- 45 después de la modificación de los péptidos con el TBMB:

Clon original	Secuencia de aminoácidos	Masa (Da)		CI_{50} (nM)	
		Péptido lineal	Péptido bicíclico	Péptido lineal	Péptido bicíclico
PK2	H-ACSDRFRNCPLWSGTCG-NH ₂	1871,2	1985,3	> 10 000	28,6

PK4	H-ACSTERRYCPIEIFPCG-NH ₂	1942,9	2055,9	7181	33
PK6	H-ACAPWRTACYEDLMWCG-NH ₂	1974,8	2088,7	5707	21,2
PK13	H-ACGTGEGRCRVNWTPCG-NH ₂	1764,8	1879,1	> 10 000	39,1
PK15	H-ACSDRFRNCPADEALCG-NH ₂	1825	1939,4	> 10 000	1,7

Eran inhibidores más de 10 veces más potentes que los conjugados péptido-D1-D2 (tabla 1; figura 13); presumiblemente sea debido a la unión del resto peptídico conjugado con el resto D1-D2. La constante de inhibición aparente (K_i) del péptido conjugado PK15 (figura 9b) se calculó que era de 1,5 nM usando la ecuación de Cheng y Prusoff¹⁷. La incubación del conjugado PK15 con la calicreína conduce a la hidrólisis de un enlace peptídico después de una incubación prolongada (escisión del 90% después de 24 h a 37 °C), como se muestra por la ganancia de masa de 18 Da, pero la actividad inhibidora de las muestras escindidas y sin escindir resultaron ser similares (Cl_{50} de 2,2 nM y 1,6 nM, respectivamente).

- 10 También se analizaron los cinco inhibidores contra la calicreína plasmática de ratón (identidad de secuencia del 79%) o las serina proteasas homólogas factor Xla (identidad de secuencia del 63%) y la trombina (identidad de secuencia del 36%). Ninguno inhibió a estas enzimas a la concentración más alta utilizada (10 μ M).

Interrupción de la vía de coagulación intrínseca mediante un inhibidor de la calicreína plasmática humana

La calicreína plasmática humana desempeña una función clave en las primeras etapas de la vía de coagulación intrínseca al convertir el factor XII en el factor XIIa que entonces actúa sobre la siguiente proteasa de la vía.

- 15 Analizamos si el conjugado PK15 inhibía la activación del factor XIIa en las muestras de plasma humano. La vía se desencadenó con caolín y la actividad del factor XIIa se midió con un sustrato colorimétrico. La actividad de XIIa cayó a la mitad en presencia del conjugado PK15 a 160 nM (figura 16). En comparación, para lograr el mismo efecto se necesitó la aprotinina a 5 μ M, un inhibidor bovino de 6 kDa de las serina proteasas que también se usa en la clínica como inhibidor de la calicreína plasmática (K_i = 30 nM).

20 Determinación de la estructura del péptido PK15 modificado con TBMB

Se leyó el espectro 2D por ¹H-RMN del conjugado PK15 y fue posible alinear específicamente la secuencia con los desplazamientos químicos en los espectros de TOCSY y de NOESY. En la figura 14 se muestra una conformación del inhibidor calculada por las restricciones de distancias derivadas del NOESY. Los dos lazos del péptido están dispuestos alrededor del núcleo de mesitileno al que están unidos covalentemente, pero no interaccionan entre sí. Los lazos no se empaquetan densamente contra el núcleo, pero algunos átomos de carbono del polipéptido (Cys9 CB, Cys16 CA, Gly17 CA) están a menos de 4 Å de los átomos del núcleo molecular, lo que sugiere que deben existir algún tipo de interacción hidrófoba.

Discusión del ejemplo 8

- 30 Hemos mostrado que la reacción del tris-(bromometil)benceno (TBMB)¹² con las genotecas de péptidos ricos en cisteína expuestos sobre bacteriófagos filamentosos genera conjugados (complejos de acuerdo con la presente invención) sobre los que se pueden realizar selecciones iterativas. Fue un desafío conjugar el péptido expuesto a la vez que se respetaba al fago, y tuvimos que variar la concentración de los reactivos, la composición del solvente y la temperatura de reacción, y también usamos fagos que carecían de puentes disulfuro en la proteína del gen 3. A partir de una genoteca de >10⁹ miembros y selecciones reiterativas conseguimos aislar inhibidores potentes de la calicreína plasmática humana (<2000 Da). Nuestro mejor inhibidor (PK15) con K_i = 1,5 nM interrumpe eficientemente la vía de la coagulación intrínseca en el plasma humano analizado *ex vivo*, y era muy específico: no inhibía la calicreína plasmática de ratón ni las proteasas plasmáticas homólogas humanas factor Xla y trombina.

- 40 Nuestro repertorio se construyó a partir de péptidos de 17 restos con tres cisteínas, separadas entre sí por seis aminoácidos aleatorios. Después de la conjugación con TBMB, se espera que los péptidos formen dos lazos de seis restos unidos a un núcleo de mesitileno, como se confirmó realmente gracias a la estructura del inhibidor de la calicreína PK15 resuelto mediante RMN (figura 14). Tales péptidos policíclicos deben tener ventajas tanto sobre los péptidos con puentes disulfuro como con los lineales. Las ventajas de los péptidos policíclicos sobre los péptidos con puentes disulfuro son que los enlaces covalentes azufre-carbono, una vez que se forman, son inertes para el intercambio¹⁸, y también son estables en un entorno reductor¹⁸. La ventaja de los péptidos policíclicos sobre los péptidos lineales es que están entrecruzados y más constreñidos. Esto tiene dos consecuencias principales: (a) los

péptidos constreñidos se espera que se unan más estrechamente a las dianas (debido a la menor pérdida de entropía conformacional). Nuestra revisión de la bibliografía sobre inhibidores peptídicos que se han aislado por exposición en fagos muestra que la mayoría contienen disulfuros y tienen constantes de inhibición del orden micromolar (tabla 3).

- 5 Tabla 3: Inhibidores peptídicos seleccionados en fagos. Se indican la secuencia de los péptidos, la diana de la enzima y la afinidad de unión. Los restos de cisteína que forman puentes disulfuro están subrayados.

Diana	Secuencia peptídica	Afinidad	Referencia
Antígeno prostático específico (PSA)	<u>C</u> VAY <u>C</u> IEHH <u>C</u> WT <u>C</u>	$K_D = 2,9 \mu M$	1
Calicreína 2 humana	SRFKVWWAAF	$Cl_{50} = 3,4 \mu M$	2
Activador del plasminógeno de tipo urocinasa (uPA)	<u>C</u> SWRGLENHRM <u>C</u>	$K_i = 6,7 \mu M$	3
Activador del plasminógeno de tipo urocinasa (uPA)	<u>C</u> PAYSRYL <u>D</u> <u>C</u>	$K_i = 0,4 \mu M$	4
Quimotripsina	<u>C</u> CFSWR <u>C</u> RC	$K_i = 1 \mu M$	5
TF-fVII	EEWEVL <u>C</u> WTWET <u>C</u> ER	$Cl_{50} = 1,5 nM$	6
Enzima convertidora de la angiotensina 2 (ACE2)	GDYSH <u>C</u> SPLRYPWWK <u>C</u> ITYPDP	$K_i = 2,8 nM$	7
ErbB-2	K <u>C</u> CYSL	$K_i = 30 \mu M$	8
Ureasa	YDFYWW	$Cl_{50} = 30 \mu M$	9
Lipasa pancreática	<u>C</u> QPHPGQ <u>T</u> <u>C</u>	$Cl_{50} = 16 \mu M$	10
β -Lactamasa	<u>C</u> VHSPNRE <u>C</u>	$Cl_{50} = 9 \mu M$	11
DNasa II	<u>C</u> LRLLQWFLW <u>A</u> <u>C</u>	$K_i = 0,2 \mu M$	12

- 10 Sólo dos inhibidores peptídicos fueron tan potentes como la PK15; ambos contenían un puente disulfuro y al menos dos restos de triptófano¹⁹⁻²⁰. Esto sugiere que (a) la conformación constreñida y la posibilidad de formar interacciones hidrófobas son clave para estas afinidades elevadas; (b) los péptidos constreñidos (y entrecruzados) también deben ser más resistentes a la escisión y/o inactivación que los péptidos lineales. De hecho, en nuestro trabajo se escindió uno de los lazos del inhibidor PK15 después de la incubación prolongada con la calicreína plasmática humana, pero permaneció intacta y activa.

- 15 Los conjugados policíclicos se prestan a la ingeniería genética y química. La masa molecular de la PK15 (1939,4 Da) es mayor que la de varios fármacos macrocíclicos peptídicos (tabla 2), pero sería posible utilizar lazos más cortos. Por ejemplo, alterando el espaciado de las cisteínas, se varía fácilmente la longitud del lazo, o incluso se añaden segmentos extra a los extremos de los péptidos.

Tabla 2. Comparación de tamaños de los fármacos macrocíclicos.

Nombre	Tamaño del ciclo o de los ciclos	Masa molecular (Da)	Aplicación
Actinomicina	16,16	1255,42	Antineoplásico

Anfotericina B	38	924,08	Antimicótico
Azitromicina	15	748,88	Antibiótico
Caspofungina	21	1093,31	Antimicótico
Ciclosporina	32	1202,61	Inmunodepresor
Daptomicina	31	1619,71	Antibiótico
Eritromicina	14	733,93	Antibiótico
Ixabepilona	16	506,70	Antineoplásico
Ocreótido	20	1019,24	Hormona
Oxitocina	20	1007,19	Hormona
Polimixina B	23	1301,56	Antibiótico
Rapamicina	29	914,17	Inmunodepresor
Rifabutina	27	847,01	Antibiótico
Vancomicina	16, 16, 12	1449,30	Antibiótico

Otras variaciones podrían incluir la mutagénesis de los lazos (como con la maduración de la afinidad de PK15); la escisión proteolítica en uno o ambos lazos para generar segmentos peptídicos «ramificados» en las cisteínas; conjugación química en los extremos del péptido naciente después de la escisión de lazo²¹; o el uso de variaciones en los núcleos orgánicos. Por ejemplo, un núcleo orgánico más grande, o uno con más grupos funcionales podría interaccionar más extensivamente con los lazos o con la diana, y también se podrían utilizar para introducir funciones completamente nuevas, tales como la fluorescencia. Si se realizasen estas operaciones sobre el conjugado expuesto en fagos, se podrían seleccionar las variaciones mediante un proceso reiterativo. Dado que los conjugados peptídicos también son adecuados para la síntesis química, se podrían introducir sintéticamente más variaciones (tales como la sustitución mediante aminoácidos no naturales).

Los inhibidores de la calicreína plasmática humana se han desarrollado clínicamente para el tratamiento del angioedema hereditario y para la intervención quirúrgica de derivación coronaria, pero se ha demostrado que es difícil fabricar moléculas pequeñas que sean específicas de la calicreína (revisado en^{22,23}). El hecho de que nosotros obtengamos fácilmente un inhibidor muy específico y de gran afinidad mediante la selección reiterativa de conjugados peptídicos policíclicos sobre los fagos indica que esta estrategia está bien encaminada.

Materiales y métodos

Modificación química con TBMB de los repertorios peptídicos sobre los fagos

Las genotecas de péptidos en fagos que se basan en el plásmido fdg3p0ss2166 se clonaron y se produjeron como se describe más adelante. Típicamente, los fagos purificados con PEG a 10^{11} - 10^{12} u.t. se redujeron en 20 ml de NH_4HCO_3 a 20 mM, pH 8 con TCEP a 1 mM a 42 °C durante 1 hora. Los fagos se centrifugaron a 4000 rpm en un filtro Vivaspin 20 (MWCO de 10 000) para reducir el volumen del tampón de reducción a 1 ml y se lavó dos veces con 10 ml de tampón de reacción enfriado en hielo (NH_4HCO_3 a 20 mM y EDTA a 5 mM, pH 8). El volumen del fago reducido se ajustó a 32 ml con tampón de reacción, y se le añadieron 8 ml de TBMB a 50 μM en ACN para obtener una concentración final de TBMB de 10 μM . La reacción se incubó a 30 °C durante 1 hora antes de que el TBMB sin reaccionar se retirase mediante precipitación del fago con un 1/5 de volumen de PEG al 20% y NaCl a 2,5 M en hielo y centrifugación a 4000 rpm durante 30 minutos.

Selección de fagos con calicreína plasmática y catepsina G humanas

La calicreína plasmática y la cathepsina G humanas biotiniladas (5 a 20 µg: el protocolo utilizado para la biotinilación se puede encontrar más adelante) se bloquearon por incubación en 0,5 ml de tampón de lavado (Tris-HCl a 10 mM, pH 7,4, NaCl a 150 mM, MgCl₂ a 10 mM y CaCl₂ a 1 mM) que contenía SAB al 1% y Tween 20 al 0,1% durante 30 minutos. El fago modificado químicamente (típicamente 10¹⁰-10¹¹ u.t. disueltos en 2 ml de tampón de lavado) se bloquearon con la adición de 1 ml de un tampón de lavado que contenía SAB al 3% y Tween 20 al 0,3% e incubación durante 30 minutos. Se pipetearon 3 ml de fagos bloqueados en 0,5 ml de antígeno bloqueado y se incubaron durante 30 minutos en una rueda giratoria a TA. Se bloquearon 50 µl de perlas magnéticas de estreptavidina (Dynal, M-280 de Invitrogen, Paisley, Reino Unido) mediante la incubación en 0,5 ml de tampón de lavado que contenía SAB al 1% y Tween 20 al 0,1% durante 30 minutos. Las perlas bloqueadas se añadieron a la mezcla de fagos/antígenos y se incubó durante 5 minutos a TA en una rueda giratoria. Se lavaron las perlas 8 veces con tampón de lavado que contenía Tween 20 al 0,1% y dos veces con tampón de lavado antes de la incubación con 100 µl de glicina a 50 µM, pH 2,2, durante 5 minutos. Se transfirieron los fagos eluidos a 50 µl de Tris-HCl a 1 M, pH 8, para la neutralización, se incubaron con 50 ml de células TG1 a una DO₆₀₀ = 0,4 durante 90 minutos a 37 °C, y las células se sembraron en placas grandes de cloranfenicol/2YT. Se realizaron dos rondas adicionales de cribado con los mismos procedimientos. En la segunda ronda de selección se utilizaron perlas magnéticas revestidas de neutravidina para impedir el enriquecimiento en péptidos específicos de estreptavidina. Las perlas de neutravidina se prepararon haciendo reaccionar 0,8 mg de neutravidina (Pierce, Rockford, IL, EE.UU.) con 0,5 ml de perlas magnéticas activadas con tosilo (Dynal, M-280 de Invitrogen, Paisley, Reino Unido), de acuerdo con las instrucciones del proveedor.

20 *Escrutinio de la actividad inhibidora en los clones seleccionados*

Los genes que codifican los péptidos seleccionados en la segunda y tercera rondas de biocribado se clonaron en un vector basado en pUC119 para la expresión de las proteínas de fusión péptido-D1-D2 (proteína D1-D2 sin puentes disulfuro; los procedimientos de clonación y expresión se describen más adelante). Los grupos sulfhidrilo oxidados de los péptidos se redujeron incubando la proteína (1 a 10 µM) con TCEP a 1 mM en NH₄HCO₃ a 20 mM, pH 8, a 42 °C durante 1 hora. El reductor se retiró mediante cromatografía de exclusión por tamaños en una columna PD-10 (Amersham Pharmacia, Upsala, Suecia) con el tampón de NH₄HCO₃ a 20 mM y EDTA a 5 mM, a pH 8. Los grupos tiol de las proteínas se hicieron reaccionar incubándolos con TBMB a 10 µM en tampón de reacción (NH₄HCO₃ a 20 mM, EDTA a 5 mM, pH 8, y ACN al 20%) a 30 °C durante 1 hora. Para la retirada del TBMB sin reaccionar y la concentración, se filtró la proteína con un microcon YM-30 (Millipore, Bedford, MA). Se determinó la concentración de los productos midiendo la absorbancia a 280 nm. Se midió la CI₅₀ mediante la incubación de diferentes concentraciones de las proteínas de fusión del péptido modificado (diluciones de 2 veces) con calicreína plasmática humana (0,1 nM) o cathepsina G (20 nM) y determinando la actividad residual en Tris-HCl a 10 mM, pH 7,4, NaCl a 150 mM, MgCl₂ a 10 mM, CaCl₂ a 1 mM, SAB al 0,1% y Tritón-X100 al 0,01%. La actividad de la calicreína plasmática humana se midió con el sustrato fluorógeno Z-Phe-Arg-AMC (Bachem, Bubendorf, Suiza) a una concentración de 100 µM en un lector de placas de fluorescencia Spectramax Gemini (excitación a 355 nm, lectura de la emisión a 460 nm; Molecular Devices, Sunnyvale, CA, EE.UU.). La actividad de la cathepsina G humana se midió con el sustrato colorimétrico N-Suc-Ala-Ala-Phe-Pro-pNA (Bachem, Bubendorf, Suiza) a una concentración de 1 mM con un lector de placas de absorción Spectramax (lectura a 410 nm; Molecular Devices, Sunnyvale, CA, EE.UU.).

40 *Síntesis química de péptidos bicíclicos*

Péptidos con una amina libre en el extremo amino y una amida en el extremo carboxilo se sintetizaron químicamente a una escala a 25 mg mediante química en fase sólida (JPT Peptide Technologies, Berlín, Alemania). Los péptidos brutos en 1 ml de NH₄HCO₃ al 70%, pH 8, y ACN al 30% (1 mM) se hicieron reaccionar con TBMB (1,2 mM) durante 1 hora a TA. El producto de reacción se purificó mediante cromatografía líquida de gran resolución (HPLC) de fase inversa con una columna C18 y elución en gradiente con una fase móvil compuesta de una solución de ACN y ácido trifluoroacético (TFA) acuoso al 0,1% a una velocidad de flujo de 2 ml/min. Los péptidos purificados se liofilizaron y disolvieron en DMSO o en un tampón de Tris-HCl a 50 mM, pH 7,8, y NaCl a 150 mM para las mediciones de actividad.

Clonación y expresión de las proteínas de fusión péptido-D12

Los dominios D1-D2 de la g3p (que comprende los restos de aminoácidos 2 a 217 de la fd-g3p madura) con y sin el péptido **¹ACGSGCGSGCG⁶** fusionado al extremo amino se expresaron en *E.coli*. El vector de expresión basado en pUC119 con una secuencia líder y el gen de D1-D2 con una etiqueta de hexa-histidinas en el extremo carboxilo (aquí denominado pUC119H6D12) fue proporcionado generosamente por Phil Holliger del Laboratorio de Biología Molecular (LMB) de Cambridge. Se clonó un plásmido para la expresión de D1-D2 con el péptido en el extremo amino mediante amplificación por PCR del gen de D1-D2 con los cebadores pepd12ba (que codifica la secuencia peptídica) y d12fo y la ligación en el vector pUC119H6D12 digerido con *Sfi*//*Not*I. El gen para la expresión de D1-D2

sin puentes disulfuro con un total de 20 aminoácidos se amplificó por PCR a partir del vector fdg3p0ss21 con alguna de las parejas de cebadores d120ssba/d120ssfo, pepd120ssba/d120ssfo, P2cd120ssba/d120ssfo o P1cd120ssba/d120ssfo y se ligó en *Sfil/NotI* en el pUC119H6D12 para la expresión del D1-D2 sin puentes disulfuro con y sin los péptidos ¹*ACGSGCGSGCGC*, ²*AGSGCGSGCGC* o ³*AGSGKSGSGCGC* fusionados al extremo amino.

- 5 Las 6 proteínas se expresaron en las células de *E. coli* TG1 a 30 °C durante 8 horas y se purificó la fracción periplasmática de manera escalonada mediante cromatografía de afinidad de Ni y filtración en gel en una columna Superdex 75 en HN₄HCO₃, pH 7,4.

Análisis espectrométrico de masas de las proteínas de fusión péptido-D12

- 10 La masa molecular de las proteínas (5 a 20 µM) antes y después de la modificación con TBMB se determinó desnaturizando las proteínas en 4 volúmenes de MeOH al 50% y ácido fórmico al 1%, y mediante el análisis en un espectrómetro de masas de tiempo de vuelo con ionización por electropulverización (Micromass, Milford, MA, EE.UU.). Se obtuvieron las masas moleculares por desconvolución de los espectros de masas de las proteínas con varias cargas utilizando la versión 4.1 de MassLynx.

Creación de la genoteca 1 de péptidos en fagos

- 15 Los genes que codifican un péptido semialeatorio con la secuencia Ala-Cys-(Xaa)₆-Cys-(Xaa)₆-Cys, el conector Gly-Gly-Ser-Gly y los dos dominios sin puentes disulfuro D1 y D2 se clonaron en la orientación correcta en el vector fágico fd0D12 para obtener la genoteca en fagos 1. El vector fd0D12, que carece de los genes de los dominios D1 y D2 del gen 3 y que tiene un segundo sitio de restricción *Sfil* que se creó previamente mediante la amplificación por PCR del plásmido fdg3p0ss21 completo utilizando los cebadores ecoG3pNba y pelbsfiecofo. Los
- 20 genes que codifican el repertorio de péptidos y los dos dominios del gen 3 se crearon escalonadamente en dos PCR consecutivas. Primero, los genes D1 y D2 se amplificaron por PCR con los dos cebadores prepqr y sfi2fo con el vector fdg3p0ss21 como plantilla. Segundo, el ADN que codifica los péptidos aleatorios se añadió en una PCR con los cebadores sfix6ba y sfi2fo. La ligación de 33 y 9 µg del plásmido fd0D12 digerido con *Sfil* y el producto de PCR produjo $4,4 \times 10^9$ colonias en 12 placas de 20 x 20 cm de 2YT con cloranfenicol (30 µg/ml). Se rasparon
- 25 colonias de las placas con medio 2YT, complementado con glicerol al 15% y se almacenaron a -80 °C. Se diluyeron las reservas en glicerol a una DO₆₀₀ = 0,1 en cultivos de 1 l de 2YT/cloranfenicol (30 µg/ml) y los fagos se expresaron a 30 °C durante una noche (12 a 16 horas).

Biotinilación de los antígenos

- 30 La calicreína plasmática humana (activada con el factor XIIa) se compró a Innovative Research (Southfield, MI, EE.UU.) y se biotiniló a una concentración de 1,2 µM con un exceso molar de 5 veces de Sulfo-NHS-LC-biotina (Pierce, Rockford, IL, EE.UU.) en PBS, pH 7,4, y DMSO al 5% a TA durante una hora. La proteína biotinilada se purificó en una columna PD-10 con un tampón de NaAC a 50 mM, pH 5,5, y NaCl a 200 mM. La catepsina G humana se compró ya biotinilada a Lee Biosolutions (San Louis, MI, EE.UU.).

Subclonación y escrutinio de la expresión de los clones seleccionados de fagos

- 35 El ADN plasmídico de los clones seleccionados después de la segunda y tercera rondas de biocribado se amplificó por PCR en un solo tubo con los cebadores 21seqba y flagfo, y se clonó en el vector pUC119H6D12 en los sitios *Sfil* y *NotI* para la expresión periplasmática de los péptidos fusionados a los dominios D1-D2 sin puentes disulfuro con una etiqueta FLAG y una etiqueta de hexa-histidinas en el extremo carboxilo. Los plásmidos ligados se electroporaron en las células TG1 y se sembraron en placas de 2YT/ampicilina (100 µg/ml). Los clones que
- 40 expresaban la proteína recombinante se identificaron como sigue: el medio (2YT con ampicilina a 100 µg/ml) en placas de 96 pocillos profundos (1 ml/pocillo) se inoculó con células de colonias diferentes y se incubó a 37 °C. Se indujo la expresión de las proteínas con IPTG a 1 mM cuando los cultivos eran turbios y se agitaron las placas a 300 rpm a 30 °C durante una noche. Se sedimentaron las células mediante centrifugación a 3500 rpm durante 30 minutos, se lisaron y se lavaron con tampón de lavado que contenía lisozima a 1 mg/ml y se centrifugaron a 3500
- 45 rpm para sedimentar los residuos celulares. Los sobrenadantes se transfirieron a placas Polysorp de 96 pocillos (Nunc. Roskilde, Dinamarca) para la absorción inespecífica. Los pocillos se enjuagaron dos veces con tampón de lavado que contenía Tween 20 al 0,1% y se bloquearon con tampón de lavado que contenía SAB al 1% y Tween 20 al 0,1% durante 1 hora. El conjugado anti-FLAG M2-peroxidasa (Sigma-Aldrich, San Louis, MO, EE.UU.) se diluyó 1:5000 y se bloqueó en tampón de lavado que contenía SAB al 1% y Tween 20 al 0,1%, y se añadió a las placas
- 50 durante 1 hora. Se lavaron los pocillos (5 veces con tampón de lavado que contenía Tween 20 al 0,1% y una vez sin detergente) y la peroxidasa unida se detectó con la solución de TMB, su sustrato (eBiosciences, San Diego, EE.UU.). El ADN plasmídico de los clones que expresan la proteína se secuenció (GeneServices, Cambridge, Reino Unido).

Maduración de la afinidad de los inhibidores de la calicreína plasmática humana

Se crearon tres genotecas de péptidos en fagos esencialmente como la genoteca 1 (véase más arriba), pero con los cebadores degenerados sficx6abc (genoteca 2), sficx6abb (genoteca 3) y sficx6aba (genoteca 4) en vez del sficx6ba. La electroporación de las reacciones de ligación en las células TG1 produjo $9,5 \times 10^8$ (genoteca 2), $1,1 \times 10^9$ (genoteca 3) y $1,2 \times 10^9$ (genoteca 4) transformantes. Los fagos de cada genoteca se produjeron en cultivos de 1 l, se purificaron, se agruparon y se hicieron reaccionar con TBMB. Se realizaron tres ciclos de cribado esencialmente como en las selecciones descritas en el apartado de materiales y métodos, pero con la calicreína plasmática humana biotinilada a una concentración inferior (1 nM en la primera y segunda rondas, 200 pM en la tercera ronda).

10 *Medición de la actividad y de la especificidad de los inhibidores de la calicreína plasmática humana*

Se determinaron las actividades inhibitoras (CI_{50}) con la medición de la actividad residual de la enzima tras la incubación (30 min, TA) con diferentes concentraciones del inhibidor (que oscila típicamente de 10 μ M a 0,5 nM). La actividad de la calicreína plasmática humana (0,1 nM) y del factor XIa (0,8 nM; Innovative Research, Southfield, MI, EE.UU.) se midieron con Z-Phe-Arg-AMC (100 μ M) y la actividad de la trombina humana (2 nM; Innovative Research, Southfield, MI, EE.UU.), con Boc-Phe-Ser-Arg-AMC (100 μ M) en Tris-HCl a 10 mM, pH 7,4, NaCl a 150 mM, $MgCl_2$ a 10 mM, $CaCl_2$ a 1 mM, SAB al 0,1%, Triton X-100 al 0,01% y DMSO al 5%. La calicreína plasmática recombinante de ratón de R & D Systems (Minneapolis, MN, EE.UU.) con un péptido señal se activó proteolíticamente con termolisina a 0,5 mg/ml a 37 °C durante 1 hora. Se midió la actividad de la calicreína plasmática de ratón (3 nM) con Z-Phe-Arg-AMC (100 μ M) en Tris-HCl a 50 mM, pH 7,5, $CaCl_2$ a 10 mM, NaCl a 250 mM, SAB al 0,1%, Triton X-100 al 0,01% y DMSO al 5%. El inhibidor que tiene hidrolizado un lazo de unión se generó mediante la incubación del péptido PK15 modificado con TBMB con la calicreína plasmática humana a una proporción molar de 5:1 durante 24 horas a 37 °C y la posterior inactivación por calor de la enzima a 60 °C (30 min). Se calcularon los valores de K_i aparentes de acuerdo con la ecuación de Cheng y Prusoff.

Medición de la activación del factor XII en el plasma humano

25 El plasma humano normal de donantes diferentes se adquirió a 3H Biomedical (Upsala, Suecia). El plasma se centrifugó a 1500xg a 20 °C durante 15 minutos para obtener plasma pobre en plaquetas (PPP). Se conservaron las alícuotas del PPP en tubos de polipropileno a -80 °C. Se prepararon muestras de PPP de 60 μ l con 5, 50, 500 o 5000 nM de aprotinina (Roche, Mannheim, Alemania) o el péptido PK15 modificado con TBMB. La activación del factor XIIa se midió como sigue. Se añadieron 2 μ g de caolín a las muestras plasmáticas, se mezclaron bien y se incubaron durante 20 minutos a 37 °C. Se diluyeron las muestras 250 veces en Tris-HCl a 50 mM, pH 7,8, y NaCl a 150 mM. La actividad de tipo calicreína plasmática se midió con el sustrato cromógeno H-D-Pro-Phe-Arg-pNA (100 μ M; Bachem, Bubendorf, Suiza) utilizando un lector de placas de absorción (absorción a 450 nm; Molecular Devices, Sunnyvale, CA, EE.UU.). El mismo sustrato cromógeno también lo reconoce y modifica la calicreína plasmática. Sin embargo, a las concentraciones requeridas del inhibidor para reducir la actividad del factor XIIa al 50% (160 nM para el péptido PK15 modificado con TBMB y 5 μ M para la aprotinina), la calicreína plasmática se inhibe esencialmente y no se puede medir con el sustrato.

Determinación de la estructura del péptido PK15 modificado con TBMB

40 Se disolvió 1 mg del péptido PK15 modificado con TBMB en 550 μ l de Tris-HCl deuterado a 10 mM, pH 6,6, NaCl a 150 mM, $MgCl_2$ a 10 mM y $CaCl_2$ a 1 mM para obtener una concentración de inhibidor de 1 mM. Se leyeron los espectros del inhibidor a 800 MHz (Bruker Avance con una criosonda de TCI). Las asignaciones espectrales se basaron en los espectros de TOCSY y de NOESY. Había en total 199 restricciones NOE, 77 de las cuales eran entre residuos, y 122 intraresiduales. La estructura mostrada en la figura 6 es la estructura promediada de la estructura calculada para 50 conformómeros. Se utilizó el programa PyMOL para el análisis de la estructura y la visualización de los modelos moleculares.

45

Bibliografía para el apartado de materiales y métodos:

1. Wu, P., Leinonen, J., Koivunen, E., Lankinen, H. y Stenman, U. H. Identification of novel prostate-specific antigen-binding peptides modulating its enzyme activity. *Eur J Biochem* 267, 6212-20 (2000).
2. Hekim, C. et al. Novel peptide inhibitors of human kallikrein 2. *J Biol Chem* 281, 12555-60 (2006).
- 50 3. Hansen, M. et al. A urokinase-type plasminogen activator-inhibiting cyclic peptide with an unusual P2 residue and an extended protease binding surface demonstrates new modalities for enzyme inhibition. *J Biol Chem* 280,

38424-37 (2005).

4. Andersen, L. M., Wind, T., Hansen, H. D. y Andreassen, P. A. A cyclic peptidyl inhibitor of murine urokinase-type plasminogen activator: changing species specificity by substitution of a single residue. *Biochem J* 412, 447-57 (2008).
- 5 5. Krook, M., Lindbladh, C., Eriksen, J. A. y Mosbach, K. Selection of a cyclic nonapeptide inhibitor to alpha-chymotrypsin using a phage display peptide library. *Mol Divers* 3, 149-59 (1997).
6. Dennis, M. S. et al. Peptide exosite inhibitors of factor VIIa as anticoagulants. *Nature* 404, 465-70 (2000).
7. Huang, L. et al. Novel peptide inhibitors of angiotensin-converting enzyme 2. *J Biol Chem* 278, 15532-40 (2003).
8. Karaseva, N. G., Glinisky, V. V., Chen, N. X., Komatireddy, R. y Quinn, T. P. Identification and characterization of peptides that bind human ErbB-2 selected from a bacteriophage display library. *J Protein Chem* 21, 287-96 (2002).
- 10 9. Houimel, M., Mach, J. P., Cortesy-Theuloz, I., Cortesy, B. y Fisch, I. New inhibitors of *Helicobacter pylori* urease holoenzyme selected from phage-displayed peptide libraries. *Eur J Biochem* 262, 774-80 (1999).
- 10 10. Lunder, M., Bratkovic, T., Kreft, S. y Strukelj, B. Peptide inhibitor of pancreatic lipase selected by phage display using different elution strategies. *J Lipid Res* 46, 1512-6 (2005).
- 15 11. Sanschagrin, F. y Levesque, R. C. A specific peptide inhibitor of the class B metallo-beta-lactamase L-1 from *Stenotrophomonas maltophilia* identified using phage display. *J Antimicrob Chemother* 55, 252-5 (2005).
12. Sperinde, J. J., Choi, S. J. y Szoka, F. C., Jr. Phage display selection of a peptide DNase II inhibitor that enhances gene delivery. *J Gene Med* 3, 101-8 (2001).

20

Bibliografía del ejemplo 8

1. Hüser, J. *High-Throughput Screening in Drug Discovery* (eds. Mannhold, R., Kubinyi, H. y Folkers, G.) (Wiley-VCH, Weinheim, 2006).
2. Bleicher, K. H., Bohm, H. J., Muller, K. y Alanine, A. I. Hit and lead generation: beyond high-throughput screening. *Nat Rev Drug Discov* 2, 369-78 (2003).
3. Marks, J. D., Hoogenboom, H. R., Griffiths, A. D. y Winter, G. Molecular evolution of proteins on filamentous phage. Mimicking the strategy of the immune system. *J Biol Chem* 267, 16007-10 (1992).
4. Jespers, L. S., Roberts, A., Mahler, S. M., Winter, G. y Hoogenboom, H. R. Guiding the selection of human antibodies from phage display repertoires to a single epitope of an antigen. *Biotechnology (N Y)* 12, 899-903 (1994).
- 30 5. Hudson, P. J. y Souriau, C. Engineered antibodies. *Nat Med* 9, 129-34 (2003).
6. Doyon, J. B., Snyder, T. M. y Liu, D. R. Highly sensitive in vitro selections for DNA-linked synthetic small molecules with protein binding affinity and specificity. *J Am Chem Soc* 125, 12372-3 (2003).
7. Woiwode, T. F. et al. Synthetic compound libraries displayed on the surface of encoded bacteriophage. *Chem Biol* 10, 847-58 (2003).
- 35 8. Brenner, S. y Lerner, R. A. Encoded combinatorial chemistry. *Proc Natl Acad Sci USA* 89, 5381-3 (1992).
9. Halpin, D. R. y Harbury, P. B. DNA display II. Genetic manipulation of combinatorial chemistry libraries for small-molecule evolution. *PLoS Biol* 2, E174 (2004).
10. Jespers, L., Bonnert, T. P. y Winter, G. Selection of optical biosensors from chemisynthetic antibody libraries. *Protein Eng Des Sel* 17, 709-13 (2004).
- 40 11. Jespers, L. S. A., Winter, G. P., Bonnert, T. P. y Simon, T. M. (PCT/GB94/01422).
12. Timmerman, P., Beld, J., Puijk, W. C. y Meloen, R. H. Rapid and quantitative cyclization of multiple peptide

loops onto synthetic scaffolds for structural mimicry of protein surfaces. *ChemBiochem* 6, 821-4 (2005).

13. Driggers, E. M., Hale, S. P., Lee, J. y Terrett, N. K. The exploration of macrocycles for drug discovery: an underexploited structural class. *Nat Rev Drug Discov* 7, 608-24 (2008).

14. Wessjohann, L. A., Ruijter, E., Garcia-Rivera, D. y Brandt, W. What can a chemist learn from nature's
5 macrocycles? A brief, conceptual view. *Mol Divers* 9, 171-86 (2005).

15. Kemp, D. S. y McNamara, P. E. Conformationally restricted cyclic nonapeptides derived from L-cysteine and LL-3-amino-2-piperidino-6-carboxylic acid (LL-acp), a potent b-turn-inducing dipeptide analogue. *Journal of Organic Chemistry* 50, 5834-5838 (1985).

16. Kather, I., Bippes, C. A. y Schmid, F. X. A stable disulfide-free gene-3-protein of phage fd generated by in vitro
10 evolution. *J Mol Biol* 354, 666-78 (2005).

17. Cheng, Y. y Prusoff, W. H. Relationship between the inhibition constant (K_i) and the concentration of inhibitor which causes 50 per cent inhibition (I_{50}) of an enzymatic reaction. *Biochem Pharmacol* 22, 3099-108 (1973).

18. Cremllyn, R. J. *An introduction to organosulfur chemistry* (Wiley, 1996).

19. Huang, L. et al. Novel peptide inhibitors of angiotensin-converting enzyme 2. *J Biol Chem* 278, 15532-40
15 (2003).

20. Dennis, M. S. et al. Peptide exosite inhibitors of factor VIIa as anticoagulants. *Nature* 404, 465-70 (2000).

21. Jackson, D. Y. et al. A designed peptide ligase for total synthesis of ribonuclease A with unnatural catalytic residues. *Science* 266, 243-7 (1994).

22. Abbenante, G. y Fairlie, D. P. Protease inhibitors in the clinic. *Med Chem* 1, 71-104 (2005).

20 23. Turk, B. Targeting proteases: successes, failures and future prospects. *Nat Rev Drug Discov* 5, 785-99 (2006).

24. Melkko, S., Scheuermann, J., Dumelin, C. E. y Neri, D. Encoded self-assembling chemical libraries. *Nat Biotechnol* 22, 568-74 (2004).

25. Li, S. y Roberts, R. W. A novel strategy for in vitro selection of peptide-drug conjugates. *Chem Biol* 10, 233-9 (2003).

25 26. Millward, S. W., Takahashi, T. T. y Roberts, R. W. A general route for post-translational cyclization of mRNA display libraries. *J Am Chem Soc* 127, 14142-3 (2005).

27. Millward, S. W., Fiacco, S., Austin, R. J. y Roberts, R. W. Design of cyclic peptides that bind protein surfaces with antibody-like affinity. *ACS Chem Biol* 2, 625-34 (2007).

30 Todas la publicaciones mencionadas en la especificación anterior se incorporan en la presente memoria por referencia. Las distintas modificaciones y variaciones de los aspectos descritos y de las realizaciones de la presente invención serán obvios para los expertos en la técnica sin separarse del alcance de la presente invención. Aunque la presente invención se ha descrito en conexión con realizaciones preferentes específicas, se debe entender que la invención, tal y como se reivindica, no debe estar indebidamente limitada a tales realizaciones específicas. De hecho, se pretende que se encuentren dentro del alcance de las reivindicaciones anejas las
35 diferentes modificaciones de los modos descritos para realizar la invención, los cuales son obvios para los expertos en la técnica.

50

**5'TATGCGGCCAGCCGGCCATGGCAGCGTGTNNKNNKNNKNNKNNKNNKTGCCGCGTT
AATTGGACCCCGTGTGGCGGTTCTGGCGCTG-3'**

sficx6abb

**5'TATGCGGCCAGCCGGCCATGGCAGCGTGTGCGCCGTGGCGCACCCGCGTGCNN
KNNKNNKNNKNNKNNKTGTGGCGGTTCTGGCGCTG-3'**

sficx6abc

**5'TATGCGGCCAGCCGGCCATGGCAGCGTGTAGCGATCGTTTTCGTAATTGCNNKNNK
NNKNNKNNKNNKTGTGGCGGTTCTGGCGCTG-3'**

5

Las secuencias de péptidos son tal y como se describe en la sección de ejemplos

Secuencias de cebadores de ADN del Ejemplo 8

Los sitios de restricción de nucleasa están subrayados

10

pepd12ba	5'-CTCGCGGCCAGCCGGCCATGGCAGCCTGTGGTC TGGCTGGGTTCGGCTGTGGCAGAAAGAGTTG-3'
d12ba	5'-GAGTCATTCTGCGGCCGATTGACAGGAGGTTGAGGCAGGTA-3'
d120ssba	5'-CATGCCATGACTCGGGGCCAGCCGGCCATGGCTGAACTGTTGAAAGTAG-3'
d120ssba	5'-GAGTCATTCTGCGGCCGATTGACAAGGAGGTTGAGGCAGGTA-3'
pepd120ssba	5'- CTCGCGGCCAGCCGGCCATGGCAGCGTGTCCCTCTCGGCTGCCCTTCCGGCTGTCTGGCTGAACTGTTGAAAGTAG -3'
p2cd120ssba	5'-CTCGCGGCCAGCCGGCCATGGCAGCGGGCTCTGGCTTCCGGCTGTGCTGAAACGTTGAAAGTAG-3'
plod120ssba	5'-CTCGCGGCTAGCCGCCATGGCATGGCAGCCCCCTCTGGAAGGTTCCGGCTCTGGGTCTTGAAGTA-3'
ocog3pttoo	5'-GCATGAATTCCAGTCAGTACGGCCTCGGGCGCCATGGCTTCTGGTACCCCGTTAAC-3'
po	5'GCATGAATTCCGATCCGAGCCGGTGGGCCCATAGGGAAGGCAACTAAAGGAAT-3'
propotoo	5'-GGCGTTCTGGCGTGAACTCTTGAAAGTAG-3'
d121o	5'-GAAGCCATGGCCCCGAGGCCCGAGCCGAGCATTGACAGG-3'
uficarbo	5'-TATGCGGCCCGCCAGCCGCCATGGCTCGAGCGTGC5'- GAGTCATTCTGCGGCCGATTGACAGGAGGTTGAGGCAGGTG-3'
nopto	5'-GATTCTGCGGCCGATTGATAGGAGGTTGAGGCAGGTA-3'
dic	5'-TATGCGGCCAGCCGGCCATGGCACCTGC GGCCGATTGACAGGAG GCCCCGATTAATTGGACCCCGTGTGGCCGGCCCTTCTGGCCGCTG-3'
disc	5'-TATGCGGCCAGCCGGCCATGGCACCTGCAGCGTTGTGGCGCCGTGGCCAGCGGCTGGC CTGTGGCGCCGTTCTGGCGCCTGG-3'
eficxdssba	5'TATGCGGCCAGCCGGCCATGGCAGTGTCTAGCGATCGTTTTCGTAAT

REIVINDICACIONES

1. Complejo que comprende una partícula de fago, comprendiendo dicha partícula de fago
 - (i) un ácido nucleico que codifica un polipéptido;
 - (ii) el polipéptido expresado por el ácido nucleico de (i) y expuesto en la superficie del fago;
- 5 (iii) un compuesto conector unido a dicho polipéptido

en el que dicho compuesto conector está unido al polipéptido mediante al menos tres enlaces covalentes independientes.
2. Complejo de acuerdo con la reivindicación 1, en el que el compuesto conector tiene una simetría molecular que corresponde al número de enlaces covalentes mediante los cuales se une al polipéptido.
- 10 3. Complejo de acuerdo con la reivindicación 2, en el que el compuesto conector posee simetría molecular ternaria y el compuesto conector está unido al polipéptido mediante tres enlaces covalentes.
4. Complejo de acuerdo con una cualquiera de las reivindicaciones anteriores, en el que el compuesto conector comprende un grupo químico estructuralmente rígido.
5. Complejo de acuerdo con la reivindicación 4, en el que el compuesto conector comprende

15 tris(bromometil)benceno (TBMB).
6. Complejo de acuerdo con una cualquiera de las reivindicaciones anteriores, en donde dicho polipéptido comprende un resto de cisteína, y en donde al menos uno de dichos tres enlaces covalentes independientes para la unión de dicho compuesto conector al polipéptido comprende un enlace a dicho resto de cisteína.
7. Genoteca de polipéptidos codificados genéticamente, que comprende al menos dos complejos diferentes de

20 acuerdo con una cualquiera de las reivindicaciones anteriores.
8. Método para fabricar un complejo, comprendiendo dicho método
 - (i) el aporte de una partícula de fago que comprende un ácido nucleico, y un polipéptido expresado por el ácido nucleico y expuesto sobre la partícula del fago;
 - (ii) el aporte de un compuesto conector
- 25 (iii) la unión de dicho compuesto conector a dicho polipéptido mediante la formación de al menos tres enlaces covalentes entre dicho compuesto conector y el polipéptido.
9. Método de acuerdo con la reivindicación 8, en el que los grupos reactivos del polipéptido expuesto en la partícula de fago están reducidos, y en el que las partículas de fagos que exponen el polipéptido que comprende grupos reactivos reducidos se purifica por filtración antes de la etapa (iii).
- 30 10. Método de acuerdo con la reivindicación 9, en el que después de la etapa de purificación por filtración se mantiene el polipéptido en el estado reducido para formar un enlace con el compuesto conector mediante la incubación en tampón desgasificado y en presencia de un quelante.
11. Método de acuerdo con una cualquiera de las reivindicaciones 8 a 10, en el que la etapa (iii) comprende la incubación del polipéptido y del compuesto conector juntos a 30 °C a pH 8 en tampón acuoso que comprende

35 acetonitrilo.
12. Método de acuerdo con una cualquiera de las reivindicaciones 8 a 11, en el que el compuesto conector comprende tris-(bromometil)benceno (TBMB).
13. Método de acuerdo con la reivindicación 12, en el que el tris-(bromometil)benceno está presente a 10 µM.
14. Método de acuerdo con la reivindicación 13, en el que el quelante es el ácido etilendiaminotetraacético (EDTA),

40 el acetonitrilo está presente al 20% y la etapa de incubación (iii) se realiza durante 1 hora.
15. Método de acuerdo con una cualquiera de las reivindicaciones 8 a 14, comprendiendo dicho método la etapa adicional de (iv) escindir uno o más enlaces de la cadena polipeptídica.
16. Método de acuerdo con la reivindicación 15, en el que dicha etapa de escisión comprende poner en contacto

dicho polipéptido con una proteasa.

17. Complejo obtenido mediante el método de una cualquiera de las reivindicaciones 8 a 16.

18. Método para identificar un complejo de acuerdo con una cualquiera de las reivindicaciones anteriores, que es capaz de fijarse a un ligando, comprendiendo el método

- 5 (i) el aporte de un complejo de acuerdo con una cualquiera de las reivindicaciones anteriores
- (ii) puesta en contacto de dicho complejo con el ligando, y
- (iii) selección de los complejos que se fijan a dicho ligando.

19. Método de acuerdo con la reivindicación 18, que además comprende la determinación de la secuencia del ácido nucleico de dicho complejo.

- 10 20. Método de acuerdo con la reivindicación 18, que además comprende la etapa de fabricar el complejo aislado con capacidad para fijarse a dicho ligando.

21. Método de acuerdo con la reivindicación 18, que además comprende la etapa de fabricar un conjugado entre el compuesto conector y el polipéptido que se aísla o identifica mediante un método de la invención, comprendiendo dicha fabricación la unión del compuesto conector al polipéptido, en donde dicho polipéptido se expresa

- 15 recombinantemente o se sintetiza químicamente.

22. Método de acuerdo con la reivindicación 21, que además comprende la etapa de extender el polipéptido en uno o más del extremo amino o del extremo carboxilo del polipéptido.

23. Método de acuerdo con una cualquiera de las reivindicaciones 21 o 22, que además comprende la etapa de conjugar a un polipéptido adicional dicho conjugado de polipéptido y compuesto conector.

- 20 24. Método de acuerdo con la reivindicación 23, en el que dicha conjugación se realiza

- (i) añadiendo otra cisteína al polipéptido después de enlazarse con el compuesto conector, y
- (ii) conjugando dicho polipéptido a dicho polipéptido adicional mediante la formación de puente disulfuro con dicha cisteína adicional.

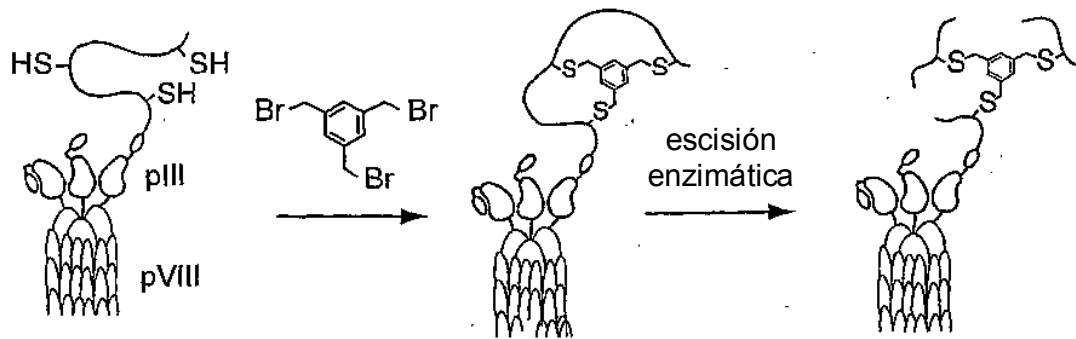


Fig. 1

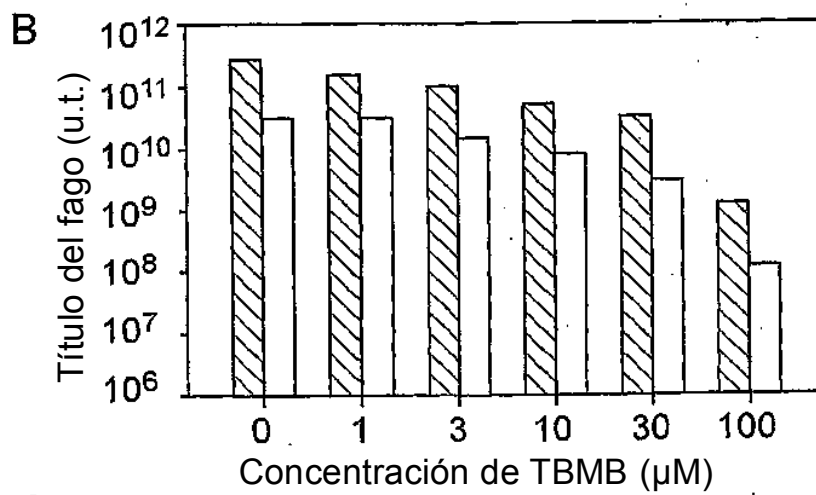
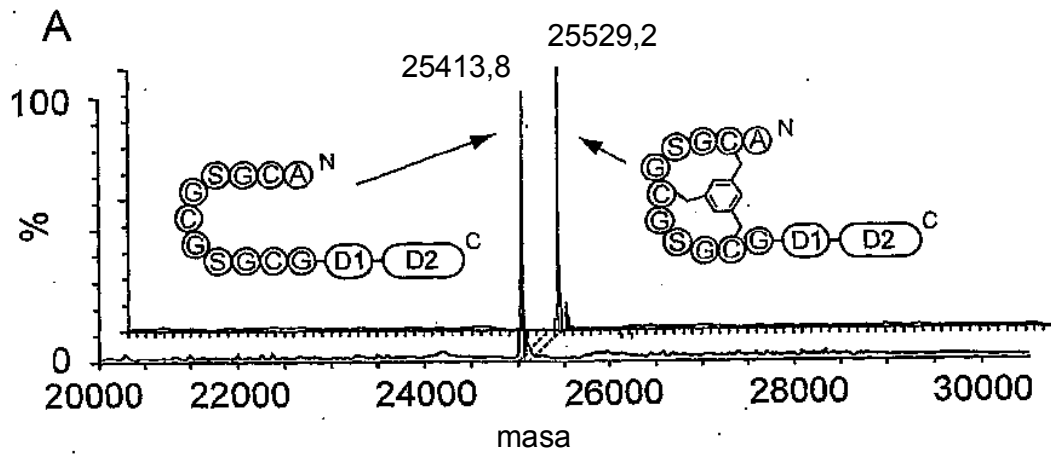
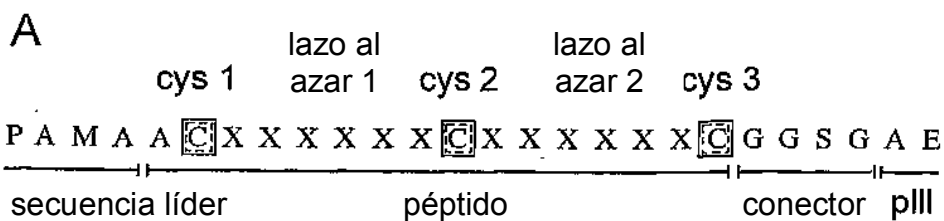


Fig. 2



B

Mutante:	Secuencia de aminoácidos:	Cl ₅₀ (nM):
PK1	A [C] [D] S R F R N [C] P W S M S G [C] G	1500
PK2	A [C] S D R F R N [C] P L W S G T [C] G	400
PK3	A [C] R D I R F R C N Y D V A V [C] G	
PK4	A [C] S T E R R Y [C] P I E I F P [C] G	400
PK5	A [C] A P W R T M [C] L N I D G P [C] G	
PK6	A [C] A P W R T A C Y E D L M W [C] G	500
PK7	A [C] R V W R T L C M E S E G V [C] G	
PK8	A [C] F P L W R T [C] V H E P T M [C] G	
PK9	A [C] W Q V Q V N [C] R V N F G K [C] G	800
PK10	A [C] G G N S D R [C] R V N N I S [C] G	2000
PK11	A [C] G R G D Q T [C] R V N W H [C] G	
PK12	A [C] - - V H N Y [C] R V N W V T [C] G	1800
PK13	A [C] G T G E G R [C] R V N W T [C] G	500

C

Mutante:	Secuencia de aminoácidos:	Cl ₅₀ (nM):
CG1	A [C] E Y G D L W [C] G W D P P V [C] G	
CG2	A [C] E Y D V G F [C] W D G F G Q [C] G	100
CG3	A [C] L F D A G F [C] Q Q H S T E [C] G	
CG4	A [C] I F D L G F [C] H N D W W N [C] G	100
CG5	A [C] L F D L G F [C] G G G E G P [C] G	150
CG6	A [C] P R I E G F [C] L P I F S D [C] G	1000
CG7	A [C] L R A Q E D [C] V Y D R G F [C] G	200
CG8	A [C] T R G S G D [C] T Y D F G F [C] G	200

Fig. 3

A

Genoteca 2

P A M A A [C] S D R F R N [C] X X X X X X [C] G G S G A E

Genoteca 3

P A M A A [C] A P W R T A [C] X X X X X X [C] G G S G A E

Genoteca 4

P A M A A [C] X X X X X X [C] R V N W T P [C] G G S G A E

secuencia líder

péptido

conector pIII

B

Mutante: Secuencia de aminoácidos: Cl₅₀ (nM):

PK14	A	[C]	S	D	R	F	R	N	[C]	P	A	D	E	R	E	[C]	G
PK15	A	[C]	S	D	R	F	R	N	[C]	P	A	D	E	A	L	[C]	G
PK16	A	[C]	S	D	R	F	R	N	[C]	P	V	D	E	A	L	[C]	G
PK17	A	[C]	S	D	R	F	R	N	[C]	P	V	D	E	W	L	[C]	G
PK18	A	[C]	S	D	R	F	R	N	[C]	P	G	D	E	S	L	[C]	G
PK19	A	[C]	S	D	R	F	R	N	[C]	P	Y	T	L	H	D	[C]	G
PK20	A	[C]	S	D	R	F	R	N	[C]	P	Y	V	S	S	D	[C]	G
PK21	A	[C]	S	D	R	F	R	N	[C]	P	Y	S	E	G	D	[C]	G
PK22	A	[C]	S	D	R	F	R	N	[C]	P	V	W	D	S	S	[C]	G
PK23	A	[C]	S	D	R	F	R	N	[C]	P	V	S	E	S	A	[C]	G

Fig. 4

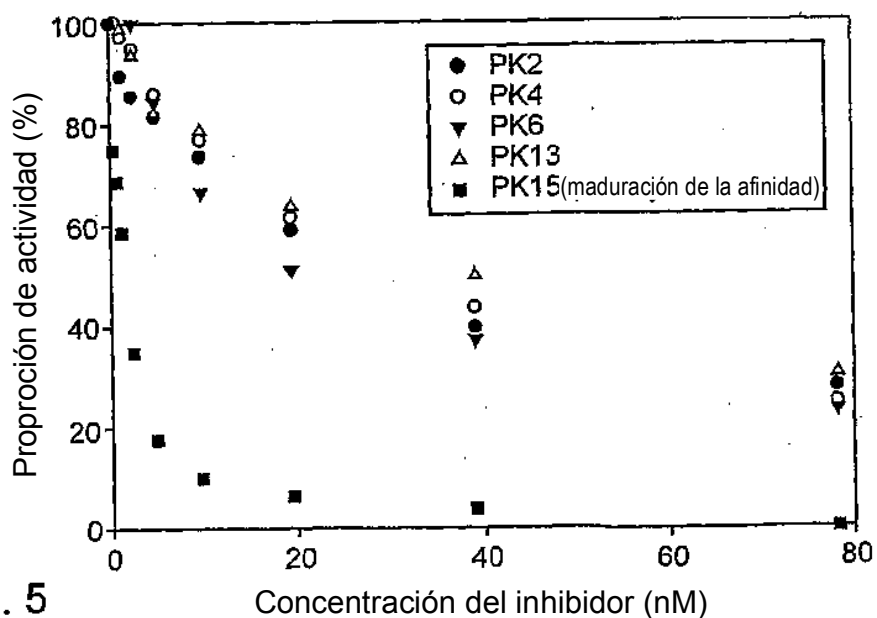
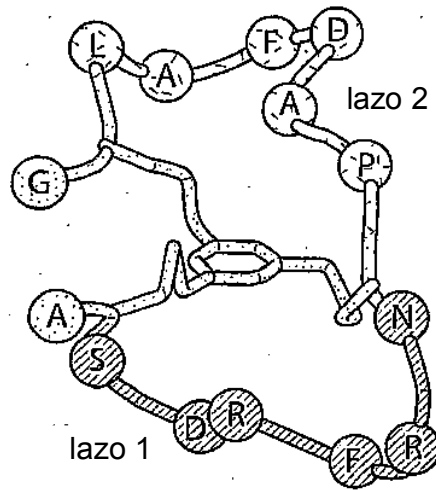
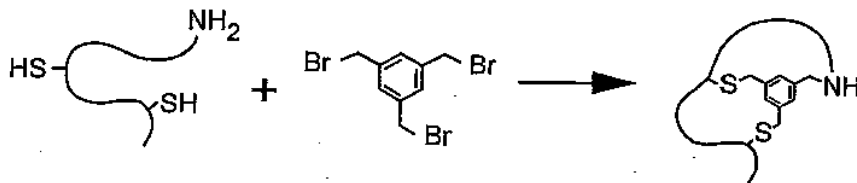


Fig. 5

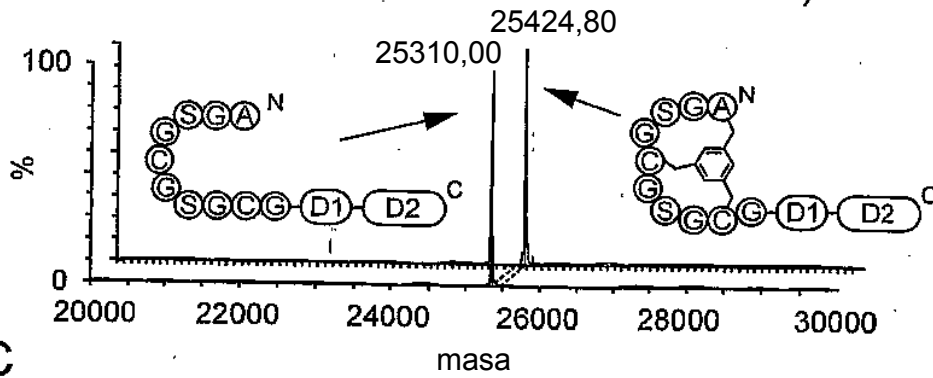
Fig. 6



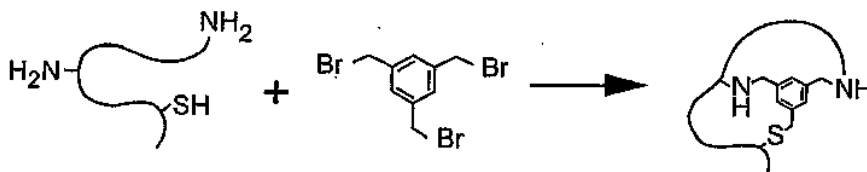
A



B



C



D

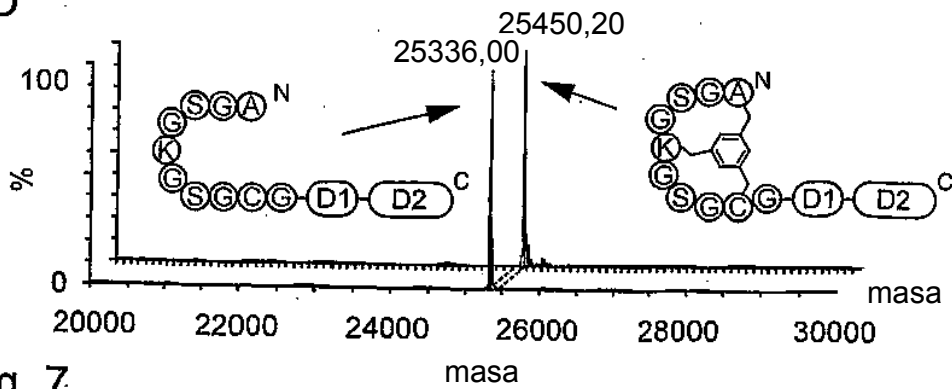


Fig. 7

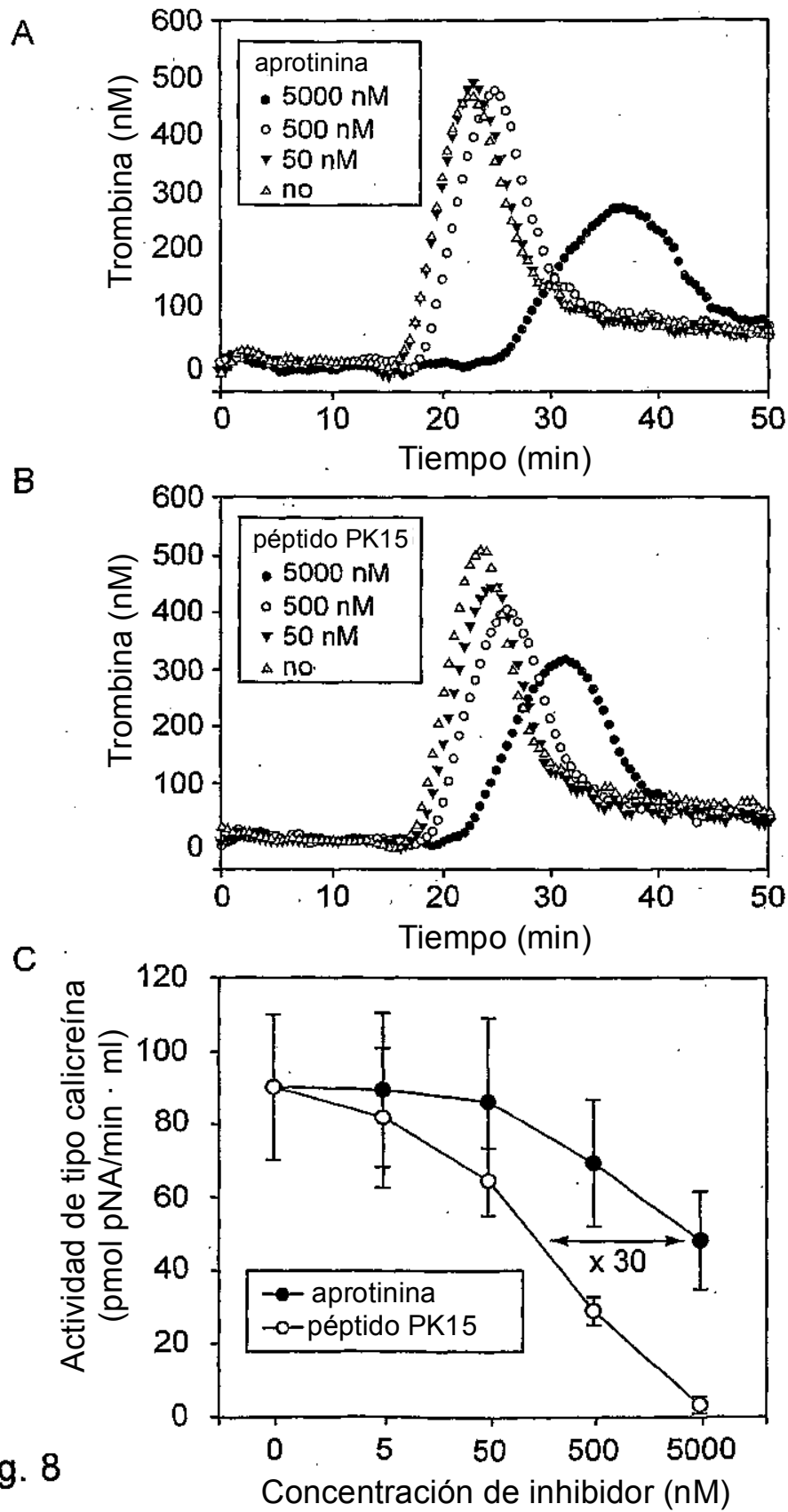
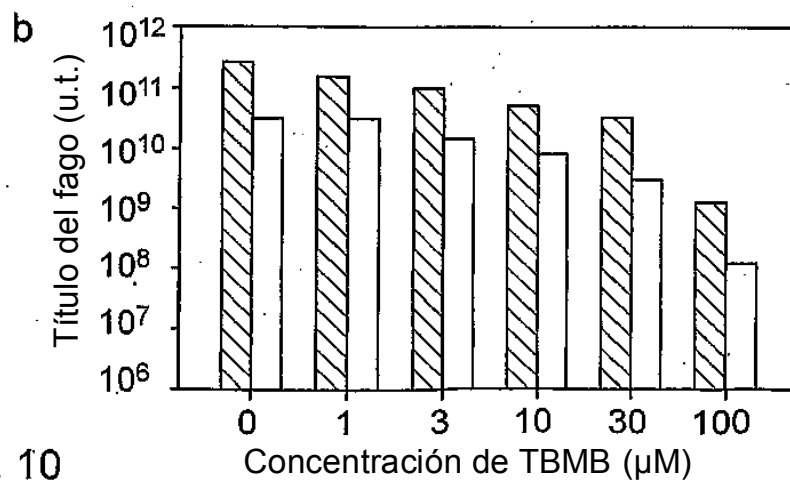
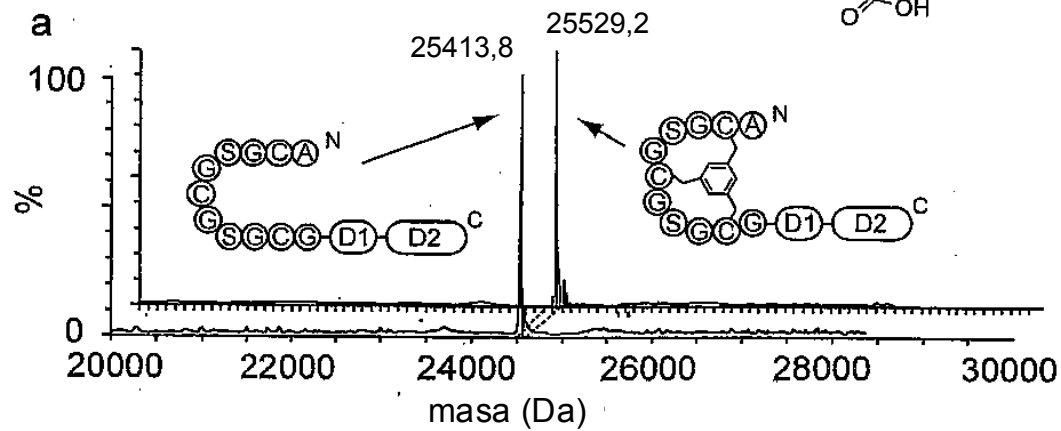
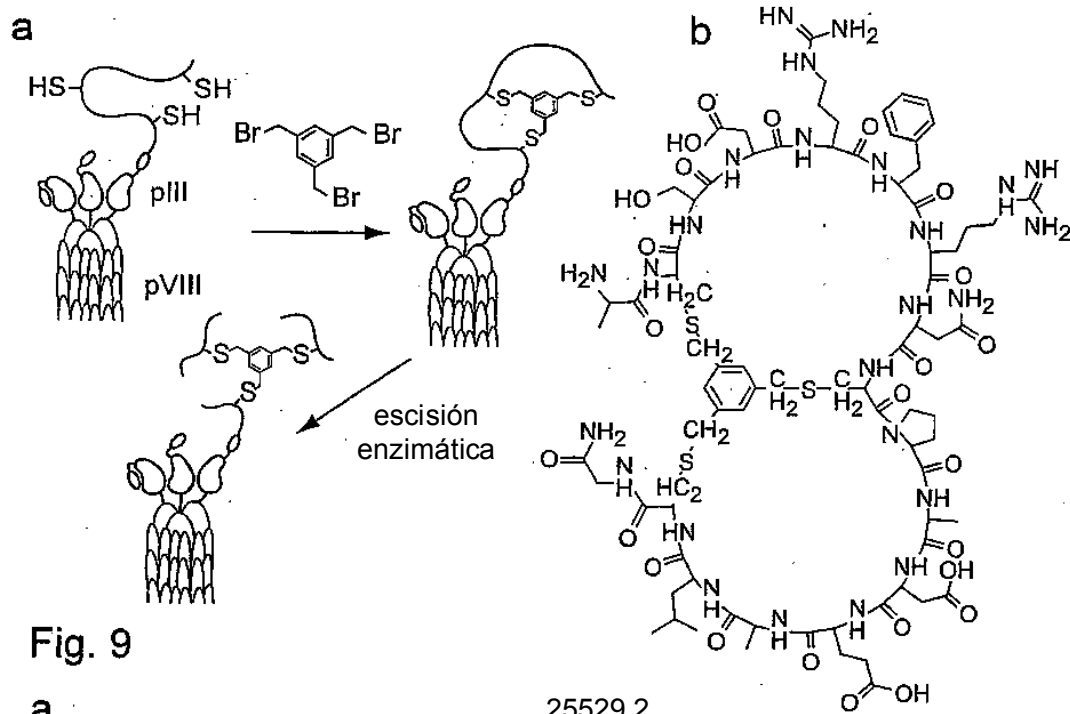
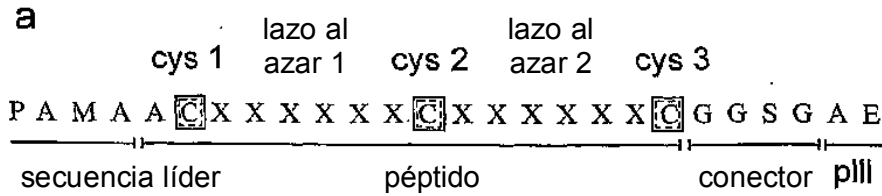


Fig. 8





b

Mutante: Secuencia de aminoácidos: Cl₅₀ (nM):

PK1	A C D S R F R N C P W S M S G C G	1500
PK2	A C S D R F R N C P L W S G T C G	400
PK3	A C R D I R F R C N Y D V A V C G	
PK4	A C S T E R R Y C P I E I F P C G	400
PK5	A C A P W R T M C L N I D G P C G	500
PK6	A C A P W R T A C Y E D L M W C G	
PK7	A C P V W R T L C M E S E G V C G	
PK8	A C F P L W R T C V H E P T M C G	
PK9	A C W Q V Q V N C R V N F G K C G	800
PK10	A C G G N S D R C R V N N I S C G	2000
PK11	A C G R G D Q T C R V N W H P C G	
PK12	A C - - V H N Y C R V N W V T C G	1800
PK13	A C G T G E G R C R V N W T P C G	500

c

Mutante: Secuencia de aminoácidos: Cl₅₀ (nM):

CG1	A C E Y G D L W C G W D P P V C G	
CG2	A C E Y D V G F C W D G F G Q C G	100
CG3	A C L F D A G F C Q Q H S T E C G	
CG4	A C I F D L G F C H N D W W N C G	100
CG5	A C L F D L G F C G G G E G P C G	150
CG6	A C P R I E G F C L P I F S D C G	1000
CG7	A C L R A Q E D C V Y D R G F C G	200
CG8	A C T R G S G D C T Y D F G F C G	200

Fig. 11

a

Genoteca 2

P A M A A [C] S D R F R N [C] X X X X X X [C] G G S G A E

Genoteca 3

P A M A A [C] A P W R T A [C] X X X X X X [C] G G S G A E

Genoteca 4

P A M A A [C] X X X X X X [C] R V N W T P [C] G G S G A E

secuencia líder

péptido

conector pIII

b

Mutante: Secuencia de aminoácidos:

CI₅₀ (nM):

PK14	A	[C]	S	D	R	F	R	N	[C]	P	A	D	E	R	E	[C]	G		
PK15	A	[C]	S	D	R	F	R	N	[C]	P	A	D	E	A	L	[C]	G	20	
PK16	A	[C]	S	D	R	F	R	N	[C]	P	V	D	E	A	L	[C]	G	20	
PK17	A	[C]	S	D	R	F	R	N	[C]	P	V	D	E	W	L	[C]	G	30	
PK18	A	[C]	S	D	R	F	R	N	[C]	P	G	D	E	S	L	[C]	G		
PK19	A	[C]	S	D	R	F	R	N	[C]	P	Y	T	L	H	D	[C]	G	30	
PK20	A	[C]	S	D	R	F	R	N	[C]	P	Y	V	S	S	D	[C]	G		
PK21	A	[C]	S	D	R	F	R	N	[C]	P	Y	S	E	G	D	[C]	G		
PK22	A	[C]	S	D	R	F	R	N	[C]	R	V	W	D	S	S	[C]	G		
PK23	A	[C]	S	D	R	F	R	N	[C]	R	V	S	E	S	A	[C]	G	50	

Fig. 12

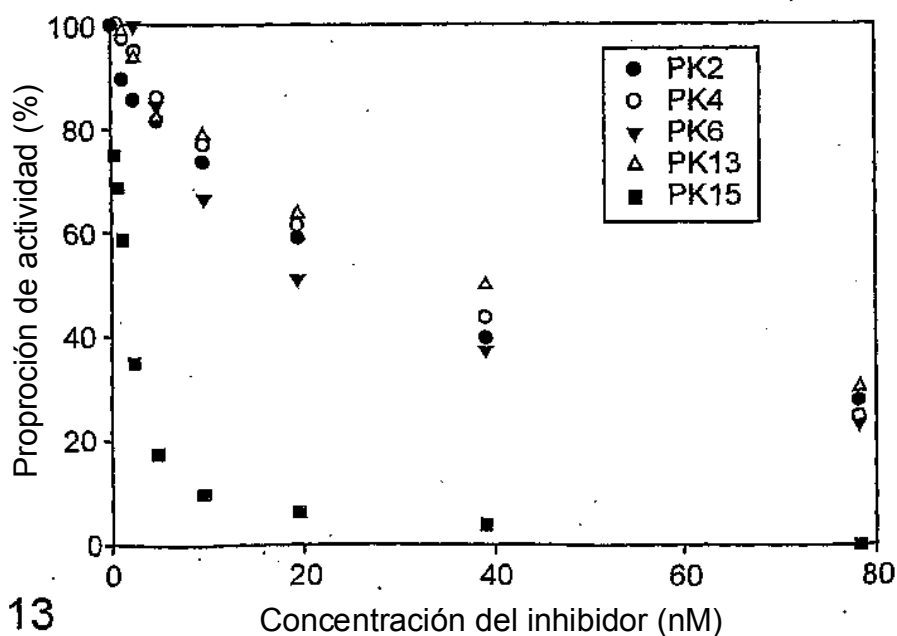


Fig. 13

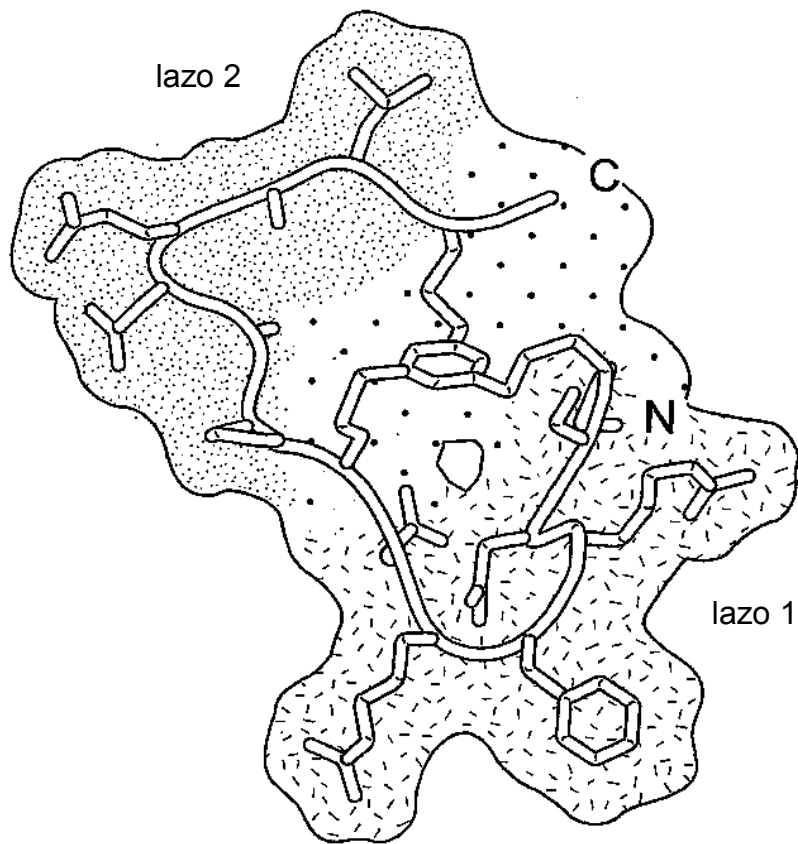


Fig. 14

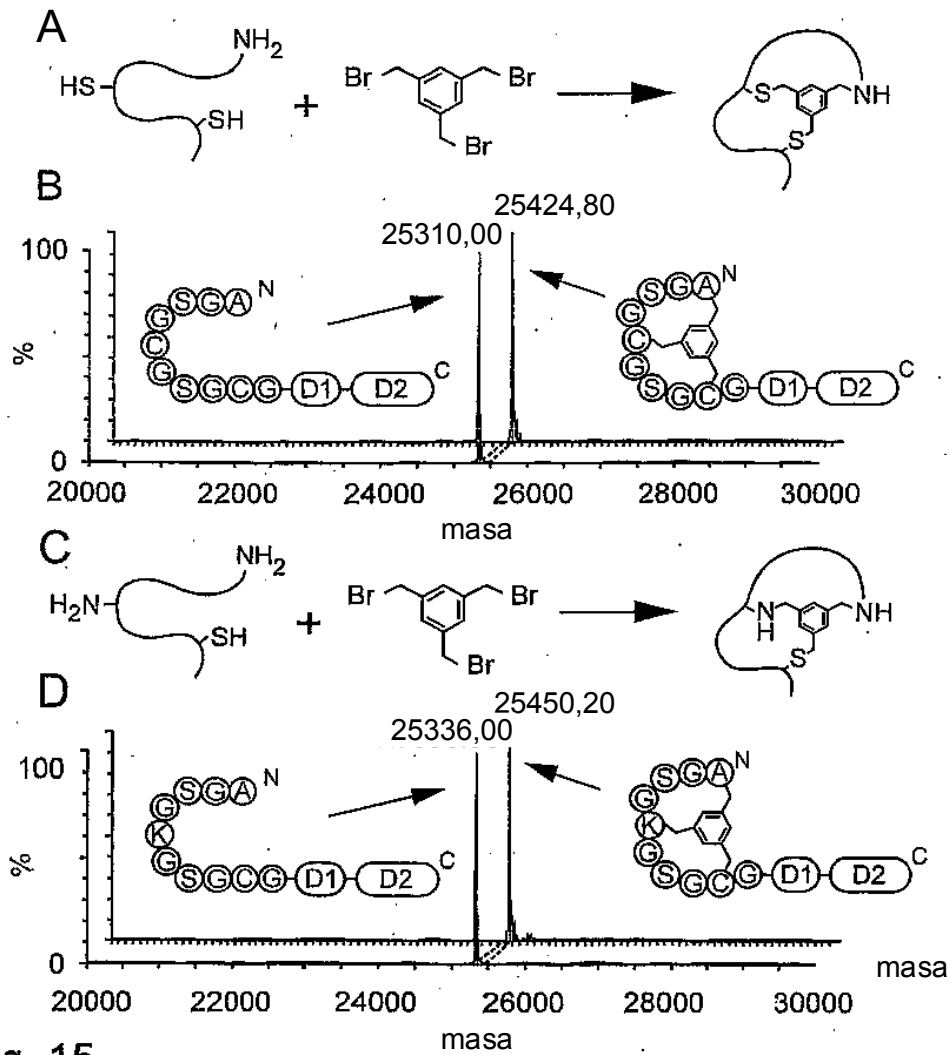


Fig. 15

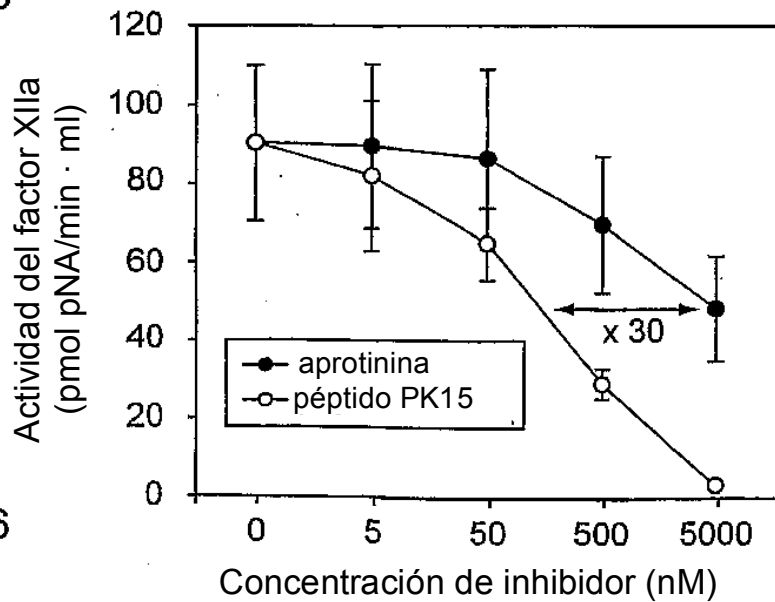


Fig. 16