

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 529 315**

51 Int. Cl.:

**G01N 33/53** (2006.01)

**G01N 33/567** (2006.01)

**C12N 15/06** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **06.03.2007 E 11172812 (7)**

97 Fecha y número de publicación de la concesión europea: **19.11.2014 EP 2402754**

54 Título: **Polímeros recombinantes no estructurados y usos de los mismos**

30 Prioridad:

**06.03.2006 US 743410 P**

**21.03.2006 US 743622 P**

**27.09.2006 US 528927**

**27.09.2006 US 528950**

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

**19.02.2015**

73 Titular/es:

**AMUNIX OPERATING INC. (100.0%)**

**500 Ellis Street, Suite B  
Mountain View, CA 94043, US**

72 Inventor/es:

**SHELLENBERGER, VOLKER;**

**STEMMER, WILLEM P.;**

**WANG, CHIA-WEI;**

**SCHOLLE, MICHAEL D.;**

**POPKOV, MIKHAIL;**

**GORDON, NATHANIEL C. y**

**CRAMERI, ANDREAS**

74 Agente/Representante:

**PONTI SALES, Adelaida**

**ES 2 529 315 T3**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

## DESCRIPCIÓN

Polímeros recombinantes no estructurados y usos de los mismos

## 5 REFERENCIA CRUZADA

## ANTECEDENTES DE LA INVENCION

**[0001]** Se ha documentado bien que las propiedades de proteínas, en particular eliminación del plasma e inmunogenicidad, pueden mejorarse uniendo polímeros hidrófilos a estas proteínas (Kochendoerfer, G. (2003) Expert Opin Biol Ther, 3: 1253-61), (Greenwald, R. B. y col. (2003) Adv Drug Deliv Rev, 55: 217-50), (Harris, J. M. y col. (2003) Nat Rev Drug Discov, 2: 214-21). Ejemplos de proteínas modificadas con polímeros que han sido autorizada por la FDA para el tratamiento de pacientes son Adagen, Oncaspar, PEG-Intron, Pegasys, Somavert y Neulasta. Muchas más proteínas modificadas con polímeros están en ensayos clínicos. Estos polímeros ejercen su efecto aumentando el radio hidrodinámico (también llamado radio de Stoke) de la proteína modificada con respecto a la proteína sin modificar, que reduce la tasa de eliminación por filtración del riñón (Yang, K. y col. (2003) Protein Eng, 16: 761-70). Además, la unión del polímero puede reducir la interacción de la proteína modificada con otras proteínas, células o superficies. En particular, la unión del polímero puede reducir interacciones entre la proteína modificada y anticuerpos y otros componentes del sistema inmunitario, reduciendo así la formación de una respuesta inmunitaria del huésped a la proteína modificada. Es de particular interés la modificación de proteínas por PEGilación, es decir, uniendo polímeros lineales o ramificados de polietilenglicol. Se mostró inmunogenicidad reducida tras la PEGilación, por ejemplo, para fenilalanina amoniaco liasa (Gamez, A. y col. (2005) Mol Ther, 11: 986-9), anticuerpos (Deckert, P. M. y col. (2000) Int J Cancer, 87: 382-90.), estafilocinasa (Collen, D. y col. (2000) Circulation, 102: 1766-72) y hemoglobina (Jin, C. y col. (2004) Protein Pept Lett, 11: 353-60). Normalmente, tales polímeros se conjugan con la proteína de interés mediante una etapa de modificación química después de que se haya purificado la proteína sin modificar.

**[0002]** Pueden unirse diversos polímeros a proteínas. Son de particular interés polímeros hidrófilos que tienen conformaciones flexibles y se hidratan bien en disoluciones acuosas. Un polímero frecuentemente usado es polietilenglicol (PEG). Estos polímeros tienden a tener grandes radios hidrodinámicos con respecto a su peso molecular (Kubetzko, S. y col. (2005) Mol Pharmacol, 68: 1439-54). Los polímeros unidos tienen a tener interacciones limitadas con la proteína a la que se han unido y así la proteína modificada con polímero retiene sus funciones relevantes.

**[0003]** La conjugación química de polímeros con proteínas requiere complejos procedimientos de múltiples etapas. Normalmente, el componente de proteína necesita producirse y purificarse antes de la etapa de conjugación química. La etapa de conjugación puede producir la formación de mezclas de productos que necesitan separarse, conduciendo a una pérdida significativa de producto. Alternativamente, tales mezclas pueden usarse como producto farmacéutico final. Algunos ejemplos son productos de interferón-alfa PEGilados actualmente comercializados que se usan como mezclas (Wang, B. L. y col. (1998) J Submicrosc Cytol Pathol, 30: 503-9; Dhalluin, C. y col. (2005) Bioconjug Chem, 16: 504-17). Tales mezclas son difíciles de fabricar y caracterizar y contienen isómeros con actividad reducida o no terapéutica.

**[0004]** Se han descrito procedimientos que permiten la adición específica para sitio de polímeros como PEG. Ejemplos son la PEGilación selectiva en un único sitio de glicosilación de la proteína diana o la PEGilación selectiva de un aminoácido no natural que ha sido manipulado en las proteínas diana. En algunos casos ha sido posible PEGilar selectivamente el extremo N de una proteína, a la vez que se evita la PEGilación de cadenas laterales de lisina en la proteína diana controlando cuidadosamente las condiciones de reacción. Todavía otro enfoque para la PEGilación específica para sitio de proteínas diana es la introducción de residuos de cisteína que permiten la conjugación selectiva. Todos estos procedimientos tienen limitaciones significativas. La PEGilación selectiva del extremo N requiere un cuidado control del procedimiento y las reacciones secundarias son difíciles de eliminar. La introducción de cisteínas para la PEGilación puede interferir con la producción y/o purificación de proteínas. La introducción específica de aminoácidos no naturales requiere organismos huésped específicos para la producción de proteínas. Otra limitación de la PEGilación es que el PEG se fabrica normalmente como una mezcla de polímeros con longitud similar, pero no uniforme. Las mismas limitaciones son inherentes en muchos otros polímeros químicos.

**[0005]** La conjugación química usando polímeros multifuncionales que permitirían la síntesis de productos con módulos de múltiples proteínas es incluso más compleja que la conjugación de polímeros de un dominio de una sola proteína.

**[0006]** Recientemente se ha observado que algunas proteínas de organismos patógenos contienen secuencias de péptidos repetitivas que parece que conducen a una semivida en suero relativamente larga de las proteínas que contienen estas secuencias (Alvarez, P. y col. (2004) J Biol Chem, 279: 3375-81). Se ha demostrado que secuencias oligoméricas que se basan en tales secuencias repetitivas derivadas de patógeno pueden fusionarse con otras proteínas produciendo elevada semivida en suero. Sin embargo, estos oligómeros derivados de patógeno tienen varias deficiencias. Las secuencias derivadas de patógeno tienden a ser inmunogénicas. Se ha descrito que las secuencias pueden modificarse para reducir su inmunogenicidad. Sin embargo, no se ha informado de intentos para eliminar epítomos de linfocitos T de las secuencias que contribuyen a la formación de reacciones inmunitarias. Además, las secuencias derivadas de patógeno no se han optimizado para aplicaciones farmacológicas que requieren secuencias con buena solubilidad y una afinidad muy baja por otras proteínas diana.

**[0007]** Así, hay una necesidad significativa de composiciones y procedimientos que permitirían que se combinaran módulos de múltiples polímeros y módulos de múltiples proteínas en productos de múltiples dominios definidos.

### DESCRIPCIÓN RESUMIDA DE LA INVENCION

**[0008]** La presente divulgación proporciona un polímero recombinante no estructurado (URP) que comprende al menos 40 aminoácidos contiguos, en el que dicho URP es sustancialmente incapaz de unirse no específicamente a una proteína del suero, y en el que (a) la suma de los residuos de glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) contenidos en el URP constituye más de aproximadamente el 80% de los aminoácidos totales del URP; y/o (b) al menos el 50% de los aminoácidos carecen de estructura secundaria como se ha determinado por el algoritmo de Chou-Fasman. En una realización relacionada, la presente divulgación describe un polímero recombinante no estructurado (URP) que comprende al menos 40 aminoácidos contiguos, en el que dicho URP tiene una semivida de degradación en suero *in vivo* superior a aproximadamente 24 horas, y en el que (a) la suma de los residuos de glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) contenidos en el URP constituye más de aproximadamente el 80% de los aminoácidos totales del URP; y/o (b) al menos el 50% de los aminoácidos carecen de estructura secundaria como se ha determinado por el algoritmo de Chou-Fasman. El URP de la invención puede comprender una secuencia de aminoácidos no naturales. Si se desea, el URP se selecciona para la incorporación en una proteína heteróloga, y en el que tras la incorporación del URP en una proteína heteróloga, dicha proteína heteróloga presenta una semivida de secreción en suero prolongada y/o mayor solubilidad con respecto a la proteína correspondiente que es deficiente en dicho URP. La semivida puede prolongarse dos veces, tres veces, cinco veces, diez veces o más. En algunos aspectos, la incorporación del URP en una proteína heteróloga produce al menos un aumento de 2 veces, 3 veces, 4 veces, 5 veces o más en el peso molecular aparente de la proteína como se aproxima por cromatografía de exclusión por tamaño. En algunos aspectos, los URP tienen una puntuación de epítomo T inferior a -3,5 (por ejemplo, -4 o menos, -5 o menos). En algunos aspectos, los URP pueden contener predominantemente residuos hidrófilos. Si se desea, al menos el 50% de los aminoácidos del URP carecen de estructura secundaria como se ha determinado por el algoritmo de Chou-Fasman. Los residuos de glicina contenidos en el URP pueden constituir al menos aproximadamente el 50% de los aminoácidos totales del URP. En algún aspecto, un tipo cualquiera de los aminoácidos solos seleccionados del grupo constituido por glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) contenidos en el URP constituye más de aproximadamente el 20%, 30%, 40%, 50%, 60% o más de los aminoácidos totales del URP. En algunos aspectos, el URP comprende más de aproximadamente 100, 150, 200 o más aminoácidos contiguos.

**[0009]** La presente divulgación también proporciona una proteína que comprende uno o más de los URP de la invención, en la que los URP de la invención son heterólogos con respecto a la proteína. La longitud total de los URP en agregación puede superar aproximadamente 40, 50, 60, 100, 150, 200 o más aminoácidos. La proteína puede comprender uno o más módulos funcionales seleccionados del grupo constituido por módulo efector, módulo de unión, módulo del extremo N, módulo del extremo C y cualquier combinación de los mismos. Si se desea, la proteína objeto comprende una pluralidad de módulos de unión, en los que los módulos de unión individuales presentan especificidades de unión por las mismas dianas o dianas diferentes. El módulo de unión puede comprender un andamiaje que contiene disulfuro formado por apareamiento entre andamiajes de cisteínas. El módulo de unión pueden unirse a una molécula diana, diana que está seleccionada del grupo constituido por proteína de la superficie celular, proteína secretada, proteína citosólica y proteína nuclear. La diana puede ser un canal de iones y/o GPCR. Si se desea, el módulo efector puede ser una toxina. La proteína que contiene el URP de la invención tiene normalmente una semivida de secreción en suero prolongada de al menos 2, 3, 4, 5, 10 o más veces con respecto a una proteína correspondiente que es deficiente en dicho URP.

**[0010]** En una divulgación separada, la presente invención proporciona una proteína que se produce naturalmente que comprende al menos 3 unidades de repetición de secuencias de aminoácidos, comprendiendo cada una de las unidades de repetición al menos 6 aminoácidos, en la que la mayoría de los segmentos que comprenden aproximadamente 6 a aproximadamente 15 aminoácidos contiguos de las al menos 3 unidades de repetición están presentes en una o más proteínas humanas nativas. En un aspecto, la mayoría de los segmentos, o cada segmento que comprende aproximadamente 9 a aproximadamente 15 aminoácidos contiguos dentro de las unidades de repetición, están presentes en una o más proteínas humanas nativas. Los segmentos pueden comprender aproximadamente 9 a aproximadamente 15 aminoácidos. Las tres unidades de repetición pueden compartir homología de secuencias sustancial, por ejemplo, compartir identidad de secuencias superior a aproximadamente el 50%, 60%, 70%, 80%, 90% o el 100% cuando se alinean. Tal proteína no natural también puede comprender uno o más módulos seleccionados del grupo constituido por módulos de unión, módulos efectores, módulos de multimerización, módulos del extremo C y módulos del extremo N. Si se desea, la proteína no natural puede comprender unidades de repetición individuales que tienen el polímero recombinante no estructurado (URP) de la invención.

**[0011]** La presente divulgación también proporciona polinucleótidos recombinantes que comprenden secuencias codificantes que codifican los URP de la invención, proteínas que contienen URP, microproteínas y toxinas. También se proporcionan en la presente divulgación vectores que contienen los polinucleótidos de la invención, células huésped que alojan los vectores, paquetes genéticos que expresan los URP de la invención, proteínas que contienen URP, toxinas y cualquier otra entidad proteínica descrita en el presente documento. También se proporciona bibliotecas de selección de vectores de expresión.

**[0012]** La presente divulgación también proporciona un procedimiento de producción de una proteína que comprende un polímero recombinante no estructurado (URP). El procedimiento implica (i) proporcionar una célula huésped que comprende un polinucleótido recombinante que codifica la proteína, comprendiendo dicha proteína uno o más URP, comprendiendo dicho URP al menos 40 aminoácidos contiguos, en el que dicho URP es sustancialmente incapaz de unirse no específicamente a una proteína del suero, y en el que (a) la suma de los residuos de glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) contenidos en el URP constituye más de aproximadamente el 80% de los aminoácidos totales del URP; y/o (b) al menos el 50% de los aminoácidos carecen de estructura secundaria como se ha determinado por el algoritmo de Chou-Fasman; y (ii) cultivar dicha célula huésped en un medio de cultivo adecuado en condiciones para efectuar la expresión de dicha proteína de dicho polinucleótido. Células huésped adecuadas son células eucariotas (por ejemplo, células CHO) y procariontas.

**[0013]** La presente divulgación también proporciona un procedimiento de aumento de la semivida en secreción de suero de una proteína que comprende: fusionar dicha proteína con uno o más polímeros recombinantes no estructurados (CTRP), en el que el URP comprende al menos aproximadamente 40 aminoácidos contiguos, y en el que (a) la suma de los residuos de glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) contenidos en el URP constituye más de aproximadamente el 80% de los aminoácidos totales del URP; y/o (b) al menos el 50% de los aminoácidos carecen de estructura secundaria tal como se ha determinado por el algoritmo de Chou-Fasman; y en el que dicho URP es sustancialmente incapaz de unirse no específicamente a una proteína del suero.

**[0014]** En la presente divulgación también se proporciona un procedimiento de detección de la presencia o ausencia de una interacción específica entre una proteína diana y una exógena que se expresa sobre un paquete genético, en el que dicha proteína comprende uno o más polímeros recombinantes no estructurados (URP), comprendiendo el procedimiento: (a) proporcionar un paquete genético que expresa una proteína que comprende uno o más polímeros recombinantes no estructurados (URP); (b) poner en contacto el paquete genético con la diana en condiciones adecuadas para producir un complejo de proteína-diana estable; y (c) detectar la formación del complejo de proteína-diana estable sobre el paquete genético, detectando así la presencia de una interacción específica. El procedimiento puede comprender además obtener una secuencia de nucleótidos del paquete genético que codifica la proteína exógena. En algunos aspectos, la presencia o ausencia de una interacción específica es entre el URP y una diana que comprende una proteína del suero. En algunos aspectos, la presencia o ausencia de una interacción específica es entre el URP y una diana que comprende una proteasa del suero.

**[0015]** En la presente divulgación también se incluye un paquete genético que expresa una microproteína, en el que dicha microproteína retiene la capacidad de unión a su diana nativa. En algunos aspectos, la microproteína presenta capacidad de unión hacia al menos una familia del canal de iones seleccionado del grupo constituido por

un canal de sodio, de potasio de calcio, de acetilcolina y de cloro. Si se desea, la microproteína es una microproteína de unión a canales de iones, y se modifica de forma que (a) la microproteína se une a una familia diferente del canal con respecto a la microproteína sin modificar correspondiente; (b) la microproteína se une a una subfamilia diferente de la misma familia del canal con respecto a la microproteína sin modificar correspondiente; (c) la microproteína se une a una especie diferente de la misma subfamilia de canal con respecto a la microproteína sin modificar correspondiente; (d) la microproteína se une a un sitio diferente sobre el mismo canal con respecto a la microproteína sin modificar correspondiente; y/o (e) la microproteína se une al mismo sitio del mismo canal pero da un efecto biológico diferente con respecto a la microproteína sin modificar correspondiente. En algún aspecto, la microproteína es una toxina. La presente divulgación también proporciona una biblioteca de paquetes genéticos que expresa las microproteínas y/o toxinas objeto. Si se desea, el paquete genético expresa una toxina proteínica que retiene en parte o por completo su espectro de toxicidad. La toxina puede derivarse de una proteína de una sola toxina, o derivarse de una familia de toxinas. La presente divulgación también proporciona una biblioteca de paquetes genéticos en la que la biblioteca expresa una familia de toxinas, en la que la familia retiene en parte o por completo su espectro de toxicidad nativo.

15

**[0016]** La presente divulgación también proporciona una proteína que comprende una pluralidad de dominios de unión a canales de iones, en la que dominios individuales son dominios de microproteína que han sido modificados de forma que (a) los dominios de microproteína se unen a una familia diferente del canal con respecto a los dominios de microproteína sin modificar correspondientes; (b) los dominios de microproteína se unen a una subfamilia diferente de la misma familia del canal con respecto a los dominios de microproteína sin modificar correspondientes; (c) los dominios de microproteína se unen a una especie diferente de la misma subfamilia con respecto a los dominios de microproteína sin modificar correspondientes; (d) los dominios de microproteína se unen a un sitio diferente sobre el mismo canal con respecto a los dominios de microproteína sin modificar correspondientes; (e) los dominios de microproteína se unen al mismo sitio del mismo canal, pero dan un efecto biológico diferente con respecto a los dominios de microproteína sin modificar correspondientes; y/o (f) los dominios de microproteína se unen al mismo sitio del mismo canal y dan el mismo efecto biológico con respecto a los dominios de microproteína sin modificar correspondientes.

**[0017]** En la presente divulgación también se desarrollan realizaciones con un procedimiento de obtención de una microproteína con propiedad deseada, que comprende: (a) proporcionar una biblioteca objeto; y (b) cribar la biblioteca de selección para obtener al menos una expresión de fago de una microproteína con la propiedad deseada. También se proporcionan polinucleótidos, vectores, paquetes genéticos, células huésped para su uso en uno cualquiera de los procedimientos desvelados.

### 35 INCORPORACIÓN POR REFERENCIA

#### BREVE DESCRIPCIÓN DE LOS DIBUJOS

**[0018]** Las novedosas características de la invención se exponen con particularidad en las reivindicaciones adjuntas. Un mejor entendimiento de las características y ventajas de la presente invención se obtendrán por referencia a la siguiente descripción detallada que expone realizaciones ilustrativas, en las que los principios de la invención se utilizan, y los dibujos adjuntos de los que:

**[0019]** La FIG. 1 muestra los componentes modulares de una MURP. Módulos unión, módulos efectores y módulos de multimerización se representan como círculos. Módulos de URP, módulos del extremo N y del extremo C se muestran como rectángulos.

**[0020]** La FIG. 2 muestra ejemplos de arquitecturas modulares de MURP. Módulos de unión (BM) en una MURP pueden tener especificidades diana idénticas o diferentes.

50

**[0021]** La FIG. 3 muestra que una proteína de repetición que se basa en una secuencia humana puede contener secuencias de aminoácidos novedosas, que pueden contener epítomos de linfocitos T. Estas secuencias novedosas se forman en el empalme entre unidades de repetición vecinas.

**[0022]** La FIG. 4 ilustra el diseño de una secuencia de URP que es una proteína de repetición basada en tres secuencias de donante humano D1, D2, y D3. La unidad de repetición de este URP se eligió de forma que incluso secuencias 9-meras que atraviesan el empalme entre unidades vecinas pudieran encontrarse en al menos una de las secuencias de donante humano.

**[0023]** FIG. 5 Ejemplo de una secuencia de URP que es una proteína de repetición basada en las secuencias de tres proteínas humanas. La porción inferior de la figura ilustra que todas las subsecuencias 9-meras en el URP se producen en al menos una de las proteínas de donante humano.

5 **[0024]** FIG. 6 Ejemplo basado en secuencia de URP basada en los residuos del dominio POU humano 146-182.

**[0025]** La FIG. 7 muestra la ventaja de separar módulos con secuencias ricas en información insertando módulos de URP entre tales secuencias. El lado izquierdo de la figura muestra que la fusión directa de los módulos A y B conduce a secuencias novedosas en la región de empalme. Estas secuencias de empalme pueden ser epítomos. La mitad derecha de la figura muestra que la inserción de un módulo de URP entre el módulo A y B previene la formación de tales secuencias de empalme que contienen secuencias parciales de los módulos A y B. En su lugar, los extremos de los módulos A y B dan secuencias de empalme que contienen secuencias de URP y así se predice que tienen baja inmunogenicidad.

15 **[0026]** La FIG. 8 muestra construcciones de administración de fármaco que se basan en URP. Las moléculas de fármaco representadas como hexágonos están químicamente conjugadas con la MURP.

**[0027]** La FIG. 9 muestra y MURP que contiene un sitio sensible a proteasas. El módulo de URP se diseña de forma que bloquee la función del módulo efector. La escisión con proteasa elimina una parte del módulo de URP y produce elevada actividad de la función efectora.

**[0028]** La FIG. 10 muestra cómo un módulo de URP puede actuar de ligador entre un módulo de unión y un módulo efector. El módulo de unión pueden unirse a una diana y como consecuencia aumenta la concentración local de módulo efector en la proximidad de la diana.

**[0029]** La FIG. 11 muestra un procedimiento para construir genes que codifican secuencias de URP de bibliotecas de módulos de URP cortos. La biblioteca de módulos de URP puede insertarse en un vector de relleno que contiene proteína verde fluorescente (GFP) como indicador para facilitar la identificación de secuencias de URP con alta expresión. La figura ilustra que los genes que codifican secuencias de URP largas pueden construirse por dimerización iterativa.

**[0030]** La FIG. 12 muestra MURP que contienen múltiples módulos de unión para receptores de muerte. Los receptores de muerte son desencadenados por la trimerización y así MURP que contienen al menos tres elementos de unión para un receptor de muerte particularmente potente en inducir muerte celular. La porción inferior de la figura ilustra que puede aumentarse la especificidad de la MURP por tejido enfermo añadiendo uno o más módulos de unión con especificidad por tejido tumoral.

**[0031]** La FIG. 13 muestra una MURP que comprende cuatro módulos de unión (rectángulos) con especificidad por un antígeno de tumor con un módulo efector como interleucina 2.

**[0032]** La FIG. 14 muestra el diagrama de flujo para la construcción de módulos de URP con 288 residuos. Los módulos de URP se construyeron como proteínas de fusión con GFP. Las bibliotecas de módulos de URP con 36 aminoácidos se construyeron primero, seguido de dimerización iterativa para dar módulos de URP con 288 aminoácidos (rPEG\_H288 y rPEG\_J288).

**[0033]** FIG. 15 Secuencia de aminoácidos y de nucleótidos de un módulo de URP con 288 aminoácidos (rPEG\_J288).

50 **[0034]** FIG. 16 Secuencia de aminoácidos y de nucleótidos de un módulo de URP con 288 aminoácidos (rPEG\_H288).

**[0035]** FIG. 17 Secuencia de aminoácidos de una región de secuencia rica en serina de la proteína humana dentina sialofosfoproteína.

55 **[0036]** La FIG. 18 muestra un derivado de liberación prolongada de una MURP. La proteína contiene dos residuos de cisteína que pueden formar un débil puente SS. La proteína puede fabricarse con el puente SS intacto. Puede formularse e inyectarse en pacientes en forma reducida. Después de la inyección se oxidará en proximidad al sitio de inyección y como resultado puede formar un polímero de alto peso molecular con difusividad muy limitada.

La MURP activa puede filtrarse lentamente del sitio de inyección por proteólisis limitada o reducción limitada del enlace SS de reticulación.

**[0037]** La FIG. 19 muestra una forma de liberación prolongada de una MURP. La MURP tiene difusividad muy limitada en el sitio de inyección y puede liberarse del sitio de inyección por proteólisis limitada.

**[0038]** La FIG. 20 muestra una forma de liberación prolongada de una MURP que contiene una secuencia rica en histidina. La MURP puede formularse e inyectarse en combinación con perlas insolubles que contienen níquel inmovilizado. La MURP se une a las perlas de níquel en el sitio de inyección y es lentamente liberada en la circulación.

**[0039]** La FIG. 21 muestra MURP que contienen módulos de multimerización. La parte superior de la figura muestra una MURP que contiene una secuencia de dimerización. Como resultado, forma un dímero que duplica eficazmente su peso molecular. El centro de la figura muestra tres diseños de MURP que comprenden dos secuencias de multimerización. Tales MURP pueden formar multímeros con peso molecular eficaz muy alto. La parte inferior de la figura ilustró una MURP que contiene múltiples secuencias de RGD que son conocidas por unirse a receptores de la superficie celular y así conferir semivida.

**[0040]** La FIG. 22 muestra una variedad de MURP que se diseñan para bloquear o modular la función de canales de iones. Los círculos indican módulos de unión con especificidad por canales de iones. Estos módulos de unión pueden derivarse o ser idénticos a toxinas naturales con afinidad por receptores de canales de iones. La figura ilustra que otros dominios de unión pueden añadirse a cualquier lado de los módulos específicos para unión a canales de iones, confiriendo así a las MURP eficacia o especificidad elevada por un tipo de célula particular.

**[0041]** La FIG. 23 muestra varios diseños de MURP para semivida elevada. El peso molecular eficaz elevado puede lograrse aumentando la longitud de cadena (A), multimerización química (B), añadiendo múltiples copias de módulos de unión a una molécula separada por sitios de no unión (C), construcción de multímeros químicos similares a C (D, E), que incluyen secuencias de multimerización (F).

**[0042]** La FIG. 24 muestra MURP que pueden formarse por conjugación química de módulos de unión con una secuencia de URP recombinante. La secuencia de URP se diseña para contener múltiples residuos de lisina (K) como sitios de conjugación.

**[0043]** La FIG. 25 muestra el diseño de una biblioteca de módulos de unión a 2SS. Las secuencias contienen una secuencia de 1SS constante en el centro que está flanqueada por secuencias aleatorias que contienen residuos de cisteína a distancia variable del núcleo de 1SS.

**[0044]** La FIG. 26 muestra el diseño de una biblioteca de módulos de unión a 2SS. Las secuencias contienen una secuencia de 1SS constante en el centro que está flanqueada por secuencias aleatorias que contienen residuos de cisteína a distancia variable del núcleo de 1SS.

**[0045]** La FIG. 27 muestra el diseño de una biblioteca de dímeros de módulos de unión 1SS. Inicialmente, una colección de módulos de unión 1SS se amplifica por dos reacciones de PCR. Los productos de PCR resultantes se combinan y generan dímeros en una posterior etapa de PCR.

**[0046]** La FIG. 28 muestra los análisis de Western de una proteína de fusión que contiene la secuencia de URP de 288 aminoácidos rPEG\_J288 después de la incubación de hasta 3 días en 50% de suero de ratón.

**[0047]** La FIG. 29 muestra los resultados de una prueba de ensayo de unión para anticuerpos preexistentes contra una secuencia de URP de 288 aminoácidos.

**[0048]** La FIG. 30 muestra la unión de MURP que contienen un (monómero), dos (dímero), cuatro (tetrámero) o cero (rPEG36) módulos de unión con especificidad por VEGF que se recubrió para placas de microtitulación.

**[0049]** La FIG. 31 muestra la secuencia de aminoácidos de una MURP con especificidad por EpCAM. La secuencia contiene cuatro módulos de unión con afinidad por EpCAM (subrayado). La secuencia contiene una secuencia de Flag del extremo N que contiene los dos únicos residuos de lisina de la secuencia entera.

**[0050]** La FIG. 32 muestra el diseño de bibliotecas de adición de 1SS. Módulos 1SS aleatorios pueden

añadirse al extremo N o C de un módulo de unión preseleccionado o simultáneamente a ambos lados.

- 5 **[0051]** La FIG. 33 muestra el alineamiento de secuencias relacionadas con toxina de tres dedos. La figura también muestra una estructura 3D que se resolvió por RMN.
- [0052]** La FIG. 34 muestra el diseño de una biblioteca basada en la toxina de tres dedos. Los residuos designados X se aleatorizaron. Se indica la elección de codón para cada posición al azar.
- 10 **[0053]** La FIG. 35 muestra el alineamiento de secuencias relacionadas con plexina.
- [0054]** La FIG. 36 muestra el diseño de una biblioteca basada en plexina. Los residuos designados X se aleatorizaron. Se indica la elección de codón para cada posición al azar.
- 15 **[0055]** FIG. 37 Secuencias de módulos de unión relacionados con plexina con especificidad por DR4, ErbB2 y HGFR.
- [0056]** La FIG. 38 muestra un ensayo de unión para dominios de unión basados en microproteína con especificidad por VEGF.
- 20 **[0057]** La FIG. 39 muestra secuencias de módulos de unión de 2SS y 3SS que se aislaron de bibliotecas de elaboración con especificidad por VEGF. La parte superior de la proteína muestra análisis de gel de PAGE de las proteínas purificadas por calor-lisis.
- [0058]** La FIG. 40 muestra etapas de clonación para construir la secuencia de URP rPEG\_J72.
- 25 **[0059]** La FIG. 41 muestra la construcción de una biblioteca de módulos de URP con 36 aminoácidos llamada rPEG\_J36. La región que codifica rPEG\_J36 se ensambló ligando tres segmentos más cortos que codifican rPEG\_J12 y un módulo tapón.
- 30 **[0060]** La FIG. 42 muestra la secuencia de nucleótidos y la traducción del vector de relleno pCW0051. La región de relleno está flanqueada por sitios Bsal y BbsI y contiene múltiples codones de terminación.
- [0061]** La FIG. 43 muestra un gel de PAGE de la purificación del URP rPEG\_J288 fusionado con GFP. El carril 2 muestra el lisado celular; carril 3: producto purificado por IMAC; carril 4: producto purificado por anti-Flag.
- 35 **[0062]** FIG. 44 Secuencia de aminoácidos de proteínas de fusión entre rPEG\_J288 y los dominios efectores humanos interferón alfa, G-CSF y hormona de crecimiento humana.
- [0063]** La FIG. 45 muestra el análisis Western de expresión de proteínas de fusión entre rPEG\_J288 y hormona de crecimiento humana (carriles 1 y 2), interferón alfa (carriles 3 y 4) y GFP (carriles 5 y 6). Se analizó tanto material soluble como insoluble para cada proteína.
- 40 **[0064]** La FIG. 46 muestra el diseño de MURP basado en la toxina OSK1. La figura muestra que secuencias de URP y/o módulos de unión pueden añadirse a cualquier lado de OSK1
- 45 **[0065]** La FIG. 47 representa a modo de ejemplo formatos de producto que comprenden los URP de la invención.

#### DESCRIPCIÓN DETALLADA DE LA INVENCION

- 50 **[0066]** Aunque en el presente documento se han mostrado y descrito realizaciones preferidas de la presente invención, será obvio para aquellos expertos en la materia que tales realizaciones solo se proporcionan a modo de ejemplo. Numerosas variaciones, cambios y sustituciones se les ocurrirán ahora a aquellos expertos en la materia sin apartarse de la invención. Debe entenderse que pueden emplearse diversas alternativas a las realizaciones de la invención descritas en el presente documento en la práctica de la invención. Se pretende que las siguientes reivindicaciones definan el alcance de la invención y que así se cubran procedimientos y estructuras dentro del alcance de estas reivindicaciones y sus equivalentes.
- 55

Técnicas generales:

**[0067]** La práctica de la presente invención emplea, a menos que se indique lo contrario, técnicas convencionales de inmunología, bioquímica, química, biología molecular, microbiología, biología celular, genómica y ADN recombinante, que están dentro de la habilidad en la técnica. Véase Sambrook, Fritsch y Maniatis, **5 MOLECULAR CLONING: A LABORATORY MANUAL**, 2ª edición (1989); **CURRENT PROTOCOLS IN MOLECULAR BIOLOGY** (F. M. Ausubel y col. eds., (1987)); las series **METHODS IN ENZYMOLOGY** (Academic Press, Inc.); **PCR 2: A PRACTICAL APPROACH** (M.J. MacPherson, B.D. Hames y G.R. Taylor eds. (1995)), Harlow y Lane, eds. (1988) **ANTIBODIES, A LABORATORY MANUAL**, y **ANIMAL CELL CULTURE** (R.I. Freshney, ed. (1987)).

#### 10 Definiciones:

**[0068]** Como se usa en la memoria descriptiva y reivindicaciones, la forma singular “un”, “una”, “el” y “la” incluyen referencias plurales, a menos que el contexto dicte claramente de otro modo. Por ejemplo, el término “una célula” incluye una pluralidad de células, que incluye mezclas de las mismas.

15

**[0069]** Los términos “polipéptido”, “péptido”, “secuencia de aminoácidos” y “proteína” se usan indistintamente en el presente documento para referirse a polímeros de aminoácidos de cualquier longitud. El polímero puede ser lineal o ramificado, puede comprender aminoácidos modificados y puede estar interrumpido por no aminoácidos. Los términos también engloban un polímero de aminoácidos que ha sido modificado, por ejemplo, formación de enlaces disulfuro, glicosilación, lipidación, acetilación, fosforilación, o cualquier otra manipulación tal como conjugación con un componente de marca. Como se usa en el presente documento, el término “aminoácido” se refiere tanto a aminoácidos naturales y/o no naturales como a sintéticos que incluyen, pero no se limitan a, glicina y tanto los isómeros ópticos D como L, y análogos de aminoácidos y peptidomiméticos. Se usan códigos de una sola letra o de tres letras convencionales para designar los aminoácidos.

25

**[0070]** Una “secuencia repetitiva” se refiere a una secuencia de aminoácidos que puede describirse como un oligómero de secuencias de péptidos de repetición, que forman repeticiones directas, o repeticiones inversas o repeticiones alternantes de motivos de múltiples secuencias. Estas secuencias de oligómeros de repetición pueden ser idénticas u homólogas entre sí, pero también pueden ser motivos repetidos múltiples. Las secuencias repetitivas se caracterizan por un contenido de información muy bajo. Una secuencia repetitiva no es una característica requerida de un URP y en algunos casos una secuencia no repetitiva será de hecho preferida.

30

**[0071]** Los aminoácidos pueden caracterizarse basándose en su hidrofobia. Se han desarrollado varias escalas. Un ejemplo es una escala desarrollada por Levitt, M y col. (véase Levitt, M (1976) *J Mol Biol* 104, 59, nº 3233, que se enumera en Hopp, TP y col. (1981) *Proc Natl Acad Sci U S A* 78, 3824, nº 3232). Ejemplos de “aminoácidos hidrófilos” son arginina, lisina, treonina, alanina, asparagina y glutamina. Son de particular interés los aminoácidos hidrófilos aspartato, glutamato y serina, y glicina. Ejemplos de “aminoácidos hidrófobos” son triptófano, tirosina, fenilalanina, metionina, leucina, isoleucina y valina.

35

**[0072]** El término “conformación desnaturalizada” describe el estado de un péptido en disolución que se caracteriza por una gran libertad conformacional del esqueleto del péptido. La mayoría de los péptidos y proteínas adoptan una conformación desnaturalizada en presencia de altas concentraciones de desnaturalizantes o a temperaturas elevadas. Los péptidos en conformación desnaturalizada tienen espectros de CD característicos y se caracterizan generalmente por una falta de interacciones de largo intervalo como se ha determinado por, por ejemplo, RMN. Conformación desnaturalizada y conformación no plegada se usarán sinónimamente.

40

**[0073]** Los términos “secuencias de proteínas no estructuradas (UNP)” y polímero recombinante no estructurado” (URP) se usan en el presente documento indistintamente. Los términos se refieren a secuencias de aminoácidos que comparten características comunes con secuencias de péptidos desnaturalizadas, por ejemplo, que presentan un comportamiento típico similar a secuencias de péptidos desnaturalizadas, bajo condiciones fisiológicas, como se detalla en el presente documento. Las secuencias de URP carecen de una estructura terciaria definida y tienen estructura limitada o no secundaria como se detecta, por ejemplo, por el algoritmo de Chou-Fasman.

50

**[0074]** Como se usa en el presente documento, el término “proteínas de la superficie celular” se refiere a los componentes de la membrana plasmática de una célula. Engloba proteínas de la membrana integrales y periféricas, glicoproteínas, polisacáridos y lípidos que constituyen la membrana plasmática. Una proteína de la membrana integral es una proteína transmembrana que se extiende a través de la bicapa de lípidos de la membrana plasmática de una célula. Una proteína de la membrana integral típica consiste en al menos un segmento que atraviesa la

55

membrana que generalmente comprende residuos de aminoácidos hidrófobos. Las proteínas de la membrana periférica no se extienden en el interior hidrófobo de la bicapa de lípidos y están unidas a la superficie de la membrana mediante interacción covalente o no covalente directamente o indirectamente con otros componentes de la membrana.

5

**[0075]** Los términos “membrana”, “citosólica”, “nuclear” y “secretada” como se aplican a proteínas celulares especifican la localización extracelular y/o subcelular en la que la proteína celular está principalmente, predominantemente o preferencialmente localizada.

10 **[0076]** Los “receptores de la superficie celular” representan un subconjunto de proteínas de la membrana, que pueden unirse a sus ligandos respectivos. Los receptores de la superficie celular son moléculas ancladas sobre o insertadas en la membrana plasmática de la célula. Constituyen una gran familia de proteínas, glicoproteínas, polisacáridos y lípidos, que sirven no solo de constituyentes estructurales de la membrana plasmática, sino también de elementos reguladores que gobiernan una variedad de funciones biológicas.

15

**[0077]** El término “módulo” se refiere a una parte de una proteína que se diferencia físicamente o funcionalmente de otras porciones de la proteína o péptido. Un módulo puede comprender uno o más dominios. En general, un módulo o dominio puede ser una única estructura tridimensional estable, independientemente del tamaño. La estructura terciaria de un dominio típico es estable en disolución y sigue igual si un miembro tal se aísla o se fusiona covalentemente con otros dominios. Un dominio tiene generalmente una estructura terciaria particular formada por relaciones espaciales de elementos de la estructura secundaria, tales como hojas beta, hélices alfa y bucles no estructurados. En dominios de la familia de las microproteínas, los puentes disulfuro son generalmente los elementos primarios que determinan la estructura terciaria. En algunos casos, los dominios son módulos que pueden conferir una actividad funcional específica, tal como avidin (sitios de múltiple unión a la misma diana), multi-especificidad (sitios de unión para diferentes dianas), semivida (usando un dominio, péptido cíclico o péptido lineal) que se une a una proteína del suero como albúmina de suero humano (HSA) o a IgG (hlgG1, 2, 3 ó 4) o a glóbulos rojos. Los dominios funcionalmente definidos tienen una función (funciones) biológica(s) distinta(s). El dominio de unión a ligando de un receptor es, por ejemplo, el dominio que se une a un ligando. Un dominio de unión a antígeno se refiere a la parte de una unidad de unión a antígeno o un anticuerpo que se une al antígeno. Los dominios funcionalmente definidos no necesitan codificarse por secuencias de aminoácidos contiguas. Los dominios funcionalmente definidos pueden contener uno o más dominios físicamente definidos. Los receptores, por ejemplo, se dividen generalmente en el dominio de unión a ligando extracelular, un dominio transmembrana y un dominio efector intracelular. Un “dominio de anclaje a membrana” se refiere a la porción de una proteína que media en la asociación a membrana. Generalmente, el dominio de anclaje a membrana está compuesto por residuos de aminoácidos hidrófobos. Alternativamente, el dominio de anclaje a membrana puede contener aminoácidos modificados, por ejemplo, aminoácidos que están unidos a una cadena de ácido graso, que a su vez ancla la proteína a una membrana.

20 **[0078]** “Que no se produce naturalmente” como se aplica a una proteína significa que la proteína contiene al menos un aminoácido que es diferente de la proteína natural o nativa correspondiente. Pueden determinarse secuencias no naturales realizando búsqueda con BLAST usando, por ejemplo, la menor probabilidad de suma mínima en la que la ventana de comparación es la longitud de la secuencia de interés (la buscada) y cuando se compara con la base de datos no redundante (“nr”) de Genbank usando BLAST 2.0. El algoritmo BLAST 2.0, que se describe en Altschul y col. (1990) J. Mol. Biol. 215:403-410, respectivamente. El software para realizar análisis con  
25  
30  
35  
40  
45 BLAST está públicamente disponible del Centro nacional para información biotecnológica.

**[0079]** Una “célula huésped” incluye una célula individual o cultivo celular que puede ser o ha sido un receptor para los vectores objeto. Células huésped incluyen progenie de una única célula huésped. La progenie puede no ser necesariamente completamente idéntica (en morfología o en genómica del complemento de ADN total) a la célula parental original debido a mutación natural, accidental o intencionada. Una célula huésped incluye células transfectadas *in vivo* con un vector descrito en el presente documento.  
50

**[0080]** Como se usa en el presente documento, el término “aislado” significa separado de constituyentes, celulares y de otro modo, en los que el polinucleótido, péptido, polipéptido, proteína, anticuerpo, o fragmentos de los mismos, están normalmente asociados en la naturaleza. Como es evidente para aquellos expertos en la materia, un polinucleótido, péptido, polipéptido, proteína, anticuerpo, o fragmentos de los mismos, que no se produce naturalmente no requiere “aislamiento” para distinguirlo de su homólogo que se produce naturalmente. Además, un polinucleótido, péptido, polipéptido, proteína, anticuerpo, o fragmentos de los mismos, “concentrado”, “separado” o “diluido” es diferenciable de su homólogo que se produce naturalmente porque la concentración o número de  
55

moléculas por volumen es mayor que “concentrado” o inferior a “separado” que el de su homólogo que se produce naturalmente.

**[0081]** “Ligado” y “fusionado” o “fusión” se usan indistintamente en el presente documento. Estos términos se refieren a unir juntos dos o más elementos o componentes químicos, por cualquier medio que incluya conjugación química o medios recombinantes. Una “fusión en marco” se refiere a la unión de dos o más marcos de lectura abiertos (OFR) para formar un OFR más largo continuo, de un modo que mantenga el marco de lectura correcto de los OFR originales. Así, la proteína de fusión recombinante resultante es una proteína individual que contiene dos o más segmentos que se corresponden con polipéptidos codificados por los OFR originales (segmentos que normalmente no están así unidos en la naturaleza.)

**[0082]** En el contexto de polipéptidos, una “secuencia lineal” o una “secuencia” es un orden de aminoácidos en un polipéptido en una dirección del extremo amino a carboxilo en la que los residuos que son vecinos entre sí en la secuencia son contiguos en la estructura primaria del polipéptido. Una “secuencia parcial” es una secuencia lineal de parte de un polipéptido que se sabe que comprende residuos adicionales en una o ambas direcciones.

**[0083]** “Heterólogo” significa derivado de una entidad genotípicamente distinta del resto de la entidad con la que está siendo comparada. Por ejemplo, una secuencia rica en glicina eliminada de su secuencia codificante nativa y operativamente ligada a una secuencia codificante distinta de la secuencia nativa es una secuencia rica en glicina heteróloga. El término “heterólogo” como se aplica a un polinucleótido, un polipéptido, significa que el polinucleótido o polipéptido se deriva de una entidad genotípicamente distinta de la del resto de la entidad con la que está siendo comparada.

**[0084]** Los términos “polinucleótidos”, “ácidos nucleicos”, “nucleótidos” y “oligonucleótidos” se usan indistintamente. Se refieren a una forma polimérica de nucleótidos de cualquier longitud, tanto desoxirribonucleótidos como ribonucleótidos, o análogos de los mismos. Los polinucleótidos pueden tener cualquier estructura tridimensional, y pueden realizar cualquier función, conocida o desconocida. Lo siguiente son ejemplos no limitantes de polinucleótidos: regiones codificantes o no codificantes de un gen o fragmento de gen, loci (locus) definidos del análisis de enlace, exones, intrones, ARN mensajero (ARNm), ARN de transferencia, ARN ribosómico, ribozimas, ADNc, polinucleótidos recombinantes, polinucleótidos ramificados, plásmidos, vectores, ADN aislado de cualquier secuencia, ARN aislado de cualquier secuencia, sondas de ácido nucleico y cebadores. Un polinucleótido puede comprender nucleótidos modificados, tales como nucleótidos metilados y análogos de nucleótidos. Si están presentes, las modificaciones a la estructura del nucleótido pueden conferirse antes o después del ensamblaje del polímero. La secuencia de nucleótidos puede interrumpirse por componentes de no nucleótido. Un polinucleótido puede modificarse adicionalmente después de la polimerización, tal como por conjugación con un componente de marca.

**[0085]** “Recombinante” como se aplica a un polinucleótido significa que el polinucleótido es el producto de diversas combinaciones de etapas de clonación, restricción y/o ligación, y otros procedimientos que producen una construcción que es distinta de un polinucleótido encontrado en la naturaleza.

**[0086]** Los términos “gen” o “fragmento génico” se usan indistintamente en el presente documento. Se refieren a un polinucleótido que contiene al menos un marco de lectura abierto que puede codificar una proteína particular después de ser transcrita y traducida. Un gen o fragmento génico puede ser ADN genómico o ADNc, en tanto que el polinucleótido contenga al menos un marco de lectura abierto que pueda cubrir la región codificante entera o un segmento de la misma. Un “gen de fusión” es un gen compuesto por al menos dos polinucleótidos heterólogos que están ligados juntos.

**[0087]** Un “vector” es una molécula de ácido nucleico, preferentemente auto-replicante, que transfiere una molécula de ácido nucleico insertada en y/o entre células huésped. El término incluye vectores que funcionan principalmente para la inserción de ADN o ARN en una célula, replicación de vectores que funcionan principalmente para la replicación de ADN o ARN y vectores de expresión que funcionan para la transcripción y/o traducción del ADN o ARN. También se incluyen vectores que proporcionan más de una de las funciones anteriores. Un “vector de expresión” es un polinucleótido que, cuando se introduce en una célula huésped apropiada, puede transcribirse y traducirse en un polipéptido(s). Un “sistema de expresión” normalmente connota una célula huésped adecuada que comprende un vector de expresión que puede funcionar para dar un producto de expresión deseado.

**[0088]** La “diana” como se usa en el contexto de MURP es una molécula o estructura bioquímica con la que el módulo de unión o el módulo de unión ligado a URP puede unirse y en la que el acontecimiento de unión produce

una actividad biológica deseada. La diana puede ser un ligando de proteína o receptor que es inhibido, activado o que actúa de otro modo por la proteína. Ejemplos de dianas son hormonas, citocinas, anticuerpos o fragmentos de anticuerpos, receptores de la superficie celular, cinasas, factores de crecimiento y otras estructuras bioquímicas con actividad biológica.

5

**[0089]** Un “módulo funcional” puede ser cualquier no URP en un producto de proteína. Así, un módulo funcional puede ser un módulo de unión (BM), un módulo efector (EM), un módulo de multimerización (MM), un módulo del extremo C (CM) o un módulo del extremo N (NM). En general, los módulos funcionales se caracterizan por un alto contenido de información de su secuencia de aminoácidos, es decir, contienen muchos aminoácidos diferentes y muchos de estos aminoácidos son importantes para la función de un módulo funcional. Un módulo funcional normalmente tiene estructura secundaria y terciaria, puede ser un dominio de proteína plegada y puede contener 1, 2, 3, 4, 5 o más enlaces disulfuro.

10

**[0090]** El término 'microproteínas' se refiere a una clasificación en la base de datos SCOP. Las microproteínas son normalmente las proteínas más pequeñas con una estructura fija y normalmente, pero no exclusivamente, tienen tan solo 15 aminoácidos con dos disulfuros o hasta 200 aminoácidos con más de diez disulfuros. Una microproteína puede contener uno o más dominios de microproteína. Algunos dominios de microproteína o familias de dominios pueden tener múltiples estructuras más o menos estables y múltiples estructuras más o menos similares que son conferidas por diferentes patrones de enlace disulfuro, así el término estable se usa de una forma relativa para diferenciar microproteínas de péptidos y dominios de no microproteína. La mayoría de las toxinas de microproteína están compuestas de un único dominio, pero las microproteínas del receptor de la superficie de la célula frecuentemente tienen múltiples dominios. Las microproteínas pueden ser tan pequeñas debido a que su plegamiento se estabiliza tanto por enlaces disulfuro como por iones tales como calcio, magnesio, manganeso, cobre, cinc, hierro, o una variedad de otros iones multivalentes, en lugar de estabilizarse por el núcleo hidrófobo típico.

25

**[0091]** El término “andamiaje” se refiere a la 'región estructural' o 'motivo de secuencia' de polipéptidos mínimo que se usa como secuencia común conservada en la construcción de bibliotecas de proteínas. Entre los residuos/posiciones fijos o conservados del andamiaje se encuentran posiciones variables e hipervariables. Se proporciona una diversidad de aminoácidos en las regiones variables entre los residuos de andamiaje fijos para proporcionar la unión específica a una molécula diana. Un andamiaje se define normalmente por los residuos conservados que se observan en un alineamiento de una familia de proteínas relacionadas con la secuencia. Los residuos fijos pueden requerirse para plegamiento o estructura, especialmente si las funciones de las proteínas alineadas son diferentes. Una descripción completa de un andamiaje de microproteína puede incluir el número, posición o separación y patrón de unión de las cisteínas, además de posición e identidad de cualquier residuo fijo en los bucles, que incluyen sitios de unión para iones tales como calcio.

30

35

**[0092]** El “pliegue” de una microproteína se define ampliamente por el patrón de enlace de los enlaces disulfuro (es decir, 1-4, 2-6, 3-5). Este patrón es una constante topológica y generalmente no es flexible a conversión en otro patrón sin no ligar y religar los disulfuros tales como mediante reducción y oxidación (agentes rédox). En general, proteínas naturales con secuencias relacionadas adoptan los mismos patrones de enlace disulfuro. Los principales determinantes son el patrón de distancia de cisteína (CDP) y algunos residuos no cis fijos, además de un sitio de unión a metal, si está presente. En algunos casos, el plegamiento de proteínas también está influido por las secuencias de alrededor (es decir, pro-péptidos) y en algunos casos por derivatización química (por ejemplo, gamma-carboxilación) de residuos que permiten que la proteína se una a iones metálicos divalentes (es decir,  $Ca^{++}$ ) que ayuda en su plegamiento. Para la gran mayoría de las microproteínas no se requiere tal ayuda en el plegamiento.

40

45

**[0093]** Sin embargo, proteínas con el mismo patrón de unión pueden todavía comprender múltiples pliegues, basándose en diferencias en la longitud y composición de los bucles que son suficientemente grandes para que la proteína dé una estructura bastante diferente. Un ejemplo son las familias de dominios de conotoxina, ciclotoxina y anato que tienen el mismo DBP, pero un CDP muy diferente, y se considera que son plegamientos diferentes. Determinantes de un pliegue de proteína son cualquier atributo que altere enormemente la estructura con respecto a un pliegue diferente, tal como el número y patrón de unión de las cisteínas, la separación de las cisteínas, diferencias en los motivos de secuencia de los bucles entre cisteínas (especialmente residuos de bucle fijos que es probable que se necesiten para el plegamiento, o en la localización o composición del sitio de unión a calcio (u otro metal o co-factor).

50

55

**[0094]** El término “patrón de enlaces disulfuro” o “DBP” se refiere al patrón de enlace de las cisteínas, que

están numeradas 1-n del extremo N con respecto al extremo C de la proteína. Los patrones de enlaces disulfuro son topológicamente constantes, que significa que solo pueden cambiarse no ligando uno o más disulfuros tales como usando condiciones rédox. Los posibles patrones de 2, 3 y 4 enlaces disulfuro se enumeran más adelante en los párrafos 0048-0075.

5

**[0095]** El término “patrón de distancia de cisteínas” o “CDP” se refiere al número de aminoácidos de no cisteína que separa las cisteínas sobre una cadena de proteína lineal. Se usan varias notaciones: C5C0C3C igual a C5CC3C igual a CxxxxCCxxxC.

10 **[0096]** El término 'Posición n6' o 'n7=4' se refiere a los bucles entre cisteínas y 'n6' se define como el bucle entre C6 y C7; 'n7=4' significa el bucle entre C7 y C8 tiene 4 aminoácidos de longitud, sin contar las cisteínas

**[0097]** Resistencia a la degradación en suero – Las proteínas pueden eliminarse por degradación en la sangre, que normalmente implica proteasas en el suero o plasma. La resistencia a la degradación en suero se mide combinando la proteína con suero o plasma humano (o ratón, rata, mono, según sea apropiado), normalmente durante un intervalo de días (es decir, 0,25, 0,5, 1, 2, 4, 8, 16 días) a 37 °C. Las muestras para estos momentos de tiempo se ejecutan entonces sobre un ensayo Western y la proteína se detecta con un anticuerpo. El anticuerpo puede ser una marca en la proteína. Si la proteína muestra una única banda sobre Western, en la que el tamaño de la proteína es idéntico al de la proteína inyectada, entonces no se ha producido degradación. El momento de tiempo al que el 50% de la proteína se degrada, como se juzga por Western, es la semivida de degradación en suero de la proteína

**[0098]** Unión a proteína del suero – Mientras que la MURP tiene normalmente varios módulos que se unen a dianas de la superficie de la célula y/o proteínas del suero, se desea que el URP carezca sustancialmente de actividades involuntarias. El URP debe diseñarse para minimizar evitar la interacción con (unión a) proteínas del suero, que incluyen anticuerpos. Pueden cribarse diferentes diseños de URP para la unión a proteínas del suero por ELISA, inmovilizando las proteínas del suero y luego añadiendo el URP, incubando, lavando y luego detectando la cantidad de URP unido. Un enfoque es detectar el URP usando un anticuerpo que reconoce una marca que se ha unido a URP. Un enfoque diferente es inmovilizar el URP (tal como mediante una fusión con GFP) y ponerlo en contacto con suero humano, incubando, lavando y luego detectando la cantidad de anticuerpos humanos que siguen unidos al URP usando anticuerpos secundarios como de cabra dirigido contra IgG humana. Usando estos enfoques, los presentes inventores han diseñado sus URP que muestran niveles muy bajos de unión a proteínas del suero. Sin embargo, en algunas aplicaciones se desea la unión a proteínas del suero o proteínas expuestas al suero, por ejemplo, debido a que puede prologar adicionalmente la semivida de secreción. En tales casos, pueden usarse estos mismos ensayos para diseñar URP que se unen a proteínas del suero o proteínas expuestas al suero tales como HSA o IgG. En otros casos, las MURP pueden ser módulos de unión dados que contienen péptidos que han sido diseñados para unirse a proteínas del suero o proteínas expuestas al suero tales como HSA o IgG.

*Polímeros recombinantes no estructurados (URP):*

40

**[0099]** La presente memoria descriptiva describe el diseño de polímeros recombinantes no estructurados (URP). Los URP de la invención son particularmente útiles para generar proteínas recombinantes de valor terapéutico y/o diagnóstico. Los URP de la invención presentan una o más de las siguientes características.

45 **[0100]** Los URP de la invención comprenden secuencias de aminoácidos que normalmente comparten características comunes con secuencias de péptidos desnaturalizadas bajo condiciones fisiológicas. Las secuencias de URP normalmente se comportan como secuencias de péptidos desnaturalizadas bajo condiciones fisiológicas. Las secuencias de URP carecen de estructuras secundarias y terciarias bien definidas bajo condiciones fisiológicas. Se ha establecido una variedad de procedimientos en la técnica para determinar las estructuras secundarias y terciarias de un polipéptido dado. Por ejemplo, la estructura secundaria de un polipéptido puede determinarse por espectroscopía de CD en la región espectral “UV lejana” (190-250 nm). Las estructuras de hélice alfa, hoja beta y de bobina al azar dan cada una lugar a una forma y magnitud característica de espectros de CD. La estructura secundaria también puede determinarse mediante ciertos programas informáticos o algoritmos tales como el algoritmo de Chou-Fasman (Chou, P. Y. y col. (1974) *Biochemistry*, 13: 222-45). Para una secuencia de URP dada, el algoritmo puede predecir si existe o no alguna estructura secundaria o absolutamente ninguna estructura secundaria. En general, las secuencias de URP tendrán espectros que se asemejan a proteínas desnaturalizadas debido a su bajo grado de estructura secundaria y terciaria. Si se desea, las secuencias de URP pueden diseñarse para tener conformaciones predominantemente desnaturalizadas bajo condiciones fisiológicas. Las secuencias de URP normalmente tienen un alto grado de flexibilidad conformacional bajo condiciones fisiológicas y tienden a tener

55

grandes radios hidrodinámicos (radio de Stokes) en comparación con proteínas globulares de peso molecular similar. Como se usa en el presente documento, condiciones fisiológicas se refiere a un conjunto de condiciones que incluyen temperatura, concentración de sales, pH que imitan a aquellas condiciones de un sujeto vivo. Se han establecido una multitud de condiciones fisiológicamente relevantes para su uso en ensayos *in vitro*. Generalmente, un tampón fisiológico contiene una concentración fisiológica de sal y ajustado a un pH neutro que oscila de aproximadamente 6,5 a aproximadamente 7,8, y preferentemente de aproximadamente 7,0 a aproximadamente 7,5. Una variedad de tampones fisiológicos se enumera en Sambrook y col. (1989), arriba, y de ahí que no se detalle en el presente documento. La temperatura fisiológicamente relevante oscila de aproximadamente 25 °C a aproximadamente 38 °C, y preferentemente de aproximadamente 30 °C a aproximadamente 37 °C.

10

**[0101]** Los URP de la invención pueden ser secuencias con baja inmunogenicidad. La baja inmunogenicidad puede ser un resultado directo de la flexibilidad conformacional de las secuencias de URP. Muchos anticuerpos reconocen los llamados epítopos conformacionales en antígenos de proteína. Los epítopos conformacionales están formado por regiones de la superficie de la proteína que están compuestas de múltiples secuencias de aminoácidos discontinuas del antígeno de proteína. El plegamiento preciso de la proteína lleva a estas secuencias a una configuración especial bien definida que puede ser reconocida por anticuerpos. Los URP preferidos se diseñan para evitar la formación de epítopos conformacionales. Por ejemplo, son de particular interés secuencias de URP que tienen una baja tendencia a adaptarse a conformaciones compactamente plegadas en disolución acuosa. En particular, la baja inmunogenicidad puede lograrse eligiendo secuencias que resisten al procesamiento de antígenos en células presentadoras de antígeno, eligiendo secuencias que no se unen a bien MHC y/o eligiendo secuencias que se derivan de secuencias humanas.

15

20

**[0102]** Los URP de la invención pueden ser secuencias con un alto grado de resistencia a proteasas. La resistencia a proteasas también puede ser un resultado de la flexibilidad conformacional de las secuencias de URP. La resistencia a proteasas puede diseñarse evitando sitios de reconocimiento de proteasas conocidos. Alternativamente, las secuencias resistentes a proteasas pueden seleccionarse por expresión de fago o técnicas relacionadas de bibliotecas de secuencias aleatorias o semialeatorias. Si se desea para aplicaciones especiales, tales como liberación lenta de una preparación de proteína de liberación controlada, pueden construirse sitios de escisión por proteasas del suero en un URP. Son de particular interés secuencias de URP con alta estabilidad (por ejemplo, larga semivida en suero, menos propensas a la escisión por proteasas presentes en fluido corporal) en sangre.

25

30

**[0103]** El URP de la invención también puede caracterizarse por el efecto por el cual tras su la incorporación en una proteína, la proteína presenta una semivida en suero prolongada y/o mayor solubilidad con respecto a la proteína correspondiente que es deficiente en URP. [En la técnica se conocen procedimientos de determinación de la semivida en suero (véase, por ejemplo, Alvarez, P. y col. (2004) J Biol Chem, 279: 3375-81). Puede determinarse fácilmente si la proteína resultante tiene una semivida en suero prolongada con respecto a la proteína sin modificar poniendo en práctica cualquier procedimiento disponible en la materia o ejemplificado en el presente documento.

35

40

**[0104]** El URP de la invención puede ser de cualquier longitud necesaria para efectuar (a) la extensión de la semivida en suero de una proteína que comprende el URP; (b) un aumento en la solubilidad de la proteína resultante; (c) un incremento de la resistencia a proteasa; y/o (d) una inmunogenicidad reducida de la proteína resultante que comprende el URP. El URP de la invención tiene al menos 40, 50, 60, 70, 80, 90, 100, 150, 200, 300, 400 o más aminoácidos contiguos. Cuando se incorpora en una proteína, el URP puede fragmentarse de forma que la proteína resultante contenga múltiples URP, o múltiples fragmentos de URP. Preferiblemente, la proteína resultante tiene una longitud combinada de secuencias de URP que supera los 40, 50, 60, 70, 80, 90, 100, 150, 200 o más aminoácidos.

45

50

**[0105]** Los URP pueden tener un punto isoeléctrico (pI) de 1,0, 1,5, 2,0, 2,5, 3,0, 3,5, 4,0, 4,5, 5,0, 5,5, 6,0, 6,5, 7,0, 7,5, 8,0, 8,5, 9,0, 9,5, 10,0, 10,5, 11,0, 11,5, 12,0, 12,5 o incluso 13,0.

55

**[0106]** En general, las secuencias de URP son ricas en aminoácidos hidrófilos y contienen un bajo porcentaje de aminoácidos hidrófobos o aromáticos. Residuos hidrófilos adecuados incluyen, pero no se limitan a, glicina, serina, aspartato, glutamato, lisina, arginina y treonina. Los residuos hidrófobos que son menos favorecidos en la construcción de URP incluyen triptófano, fenilalanina, tirosina, leucina, isoleucina, valina y metionina. Las secuencias de URP pueden ser ricas en glicina, pero las secuencias de URP también pueden ser ricas en los aminoácidos glutamato, aspartato, serina, treonina, alanina o prolina. Así, el aminoácido predominante puede ser G, E, D, S, T, A o P. La inclusión de residuos de prolina tiende a reducir la sensibilidad a degradación proteolítica.

**[0107]** La inclusión de residuos hidrófilos normalmente aumenta la solubilidad de URP en agua y medios acuosos bajo condiciones fisiológicas. Como resultado de su composición de aminoácidos, las secuencias de URP tienen una baja tendencia a formar agregados en formulaciones acuosas y la fusión de secuencias de URP a otras proteínas o péptidos tiende a potenciar su solubilidad y reducir su tendencia a formar agregados, que es un mecanismo separado para reducir la inmunogenicidad.

**[0108]** Las secuencias de URP pueden diseñarse para evitar que ciertos aminoácidos confieran propiedades no deseables a la proteína. Por ejemplo, pueden diseñarse secuencias de URP que contengan pocos o ninguno de los siguientes aminoácidos: cisteína (para evitar la formación y oxidación de disulfuro), metionina (para evitar la oxidación), asparagina y glutamina (para evitar la desamidación).  
*URP ricos en glicina:*

**[0109]** En una realización, el URP de la invención comprende una secuencia rica en glicina (GRS). Por ejemplo, la glicina puede estar presente predominantemente de forma que sean los residuos más prevalentes presentes en la secuencia de interés. En otro ejemplo, las secuencias de URP pueden diseñarse de forma que los residuos de glicina constituyan al menos aproximadamente el 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100% de los aminoácidos totales. Los URP también puede contener el 100% de glicinas. En otro ejemplo más, los URP contienen al menos 30% de glicina y la concentración total de triptófano, fenilalanina, tirosina, valina, leucina e isoleucina es entonces menos del 20%. En otro ejemplo adicional, los URP contienen al menos 40% de glicina y la concentración total de triptófano, fenilalanina, tirosina, valina, leucina e isoleucina es entonces menos del 10%. En todavía otro ejemplo, los URP contienen al menos aproximadamente 50% de glicina y la concentración total de triptófano, fenilalanina, tirosina, valina, leucina e isoleucina es entonces menos del 5%.

**[0110]** La longitud de GRS pueden variar entre aproximadamente 5 aminoácidos y 200 aminoácidos o más. Por ejemplo, la longitud de un GRS contiguo individual puede contener 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 240, 280, 320 ó 400 o más aminoácidos. La GRS puede comprender residuos de glicina en ambos extremos.

**[0111]** La GRS también puede tener un contenido significativo de otros aminoácidos, por ejemplo, Ser, Thr, Ala o Pro. La GRS puede contener una fracción significativa de aminoácidos negativamente cargados que incluyen, pero no se limitan a, Asp y Glu. La GRS puede contener una fracción significativa de aminoácidos positivamente cargados que incluyen, pero no se limitan a, Arg o Lys. Si se desea, los URP pueden diseñarse para contener solo un único tipo de aminoácido (es decir, Gly o Glu), algunas veces solo algunos tipos de aminoácidos, por ejemplo, dos a cinco tipos de aminoácidos (por ejemplo, seleccionados de G, E, D, S, T, A y P), a diferencia de proteínas típicas y ligadores típicos que generalmente están compuestos de la mayoría de los veinte tipos de aminoácidos. Los URP pueden contener residuos negativamente cargados (Asp, Glu) en el 30, 25, 20, 15, 12, 10, 9, 8, 7, 6, 5, 4, 3, 2 o el 1 por ciento de las posiciones de aminoácidos.

**[0112]** Los URP que contienen GRS son de particular interés debido a, en parte, la elevada libertad conformacional de los péptidos que contienen glicina. Los péptidos desnaturalizados en disolución tienen un alto grado de libertad conformacional. La mayor parte de esa libertad conformacional se pierde tras la unión de dichos péptidos a una diana como un receptor, un anticuerpo o una proteasa. Esta pérdida de entropía necesita compensarse por la energía de interacción entre el péptido y su diana. El grado de libertad conformacional de un péptido desnaturalizado depende de sus secuencias de aminoácidos. Los péptidos que contienen muchos aminoácidos con cadenas laterales pequeñas tienden a tener más libertad conformacional que los péptidos que están compuestos de aminoácidos con cadenas laterales mayores. Los péptidos que contienen el aminoácido glicina tienen grados particularmente grandes de libertad. Se ha estimado que los enlaces peptídicos que contienen glicina tienen aproximadamente 3,4 veces más entropía en disolución con respecto a secuencias que contienen alanina correspondientes (D'Aquino, J. A. y col. (1996) *Proteins*, 25: 143-56). Este factor aumenta con el número de residuos de glicina en una secuencia. Como resultado, tales péptidos tienden a perder más entropía tras la unión a dianas, que reduce su capacidad global para interactuar con otras proteínas, además de su capacidad para adoptar estructuras tridimensionales definidas. La gran flexibilidad conformacional de los enlaces peptídicos de glicina también es evidente cuando se analizan representaciones de Ramachandran de estructuras de proteína en las que los enlaces peptídicos de glicina ocupan áreas que están raramente ocupadas por otros enlaces peptídicos (Venkatachalam, C. M. y col. (1969) *Annu Rev Biochem*, 38: 45-82). Stites y col. estudiaron una base de datos de 12.320 residuos de 61 estructuras cristalinas de alta resolución no homólogas para determinar las preferencias conformaciones phi, psi de cada uno de los 20 aminoácidos. Las distribuciones observadas en el estado nativo de proteínas se supone que también refleja las distribuciones encontradas en el estado desnaturalizado. Las

distribuciones se usaron para aproximar la superficie de energía a cada residuo, permitiendo el cálculo de entropías conformaciones relativas para cada residuo con respecto a glicina. En el caso más extremo, la sustitución de prolina con glicina, los cambios de entropía conformacional estabilizarán el estado nativo con respecto al estado desnaturalizado  $-0,82 \pm 0,08$  kcal/mol a 20 °C (Stites, W. E. y col. (1995) *Protein*, 22: 132). Estas observaciones 5 confirman la función especial de la glicina entre los 20 aminoácidos naturales.

[0113] En el diseño de los URP de la invención pueden usarse secuencias naturales o no naturales. Por ejemplo, una multitud de secuencias naturales que contienen alto contenido de glicina se proporciona en la Tabla 1, Tabla 2, Tabla 3 y Tabla 4. Un experto en la materia puede adoptar una cualquiera de las secuencias como un URP, 10 o modificar las secuencias para lograr las propiedades previstas. Si es de interés la inmunogenicidad para el sujeto huésped, es preferible diseñar URP que contengan GRS basándose en secuencias ricas en glicina derivadas del huésped. URP que contienen GRS preferidos son secuencias de proteínas humanas o secuencias que comparten homología sustancial con las secuencias ricas en glicina correspondientes en las proteínas humanas de referencia.

15 **Tabla 1. Análisis estructural de proteínas que contienen secuencias ricas en glicina**

Archivo de PDB	Función de la proteína	Secuencias ricas en glicina
1K3V	Cápside del parvovirus porcino	sgggggggggrgagg
1FPV	Virus de la panleucopenia felina	tgsgngsggggggsgg
1IJS	Cepa D de CpV, mutante A300d	tgsgngsggggggsgg
1MVM	Virus de Mvm (cepa I)	ggsggggsgggg

**Tabla 2: Marcos de lectura abiertos que codifican GRS con 300 o más residuos de glicina**

Acceso	Organismo	Gly (%)	Longitud de GRS	Longitud del gen	Función predicha
NP_974499	<i>Arabidopsis thaliana</i>	64	509	579	desconocida
ZP_00458077	<i>Burkholderia cenocepacia</i>	66	373	518	lipoproteína putativa
XP_477841	<i>Oryza sativa</i>	74	371	422	desconocida
MP_910409	<i>Oryza sativa</i>	75	368	400	precursor putativo de la pared celular
NP_610660	<i>Drosophila melanogaster</i>	66	322	610	elemento transponible

20

**Tabla 3. Ejemplos de GRS humana**

Acceso	Gly (%)	Longitud de GRS	Longitud del gen	Hidrofobia	Función predicha
NP_000217	62	135	622	sí	queratina 9
NP_631961	61	73	592	sí	isoforma 1 del factor 15 asociado a TBP
NP_476429	65	70	629	sí	queratina 3
NP_000418	70	66	316	sí	loricrina, envuelta de la célula
NP_056932	60	66	638	sí	citoqueratina 2

**Tabla 4. Ejemplos adicionales de GRS humana**

25

Acceso	Secuencias	Número de aminoácidos
NP_006228	GPGGGGGPGGGGPPGGGGPGGGGGGGGPGGGGGGPGGG	37
NP_787059	GAGGGGGGGGGGGSSGGGGGGGAGAGGAGAG	33
NP_009060	GGGSGGGAGGGSGGGSGGGGGGAGGGGGG	32
NP_031393	GDGGGAGGGGGGGSGGGSGGGGGGG	27
NP_005850	GSGSGSGGGGGGGGGGGSGGGGGG	25
NP_061856	GGGRGRRGGRRGGRRGGRRGG	22
NP_787059	GAGGGGGGGGGGGSGGGGGGGGAGAGGAGAG	33
NP_009060	GGGSGSGGAGGGSGGGSGGGGGGAGGGGGG	32
NP_031393	GDGGGAGGGGGGGSGGGSGGGGGGG	27
NP_115818	GSGGSGSGGGPPGPPGGGGG	21
XP_376532	GEGGGGGEGGGAGGGSG	18
NP_065104	GGGGGGGGDGGG	12

- GGGSGGGAGGGSGGGSGGGGGAGGGGGSSGGGSGTAGGHS  
 dominio POU, clase 4, factor de transcripción 1 [Homo sapiens]  
GPGGGGPGGGGGPGGGPGGGGGGGPGGGGGPGGGGGPGGG
- 5 2 que contiene dominio YEATS [Homo sapiens]  
GGSGAGGGGGGGGGGGSGSGGGGSGTGGGGTAGGG  
 isoforma 3 de dominio 1B interactivo rico en AT (similar a SWI1); proteína de unión a BRG1 ELD/OSA1; Eld (párpado)/proteína Osa [Homo sapiens]  
GAGGGGGGGGGGGGGSGGGGGGGGAGAGGAGAG
- 10 isoforma 2 de dominio 1B interactivo rico en AT (similar a SWI1); proteína de unión a BRG1 ELD/OSA1; Eld (párpado)/proteína Osa [Homo sapiens]  
GAGGGGGGGGGGGGGSGGGGGGGGAGAGGAGAG  
 isoforma 1 de dominio 1B interactivo rico en AT (similar a SWI1); proteína de unión a BRG1 ELD/OSA1; Eld (párpado)/proteína Osa [Homo sapiens]
- 15 GAGGGGGGGGGGGGGSGGGGGGGGAGAGGAGAG  
 proteína A de unión al elemento rico en purina; proteína alfa de unión a ADN monocatenario rico en purina; proteína PUR-alfa del activador transcripcional [Homo sapiens]  
GHPGSGSGGGGGGGGGGGGGSGGGGGGAPGG  
 factor X1 regulador; factor 1 regulador de acción trans; factor C potenciador; factor RFX regulador del MHC de clase II [Homo sapiens]
- 20 GGGSGGGGGGGGGGGGGGGSGTGGGGSGAG  
 proteína que contiene dominio de promo alterada en leucemia [Homo sapiens]  
GGRGRGGRGRGRSRGRGGGTRGRGRGRGRGRG  
 proteína desconocida [Homo sapiens]
- 25 GSGGSGGSGGGPGPGGGGGPSGSGSGPG  
 PREDICHA: proteína hipotética XP\_059256 [Homo sapiens]  
GGGGGGGGGGRRGGGRRGGRRGGGEGGG  
 proteína de dedo de cinc 281; factor de transcripción de ZNP-99 [Homo sapiens]  
GGGGTSSSGSGSGGGGGSGGGGGGGSSG
- 30 isoforma corta de la proteína de unión a ARN (autoantigénica, asociada a hnRNP con amarillo letal); proteína de unión a ARN (autoantigénica); proteína de unión a ARN (autoantigénica, asociada a hnRNP con amarillo letal) [Homo sapiens]  
GDGGGAGGGGGGGGGSGGGGGSGGGGGGG  
 partícula de reconocimiento de señales de 68 kDa [Homo sapiens]
- 35 GGGGGGGGSGGGGGSGGGGGSGGGRGAGG  
 proteína KIAA0265 [Homo sapiens]  
GGGAAGAGGGGGSGAGGGSGSGGRGTG  
 homólogo de *engrailed 2*; *Engrailed-2* [Homo sapiens]  
GAGGGRGGGAGGEGGASGAEGGGGAGG
- 40 isoforma larga de proteína de unión a ARN (autoantigénica, asociada a hnRNP con amarillo letal), proteína de unión a ARN (autoantigénica); proteína de unión a ARN (autoantigénica, asociada a hnRNP con amarillo letal) [Homo sapiens]  
GDGGGAGGGGGGGGGSGGGGGSGGGGGGG  
 receptor de andrógenos; receptor de dihidrotestosterona [Homo sapiens]
- 45 GGGGGGGGGGGGGGGGGGGGGGGEAG  
 caja homeótica D11; caja homeótica 4F; Hox-4.6, ratón, homólogo de; proteína de caja homeótica Hox-D11 [Homo sapiens]  
GGGGGGSAGGGSSGGGPGGGGGAGG  
 frizzled 8; homólogo 8 de frizzled (*Drosophila*) [Homo sapiens]
- 50 GGGGPGGGGGGGPGGGGGPGGGGG  
 gen asociado al desarrollo ocular [Homo sapiens]  
GRGGAGSGGAGSGAAGGTGSSGGGG  
 caja homeótica B3; caja homeótica 2G; proteína de la caja homeótica Hox-B3 [Homo sapiens]  
GGGGGGGGGGGGSGGSGGGGGGGGGGG
- 55 marco de lectura abierto 29 del cromosoma 2 [Homo sapiens]  
GGSGGGRGGASGPGSGSGGPGGPAG  
 proteína DKFZP564F0522 [Homo sapiens]  
GGHHGDRGGGRGGRGGRGGRGGRAG  
 PREDICHA: similar a proteína 2 homóloga a even-skipped (EVX-2) de caja homeótica [Homo sapiens]

- GSRGGGGGGGGGGGGGGGGAGAGGG  
 familia de genes homólogos a ras, miembro U; Ryu GTPasa; homólogo Cdc42 sensible a Wnt-1; 2310026M05Rik; 1 similar a proteína de unión a GTP; GTPasa similar a CDC42 [Homo sapiens]
- 5 GGRGGRGPGEPGGRGRAGGAEGRG  
 proteína scratch 2; scratch 2 de represor transcripcional; scratch (homólogo de Drosophila) 2, proteína de dedo de cinc [Homo sapiens]
- GGGGDAGGSGDAGGAGGRAGRAG  
 familia A de la proteína nucleolar, miembro 1; proteína GAR1 [Homo sapiens]
- 10 GGGRGGRGGGRGGGGRGGGRGGG  
 queratina 1; queratina-1; citoqueratina 1; alfa-proteína del pelo [Homo sapiens]
- GGSGGGGGSSGGRGSGGGSSGG  
 proteína hipotética FLJ31413 [Homo sapiens]
- GSGPCTGGGGSGSGGGGGSSGG  
 dominio onecut, miembro 2 de la familia; onecut 2 [Homo sapiens]
- 15 GARGGGSGGGGGGGGGGGGGPG  
 dominio POU, clase 3, factor de transcripción 2 [Homo sapiens]
- GGGGGGGGGGGGGGGGGGGGDG  
 PREDICHA: similar a la subunidad 4 del completo THO (Tho4) (ARN y proteína 1 de unión a factor de exportación) (REF1-I) (Ally de AML-1 y LEF-1) (Aly/REF) [Homo sapiens]
- 20 GGTRGGTRGGTRGGDRGRGRGAG  
 PREDICHA: similar a la subunidad 4 del complejo THO (Tho4) (ARN y proteína 1 de unión a factor de exportación) (REF1-I) (Ally de AML-1 y LEF-1) (Aly/REF) [Homo sapiens]
- GGTRGGTRGGTRGGDRGRGRGAG  
 dominio POU, clase 3, factor de transcripción 3 [Homo sapiens]
- 25 GAGGGGGGGGGGGGGGAGGGGGG  
 familia A de la proteína nucleolar, miembro 1; proteína GAR1 [Homo sapiens]
- GGGRGGRGGGRGGGGRGGGRGGG  
 fibrilarina; antígeno de esclerodermia nucleolar de 34 kD; ARN, proteína 1 de interacción con nucleolar pequeña U3 [Homo sapiens]
- 30 GRGRGGGGGGGGGGGGGRGGGG  
 proteína 579 de dedo de cinc [Homo sapiens]
- GRGRGRGRGRGRGRGRGGAG  
 calpaína, subunidad 1 pequeña; proteinasa neutra activada con calcio; calpaína, polipéptido pequeño; calpaína 4, subunidad pequeña (30 K); proteasa dependiente del calcio, subunidad pequeña [Homo sapiens]
- 35 GAGGGGGGGGGGGGGGGGGGGG  
 queratina 9 [Homo sapiens]
- GGGSGGGHSGGSGGGHSGGSGG  
 caja forkhead D1; activador 4 relacionado con forkhead; Forkhead, drosophila, 8 similar a homólogo; 8 similar a forkhead (Drosophila) [Homo sapiens]
- 40 GAGAGGGGGGGGAGGGGSAGSG  
 PREDICHA: similar a ADNc de RIKEN C230094B15 [Homo sapiens]
- GGPGTGSGGGGAGTGGGAGGPG  
GGGGGGGGGAGGAQAGSAGGG  
 precursor de cadherina 22, ortólogo de cadherina de rata PB [Homo sapiens]
- 45 GGDGGGSAGGGAGGGSGGGAG  
 factor 1 de transcripción de unión a AT; factor 1 de unión a motivo AT [Homo sapiens]
- GGGGGGSGGGGGGGGGGGGGG  
 eomesodermina; caja t, cerebro, 2; homólogo de eomesodermina (Xenopus laevis) [Homo sapiens]
- GPGAGAGSAGGSSGGGGPG  
 proteína de transferencia de fosfatidilinositol, 2 asociada a membrana; receptor 3 de interacción con el dominio del extremo N de PYK2; B alfa 2 de degeneración retiniana (Drosophila) [Homo sapiens]
- 50 GGGGGGGGGGSSGGGGSSGG  
 isoforma 2 del antígeno 8 asociado a espermatozoides; proteína 1 de la membrana de espermatozoides [Homo sapiens]
- 55 GSGSGPGPGSGPGSGPGHGSG  
 PREDICHA: proteína 27 de motivo de unión a ARN [Homo sapiens]
- GPGPGPGPGPGPGPGPGPG  
 isoforma 1 de la proteína 1 de unión a la subunidad gamma de AP1; proteína 1 de unión a la subunidad gamma del complejo 1 de gamma-sinergina; proteína relacionada con adaptador [Homo sapiens]

- GAGSGGGGAAGAGAGSAGGGG  
 isoforma 2 de la proteína 1 de unión a la subunidad gamma de AP1; gamma-sinergina; proteína 1 de unión a la subunidad gamma del complejo 1 de proteína relacionada con adaptador [Homo sapiens]
- 5 GAGSGGGGAAGAGAGSAGGGG  
 repetición de anquirina y 1 que contiene dominio del motivo alfa estéril; repetición de anquirina y 1 que contiene dominio SAM [Homo sapiens]
- GGGGGGGSGGGGGGSGGGGGG  
 isoforma 1 de la proteína 2 de dominio de unión a metil-CpG [Homo sapiens]
- 10 GRGRGRGRGRGRGRGRGRGRG  
 dominio funcional triple (que interacciona con PTPRF) [Homo sapiens]
- GGGGGGGSGGGGGGSGGGGGG  
 caja forkhead D3 [Homo sapiens]
- GGEEGGASGGGPGAGSGSAGG  
 isoforma 1 del antígeno 8 asociado a espermatozoides; proteína 1 de la membrana de espermatozoides [Homo sapiens]
- 15 GSGSGPGPGSGPGSGPGHGSG  
 isoforma específica de testículos de la proteína 2 del dominio de unión a metil-CpG [Homo sapiens]
- GRGRGRGRGRGRGRGRGRGRG  
 regulador de muerte celular aven; muerte celular programada 12 [Homo sapiens]
- 20 GGGGGGGGDGGGRRGRGRGRG  
 regulador de transcritos 1 de terminación; delta helicasa; homólogo de la mutación 1 de Up-Frameshift (S. cerevisiae); factor 1 reductor de ARNm de terminación; homólogo de levadura Upflp [Homo sapiens]
- GGPGPGGGGAGGPGGAGAG  
 isoforma a de la proteína 2 de canales de potasio activados por calcio de pequeña conductancia; canal de potasio
- 25 GTGGGGSTGGGGGGGGSGHG  
 activado por Ca<sup>2+</sup> de pequeña conductancia sensible a apamina [Homo sapiens]
- GTGGGGSTGGGGGGGGSGHG  
 caja 1 de SRY (región Y determinante del sexo); gen 1 de la caja HMG relacionado con SRY [Homo sapiens]
- GPAGAGGGGGGGGGGGGGGGG  
 isoforma 2 del factor de transcripción 20; proteína de unión a elementos sensibles al factor de crecimiento derivado
- 30 GGTGGSSGSSGSGSGGGRRG  
 de plaquetas de estromelina-1; proteína de unión a elementos sensibles a PDGF de estromelina 1; proteína de unión a SPRE; factor SPBP nuclear [Homo sapiens]
- GGTGGSSGSSGSGSGGGRRG  
 isoforma 1 del factor de transcripción 20; proteína de unión a elementos sensibles al factor de crecimiento derivado
- 35 GGTGGSSGSSGSGSGGGRRG  
 de plaquetas de estromelina-1; proteína de unión a elementos sensibles a PDGF de estromelina 1; proteína de unión a SPRE; factor SPBP nuclear [Homo sapiens]
- GSQTGTTGSSGAGGPGTPGG  
 proteína 1 de interacción con ras [Homo sapiens]
- 40 GGSGGGAAGGGAGGAGAGAG  
 isoforma b de cinasa inducible por BMP-2 [Homo sapiens]
- GGSGGGAAGGGAGGAGAGAG  
 isoforma a de cinasa inducible por BMP-2 [Homo sapiens]
- GSSGGGGGAGAAAGGAGGAG  
 caja forkhead C1; activador 3 relacionado con forkhead; Forkhead, drosophila, 7 similar a homólogo; 7 similar a forkhead (Drosophila); iridogoniodisgénesis tipo 1 [Homo sapiens]
- 45 GPGPSGGPGGGGGGGGGGGGG  
 factor p54 de corte y empalme; proteína nuclear de 54 kDa rica en arginina [Homo sapiens]
- GGGGGGGGGGGGGGAAGAGG  
 homólogo oncogénico del fibrosarcoma musculoponeurótico de v-maf; protooncogén de fibrosarcoma musculoponeurótico aviar (MAF); homólogo oncogénico del fibrosarcoma musculoponeurótico de v-maf (aviar)
- 50 GRGRGRGRGRGRGRGRGRGG  
 [Homo sapiens]
- GSAGGSSGAAGAAGGGAGAG  
 polipéptido de ribonucleoproteína D1 nuclear pequeña 16 kDa; proteína D1 de núcleo de snRNP; autoantígeno Sm-D; polipéptido de ribonucleoproteína D1 nuclear pequeña (16kD) [Homo sapiens]
- 55 GSAGGSSGAAGAAGGGAGAG  
 proteína hipotética H41 [Homo sapiens]

**URP que contienen residuos que no son glicina (NGR):**

**[0114]** Las secuencias de residuos de no glicina en estas GRS pueden seleccionarse para optimizar las propiedades de URP y de ahí las proteínas que contienen los URP deseados. Por ejemplo, pueden optimizarse las secuencias de URP para potenciar la selectividad de la proteína resultante para un tejido particular, tipo de célula específica o linaje celular. Por ejemplo, pueden incorporarse secuencias de proteínas que no se expresan 5 ubicuamente, sino que se expresan diferencialmente en uno o más de los tejidos del cuerpo que incluyen corazón, hígado, próstata, pulmón, riñón, médula ósea, sangre, piel, vejiga, cerebro, músculos, nervios y tejidos seleccionados que son afectados por enfermedades tales como enfermedades infecciosas, enfermedad autoinmunitaria, trastornos renales, neuronales, cardíacos y cánceres. Pueden emplearse secuencias representativas de un origen de desarrollo específico, tales como aquellas expresadas en un embrión o un adulto, 10 durante la formación de ectodermo, endodermo o mesodermo en un organismo pluricelular. También pueden utilizarse secuencias que participan en un proceso biológico específico que incluye, pero no se limita a, regulación del ciclo celular, diferenciación celular, apoptosis, quimiotaxis, motilidad celular y transposición citoesquelética. También pueden utilizarse otras secuencias de proteínas no ubicuamente expresadas para dirigir la proteína resultante a localizaciones subcelulares específicas: matriz extracelular, núcleo, citoplasma, citoesqueleto, plasma 15 y/o estructuras membranosas intracelulares que incluyen, pero no se limitan a, depresiones recubiertas, aparato de Golgi, retículo endoplásmico, endosoma, lisosoma y mitocondria.

**[0115]** Se conoce una variedad de estas secuencias específicas de tejido, específicas de tipo de célula, específicas de localización subcelular y están disponibles de numerosas bases de datos de proteínas. Tales 20 secuencias de URP selectivas pueden obtenerse generando bibliotecas de secuencias de URP aleatorias o semi-aleatorias, inyectándolas en animales o pacientes, y determinando secuencias con la selectividad por tejido deseada en muestras de tejido. La determinación de secuencias puede realizarse por espectrometría de masas. Usando procedimientos similares pueden seleccionarse secuencias de URP que facilitan la captación oral, bucal, intestinal, nasal, tecal, peritoneal, pulmonar, rectal o dérmica.

**[0116]** Son de particular interés secuencias de URP que contienen regiones que son relativamente ricas en 25 los aminoácidos positivamente cargados arginina o lisina que favorecen la captación celular o transporte por las membranas. Pueden diseñarse secuencias de URP que contienen una o varias secuencias sensibles a proteasas. Tales secuencias de URP pueden escindirse una vez que el producto de la invención ha alcanzado su localización 30 diana. Esta escisión puede desencadenar un aumento en la potencia del dominio farmacéuticamente activo (activación de pro-fármacos) o puede potenciar la unión del producto de escisión a un receptor. Las secuencias de URP pueden diseñarse para llevar un exceso de cargas negativas introduciendo residuos de ácido aspártico o ácido glutámico. Son de particular interés URP que contienen más del 5%, más del 6%, 7%, 8%, 9%, 10%, 15%, 30% o 35 más ácido glutámico y menos del 2% de lisina o arginina. Tales URP llevan un exceso de carga negativa y como resultado tienen una tendencia a adoptar conformaciones abiertas debido a la repulsión electrostática entre cargas negativas individuales del péptido. Un exceso de carga negativa tal conduce a un aumento eficaz en su radio hidrodinámico y como resultado puede conducir a eliminación renal reducida de tales moléculas. Así, puede modularse la carga neta eficaz y el radio hidrodinámico de una secuencia de URP controlando la frecuencia y 40 distribución de aminoácidos negativamente cargados en las secuencias de URP. La mayoría de los tejidos y superficies en un ser humano o animal llevan exceso de cargas negativas. Diseñando las secuencias de URP para llevar un exceso de cargas negativas pueden minimizarse las interacciones no específicas entre la proteína resultante que comprende el URP y diversas superficies tales como vasos sanguíneos, tejidos sanos o diversos receptores.

**[0117]** Los URP pueden tener una secuencia repetitiva de aminoácidos del formato (Motivo)<sub>x</sub> en el que un 45 motivo de secuencia forma una repetición directa (es decir, ABCABCABCABC) o una repetición invertida (ABCCBAABCCBA) y el número de estas repeticiones puede ser 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 35, 40, 50 o más URP o las repeticiones dentro de URP frecuentemente contienen solo 1, 2, 3, 4, 5 ó 6 50 tipos diferentes de aminoácidos. Los URP normalmente consisten en repeticiones de secuencias de aminoácidos humanas que tienen 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 24, 26, 28, 30, 32, 34, 36 o más aminoácidos de longitud, pero los URP también pueden consistir en secuencias de aminoácidos no humanas que tienen 20, 22, 24, 26, 28, 30, 32, 34, 36, 38 40, 42, 44, 46, 48, 50 aminoácidos de longitud.

#### **URP derivados de secuencias humanas:**

**[0118]** Los URP pueden derivarse de secuencias humanas. El genoma humano contiene muchas 55 subsecuencias que son ricas en un aminoácido particular. Son de particular interés aquellas secuencias de aminoácidos que son ricas en un aminoácido hidrófilo como serina, treonina, glutamato, aspartato o glicina. Son de particular interés aquellas subsecuencias que contienen pocos aminoácidos hidrófobos. Se predice que tales

subsecuencias son no estructuradas y altamente solubles en disolución acuosa. Tales subsecuencias humanas pueden modificarse para mejorar adicionalmente su utilidad. La Figura 17 muestra una secuencia humana a modo de ejemplo que es rica en serina y que puede aislarse como URP de la invención. La dentina sialofosfoproteína ejemplificada contiene una subsecuencia de 670 aminoácidos en la que el 64% de los residuos son serina y la mayoría de las otras posiciones son aminoácidos hidrófilos tales como aspartato, asparaginas y glutamato. La secuencia es extremadamente repetitiva y como resultado tiene un bajo contenido de información. Pueden usarse directamente subsecuencias de una proteína humana tal. Si se desea, la secuencia puede modificarse de una forma que preserve su carácter global, pero que la haga más adecuada para aplicaciones farmacéuticas. Ejemplos de secuencias que están relacionadas con dentina sialofosfoproteína son  $(SSD)_n$ ,  $(SSDSSN)_n$ ,  $(SSE)_n$  en las que n es entre aproximadamente 4 y 200.

**[0119]** El uso de secuencias de proteínas humanas es particularmente deseable en el diseño de URP con inmunogenicidad reducida en un sujeto humano. Una etapa clave para provocar una respuesta inmunitaria a una proteína extraña es la presentación de fragmentos de péptido de dicha proteína por receptores de MHC de clase II. Estos fragmentos unidos a MHCII pueden luego detectarse por receptores de linfocitos T, que desencadenan la proliferación de linfocitos T colaboradores e inicia una respuesta inmunitaria. La eliminación de epítomos de linfocitos T de proteínas farmacéuticas se ha reconocido como un medio para reducir el riesgo de provocación de una reacción inmunitaria (Stickler, M. y col. (2003) *J Immunol Methods*, 281: 95-108). Los receptores de MHCII normalmente interaccionan con un epítipo que tiene, por ejemplo, una región de 9 aminoácidos de longitud de los péptidos expresados. Así, puede reducirse el riesgo de provocar una respuesta inmunitaria a una proteína en pacientes si todas las subsecuencias o la mayoría de las posibles subsecuencias 9-meras de la proteína pueden encontrarse en proteínas humanas y, si es así, estas secuencias y repeticiones de estas secuencias no serán reconocidas por el paciente como secuencias extrañas. Pueden incorporarse secuencias humanas en el diseño de secuencias de URP oligomerizando o concatenando secuencias humanas que tienen composiciones de aminoácidos adecuadas. Estas pueden ser repeticiones directas o repeticiones invertidas o mezclas de diferentes repeticiones. Por ejemplo, las secuencias pueden oligomerizarse como se muestra en la Tabla 2. Tales oligómeros tienen un reducido riesgo de ser inmunogénicas. Sin embargo, las secuencias de empalme entre las unidades de monómeros todavía pueden contener epítomos de linfocitos T que pueden desencadenar una reacción inmunitaria, que se ilustra en la Figura 3. Adicionalmente puede reducirse el riesgo de provocar una respuesta inmunitaria diseñando secuencias de URP basadas en múltiples secuencias humanas de solapamiento. Este enfoque se ilustra en la Figura 4. La secuencia de URP en la Figura 2 se diseñó como un oligómero basado en múltiples secuencias humanas de forma que cada subsecuencia 9-mera del oligómero pueda encontrarse en una proteína humana. En estos diseños, cada subsecuencia 9-mera es una secuencia humana. Un ejemplo de una secuencia de URP basada en tres secuencias humanas se muestra en la Figura 5. También es posible diseñar secuencias de URP basadas en una única secuencia humana de forma que todas las posibles subsecuencias 9-meras en las secuencias de URP oligoméricas se produzcan en la misma proteína humana. Un ejemplo se muestra en la Figura 6 basándose en el dominio POU que es rico en glicina y prolina. El monómero de repetición en la secuencia de URP es solo un fragmento de la proteína humana y sus secuencias flanqueantes son idénticas a la unidad de repetición como se ilustra en la Figura 6. También pueden diseñarse secuencias de URP no oligoméricas basándose en proteínas humanas. Las condiciones primarias son que todas las subsecuencias 9-meras puedan encontrarse en secuencias humanas. La composición de aminoácidos de las secuencias contienen preferentemente pocos residuos hidrófobos. Son de particular interés secuencias de URP que se diseñan basándose en secuencias humanas y que contienen una gran fracción de residuos de glicina.

**[0120]** Utilizando este esquema o similar puede diseñarse una clase de URP que comprenda secuencias de repetición con baja inmunogenicidad para el huésped de interés. El huésped de interés puede ser cualquier animal, que incluye vertebrados e invertebrados. Huéspedes preferidos son mamíferos tales como primates (por ejemplo chimpancés y seres humanos), cetáceos (por ejemplo, ballenas y delfines), quirópteros (por ejemplo, murciélagos), perisodáctilos (por ejemplo, caballos y rinocerontes), roedores (por ejemplo, ratas), y ciertos tipos de insectívoros tales como musarañas, topos y erizos. Si se selecciona el ser humano como huésped, los URP normalmente contienen múltiples copias de las secuencias o unidades de repetición, en los que la mayoría de los segmentos que comprenden aproximadamente 6 a aproximadamente 15 aminoácidos contiguos están presentes en una o más proteínas humanas nativas. También pueden diseñarse URP en los que la mayoría de los segmentos que comprenden entre aproximadamente 9 y aproximadamente 15 aminoácidos contiguos se encuentran en una o más proteínas humanas nativas. Como se usa en el presente documento, la mayoría de los segmentos se refieren a más de aproximadamente el 50%, preferentemente el 60%, preferentemente el 70%, preferentemente el 80%, preferentemente el 90%, preferentemente el 100%. Si se desea, cada uno de los posibles segmentos entre aproximadamente 6 y 15 aminoácidos, preferentemente entre aproximadamente 9 y 15 aminoácidos dentro de las unidades de repetición, están presentes en una o más proteínas humanas nativas. Los URP pueden comprender

múltiples unidades o secuencias de repetición, por ejemplo, teniendo 2, 3, 4, 5, 6, 7, 8, 9, 10 o más unidades de repetición.

**Diseño de URP que están sustancialmente libres de epítomos de linfocitos T humanos:**

5

**[0121]** Las secuencias de URP pueden diseñarse para estar sustancialmente libres de epítomos reconocidos por linfocitos T humanos. Por ejemplo, puede sintetizarse una serie de secuencias semi-aleatorias con composiciones de aminoácidos que favorecen conformaciones no estructuradas desnaturalizadas y evaluar estas secuencias para la presencia de epítomos de linfocitos T humanos y si son secuencias humanas o no. Se han descrito ensayos para epítomos de linfocitos T humanos (Stickler, M. y col. (2003) J Immunol Methods, 281: 95-108). Son de particular interés secuencias de péptidos que pueden oligomerizarse sin generar epítomos de linfocitos T o secuencias no humanas. Esto puede lograrse probando repeticiones directas de estas secuencias para la presencia de epítomos de linfocitos T y para la aparición de subsecuencias 6 a 15-meras y en particular 9-meras que no son humanas. Una alternativa es evaluar múltiples secuencias de péptidos que puedan ensamblarse en unidades de repetición como se describe en la sección previa para el ensamblaje de secuencias humanas. Otra alternativa es diseñar secuencias de URP que produzcan bajas puntuaciones usando algoritmos de predicción de epítomos como TEPITOPE (Sturniolo, T. y col. (1999) Nat Biotechnol, 17: 555-61). Otro enfoque para evitar epítomos de linfocitos T es evitar aminoácidos que puedan servir de residuos de anclaje durante la expresión de péptidos sobre MHC, tal como M, I, L, V, F. Aminoácidos hidrófobos y aminoácidos positivamente cargados pueden servir frecuentemente de tales residuos de anclaje y el minimizar su frecuencia en una secuencia de URP reduce la probabilidad de generación de epítomos de linfocitos T y así provocar una reacción inmunitaria. Los URP seleccionados contienen generalmente subsecuencias que se encuentran en al menos una proteína humana, y tienen un menor contenido de aminoácidos hidrófobos.

10

15

20

25

30

**[0122]** Las secuencias de URP pueden diseñarse para optimizar la producción de proteínas. Esto puede lograrse evitando o minimizando la repetitividad del ADN codificante. Secuencias de URP tales como poli-glicina pueden tener propiedades farmacéuticas muy deseables, pero su fabricación puede ser difícil debido al alto contenido de GC de secuencias de ADN que codifican GRS y debido a la presencia de secuencias de ADN de repetición que pueden conducir a recombinación.

35

40

**[0123]** Como se observa anteriormente, las secuencias de URP pueden diseñarse para ser altamente repetitivas al nivel de aminoácidos. Como resultado, las secuencias de URP tienen contenido de información muy bajo y pueden reducir el riesgo de provocar una reacción inmunitaria.

**[0124]** Ejemplos no limitantes de URP que contienen aminoácidos de repetición son: poli-glicina, poli-ácido glutámico, poli-ácido aspártico, poli-serina, poli-treonina,  $(GX)_n$  en la que G es glicina y X es serina, ácido aspártico, ácido glutámico, treonina o prolina y n es al menos 20,  $(GGX)_n$  en la que X es serina, ácido aspártico, ácido glutámico, treonina o prolina y n es al menos 13,  $(GGGX)_n$  en la que X es serina, ácido aspártico, ácido glutámico, treonina o prolina y n es al menos 10,  $(GGGGX)_n$  en la que X es serina, ácido aspártico, ácido glutámico, treonina o prolina y n es al menos 8,  $(G_xC)_n$  en la que X es serina, ácido aspártico, ácido glutámico, treonina o prolina, n es al menos 15, y z es entre 1 y 20.

45

50

55

**[0125]** El número de estas repeticiones puede ser cualquier número entre 10 y 100. Los productos de la invención pueden contener secuencias de URP que son secuencias semi-aleatorias. Ejemplos son secuencias semi-aleatorias que contienen al menos el 30, 40, 50, 60 o el 70% de glicina en las que las glicinas están bien dispersas y en las que la concentración total de triptófano, fenilalanina, tirosina, valina, leucina e isoleucina es entonces menos del 70, 60, 50, 40, 30, 20 o el 10% cuando se combinan. Una secuencia de URP semi-aleatoria preferida contiene al menos el 40% de glicina y la concentración total de triptófano, fenilalanina, tirosina, valina, leucina e isoleucina es entonces menos del 10%. Una secuencia de URP aleatoria más preferida contiene al menos el 50% de glicina y la concentración total de triptófano, fenilalanina, tirosina, valina, leucina e isoleucina es entonces menos del 5%. Pueden diseñarse secuencias de URP combinando las secuencias de dos o más secuencias de URP más cortas o fragmentos de secuencias de URP. Una combinación tal permite modular mejor las propiedades farmacéuticas del producto que contiene las secuencias de URP y permite que se reduzca la repetitividad de las secuencias de ADN que codifican las secuencias de URP, que pueden mejorar la expresión y reducir la recombinación de secuencias codificantes de URP.

**[0126]** Pueden diseñarse secuencias de URP y seleccionarse para que posean varias de las siguientes propiedades deseadas: a) alta estabilidad genética de las secuencias codificantes en el huésped de producción, b) alto nivel de expresión, c) baja inmunogenicidad (predicha/calculada), d) alta estabilidad en presencia de proteasas

del suero y/u otras proteasas de tejido, e) gran radio hidrodinámico bajo condiciones fisiológicas. Un enfoque a modo de ejemplo para obtener secuencias de URP que cumplen múltiples criterios es construir una biblioteca de secuencias candidatas e identificar a partir de la biblioteca las subsecuencias adecuadas. Las bibliotecas pueden comprender secuencias aleatorias y/o semi-aleatorias. Son de particular utilidad bibliotecas de codones, que es una biblioteca de moléculas de ADN que contiene múltiples codones para el residuo de aminoácido idéntico. Puede aplicarse aleatorización de codones a posiciones de aminoácidos seleccionadas de un cierto tipo o a la mayoría o a todas las posiciones. Bibliotecas de codones verdaderas solo codifican una secuencia de un solo aminoácido, pero pueden combinarse fácilmente con bibliotecas de aminoácidos, que es una población de moléculas de ADN que codifica una mezcla de aminoácidos (relacionados o sin relacionar) en la posición de residuo. Las bibliotecas de codones permiten la identificación de genes que tienen repetitividad relativamente baja al nivel de ADN, pero que codifican secuencias de aminoácidos altamente repetitivas. Esto es útil debido a que secuencias de ADN repetitivas tienden a recombinarse, conduciendo a inestabilidad. También pueden construirse bibliotecas de codones que codifican diversidad de aminoácidos limitada. Tales bibliotecas permiten la introducción de un número limitado de aminoácidos en algunas posiciones de la secuencia, mientras que otras posiciones permiten variación de codones, pero todos los codones codifican el mismo aminoácido. Pueden sintetizarse oligonucleótidos parcialmente aleatorios incorporando mezclas de nucleótidos en la misma posición durante la síntesis de oligonucleótidos. Tales oligonucleótidos parcialmente aleatorios pueden fusionarse por PCR de solapamiento o enfoques basados en ligación. En particular, pueden multimerizarse oligonucleótidos semi-aleatorios que codifican secuencias ricas en glicina. Estos oligonucleótidos pueden diferenciarse en longitud y secuencias y uso de codones. Como resultado, se obtiene una biblioteca de secuencias de URP candidatas. Otro procedimiento para generar bibliotecas es sintetizar una secuencia de partida y posteriormente someter dichas secuencias a aleatorización parcial. Esto puede hacerse por cultivo del gen que codifica las secuencias de URP en una cepa de mutador o por amplificación del gen codificante bajo condiciones mutagénicas (Leung, D. y col. (1989) *Technique*, 1: 11-15). Secuencias de URP con propiedades deseables pueden identificarse a partir de bibliotecas usando una variedad de procedimientos. Las secuencias que tienen un alto grado de estabilidad genética pueden enriquecerse cultivando la biblioteca en un huésped de producción. Secuencias que son inestables acumularán mutaciones, que pueden identificarse por secuenciación de ADN. Variantes de secuencias de URP que pueden expresarse a alto nivel pueden identificarse por cribado o selección usando múltiples protocolos conocidos para cualquier experto en la materia. Por ejemplo, pueden cultivarse múltiples cepas aisladas de una biblioteca y comparar niveles de expresión. Los niveles de expresión pueden medirse por análisis de gel, cromatografía analítica o diversos procedimientos basados en ELISA. La determinación de los niveles de expresión de variantes de secuencia individuales puede facilitarse fusionando la biblioteca de secuencias de URP candidatas con marcas de secuencia como marca myc, marca His, marca HA. Otro enfoque es fusionar la biblioteca con una enzima u otra proteína indicadora como proteína verde fluorescente. Es de particular interés la fusión de la biblioteca con un marcador de selección como beta-lactamasa o kanamicina-acil transferasa. Puede usarse selección con antibiótico para enriquecer en variantes con alto nivel de expresión y buena estabilidad genética. Pueden identificarse variantes con buena resistencia a proteasas cribando secuencias intactas después de la incubación con proteasas. Una forma eficaz para identificar secuencias de URP resistentes a proteasas es la expresión de fago bacteriano o procedimientos de expresión relacionados. Se han descrito múltiples sistemas en los que las secuencias que experimentan rápida proteólisis pueden enriquecerse por expresión de fago. Estos procedimientos pueden adoptarse fácilmente para enriquecer en secuencias resistentes a proteasas. Por ejemplo, puede clonarse una biblioteca de secuencias de URP candidatas entre una marca de afinidad y la proteína pIII del fago M13. La biblioteca puede entonces exponerse a proteasas o muestras biológicas que contienen proteasas como sangre o preparaciones lisosómicas. Las secuencias resistentes a proteasas que contienen fago pueden capturarse después del tratamiento con proteasas por unión a la marca de afinidad. Secuencias que resisten a la degradación por preparaciones lisosómicas son de particular interés debido a que la degradación lisosómica es una etapa clave durante la presentación de antígeno en células dendríticas y otras células presentadoras de antígeno. La expresión de fago puede utilizarse para identificar secuencias de URP candidatas que no se unen a un suero inmune particular con el fin de identificar secuencias de URP con baja inmunogenicidad. Pueden inmunizarse animales con una secuencia de URP candidata o con una biblioteca de secuencias de URP para producir anticuerpos contra las secuencias de URP en la biblioteca. El suero resultante puede entonces usarse para la inmunopurificación de fagos para eliminar o identificar secuencias que son reconocidas por anticuerpos en el suero inmune resultante. Otros procedimientos como expresión bacteriana, expresión de levadura, expresión ribosómica pueden utilizarse para identificar variantes de secuencias de URP con propiedades deseables. Otro enfoque es la identificación de secuencias de URP de interés por espectrometría de masas. Por ejemplo, puede incubarse una biblioteca de secuencias de URP candidatas con una proteasa o muestra biológica de interés e identificar secuencias que resistan a la degradación por espectrometría de masas. En un enfoque similar, pueden identificarse secuencias de URP que facilitan la captación oral. Puede alimentarse una mezcla de secuencias de URP candidatas a animales o seres humanos e identificar variantes con la mayor eficiencia de transferencia o captación a través de algún tejido barrera (es decir, dérmico, etc.) por espectrometría de masas. De un modo similar, pueden identificarse

secuencias de URP que favorecen otros mecanismos de captación como administración pulmonar, intranasal, rectal, transdérmica. También pueden identificarse secuencias de URP que favorezcan la captación celular o secuencias de URP que resistan a la captación celular.

5 **[0127]** Las secuencias de URP pueden diseñarse combinando secuencias de URP o fragmentos de secuencias de URP que se diseñaron por cualquiera de los procedimientos descritos anteriormente. Además, pueden aplicarse enfoques semi-aleatorios para optimizar secuencias que se diseñaron basándose en las reglas descritas anteriormente. Es de particular interés la optimización de codones con el objetivo de mejorar la expresión de las proteínas potenciadas y mejorar la estabilidad genética del gen codificante en los huéspedes de producción.

10 La optimización de codones es de particular importancia para secuencias de URP que son ricas en glicina o que tienen secuencias de aminoácidos muy repetitivas. La optimización de codones puede realizarse usando programas informáticos (Gustafsson, C. y col. (2004) Trends Biotechnol, 22: 346-53), algunos de los cuales minimizan la detención ribosómica (Coda Genomics Inc.). Cuando se diseñan secuencias de URP pueden considerarse varias propiedades. Puede minimizarse la repetitividad en las secuencias de ADN codificantes. Además, puede evitarse o

15 minimizarse el uso de codones que son raramente usados por el huésped de producción (es decir, los codones de arginina AGG y AGA y un codón de leucina en *E. coli*). Las secuencias de ADN que tienen un alto nivel de glicina tienden a tener un alto contenido de GC que puede conducir a inestabilidad o bajos niveles de expresión. Así, cuando sea posible, se prefiere elegir codones de forma que el contenido de GC de la secuencia codificante de URP sea adecuado para el organismo de producción que se usará para fabricar el URP.

20 **[0128]** Pueden prepararse genes que codifican URP en una o más etapas, tanto completamente sintéticamente como por síntesis combinada con procedimientos enzimáticos tales como clonación mediada por enzimas de restricción, PCR y extensión por solapamiento. Los módulos de URP pueden construirse de forma que el gen que codifica el módulo de URP tenga baja repetitividad, mientras que la secuencia de aminoácidos codificada

25 tenga un alto grado de repetitividad. El enfoque se ilustra en la Figura 11. Como primera etapa se construye una biblioteca de secuencias de URP relativamente cortas. Ésta puede ser una biblioteca de codones pura de forma que cada miembro de la biblioteca tenga la misma secuencia de aminoácidos, pero son posibles muchas secuencias codificantes diferentes. Para facilitar la identificación de miembros de bibliotecas de buena expresión, la biblioteca puede construirse como fusión con una proteína indicadora. Ejemplos de genes indicadores adecuados son proteína

30 verde fluorescente, luciferasa, fosfatasa alcalina, beta-galactosidasa. Mediante cribado pueden identificarse secuencias de URP cortas que pueden expresarse en alta concentración en el organismo huésped de elección. Posteriormente, puede generarse una biblioteca de dímeros de URP aleatorios y repetir el cribado para alto nivel de expresión. La dimerización puede realizarse por ligación, extensión por solapamiento o técnicas de clonación similares. Este procedimiento de dimerización y posterior cribado puede repetirse múltiples veces hasta que la

35 secuencia de URP resultante haya alcanzado la longitud deseada. Opcionalmente, pueden secuenciarse clones en la biblioteca para eliminar cepas aisladas que contienen secuencias no deseables. La biblioteca inicial de secuencias de URP cortas puede permitir alguna variación en la secuencia de aminoácidos. Por ejemplo, pueden aleatorizarse algunos codones de forma que puedan producirse varios aminoácidos hidrófilos en dicha posición. Durante el procedimiento de multimerización iterativa pueden cribarse miembros de la biblioteca para otras características como

40 solubilidad o resistencia a proteasas, además de un cribado para expresión de alto nivel. En lugar de dimerizar secuencias de URP, también pueden generarse multímeros más largos. Esto permite aumentar más rápidamente la longitud de módulos de URP.

**[0129]** Muchas secuencias de URP contienen aminoácidos particulares a alta fracción. Tales secuencias

45 pueden ser difíciles de producir por técnicas recombinantes, ya que sus genes codificantes pueden contener secuencias repetitivas que están sometidas a recombinación. Además, genes que contienen codones particulares a frecuencias muy altas pueden limitar la expresión, ya que los ARNt cargados respectivos en el huésped de producción se vuelven limitantes. Un ejemplo es la producción recombinante de GRS. Los residuos de glicina están codificados por 4 tripletes, GGG, GGC, GGA y GGT. Como resultado, los genes que codifican GRS tienden a tener

50 alto contenido de GC y tienden a ser particularmente repetitivos. Un reto adicional puede resultar del sesgo de codones del huésped de producción. En el caso de *E. coli*, dos codones de glicina, GGA y GGG, se usan raramente en proteínas altamente expresadas. Así, la optimización de codones del gen que codifica secuencias de URP puede ser muy deseable. Puede optimizarse el uso de codones empleando programas informáticos que consideran sesgo de codones del huésped de producción (Gustafsson, C. y col. (2004) Trends Biotechnol, 22: 346-53). Como

55 alternativa, pueden construirse bibliotecas de codones en las que todos los miembros de la biblioteca codifican la misma secuencia de aminoácidos, pero en la que se varía el uso de codones. Tales bibliotecas pueden cribarse para miembros altamente expresantes y genéticamente estables que son particularmente adecuados para la producción a gran escala de productos que contienen URP.

**Proteínas recombinantes no estructuradas multivalentes (MURP):**

**[0130]** Como se observa anteriormente, los URP de la invención son particularmente útiles como módulos para diseñar proteínas de valor terapéutico. Por consiguiente, la presente divulgación proporciona proteínas que comprenden uno o más URP de la invención. Tales proteínas se denominan en el presente documento proteínas recombinantes no estructuradas multivalentes (MURP).

**[0131]** Para construir MURP, una o más secuencias de URP pueden fusionarse con el extremo N o extremo C de una proteína o insertarse en el centro de la proteína, por ejemplo, en bucles de una proteína o entre módulos de la proteína de interés, dando las propiedades mejoradas de proteínas modificadas resultantes con respecto a las proteínas sin modificar. La longitud combinada de secuencias de URP que están unidas a una proteína puede ser de 40, 50, 60, 70, 80, 90, 100, 150, 200 o más aminoácidos.

**[0132]** Las MURP presentan una o más propiedades mejoradas como se detalla a continuación.

**Semivida mejorada:**

**[0133]** Añadiendo secuencias de URP a una proteína farmacéuticamente activa pueden mejorarse muchas propiedades de esa proteína. En particular, añadiendo una secuencia de URP larga puede aumentarse significativamente la semivida en suero de la proteína. Tales URP contienen habitualmente secuencias de aminoácidos de al menos aproximadamente 40, 50, 60, 70, 80, 90, 100, 150, 200 o más aminoácidos.

**[0134]** Los URP pueden fragmentarse de forma que la proteína resultante contenga múltiples URP, o múltiples fragmentos de URP. En un aspecto, los URP fusionados pueden aumentar el radio hidrodinámico de una proteína y así reducir su eliminación de la sangre por el riñón. El aumento en el radio hidrodinámico de la proteína de fusión resultante con respecto a la proteína sin modificar puede detectarse por ultracentrifugación, cromatografía de exclusión por tamaño o dispersión de la luz.

**Selectividad por tejido mejorada:**

**[0135]** El aumento del radio hidrodinámico también puede conducir a penetración reducida en tejidos, que puede explotarse para minimizar efectos secundarios de una proteína farmacéuticamente activa. Está muy documentado que los polímeros hidrófilos tienen una tendencia a acumularse en tejido tumoral que se produce por el potenciado efecto de la permeabilidad y retención (EPR). La causa subyacente del efecto EPR es la naturaleza porosa de la vasculatura tumoral (McDonald, D. M. y col. (2002) Cancer Res, 62: 5381-5) y la falta de drenaje linfático en tejidos tumorales. Por tanto, la selectividad de proteínas farmacéuticamente activas por tejidos tumorales puede potenciarse añadiendo polímeros hidrófilos. Como tal, el índice terapéutico de una proteína farmacéuticamente activa dada puede aumentarse incorporando los URP de la invención.

**Protección de la degradación e inmunogenicidad reducida:**

**[0136]** La adición de secuencias de URP puede mejorar significativamente la resistencia a proteasas de una proteína. Las propias secuencias de URP pueden diseñarse para ser resistentes a proteasas y uniéndolas a una proteína puede protegerse esa proteína del acceso de enzimas degradantes. Las secuencias de URP pueden añadirse a proteínas farmacéuticamente activas con el objetivo de reducir interacciones no deseables de la proteína con otros receptores o superficies. Para lograr esto puede ser beneficioso añadir las secuencias de URP a la proteína farmacéuticamente activa en proximidad al sitio de la proteína que hace tales contactos no deseados. En particular, pueden añadirse secuencias de URP a proteínas farmacéuticamente activas con el objetivo de reducir sus interacciones con cualquier componente del sistema inmunitario para prevenir una respuesta inmunitaria contra el producto de la invención. La adición de una secuencia de URP a una proteína farmacéuticamente activa puede reducir la interacción con anticuerpos preexistentes o receptores de células B. Además, la adición de secuencias de URP puede reducir la captación y procesamiento del producto del procedimiento de la invención por células presentadoras de antígeno. La adición de una o más secuencias de URP a una proteína es una forma preferida de reducir su inmunogenicidad, ya que suprimirá una respuesta inmunitaria en muchas especies, permitiendo que se prediga la inmunogenicidad esperada de un producto en pacientes basándose en datos animales. Tal prueba independiente de inmunogenicidad de la especie no es posible para enfoques que se basan en la identificación y eliminación de epítopos de linfocitos T humanos o comparación de secuencias con secuencias humanas.

**Interrupción de epítopos de linfocitos T:**

**[0137]** Las secuencias de URP pueden introducirse en proteínas con el fin de interrumpir epítomos de linfocitos T. Esto es particularmente útil para proteínas que combinan múltiples módulos funcionales separados. La formación de epítomos de linfocitos T requiere que los fragmentos de péptido de un antígeno de proteína se unan a MHC. Las moléculas de MHC interactúan con un segmento de aminoácidos corto, normalmente 9 residuos contiguos de los péptidos presentados. La fusión directa de diferentes módulos de unión en una molécula de proteína puede conducir a epítomos de linfocitos T que atraviesan dos dominios vecinos. Separando los módulos funcionales por módulos de URP se previene la generación de tales epítomos de linfocitos T que atraviesan módulos como se ilustra en la Figura 7. La inserción de secuencias de URP entre módulos funcionales también puede interferir con el procesamiento proteolítico en células presentadoras de antígeno, que conducirá a una reducción adicional de la inmunogenicidad. Otro enfoque para reducir el riesgo de inmunogenicidad es alterar epítomos de linfocitos T dentro de módulos funcionales de un producto. En el caso de microproteínas, un enfoque es que algunos de los bucles de intercisteína (aquellos que no participan en la unión a diana) sean ricos en glicina. En microproteínas, cuya estructura es debida a un número pequeño de cisteínas, podría de hecho sustituirse la mayoría o todos los residuos que no participan en la unión a diana con glicina, serina, glutamato, treonina, reduciéndose así las posibilidades de inmunogenicidad, a la vez que no se afecta la afinidad por la diana. Por ejemplo, esto puede llevarse a cabo realizando un 'barrido de glicinas' de todos los residuos, en el que cada residuo se sustituye con una glicina, luego seleccionando los clones que retienen la unión a diana usando expresión o cribado de fago, y luego combinando todas las sustituciones de glicina que son permitidas. En general, los módulos funcionales tienen una probabilidad mucho mayor de contener epítomos de linfocitos T que los módulos de URP. Puede reducirse la frecuencia de epítomos de linfocitos T en módulos funcionales reemplazando todos o muchos residuos de aminoácidos no críticos con residuos hidrófilos pequeños como gly, ser, ala, glu, asp, asn, gln, thr. Las posiciones en un módulo funcional que permiten la sustitución pueden identificarse usando una variedad de enfoques aleatorios o de ingeniería de proteínas basada en estructura.

25 ***Solubilidad mejorada:***

**[0138]** Los módulos funcionales de una proteína pueden tener solubilidad limitada. En particular, módulos de unión tienden a llevar residuos hidrófobos sobre su superficie, que pueden limitar su solubilidad y pueden conducir a agregación. Separando o flanqueando tales módulos funcionales con módulos de URP puede mejorarse la solubilidad global del producto resultante. Esto es cierto en particular para módulos de URP que llevan un porcentaje significativo de residuos hidrófilos o cargados. Separando módulos funcionales con módulos de URP solubles pueden reducirse interacciones intramoleculares entre estos módulos funcionales.

35 ***Perfil de pH mejorado y homogeneidad de la carga del producto:***

**[0139]** Pueden diseñarse secuencias de URP para llevar un exceso de cargas negativas o positivas. Como resultado, confieren un campo electrostático a cualquier componente de fusión que pueda utilizarse para desplazar el perfil de pH de una enzima o una interacción de unión. Además, el campo electrostático de una secuencia de URP cargada puede aumentar la homogeneidad de valores de pKa de cargas superficiales de un producto de proteína, que conduce a perfiles de pH bruscos de interacciones de ligandos y a separaciones refinadas por isoelectroenfoque o cromatografía.

45 ***Propiedades de purificación mejoradas debido a pKa de producto más refinado:***

**[0140]** Cada aminoácido en disolución tiene por sí mismo un único pKa fijo, que es el pH al que sus grupos funcionales están protonados a la mitad. En una proteína típica se tienen muchos tipos de residuos y debido a la proximidad y los efectos de respiración de la proteína, también se cambian entre sí el pKa eficaz de formas variables. Debido a esto, a un amplio intervalo de condiciones de pH, proteínas típicas pueden adoptar cientos de especies diferentemente ionizadas, cada una con un peso molecular diferente y carga neta, debido a grandes números de combinaciones de residuos de aminoácidos cargados y neutros. Esto se denomina un espectro de ionización ancho y dificulta más el análisis (es decir, espectrometría de masas) y purificación de tales proteínas.

**[0141]** El PEG no tiene carga y no afecta el espectro de ionización de la proteína con la que está unido, dejándola con un amplio espectro de ionización. Sin embargo, un URP con un alto contenido de Gly y Glu solo existe en principio en dos estados: neutro (-COOH) si el pH está por debajo del pKa del glutamato y negativamente cargado (-COO<sup>-</sup>) si el pH está por encima del pKa del glutamato. Los módulos de URP pueden formar un único tipo de molécula homogéneamente ionizada y pueden dar una única masa en espectrometría de masas.

[0142] Si se desea, las MURP pueden expresarse como una fusión con un URP que tiene un único tipo de carga (Glu) distribuida a separación constante por el módulo de URP. Puede elegirse incorporar 25-50 residuos de Glu por 20 kD de URP y todos estos 25-50 residuos tendrían pKa muy similar.

5 [0143] Además, la adición de 25-50 cargas negativas a una proteína pequeña como IFN, hGH o GCSF (con solo 20 residuos cargados) aumentaría la homogeneidad de la carga del producto y aumentaría su punto isoeléctrico, que sería muy próximo al pKa del glutamato libre.

[0144] El aumento en la homogeneidad de la carga de la población de proteínas tiene propiedades de procesamiento favorables tales como en intercambio iónico, isoelectroenfoque, espectrometría de masas, etc., en comparación con la PEGilación tradicional.

**Formulación mejorada y/o administración:**

15 [0145] La adición de secuencias de URP a proteínas farmacéuticamente activas puede simplificar significativamente la formulación y o la administración de los productos resultantes. Las secuencias de URP pueden diseñarse para ser muy hidrófilas y como resultado mejoran la solubilidad de (por ejemplo) proteínas humanas, que frecuentemente contienen parches hidrófobos que se usan para unirse a otras proteínas humanas. La formulación de tales proteínas humanas, como anticuerpos, puede ser bastante exigente y frecuentemente limita sus opciones de concentración y administración. Los URP pueden reducir la precipitación y agregación de productos y permite que se usen formulaciones más simples que contienen menos componentes, que son normalmente necesarios para estabilizar un producto en disolución. La solubilidad mejorada de productos que contienen secuencias de URP permite formular estos productos a mayor concentración y como resultado puede reducirse el volumen de inyección para productos inyectables, que pueden permitir la inyección en casa, que está limitado a un volumen inyectado muy bajo. La adición de una secuencia de URP también puede simplificar el almacenamiento de los productos formulados resultantes. Las secuencias de URP pueden añadirse a proteínas farmacéuticamente activas para facilitar su captación oral, pulmonar, rectal o intranasal. Las secuencias de URP pueden facilitar diversos modos de administración debido a que permiten mayores concentraciones de producto y estabilidad de producto mejorada. Pueden lograrse mejoras adicionales diseñando secuencias de URP que facilitan la penetración en la membrana.

30

**Producción mejorada:**

[0146] La adición de secuencias de URP puede tener beneficios significativos para la producción del producto resultante. Muchos productos recombinantes, especialmente proteínas humanas nativas, tienen una tendencia a formar agregados durante la producción que pueden ser difíciles o imposibles de disolver e incluso cuando se eliminan del producto final pueden volver a producirse. Éstos son normalmente debidos a parches hidrófobos por los cuales estas proteínas (humanas nativas) se pusieron en contacto con otras proteínas (humanas nativas) y la mutación de estos residuos se considera arriesgado debido a la inmunogenicidad. Sin embargo, los URP pueden aumentar la hidrofilia de tales proteínas y permitir su formulación sin mutar la secuencia de la proteína humana. Las secuencias de URP pueden facilitar el plegamiento de una proteína para alcanzar su estado nativo. Muchas proteínas farmacéuticamente activas se producen mediante procedimientos recombinantes en un estado agregado no nativo. Estos productos necesitan desnaturalizarse y posteriormente se incuban en condiciones que permitan que las proteínas se plieguen en su estado activo nativo. Una reacción secundaria frecuente durante la renaturalización es la formación de agregados. La fusión de secuencias de URP con una proteína reduce significativamente su tendencia a formar agregados y así facilita el plegamiento del componente farmacéuticamente activo del producto. Productos que contienen URP son mucho más fáciles de preparar con respecto a proteínas modificadas con polímeros. La modificación con polímeros químicos requiere etapas de modificación y de purificación adicionales después de purificarse la proteína activa. A diferencia, pueden fabricarse secuencias de URP usando procedimientos de ADN recombinante junto con la proteína farmacéuticamente activa. Los productos de la invención también son significativamente más fáciles de caracterizar en comparación con productos modificados con polímeros. Debido al procedimiento de producción recombinante, pueden obtenerse productos más homogéneos con características moleculares definidas. Las secuencias de URP también pueden facilitar la purificación de un producto. Por ejemplo, secuencias de URP pueden incluir subsecuencias que pueden ser capturadas por cromatografía de afinidad. Un ejemplo son secuencias ricas en histidina, que pueden ser capturadas sobre resinas con metales inmovilizados como níquel. Las secuencias de URP también pueden diseñarse para tener un exceso de aminoácidos negativamente o positivamente cargados. Como resultado, pueden impactar significativamente en la carga neta de un producto, que puede facilitar la purificación de productos por cromatografía de intercambio iónica o electroforesis preparativa.

55

**[0147]** Las MURP de la invención pueden contener una variedad de módulos que incluyen, pero no se limitan a, módulos de unión, módulos efectores, módulos de multimerización, módulos del extremo C y módulos del extremo N. La Figura 1 representa una MURP a modo de ejemplo que tiene múltiples módulos. Sin embargo, las MURP también pueden tener arquitecturas relativamente simples que se ilustran en la Fig. 2. Las MURP también pueden contener sitios de fragmentación. Éstos pueden ser secuencias sensibles a proteasas o secuencias químicamente sensibles que pueden escindirse preferencialmente cuando las MURP llegan a su sitio diana.

**Módulo de unión (BM):**

10 **[0148]** Las MURP pueden comprender uno o más módulos de unión. Módulo de unión (BM) se refiere a una secuencia de péptidos o de proteínas que puede unirse específicamente a una o varias dianas, que pueden ser una o más dianas terapéuticas o dianas accesorias, tales como para la elección de diana de células, tejidos u órganos. Los BM pueden ser péptidos lineales o cíclicos, péptidos restringidos a cisteína, microproteínas, proteínas de andamiaje (por ejemplo, fibronectina, anquirinas, cristalina, estreptavidina, fragmentos de anticuerpos, anticuerpos de dominio), hormonas peptídicas, factores de crecimiento, citocinas, o cualquier tipo de dominio de proteína, humano o no humano, natural o no natural, y pueden basarse en un andamiaje natural o no basarse en un andamiaje natural, o basarse en combinaciones o pueden ser fragmentos de cualquiera de lo anterior. Opcionalmente, estos BM pueden manipularse añadiendo, quitando o sustituyendo uno o múltiples aminoácidos con el fin de potenciar sus propiedades de unión, su estabilidad, u otras propiedades. Los módulos de unión pueden obtenerse de proteínas naturales, por diseño o por expresión de paquetes genéticos, que incluyen expresión de fago, expresión celular, expresión ribosómica u otros procedimientos de expresión. Los módulos de unión pueden unirse a la misma copia de la misma diana, que produce avidéz, o pueden unirse a diferentes copias de la misma diana (que puede producir avidéz si estas copias están conectadas o ligadas de alguna forma, tal como por una membrana celular), o pueden unirse a dos dianas sin relacionar (que da avidéz si estas dianas están ligadas de alguna forma, tal como por una membrana). Los módulos de unión pueden identificarse cribando o analizando de otro modo bibliotecas aleatorias de péptidos o proteínas.

30 **[0149]** Módulos de unión particularmente deseables son aquellos que tras la incorporación en una MURP, la MURP da una puntuación de epítipo T deseable. La puntuación de epítipo T de una proteína es el log de Kd (constante de disociación, afinidad, tasa de disociación) de la unión de esa proteína a múltiples de los alelos de MHC humanos más comunes, como se desvela en Sturniolo T. y col. (1999) Nature Biotechnology 17:555). La puntuación oscila sobre al menos 15 logaritmos, de aproximadamente 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0, -1, -2, -3, -4, -5 (10e<sup>10</sup> Kd) a aproximadamente -5. MURP preferidas dan una puntuación inferior a aproximadamente -3,5 [KKW: ¿En escala absoluta?]

35 **[0150]** También son de particular interés módulos de unión que comprenden enlaces disulfuro formados por apareamiento de dos residuos de cisteína. En ciertas realizaciones, los módulos de unión comprenden polipéptidos que tienen alto contenido de cisteína o alta densidad de disulfuro (HDD). Los módulos de unión de la familia HDD normalmente tienen 5-50% (5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 45 ó 50%) de residuos de cisteína y cada dominio contiene normalmente al menos dos disulfuros y opcionalmente un co-factor tal como calcio u otro ión.

45 **[0151]** La presencia del andamiaje de HDD permite que estos módulos sean pequeños, pero que todavía adopten una estructura relativamente rígida. La rigidez es importante para obtener altas afinidades de unión, resistencia a proteasas y calor, que incluyen las proteasas que participan en el procesamiento de antígenos, y así contribuye a la baja inmunogenicidad o no inmunogenicidad de estos módulos. La región estructural de disulfuro pliega los módulos sin la necesidad de un gran número de interacciones de cadenas laterales hidrófobas en el interior de la mayoría de los módulos. El pequeño tamaño también es ventajoso para la rápida penetración en el tejido y para administración alternativa tal como oral, nasal, intestinal, pulmonar, barrera hematoencefálica, etc. Además, el pequeño tamaño también ayuda a reducir la inmunogenicidad. Puede obtenerse una mayor densidad de disulfuro, tanto aumentando el número de disulfuros como usando dominios con el mismo número de disulfuros, pero menos aminoácidos. También se desea disminuir el número de residuos fijos de no cisteína, de manera que esté disponible un mayor porcentaje de aminoácidos para la unión a diana.

50 **[0152]** Los módulos de unión que contienen cisteína pueden adoptar una amplia gama de patrones de enlaces disulfuro (DBP). Por ejemplo, módulos de dos disulfuros pueden tener tres patrones de enlaces disulfuro (DBP) diferentes, módulos de tres disulfuros pueden tener 15 DBP diferentes y módulos de cuatro disulfuros tienen hasta 105 DBP diferentes. Ejemplos naturales existen para todos los DBP de 2SS, la mayoría de los DBP de 3SS y menos de la mitad de los DBP de 4SS.

- [0153]** En un aspecto, el módulo que contiene cisteína (C) natural o que no se produce naturalmente comprende un polipéptido que tiene dos enlaces disulfuro formados por apareamiento de cisteínas contenidos en el polipéptido según un patrón seleccionado del grupo constituido por C<sup>1-2, 3-4</sup>, C<sup>1-3, 2-4</sup> y C<sup>1-4, 2-3</sup>, en los que los dos números numéricos unidos por un guión indican que dos cisteínas contando desde el extremo N del polipéptido se aparean para formar un enlace disulfuro. En otro aspecto, el módulo que contiene cisteína (C) natural o que no se produce naturalmente comprende un polipéptido que tiene tres enlaces disulfuro formados por apareamiento de cisteínas entre andamiajes según un patrón seleccionado del grupo constituido por C<sup>1-2, 3-4 5-6</sup>, C<sup>1-2, 3-5, 4-6</sup>, C<sup>1-2, 3-6, 4-5</sup>, C<sup>1-3, 2-4, 5-6</sup>, C<sup>1-3, 2-5, 4-6</sup>, C<sup>1-3, 2-6, 4-4</sup>, C<sup>1-4, 2-3, 5-6</sup>, C<sup>1-4, 2-6, 3-5</sup>, C<sup>1-5, 2-2, 4-6</sup>, C<sup>1-5, 2-4, 3-6</sup>, C<sup>1-5, 2-6, 3-4</sup>, C<sup>1-6, 2-3, 4-5</sup> y C<sup>1-6, 2-5, 3-4</sup>, en los que los dos números numéricos unidos por un guión indican que dos cisteínas contando desde el extremo N del polipéptido se aparean para formar un enlace disulfuro. En otro aspecto más, el módulo que contiene cisteína (C) natural o que no se produce naturalmente comprende un polipéptido que tiene al menos cuatro enlaces disulfuro formados por apareamiento de cisteínas contenidas en el polipéptido según un patrón seleccionado del grupo de permutaciones definidas por la fórmula anterior. En otro aspecto más, el módulo que contiene cisteína (C) natural o que no se produce naturalmente comprende un polipéptido que tiene al menos cinco, seis o más enlaces disulfuro formados por apareamiento de cisteínas entre proteínas según un patrón seleccionado del grupo de permutaciones representado por la fórmula anterior. Cualquiera de las proteínas que contienen cisteína o andamiajes desvelados en las solicitudes de patente en tramitación junto con la presente [números de serie 11/528.927 y 11/528.950] son módulos de unión candidatos.
- [0154]** Los módulos de unión también pueden seleccionarse de bibliotecas de péptidos cíclicos restringidos a cisteína con 4, 5, 6, 7, 8, 9, 10, 11 y 12 aminoácidos aleatorizados o parcialmente aleatorizados entre las cisteínas unidas por disulfuro (por ejemplo, en un modo de formación), y en algunos casos aminoácidos aleatorizados adicionales sobre el exterior del par de cisteínas puede construirse usando una variedad de procedimientos. Los miembros de las bibliotecas con especificidad por una diana de interés pueden identificarse usando diversos procedimientos que incluyen expresión de fago, expresión ribosómica, expresión de levadura y otros procedimientos conocidos en la técnica. Tales péptidos cíclicos pueden utilizarse como módulos de unión en MURP. En una realización preferida, pueden manipularse adicionalmente péptidos restringidos a cisteína para aumentar la afinidad de unión, estabilidad proteolítica y/o especificidad usando enfoques de formación que conducen a módulos de unión que contienen más de un enlace disulfuro. Un enfoque de formación particular se ilustra en la Fig. 25. Se basa en la adición de una única cisteína más múltiples residuos aleatorizados sobre el lado del extremo N del péptido cíclico previamente seleccionado, además de sobre el lado del extremo C. Pueden generarse bibliotecas que han sido diseñadas como se ilustra en la Fig. 25. Los módulos de unión con propiedades mejoradas pueden identificarse por expresión de fago o procedimientos similares. Tales bibliotecas de formación pueden contener entre 1 y 12 posiciones aleatorias sobre el extremo N, además de sobre el lado del extremo C de un péptido cíclico. La distancia entre los residuos de cisteína en los flancos aleatorios recientemente añadidos y los residuos de cisteína en el péptido cíclico puede variarse entre 1 y 12 residuos. Tales bibliotecas contendrán cuatro residuos de cisteína por miembro de biblioteca, con dos cisteínas resultantes del péptido cíclico original y dos residuos de cisteína en los flancos recientemente añadidos. Este enfoque favorece un DBP 1-4 2-3 o un cambio en DBP, rompiendo el disulfuro 1-2 preexistente (= 2-3 en la construcción de 4 cisteínas) para formar un DBP 1-2 3-4 ó 1-3 2-4. Tales enfoques de formación pueden realizarse con cebadores específicos para clon de manera que no quede secuencia fija entre las áreas de la biblioteca como se muestra en la Fig. 25, o pueden realizarse con cebadores que usan (y entonces queda) una secuencia fija en ambos lados del péptido previamente seleccionado y, por tanto, estos mismos cebadores pueden usarse para cualquier clon previamente seleccionado como se ilustra en la Fig. 26. El procedimiento ilustrado en la Fig. 26 puede aplicarse a un conjunto de péptidos cíclicos con especificidad por una diana de interés. Se mostró que ambos enfoques de formación funcionaban para maduración por afinidad anti-VEGF por formación. Este enfoque puede repetirse para generar módulos de unión con seis o más residuos de cisteína.
- [0155]** Otra formación de un disulfuro en una secuencia de 2 disulfuros se ilustra en la Fig. 27. Implica la dimerización de un conjunto previamente seleccionado de péptidos de 1 disulfuro consigo mismo de manera que el conjunto de péptidos preseleccionado termine en el extremo N, además de en la posición del extremo C. Este enfoque favorece la formación de secuencias de 2 disulfuros que reconocen dos epítomos separados sobre una diana.
- [0156]** Otro enfoque de formación implica la adición de una secuencia (parcialmente) aleatorizada de 6-15 residuos que contiene dos cisteínas que están separadas 4, 5, 6, 7, 8, 9 ó 10 aminoácidos, con posiciones aleatorizadas opcionalmente adicionales fuera de las cisteínas ligadas. Esta secuencia aleatoria de 2 cisteínas se añade sobre el lado del extremo N del péptido previamente seleccionado, o sobre el lado del extremo C. Este enfoque favorece un DBP 1-2 3-4, aunque pueden formarse otros DBP. Este enfoque puede repetirse para generar módulos de unión con seis o más residuos de cisteína.

**[0157]** Los módulos de unión pueden construirse basándose en andamiajes de estructuras naturales. Tales andamiajes pueden identificarse por búsqueda en bases de datos. Bibliotecas que se basan en andamiajes naturales pueden someterse a inmunopurificación por expresión de fago, seguido de cribado para identificar 5 secuencias que se unen específicamente a una diana de interés.

**[0158]** Está disponible una amplia selección de andamiajes naturales para construir los módulos de unión. La elección de un andamiaje particular dependerá de la diana prevista. Ejemplos no limitantes de andamiajes naturales incluyen proteínas similares a toxinas de serpiente tales como toxinas del veneno de serpiente y dominio extracelular 10 de receptores de la superficie celular humana. Ejemplos no limitantes de toxinas del veneno de serpiente son erabutoxina B, gamma-cardiotoxina, faciculina, toxina muscarínica, erabutoxina A, neurotoxina I, cardiotoxina V4II (toxina III), cardiotoxina V, alfa-cobratoxina, neurotoxina larga 1, FS2 axina, bungarotoxina, bucardina, cardiotoxina CTXI, cardiotoxina CTX IIB, cardiotoxina II, cardiotoxina III, cardiotoxina IV, cobrotoxina 2, alfa-toxinas, neurotoxina II (cobrotoxina B), toxina B (neurotoxina), candotoxina, bucaína. Ejemplos no limitantes de dominio extracelular de 15 receptores de la superficie celular (humana) incluyen CD59, receptor de activina tipo II, ectodominio del receptor la de BMP, dominio extracelular del receptor de tipo II de TGF-beta. Otros andamiajes naturales incluyen, pero no se limitan a, A-dominios, EGF, Ca-EGF, TNF-R, Notch, DSL, Trefoil, PD, TSP1, TSP2, TSP3, anato, integrina beta, tiroglobulina, defensina 1, defensina 2, ciclotida, SHKT, desintegrinas, miotoxinas, gamma-tioneínas, conotoxina, mu-conotoxina, omega-atracotoxinas, delta-atracotoxinas, además de familias adicionales desveladas en las 20 solicitudes de patente en tramitación junto con la presente número de serie 11/528.927 y 11/528.950, que se incorporan en el presente documento en su totalidad.

**[0159]** Se ha descrito una gran variedad de procedimientos que permiten identificar moléculas de unión en una gran biblioteca de variantes. Un procedimiento es la síntesis química. Los miembros de las bibliotecas pueden 25 sintetizarse sobre perlas de forma que cada perla lleve una secuencia de péptidos diferente. Las perlas que llevan ligandos con una especificidad deseable pueden identificarse usando componentes de unión marcados. Otro enfoque es la generación de sub-bibliotecas de péptidos que permiten identificar secuencias de unión específicas en un procedimiento iterativo (Pinilla, C. y col. (1992) *BioTechniques*, 13: 901-905). Más comúnmente usados son los procedimientos de expresión en los que una biblioteca de variantes se expresa sobre la superficie de un fago, 30 proteína o célula. Estos procedimientos tienen en común que el ADN o ARN que codifica cada variante en la biblioteca está físicamente ligado al ligando. Esto permite detectar o recuperar el ligando de interés y luego determinar su secuencia de péptidos secuenciando el ADN o ARN unido. Los procedimientos de expresión permiten que un experto en la técnica enriquezca miembros de las bibliotecas con propiedades de unión deseables de grandes bibliotecas de variantes aleatorias. Frecuentemente, variantes con propiedades de unión deseables pueden 35 identificarse de bibliotecas enriquecidas cribando cepas individuales aisladas de una biblioteca enriquecida para propiedades deseables. Ejemplos de procedimientos de expresión son fusión con el represor lac (Cull, M. y col. (1992) *Proc. Natl. Acad. Sci. USA*, 89: 1865-1869), expresión de superficie celular (Wittrup, K. D. (2001) *Curr Opin Biotechnol*, 12: 395-9). Son de particular interés procedimientos en los que los péptidos o proteínas aleatorios se 40 ligan a partículas de fago. Comúnmente se usan fago M13 (Smith, G. P. y col. (1997) *Chem Rev*, 97: 391-410) y fago T7 (Danner, S. y col. (2001) *Proc Natl Acad Sci U S A*, 98: 12954-9). Hay múltiples procedimientos disponibles para expresar péptidos o proteínas sobre el fago M13. En muchos casos, la secuencia de la biblioteca está fusionada con el extremo N del péptido pIII del fago M 13. El fago lleva normalmente 3-5 copias de esta proteína y así el fago en una biblioteca tal llevará en la mayoría de los casos entre 3-5 copias de un miembro de la biblioteca. Este enfoque se denomina expresión multivalente. Una alternativa es la expresión en fagémido en la que la biblioteca está 45 codificada sobre un fagémido. Las partículas de fago pueden formarse por infección de células que llevan un fagémido con un fago colaborador (Lowman, H. B. y col. (1991) *Biochemistry*, 30: 10832-10838). Este procedimiento normalmente conduce a expresión monovalente. En algunos casos se prefiere la expresión monovalente para obtener ligantes de alta afinidad. En otros casos se prefiere la expresión multivalente (O'Connell, D. y col. (2002) *J Mol Biol*. 321: 49-56).

**[0160]** Se han descrito una variedad de procedimientos para enriquecer secuencias con características deseables por expresión de fago. Puede inmovilizarse una diana de interés por unión a inmunotubos, placas de 50 microtitulación, perlas magnéticas u otras superficies. Posteriormente, una biblioteca de fagos se pone en contacto con la diana inmovilizada, el fago que carece de un ligando de unión se lava, y el fago que lleva un ligando específico para diana puede eluirse mediante una variedad de condiciones. La elución puede realizarse a pH bajo, 55 pH alto, urea u otras condiciones que tienden a romper contactos proteína-proteína. El fago unido también puede eluirse añadiendo células de *E. coli* de forma que la elución del fago pueda infectar directamente el huésped de *E. coli* añadido. Un protocolo interesante es la elución con proteasa que puede degradar el ligando unido a fago o la diana inmovilizada. También pueden utilizarse proteasas como herramientas para enriquecer ligandos unidos a fago

resistentes a proteasas. Por ejemplo, puede incubarse una biblioteca de ligandos unidos a fago con una o más proteasas (humanas o de ratón) antes de la inmunopurificación sobre la diana de interés. Este procedimiento degrada y elimina ligandos lábiles a proteasas de la biblioteca (Kristensen, P. y col. (1998) *Fold Des*, 3: 321-8). También pueden enriquecerse bibliotecas de expresión de fago de ligandos para unirse a muestras biológicas complejas. Ejemplos son la inmunopurificación sobre fracciones de membrana celular inmovilizadas (Tur, M. K. y col. (2003) *Int J Mol Med*, 11: 523-7), o células enteras (Rasmussen, U. B. y col. (2002) *Cancer Gene Ther*, 9: 606-12; Kelly, K. A. y col. (2003) *Neoplasia*, 5: 437-44). En algunos casos tienen que optimizarse las condiciones de inmunopurificación para el enriquecimiento de ligandos específicos para células de bibliotecas de fagos (Watters, J. M. y col. (1997) *Immunotechnology*, 3: 21-9). La inmunopurificación de fagos también puede realizarse en pacientes o animales vivos. Este enfoque es de particular interés para la identificación de ligandos que se unen a dianas vasculares (Arap, W. y col. (2002) *Nat Med*, 8: 121-7).

**[0161]** Están disponibles una variedad de procedimientos de clonación que permiten que un experto en la materia genere bibliotecas de secuencias de ADN que codifican bibliotecas de péptidos. Pueden utilizarse mezclas aleatorias de nucleótidos para sintetizar oligonucleótidos que contienen una o múltiples posiciones aleatorias. Este procedimiento permite controlar el número de posiciones aleatorias, además del grado de aleatorización. Además, pueden obtenerse secuencias aleatorias o semi-aleatorias de ADN por digestión parcial de ADN de muestras biológicas. Pueden usarse oligonucleótidos aleatorios para construir bibliotecas de plásmidos o fagos que se aleatorizan en localizaciones predefinidas. Esto puede hacerse por fusión de PCR como se describe en (de Kruijff, J. y col. (1995) *J Mol Biol*, 248: 97-105). Otros protocolos se basan en ligación de ADN (Felici, F. y col. (1991) *J Mol Biol*, 222: 301-10; Kay, B. K. y col. (1993) *Gene*, 128: 59-65). Otro enfoque comúnmente usado es mutagénesis de Kunkel en la que una hebra mutagenizada de un plásmido o fagémido se sintetiza usando ADN cíclico monocatenario como molde. Véase, Sidhu, S. S. y col. (2000) *Methods Enzymol*, 328: 333-63; Kunkel, T. A. y col. (1987) *Methods Enzymol*, 154: 367-82.

**[0162]** La mutagénesis de Kunkel usa moldes que contienen bases de uracilo aleatoriamente incorporadas que pueden obtenerse de cepas de *E. coli* como CJ236. La hebra molde que contiene uracilo se degrada preferencialmente tras la transformación en *E. coli*, mientras que se retiene la hebra mutagenizada sintetizada *in vitro*. Como resultado, la mayoría de las células transformadas llevan la versión mutagenizada del fagémido o fago. Un enfoque valioso para aumentar la diversidad en una biblioteca es combinar múltiples sub-bibliotecas. Estas sub-bibliotecas pueden generarse por cualquiera de los procedimientos descritos anteriormente y pueden basarse en el mismo andamiaje o en andamiajes diferentes.

**[0163]** Recientemente se ha descrito un procedimiento útil para generar grandes bibliotecas de fagos de péptidos cortos (Scholle, M. D. y col. (2005) *Comb Chem High Throughput Screen*, 8: 545-51). Este procedimiento está relacionado con el enfoque de Kunkel, pero no requiere la generación de ADN mensajero monocatenario que contiene bases de uracilo aleatorias. En su lugar, el procedimiento empieza con un fago molde que lleva una o más mutaciones próximas al área que va a mutagenizarse y dicha mutación hace que el fago sea no infeccioso. El procedimiento usa un oligonucleótido mutagénico que lleva codones aleatorizados en algunas posiciones y que corrige la mutación inactivante del fado en el molde. Como resultado, solo partículas de fago mutagenizadas son infecciosas después de la transformación y muy pocos fagos parentales están contenidos en tales bibliotecas. Este procedimiento puede modificarse adicionalmente de varias formas. Por ejemplo, pueden utilizarse múltiples oligonucleótidos mutagénicos para mutagenizar simultáneamente múltiples regiones discontiguas de un fago. Los presentes inventores han dado un paso más con este enfoque aplicándolo a microproteínas completas de >25, 30, 35, 40, 45, 50, 55 y 60 aminoácidos, en lugar de a péptidos cortos de <10, 15 ó 20 aminoácidos, que plantea un desafío adicional. Este enfoque da ahora bibliotecas de más de  $10e^{10}$  transformantes (hasta  $10e^{11}$ ) con una única transformación, de manera que de 10 transformaciones se espera una única biblioteca con una diversidad de  $10e^{12}$ .

**[0164]** Otra variación del procedimiento de Scholle es diseñar el oligonucleótido mutagénico de forma que un codón de terminación ámbar en el molde se convierta en un codón de terminación ocre, y uno ocre en uno ámbar en el siguiente ciclo de mutagénesis. En este caso, el fago molde y los miembros de la biblioteca mutagenizada deben cultivarse en diferentes cepas supresoras de *E. coli*, alternando un supresor ocre con cepas de supresor ámbar. Esto permite realizar rondas sucesivas de mutagénesis de un fago alternando entre estos dos tipos de codones de terminación y dos cepas supresoras.

**[0165]** Todavía otra variación del enfoque de Scholle implica el uso de megacebadores con un molde de ADN de fago monocatenario. El megacebador es un ADNmc largo que se generó a partir de los insertos de biblioteca del conjunto seleccionado del fago de la ronda previa de inmunopurificación. El objetivo es capturar la diversidad completa de los insertos de biblioteca del conjunto previo, que se mutagenizó en una o más áreas, y transferirla a

una nueva biblioteca de tal forma que pueda mutagenizarse un área adicional. El procedimiento de megacebadores puede repetirse para múltiples ciclos usando el mismo molde que contiene un codón de terminación en el gen de interés. El megacebador es un ADNmc (opcionalmente generado por PCR) que contiene 1) áreas de solapamiento de 5' y 3' de al menos 15 bases para complementariedad con el molde de ADNmc, y 2) una o más áreas de la biblioteca previamente seleccionadas (1, 2, 3, 4 o más) que se copiaron (opcionalmente por PCR) del conjunto de clones previamente seleccionados, y 3) un área de biblioteca recientemente mutagenizada que va a seleccionarse en la siguiente ronda de inmunopurificación. El megacebador se prepara opcionalmente 1) sintetizando uno o más oligonucleótidos que codifican el área de la biblioteca recientemente sintetizada y 2) fusionado ésta, opcionalmente usando PCR de solapamiento, con un fragmento de ADN (opcionalmente obtenido por PCR) que contiene cualquier otra área de biblioteca que se optimizó previamente. La ronda final o PCR monocatenaria del producto de PCR (solapado) combinado se usa para generar el megacebador monocatenario que contiene todas las áreas previamente optimizadas de la nueva biblioteca para un área que va a optimizarse en el siguiente experimento de inmunopurificación. Se espera que este enfoque permita afinidad por maduración de proteínas usando múltiples ciclos rápidos de creación de bibliotecas que generan  $10e^{11}$  a  $10e^{12}$  diversidades por ciclo, cada uno seguido de inmunopurificación.

**[0166]** Puede aplicarse una variedad de procedimientos para introducir diversidad de secuencias en bibliotecas (previamente seleccionadas o sin tratamiento previo) de microproteínas o para mutar clones de microproteínas individuales con el objetivo de potenciar su unión u otras propiedades como fabricación, estabilidad o inmunogenicidad. En principio, todos los procedimientos que pueden usarse para generar bibliotecas también pueden usarse para introducir diversidad en bibliotecas enriquecidas (previamente seleccionadas) de microproteínas. En particular, pueden sintetizarse variantes con unión deseable u otras propiedades y diseñar oligonucleótidos parcialmente aleatorizados basándose en estas secuencias. Este procedimiento permite controlar las posiciones y grado de aleatorización. Puede deducirse la utilidad de mutaciones individuales en una proteína de datos de secuencia de múltiples variantes usando una variedad de algoritmos informáticos (Jonsson, J. y col. (1993) *Nucleic Acids Res.* 21: 733-9; Amin, N. y col. (2004) *Protein Eng Des Sel*, 17: 787-93). De particular interés para la re-mutagénesis de bibliotecas enriquecidas es el barajado de ADN (Stemmer, W. P. C. (1994) *Nature*, 370: 389-391), que genera recombinantes de secuencias individuales en una biblioteca enriquecida. El barajado puede realizarse usando una variedad de condiciones de PCR modificadas y los moldes pueden degradarse parcialmente para potenciar la recombinación. Una alternativa es la recombinación en posiciones predefinidas usando clonación basada en enzimas de restricción. Son de particular interés procedimientos que utilizan enzimas de restricción tipo IIS que escinden ADN fuera de su sitio de reconocimiento de secuencias (Collins, J. y col. (2001) *J Biotechnol.* 74: 317-38). Pueden utilizarse enzimas de restricción que generan residuos protuberantes no palindrómicos para escindir plásmidos u otras mezclas de variantes que codifican ADN en múltiples localizaciones y plásmidos completos pueden reensamblarse por ligación (Berger, S. L. y col. (1993) *Anal Biochem*, 214: 571-9). Otro procedimiento para introducir diversidad es mutagénesis por PCR en la que las secuencias de ADN que codifican miembros de las bibliotecas se someten a PCR bajo condiciones mutagénicas. Se ha descrito que las condiciones de PCR conducen a mutaciones a frecuencias de mutación relativamente altas (Leung, D. y col. (1989) *Technique*, 1: 11-15). Además, puede emplearse una polimerasa con fidelidad reducida (Vanhercke, T. y col. (2005) *Anal Biochem*, 339: 9-14). Un procedimiento de particular interés se basa en cepas de mutador (Irving, R. A. y col. (1996) *Immunotechnology*, 2: 127-43; Coia, G. y col. (1997) *Gene*, 201: 203-9). Éstas son cepas que llevan defectos en uno o más genes de reparación de ADN. Los plásmidos o fago u otro ADN en estas cepas acumulan mutaciones durante la replicación normal. Pueden propagarse clones individuales o poblaciones enriquecidas en cepas de mutador para introducir diversidad genética. Muchos de los procedimientos descritos anteriormente pueden utilizarse en un procedimiento iterativo. Pueden aplicarse múltiples rondas de mutagénesis y cribado o inmunopurificación a genes enteros, o a porciones de un gen, o pueden mutagenizarse diferentes porciones de una proteína durante cada ronda posterior (Yang, W. P. y col. (1995) *J Mol Biol*, 254: 392-403).

**[0167]** Las bibliotecas pueden tratarse adicionalmente para reducir artefactos. Artefactos conocidos de inmunopurificación de fagos incluyen 1) unión no específica basada en hidrofobia, y 2) unión multivalente a la diana, tanto debido a a) la pentavalencia de la proteína del fago pIII como b) debido a la formación de disulfuros entre diferentes microproteínas, produciendo multímeros, o c) debido a recubrimiento de alta densidad de la diana sobre un soporte sólido y 3) unión a diana dependiente de contexto, en la que el contexto de la diana o el contexto de las microproteínas se vuelve crítico para la actividad de unión o de inhibición. Pueden realizarse diferentes etapas de tratamiento para minimizar la magnitud de estos problemas. Por ejemplo, tales tratamientos se aplican a la biblioteca completa, pero algunos tratamientos útiles que eliminan clones malos solo pueden aplicarse a conjuntos de proteínas solubles o solo a proteínas solubles individuales.

**[0168]** Es probable que las bibliotecas de andamiajes que contienen cisteína contengan tioles libres, que

pueden complicar la evolución dirigida por reticulación a otras proteínas. Un enfoque es eliminar los peores clones de la biblioteca pasándolos por una columna sin tior, eliminando así todos los clones que tienen uno más sulfhidrilos libres. Los clones con grupos SH libres también pueden hacerse reaccionar con reactivos de biotina-SH, permitiendo la eliminación eficiente de clones con grupos SH reactivos usando columnas de estreptavidina. Otro enfoque es no eliminar los tioles libres, pero inactivarlos tapando los extremos con compuestos químicos reactivos con sulfhidrilo tales como ácido yodoacético. Son de particular interés reactivos de sulfhidrilo voluminosos o hidrófilos que reducen la unión a diana no específica o variantes modificadas.

**[0169]** Ejemplos de dependencia del contexto son todas las secuencias constantes, que incluyen la proteína pIII, ligadores, marcas de péptidos, biotina-estreptavidina, Fc y otras proteínas de fusión que contribuyen a la interacción. El enfoque típico para evitar la dependencia del contexto implica cambiar el contexto tan frecuentemente como sea práctico con el fin de evitar la formación. Esto puede implicar alternar entre diferentes sistemas de expresión (es decir, M13 frente a T7, o M13 frente a levadura), alternar las marcas y ligadores que se usan, alternar el soporte (sólido) usado para la inmovilización (es decir, química de inmovilización) y alternar las propias proteínas diana (diferentes vendedores, diferentes versiones de fusión).

**[0170]** También puede usarse tratamiento de bibliotecas para seleccionar proteínas con cualidades preferidas. Una opción es el tratamiento de bibliotecas con proteasas con el fin de eliminar variantes inestables de la biblioteca. Las proteasas usadas son normalmente aquellas que se encontrarían en la aplicación. Para administración pulmonar se usarían proteasas pulmonares, por ejemplo, obtenidas por un lavado pulmonar. Similarmente, se obtendrían mezclas de proteasas de suero, saliva, estómago, intestino, piel, nariz, etc. Sin embargo, también es posible usar mezclas de proteasas purificadas individuales. Una amplia lista de proteasas se muestra en [Apéndice E]. Los fagos por sí mismos son excepcionalmente resistentes a la mayoría de las proteasas y otros tratamientos rigurosos.

**[0171]** Por ejemplo, es posible seleccionar la biblioteca para las estructuras más estables, es decir, aquellas con los enlaces disulfuro más fuertes, por exposición a concentraciones crecientes de agentes reductores (es decir, DTT o beta-mercaptoetanol), eliminando así primero las estructuras menos estables. Normalmente se usaría agente reductor (es decir, DTT, BME, otro) a concentraciones de 2,5 mM, a 5 mM, 10 mM, 20 mM, 30 mM, 40 mM, 50 mM, 60 mM, 70 mM, 80 mM, 90 mM o incluso 100 mM, dependiendo de la estabilidad deseada.

**[0172]** También es posible seleccionar clones que pueden replegarse eficientemente *in vitro*, reduciendo la biblioteca de expresión entera con un alto nivel de agente reductor, seguido de reoxidando gradualmente la biblioteca de proteínas para volver a formar los disulfuros, seguido de la eliminación de clones con grupos SH libres, como se ha descrito anteriormente. Este procedimiento puede aplicarse una vez o múltiples veces para eliminar clones que tienen baja eficiencia de replegamiento *in vitro*.

**[0173]** Un enfoque es aplicar una selección genética para el nivel de expresión de proteínas, plegamiento y solubilidad como se describe por A. C. Fisher y col. (2006) Genetic selection for protein solubility enabled by the folding quality control feature of the twin-arginine translocation pathway. Protein Science (en línea). Después de la inmunopurificación de bibliotecas de expresión (opcional), se querría evitar cribar miles de clones al nivel de proteínas para la unión a diana, nivel de expresión y plegamiento. Una alternativa es clonar el conjunto completo de insertos seleccionados en un vector de fusión de beta-lactamasa que, cuando se siembra sobre beta-lactama, los autores demostraron que era selectivo para proteínas bien expresadas, completamente unidas por disulfuro y solubles.

**[0174]** Tras la expresión de fago M13 de bibliotecas de proteínas e inmunopurificación sobre dianas para uno o más ciclos, hay una variedad de formas de proceder, que incluyen (1) cribar clones de fago individuales por ELISA de fago, que mide el número de partículas de fago (usando anticuerpos anti-M13) que se unen a una diana inmovilizada; (2) transferir bibliotecas de expresión de fago de M13 a T7. El segundo enfoque es particularmente útil en reducir la aparición de positivos falsos basándose en la valencia. Cualquier formato de biblioteca individual tiende a favorecer clones que puedan formar contacto de alta avidéz con la diana. Este es el motivo por el cual el cribado de proteínas solubles es importante, aunque esto es una solución tediosa. La multivalencia alcanzada en la expresión de fago T7 es probablemente muy diferente de la alcanzada en la expresión de M13, y la realización en ciclos entre T7 y M13 puede ser un enfoque excelente para reducir la aparición de positivos falsos basándose en la valencia.

**[0175]** La elevación del filtro es otra metodología que puede ser con colonias bacterianas cultivadas a alta densidad sobre placas de agar grandes ( $10^2$ - $10^5$ ). Pequeñas cantidades de algunas proteínas son secretadas en

los medios y terminan unidas a la membrana de filtración (nitrocelulosa o nailon). Entonces, los filtros se bloquean en leche desnatada, 1% de hidrolizado de caseína o una disolución de 1% de BSA y se incuban con la proteína diana que se ha marcado con un colorante fluorescente o una enzima indicadora (directamente o indirectamente mediante anticuerpos o mediante biotina-estreptavidina). La localización de la colonia se determina por revestimiento del filtro sobre el fondo de la placa y todas las colonias positivas se seleccionan y se usan para caracterización adicional. La ventaja de las elevaciones del filtro es que puede hacerse que sean selectivas por afinidad leyendo la señal después de lavar durante diferentes periodos de tiempo. La señal de clones de alta afinidad se 'atenúa' lentamente, mientras que la señal de clones de baja afinidad se atenúa rápidamente. Tal caracterización por afinidad normalmente requiere un ensayo de 4 puntos con un ensayo fundamentado y puede proporcionar mejor comparabilidad clon a clon que ensayos fundamentados. La molienda de colonias en una matriz es útil, ya que minimiza diferencias debido al tamaño o localización de colonias.

#### **Módulos del extremo N:**

15 **[0176]** Las MURP de la invención pueden contener módulos del extremo N (NM), que son particularmente útiles, por ejemplo, en facilitar la producción de las MURP. El NM puede ser un único residuo de metionina si los productos se expresan en el citoplasma de *E. coli*. Un formato de producto típico es un URP fusionado con una proteína terapéutica, que se expresa en el citoplasma bacteriano de manera que el extremo N sea formil-metionina. La formil-metionina puede tanto ser permanente como temporal, si se elimina por procesamiento biológico o químico.

20 **[0177]** El NM también puede ser una secuencia de péptidos que se ha manipulado para procesamiento proteolítico, que puede usarse para eliminar marcas o para eliminar proteínas de fusión. El módulo del extremo N puede manipularse para facilitar la purificación de la MURP, incluyendo una marca de afinidad tal como la marca Flag, Myc, HA o His. El módulo del extremo N también puede incluir una marca de afinidad que puede usarse para la detección de la MURP. Un NM puede manipularse o seleccionarse para expresión de alto nivel de la MURP. También puede manipularse o seleccionarse para potenciar la resistencia a proteasas de la MURP resultante. Las MURP pueden producirse con un módulo del extremo N que facilita la expresión y/o purificación. Este módulo del extremo N puede separarse por escisión durante el procedimiento de producción con una proteasa, de forma que el producto final no contiene un módulo del extremo N.

30 **[0178]** Optimizando la elección de aminoácidos y codones del módulo del extremo N puede aumentarse la producción recombinante. El módulo del extremo N también puede contener un sitio de procesamiento que puede escindirse por una proteasa específica como factor Xa, trombina o enterocinasa. Proteasa del virus del grabado del tomate (VGT). También pueden diseñarse sitios de procesamiento para ser escindibles por hidrólisis química. Un ejemplo es la secuencia de aminoácidos asp-pro que puede escindirse en condiciones ácidas. También puede diseñarse un módulo del extremo N para facilitar la purificación de una MURP. Por ejemplo, pueden diseñarse módulos del extremo N para contener múltiples residuos his que permiten la captura de producto por cromatografía de metal inmovilizado. Los módulos del extremo N pueden contener secuencias de péptidos que pueden capturarse o detectarse específicamente por anticuerpos. Ejemplos son FLAG, HA, c-myc.

#### **Módulos del extremo C:**

45 **[0179]** Las MURP pueden contener un módulo del extremo C, que son particularmente útiles, por ejemplo, en facilitar la producción de las MURP. Por ejemplo, el módulo del extremo C puede comprender un sitio de escisión para efectuar el procesamiento proteolítico para eliminar secuencias que están fusionadas y de ahí aumentar la expresión de proteínas o facilitar la purificación. En particular, el módulo del extremo C también puede contener un sitio de procesamiento que puede escindirse por una proteasa específica como factor Xa, trombina, proteasa de VGT o enterocinasa. Los sitios de procesamiento también pueden diseñarse para ser escindibles por hidrólisis química. Un ejemplo es la secuencia de aminoácidos asp-pro que puede escindirse en condiciones ácidas. El módulo del extremo C puede ser una marca de afinidad que tiene como objetivo facilitar la purificación de la MURP. Por ejemplo, los módulos del extremo C pueden diseñarse para contener múltiples residuos his que permiten la captura de producto por cromatografía de metal inmovilizado. Los módulos del extremo C pueden contener secuencias de péptidos que pueden capturarse o detectarse específicamente por anticuerpos. Ejemplos no limitantes de las marcas incluyen marca FLAG, HA, c-myc o His. El módulo del extremo C también puede manipularse o seleccionarse para potenciar la resistencia a proteasas de la MURP resultante.

**[0180]** Si se desea, el extremo N de la proteína puede ligarse a su propio extremo C. Por ejemplo, la ligación de estos dos módulos puede llevarse a cabo creando un enlace natural similar a aminoácido (enlace peptídico) o usando una entidad de enlace exógena. Son de particular interés ciclotidas, una familia de proteínas pequeñas en

que esto se produce naturalmente. Adoptando un formato estructural como ciclotidas se espera proporcionar estabilidad adicional contra exo-proteasas. Tal enlace intramolecular normalmente funciona mejor a menores concentraciones de proteína.

## 5 Módulos efectores:

**[0181]** Las MURP pueden comprender uno o múltiples módulos efectores (EM), o ninguno. Los módulos efectores normalmente no proporcionan la elección de diana, pero proporcionan una actividad requerida para efecto terapéutico, como destrucción de células. Los EM pueden ser moléculas pequeñas farmacéuticamente activas (es decir, fármacos tóxicos), péptidos o proteínas. Ejemplos no limitantes son citocinas, anticuerpos, enzimas, factores de crecimiento, hormonas, receptores, agonistas o antagonistas de receptores, tanto si están completos como un fragmento o dominio de los mismos. Los módulos efectores también pueden comprender secuencias de péptidos que llevan fármacos de molécula pequeña químicamente ligados, tanto si son sintéticos como naturales. Opcionalmente, estas moléculas efectoras pueden ligarse al módulo efector mediante ligadores químicos, que pueden o pueden no escindirse bajo condiciones seleccionadas conduciendo a una liberación de la actividad tóxica. Los EM también pueden incluir radioisótopos y sus quelatos, además de diversas marcas para TEP y RMN. Los módulos efectores también pueden ser tóxicos para una célula o un tejido. Son de particular interés MURP que contienen módulos efectores tóxicos y módulos de unión con especificidad por un tejido enfermo o tipo de célula enferma. Tales MURP pueden acumularse específicamente en un tejido enfermo o en células enfermas y pueden ejercer su acción tóxica preferencialmente en las células o tejidos enfermos. A continuación se enumeran módulos efectores a modo de ejemplo.

**[0182]** *Enzimas* – Los módulos efectores pueden ser enzimas. Son de particular interés enzimas que degradan metabolitos que son críticos para el crecimiento celular como hidratos de carbono o aminoácidos o lípidos o co-factores. Otros ejemplos de módulos efectores con actividad enzimática son RNasa, ADNasa y fosfatasa, asparaginasa, histidinasa, arginasa, beta-lactamasa. Los módulos efectores con actividad enzimática pueden ser tóxicos cuando se administran a un tejido o célula. Son de particular interés MURP que combinan módulos efectores que son tóxicos y módulos de unión que se unen específicamente a un tejido enfermo. También son módulos efectores posibles enzimas que convierten un profármaco inactivo en un fármaco activo en el sitio tumoral.

**[0183]** *Fármaco* – La MURP de la invención puede contener un efector que es un fármaco. Si se desea, pueden diseñarse secuencias para la administración selectiva al órgano de moléculas de fármaco. Un ejemplo se ilustra en la Figura 8. Una secuencia de URP puede fusionarse con una proteína que se une preferencialmente a un tejido enfermo. La misma secuencia de URP puede contener uno o más residuos de aminoácidos que pueden modificarse para la unión de moléculas de fármaco. Dicho conjugado puede unirse a tejido enfermo con alta especificidad y las moléculas de fármaco unidas pueden dar lugar a una acción local, a la vez que mantienen la exposición sistémica a fármaco. La MURP puede diseñarse para facilitar la liberación de moléculas de fármaco al tamaño de la diana introduciendo sitios sensibles a proteasas que pueden escindirse por proteasas nativas en el sitio de acción deseada. Una ventaja significativa de uso de secuencias de URP para el diseño de construcciones de administración de fármaco es que pueden evitarse interacciones no deseables entre la molécula de fármaco y el dominio de reconocimiento de la construcción. Muchas moléculas de fármaco que pueden conjugarse con dominios de reconocimiento tienen una hidrofobia significativa y los conjugados resultantes tienden a agregarse. Añadiendo secuencias de URP hidrófilas a tales construcciones puede mejorarse la solubilidad de las construcciones de administración resultantes y, como consecuencia, reducirse la tendencia a la agregación. Además, puede aumentarse el número de moléculas de fármaco que pueden fusionarse con un dominio de reconocimiento añadiendo secuencias de URP largas. Además, el uso de secuencias de URP permite optimizar la distancia entre los sitios de conjugación del fármaco para facilitar la completa conjugación. La lista de fármacos adecuados incluye, pero no se limita a, agentes quimioterapéuticos tales como tiotepa y ciclofosfamida (CYTOXAN™); alquilsulfonatos tales como busulfan, improsulfan y piposulfan; aziridinas tales como benzodopa, carbocouona, meturedopa, y uredopa; etileniminas y metilamelaminas que incluyen altretamina, trietilenmelamina, trietilenfosforamida, trietilfosforamida y trimetilolomelamina; mostazas de nitrógeno tales como clorambucilo, clornafazina, ciclofosfamida, estramustina, ifosfamida, mecloretamina, clorhidrato de óxido de mecloretamina, melfalan, novembiquina, fenesterina, prednimustina, trofosfamida, mostaza de uracilo; nitrosoureas tales como carmustina, clorozotocina, fotemustina, lomustina, nimustina, ranimustina; antibióticos tales como aclacinomisininas, actinomycinas, autramicina, azaserina, bleomicinas, cactinomycinas, carabicina, carminomicina, carzinofilina, cromomicinas, dactinomycinas, daunorubicina, detorubicina, 6-diazo-5-oxo-L-norleucina, doxorubicina, epirubicina, esorubicina, idarubicina, marcelomicina, mitomicinas, ácido micofenólico, nogalamicina, olivomicinas, peplomycinas, potfiromycinas, puromycinas, quelamicinas, rodorubicina, estreptonigrina, estreptozocina, tubercidina, ubenimex, zinostatina, zorubicina; antimetabolitos tales como metotrexato y 5-fluorouracilo (5-FU); análogos de ácido fólico tales como

denopterina, metotrexato, pteropterina, trimetrexato; análogos de purina tales como fludarabina, 6-mercaptopurina, tiampirina, tioguanina; análogos de pirimidina tales como ancitabina, azacitidina, 6-azauridina, carmofur, citarabina, didesoxiuridina, doxilfluridina, enocitabina, floxuridina, andrógenos tales como calusterona, propionato de dromostanolona, epitioestanol, mepitioestano, testolactona; antiadrenérgicos tales como aminoglutetimida, mitotano, 5 trilostano; reparadores de ácido fólico tales como ácido frolínico; aceglatona; glucósido aldofosfamida; ácido aminolevulínico; amsacrina; bestrabucilo; bisantreno; edatraxato; defofamina; demecolcina; diazicuona; duocarmicina, maitansina, auristatina, elfornitina; acetato de eliptinio; etoglúcido, nitrato de galio; hidroxiurea; lentinan; lonidamina; mitoguzona; mitoxantrona; mopidamol; nitracrina; pentostatina; fenamet; pirarubicina; ácido podofilínico; 2-etilhidrazida; procarbazona; PSK.R<sup>TM</sup>; razoxana; sizofiran; espirogermanio; ácido tenuazonico; 10 triazicuona; 2,2',2"-trichlorotrietilamina; uretano; vindesina; dacarbazina; manomustina; mitobronitol; mitolactol; pipobromano; gacitosina; arabinósido ("Ara-C"); ciclofosfamida; tiotepa; taxanos, por ejemplo, paclitaxel (TAXOL<sup>TM</sup>, Bristol-Myers Squibb Oncology, Princeton, N.J.) y doxetaxel (TAXOTERE<sup>TM</sup>, Rhône-Poulenc Rorer, Antony, Francia); clorambucilo; gemcitabina; 6-tioguanina; mercaptopurina; metotrexato; análogos de platino tales como cisplatino y carboplatino; vinblastina; platino; etopósido (VP-16); ifosfamida; mitomicina C; mitoxantrona; vincristina; vinorelbina; 15 navelbina; novantrona; tenipósido; daunomicina; aminopterina; xeloda; ibandronato; camptotecina-11 (CPT-11); inhibidor de la topoisomerasa RFS 2000; difluorometilornitina (DMFO); ácido retinoico; esperamicinas; capecitabina; y sales, ácidos o derivados farmacéuticamente aceptables de cualquiera de los anteriores. Como acondicionadores celulares quimioterapéuticos adecuados también están incluidos agentes antihormonales que actúan para regular o inhibir la acción de hormonas sobre tumores tales como antiestrógenos que incluyen, por ejemplo, tamoxifeno, 20 raloxifeno, 4(5)-imidazoles inhibidores de aromatasas, 4-hidroxitamoxifeno, trioxifeno, keoxifeno, LY117018, onapristona y toremifeno (Fareston); y antiandrógenos tales como flutamida, nilutamida, bicalutamida, leuprolida, goserelina, doxorubicina, daunomicina, duocarmicina, vincristina y vinblastina.

**[0184]** Otros fármacos que pueden usarse como módulos efectores incluyen aquellos que son útiles para 25 tratar afecciones inflamatorias, enfermedades cardíacas, enfermedades infecciosas, enfermedades respiratorias, enfermedades autoinmunitarias, trastornos neuronales y musculares, trastornos metabólicos y cánceres.

**[0185]** Fármacos adicionales que pueden usarse como efectores en MURP incluyen agentes para el dolor e inflamación tales como histamina y antagonistas de la histamina, bradiquinina y antagonistas de bradiquinina, 5- 30 hidroxitriptamina (serotonina), sustancias lipídicas que se generan por biotransformación de los productos de la hidrólisis selectiva de fosfolípidos de membrana, eicosanoides, prostaglandinas, tromboxanos, leucotrienos, aspirina, agentes antiinflamatorios no esteroideos, agentes analgésico-antipiréticos, agentes que inhiben la síntesis de prostaglandinas y tromboxanos, inhibidores selectivos de la ciclooxigenasa inducible, inhibidores selectivos de la ciclooxigenasa-2 inducible, autacoides, hormonas paracrinas, somatostatina, gastrina, citocinas que median en 35 interacciones implicadas en respuestas inmunitarias humorales y celulares, autacoides derivados de péptidos, eicosanoides, agonistas β-adrenérgicos, ipratropio, glucocorticoides, metilxantinas, bloqueantes de los canales de sodio, agonistas de receptores de opioides, bloqueantes de canales de calcio, estabilizadores de la membrana e inhibidores de leucotrieno.

40 **[0186]** Otros fármacos que pueden usarse como efector incluyen agentes para el tratamiento de úlceras peptídicas, agentes para el tratamiento de enfermedad por reflujo gastroesofágico, agentes procinéticos, antieméticos, agentes usados en síndrome del intestino irritable, agentes usados para diarrea, agentes usados para estreñimiento, agentes usados para enfermedad inflamatoria del intestino, agentes usados para enfermedad biliar, 45 agentes usados para enfermedad pancreática.

**[0187]** *Radionúclidos* – Las MURP pueden diseñarse para la administración que elige tejido como diana de radionúclidos, además de para la obtención de imágenes con radionúclidos. Las URP son ideales para la obtención de imágenes debido a que la semivida puede optimizarse cambiando la longitud del URP. Para la mayoría de las aplicaciones de obtención de imágenes es probable que se prefiera un URP moderadamente largo, proporcionando 50 una semivida de 5 minutos a algunas horas, no días o semanas. Las MURP pueden diseñarse de forma que solo contengan un único grupo amino o un pequeño número definido de grupos amino que pueden modificarse con agentes quelantes (tales como DOTA) para radioisótopos tales como tecnecio, indio e itrio (EXPAND). Procedimientos alternativos de conjugación son mediante cadenas laterales de cisteína reservadas. Tales MURP que llevan radionúclidos pueden emplearse para el tratamiento de tumores u otros tejidos enfermos, además de para 55 la obtención de imágenes.

**[0188]** Muchas proteínas farmacéuticamente activas o dominios de proteína pueden usarse como modelos efectores en MURP. Ejemplos son las siguientes proteínas, además de fragmentos de estas proteínas: citocinas, factores de crecimiento, enzimas, receptores, microproteínas, hormonas, eritropoyetina, adenosina desiminasa,

asparaginasa, arginasa, interferón, hormona de crecimiento, hormona liberadora de hormona de crecimiento, G-CSF, GM-CSM, insulina, hirudina, receptor de TNF, uricasa, rasburicasa, axokine, RNAsa, ADNsa, fosfatasa, exotoxina de *Pseudomonas*, ricina, gelonina, desmoteplasa, laronidasa, trombina, enzima de coagulación de la sangre, VEGF, protropina, somatropina, alteplasa, interleucina, factor IIV, factor VIII, factor X, factor IX, dornasa, glucocerebrosidasa, folitropina, glucagón, tiotropina, nesiritida, alteplasa, teriparatida, agalsidasa, laronidasa, metioninasa.

**[0189]** *MURP activada por proteasa:* Para potenciar el índice terapéutico de un módulo efector pueden insertarse secuencias lábiles a proteasas en secuencias de URP que son sensibles a proteasas que se encuentran preferencialmente en suero o en el tejido diana que va a tratarse por la MURP. Este enfoque se ilustra en la Figura 9. Algunos diseños permiten construir proteínas que son selectivamente activadas cuando llegan a un tejido diana. Son de particular interés MURP que se activan en un sitio de enfermedad. Para facilitar tal activación específica para diana, secuencias de URP pueden unirse en estrecha proximidad al sitio activo o sitio de unión a receptor del módulo efector de forma que la proteína de fusión resultante tenga actividad biológica limitada. Es de particular interés la activación de un módulo efector en un sitio tumoral. Muchos tejidos tumorales expresan proteasas a concentraciones relativamente altas y secuencias que son específicamente escindidas por estas proteasas tumorales pueden insertarse en secuencias de URP. Por ejemplo, la mayoría de los tejidos tumorales de la próstata contienen altas concentraciones de antígeno específico de la próstata (PSA), que es una serina proteasa. Profármacos constituidos por un péptido lábil a PSA conjugado con el fármaco contra el cáncer doxorubicina han mostrado activación selectiva en tejido de próstata [DeFeo-Jones, D. y col. (2000) Nat Med, 6: 1248]. De particular interés para activación específica para enfermedad son proteínas con actividad citostática o citotóxica como TNF-alfa, y muchas citocinas e interleucinas. Otra aplicación es la activación selectiva de proteínas en el sitio de inflamación o en el sitio de virus o infección bacteriana.

**[0190]** *Procedimientos de producción* – Las MURP que contienen secuencias de URP pueden producirse usando enfoques de biología molecular que son muy conocidos en la materia. Están disponibles una variedad de vectores de clonación para diversos sistemas de expresión como células de mamífero, levadura y microbios. De particular interés como huéspedes de expresión son *E. coli*, *S. cerevisiae*, *P. pastoris* y células de ovario de hámster chino. Son de particular interés huéspedes que han sido optimizados para ampliar su uso de codones. Es de particular interés un huésped que ha sido modificado para potenciar la expresión de GRS. Esto puede hacerse proporcionando ADN que codifica ARNt específicos para glicina. Además, puede manipularse el huésped de forma que se potencie la carga de ARNt específico para glicina. El ADN que codifica la proteína potenciada puede ligarse operacionalmente a secuencias promotoras. El ADN que codifica la proteína potenciada, además del promotor operacionalmente ligado, puede ser parte de un vector de plásmido, vector vírico o puede insertarse en el cromosoma del huésped.

**[0191]** Para la producción, el huésped puede cultivarse en condiciones que faciliten la producción de la proteína potenciada. Son de particular interés condiciones que mejoran la producción de GRS.

**[0192]** Las MURP de la invención pueden adoptar una variedad de formatos. Por ejemplo, las MURP pueden contener URP que están fusionados con proteínas farmacéuticamente activas para producir productos de lenta liberación. Tales productos pueden inyectarse o implantarse localmente, por ejemplo, en o bajo la piel de un paciente. Debido a su gran radio hidrodinámico, el producto que contiene secuencias de URP se libera lentamente del sitio de inyección o implantación que conduce a una reducción de la frecuencia de inyección o implantación. Las secuencias de URP pueden diseñarse para contener regiones que se unen a superficies celulares o tejido con el fin de prolongar la retención local del fármaco en el sitio de inyección. Son de particular interés productos que contienen URP que pueden formularse como compuestos solubles, pero forman agregados o precipitados tras la inyección. Esta agregación o precipitación puede desencadenarse por un cambio en pH entre el producto formulado y el pH en el sitio de inyección. Alternativas son productos que contienen URP que precipitan o forman agregados como resultado de un cambio en las condiciones rédox. Todavía otro enfoque es un producto que contiene URP que se estabiliza en disolución mediante la adición de solutos no activos, pero que precipita o se agrega tras la inyección como resultado de la difusión de los solutos solubilizantes. Otro enfoque es diseñar productos que contienen URP que contienen uno o múltiples residuos de lisina o cisteína en su secuencia de URP y que pueden reticularse antes de la inyección.

**[0193]** Si se desea, la MURP es monomérica (que aquí significa no reticulada) cuando se fabrica y se formula y cuando se inyecta, pero después de la inyección subcutánea la proteína empieza a reticularse consigo mismo o con proteínas humanas nativas, formando un polímero bajo la piel del cual las moléculas de fármaco activas se liberan solo muy gradualmente. Tal liberación puede ser por reducción del enlace disulfuro o barajado de disulfuros

como se ilustra en la Fig. 18, o puede mediarse por la proteólisis como se muestra en la Fig. 19, liberando fragmentos activos en la circulación. Es importante que estos fragmentos cautivos sean suficientemente grandes para tener una larga semivida, debido a que cuanto más larga sea su semivida de secreción, menor podrá ser la dosis de la proteína liberada, permitiendo el uso de una menor dosis de producto que va a inyectarse o un tiempo prolongado entre inyecciones.

**[0194]** Un enfoque que ofrece estas ventajas es la reticulación mediada por disulfuro de proteínas. Por ejemplo, un fármaco de proteína se fabricaría con un péptido cíclico en él (uno o más). Este péptido cíclico puede o puede no participar en la unión a la diana. Esta proteína se fabrica con el péptido cíclico formado, es decir, en forma oxidada, para simplificar la purificación. Sin embargo, el producto se reduce luego y se formula para mantener la proteína en forma reducida. Es importante que el péptido cíclico se reduzca a una baja concentración de agente reductor, tal como ditioneitol 0,25, 0,5, 1,0, 2,0, 4,0 ó 8,0 mM o beta-mercaptoetanol o cisteína o agente reductor equivalente, de manera que el péptido cíclico pueda reducirse sin reducir otros módulos de proteína que contienen disulfuro en el producto. Se prefiere el uso de agentes reductores aprobados por la FDA, tales como cisteína o glutatión. Después de la inyección subcutánea, el agente reductor de bajo peso molecular difunde rápidamente o se neutraliza por proteínas humanas, exponiendo el fármaco a un ambiente oxidante, a la vez que está a alta concentración molar, produciendo reticulación de cisteínas localizadas sobre diferentes cadenas de proteína, que conducen a polimerización del fármaco en el sitio de inyección. Cuanto mayor sea la distancia entre las cisteínas en el péptido cíclico, y mayor sea la concentración de fármaco, mayor será el grado de polimerización del fármaco, ya que la polimerización compite con la re-formación de péptido cíclico. Con el tiempo, la reducción y oxidación del disulfuro producirá re-barajado de disulfuro, que conducirá a la re-formación de péptido cíclico y monomerización y re-solubilización del fármaco. La liberación del fármaco del polímero también puede producirse mediante proteólisis, que podría ser elegida como diana y controlarse o aumentarse formando sitios de escisión para proteasas del suero. La reticulación de las proteínas también podría realizarse con un agente de reticulación químico de proteína-proteína tal como los enumerados en la [Tabla x]. Idealmente, éste es un agente ya aprobado por la FDA, tal como aquellos usados para la conjugación de vacunas o conjugación de productos químicos con proteínas.

**[0195]** En lugar de usar disulfuros, también pueden estabilizarse proteínas contra la degradación proteolítica usando una amplia variedad de agentes de reticulación. La mayoría de los siguientes agentes son comercializados por Pierce Chemicals bajo el mismo nombre e instrucciones para su uso están disponibles en línea ([www.piercenet.com](http://www.piercenet.com)). Los agentes que producen la misma distancia cadena a cadena como se ha obtenido con disulfuros son los que más probablemente serán útiles para la presente aplicación. Los agentes de ligador corto tales como DFDNB son los más prometedores. La distancia entre cadenas puede determinarse fácilmente a partir de las estructuras de los productos químicos como se muestra en [www.piercenet.com](http://www.piercenet.com).

**[0196]** Hay un gran número de productos químicos específicos que funcionan basándose en el siguiente número pequeño de esquemas de reacción básicos, todos los cuales se describen en detalle en [www.piercenet.com](http://www.piercenet.com). Ejemplos de agentes de reticulación útiles son imidoésteres, halógenos activos, maleimida, disulfuro de piridilo, éster de NHS. Los agentes de reticulación homobifuncionales tienen dos grupos reactivos idénticos y se usan frecuentemente en un procedimiento de reticulación química de una etapa. Ejemplos son BS3 (un análogo de DSS soluble en agua no escindible), BSOE (de bases reversibles), DMA (adipimidato de dimetilo-2HCl), DMP (pimelimidato de dimetilo-2HCl), DEM (suberimidato de dimetilo-2HCl), DSG (análogo de 5 carbonos de DSS), DSP (reactivo de Lomant), DSS (no escindible), DST (escindible por agentes de oxidación), DTBP (3,3'-ditiobispropionimidato de dimetilo-2HCl), DTSSP, EGS, Sulfo-EGS, THPP, TSAT, DFDNB (1,5-difluoro-2,4-dinitrobenzoceno) es especialmente útil para reticular entre pequeñas distancias espaciales (Kornblatt, J.A. y Lake, D.F. (1980). Cross-linking of cytochrome oxidase subunits with difluorodinitrobenzene. *Can J. Biochem.* 58, 219-224).

**[0197]** Los agentes de reticulación homofuncionales reactivos con sulfhidrilo que son reticulantes de proteínas homobifuncionales que reaccionan con sulfhidrilos se basan frecuentemente en maleimididas, que reaccionan con grupos -SH a pH 6,5-7,5, formando enlaces tioéter estables. BM[PEO]3 es un espaciador de poliéter de 8 átomos que reduce las posibilidades de precipitación de conjugados en aplicaciones de reticulación de sulfhidrilo con sulfhidrilo. BM[PEO]4 es similar, pero con un espaciador de 11 átomos. BMB es un reticulante no escindible con un espaciador de cuatro carbonos. BMDB forma un enlace que puede escindirse con peryodato. BMH es un reticulante reactivo con sulfhidrilo homobifuncional ampliamente usado. BMOE tiene un ligador especialmente corto. DPDPB y DTME son reticulantes escindibles. HVBS no tiene el potencial de hidrólisis de las maleimididas. TMEA es otra opción. Los agentes reticulantes heterobifuncionales tienen dos grupos reactivos diferentes. Ejemplos son ésteres de NHS y aminas/hidracinas mediante activación de EDC, AEDP, ASBA (fotorreactivo, yodable), EDC (carbodiimida soluble en agua). Los reticulantes bifuncionales reactivos con amina-sulfhidrilo son AMAS, APDP, BMPS, EMCA, EMCS, GMBS, KMUA, LC-SMCC, LC-SPDP, MBS, SBAP, SIA (extra-corto), SIAB, SMCC, SMPB, SMPH, SMPT, SPDP,

sulfo-EMCS, Sulfo-GMBS, Sulfo-KMUS, Sulfo-LC-SMPT, Sulfo-LC-SPDP, Sulfo-MBS, Sulfo-SIAB, Sulfo-SMCC, Sulfo-SMPB. Agentes de reticulación heterobifuncionales reactivos con grupos amino son ANB-NOS, MSA, NHS-ASA, SADP, SAED, SAND, SANPAH, SASD, SFAD, Sulfo-HSAB, Sulfo-NHS-LC-ASA, Sulfo-SADP, Sulfo-SANPAH, TFCS.

5

**[0198]** El fármaco marcado con una marca de His6 tiene un formato de liberación lenta diferente, que se mezcla y co-inyecta con perlas conjugadas con níquel-ácido nitrilotriacético (perlas Ni-NTA), una versión de GMO de las que están disponibles de Qiagen. El fármaco se liberaría lentamente de las perlas, proporcionando liberación controlada y lenta como se ilustra en la Fig. 20. Las perlas son opcionales y pueden sustituirse por un níquel-ácido nitrilotriacético polimérico reticulado que conduce a ensamblaje de un polímero incluso mayor.

10

**[0199]** Las secuencias de URP pueden contener secuencias que son conocidas por formar multímeros como alfa2D [Hill, R, y col. (1998) J Am Chem Soc, 120: 1138-1145] que se utilizó para dimerizar un fragmento de anticuerpo [Kubetzko, S. y col. (2005) Mol Pharmacol, 68: 1439-54]. Ejemplos de un péptido de homodimerización útil es la secuencia SKVILFE. Un ejemplo de secuencias de heterodimerización útiles son el péptido ARARAR que puede formar dímeros con la secuencia DADADA y secuencias relacionadas. La multimerización puede mejorar la función biológica de una molécula aumentando su avidéz y puede influir en las propiedades farmacocinéticas y distribución de tejido de las MURP resultantes.

15

**[0200]** Los “módulos de multimerización” son secuencias de aminoácidos que facilitan la formación de dímeros o multímeros de MURP. Los módulos de multimerización pueden unirse a sí mismos para formar dímeros o multímeros. Alternativamente, los módulos de multimerización pueden unirse a otros módulos de la MURP. Éstas pueden ser cremalleras de leucina o péptidos pequeños como derivados del activador de la cabeza de la hidra (similares a SKVILF) que forman homopolímeros antiparalelos, o péptidos como RARARA y DADADA, que forman heteropolímeros antiparalelos de alta afinidad. Usando una, dos o más copias de estos péptidos puede forzarse la formación de dímeros de proteína, multímeros lineales o multímeros ramificados.

20

**[0201]** La afinidad de la asociación puede confeccionarse cambiando el tipo, longitud y composición de los péptidos. Algunas aplicaciones requieren péptidos que forman homodímeros como se ilustra en la Fig. 21. Otras aplicaciones requieren heterodímeros. En algunos casos, una vez asociados, los péptidos pueden bloquearse *in situ* formando enlaces disulfuro entre las dos cadenas de proteína, normalmente a cualquier lado de los péptidos. Los módulos de multimerización son útiles para enlazar juntas dos moléculas de MURP (cabeza con cola, cabeza con cabeza, o cola con cola) como se ilustra en la Fig. 21. Los módulos de multimerización pueden localizarse en cualquiera del extremo N o C con el fin de formar dímeros. Si los módulos de multimerización están presentes en ambos extremos, se formarán multímeros lineales largos. Si más de dos módulos de multimerización están presentes por proteína, pueden formarse redes poliméricas ramificadas. Los conceptos de multimerización y conjugación química pueden combinarse conduciendo a útiles para la extensión de la semivida y formación de depósito, conduciendo a lenta liberación del fármaco activo del depósito o sitio de inyección como se ilustra en la Fig. 23.

30

**[0202]** Las MURP de la invención pueden incorporar un URP genérico o universal. Un enfoque es expresar un URP que contiene un módulo de URP largo, que proporciona semivida y contiene múltiples (normalmente 4-10) lisinas (u otros sitios) que permiten la conjugación específica para sitio de péptidos (es decir, lineales, cíclicos, 2SS, 3SS, etc.) que se unen a una diana específica. La ventaja de este enfoque es que el módulo de URP es genérico y puede conjugarse con cualquier péptido específico para diana. Idealmente, el enlace del péptido específico para diana con el URP es un enlace dirigido, de manera que los residuos sobre el URP solo pueden reaccionar con un residuo sobre el péptido específico para diana y el acoplamiento exhaustivo solo puede producir una única especie, que es, por ejemplo, un URP que está ligado a un péptido en cada lisina. Este complejo se comporta como un multímero de alta avidéz en sus propiedades de unión, pero es simple de fabricar. Este enfoque se ilustra en la Fig. 24.

40

**[0203]** Las MURP de la invención también pueden incorporar URP para efectuar la administración a través de barreras de tejido. Los URP pueden manipularse para potenciar la administración a través de las barreras dérmica, oral, bucal, intestinal, nasal, hematoencefálica, pulmonar, tecal, peritoneal, rectal, vaginal o muchas otras barreras de tejido.

50

**[0204]** Uno de los obstáculos clave para la administración de proteínas orales es la sensibilidad de la mayoría de las proteínas a proteasas en el aparato digestivo. La conjugación con secuencias de URP puede mejorar la resistencia a proteasas de proteínas farmacéuticamente activas y así facilitar su captación. Se ha mostrado que la

captación de proteínas en el aparato digestivo puede mejorarse añadiendo vehículos moleculares. La función principal de estos vehículos es una mejora de la permeabilidad de la membrana [Stoll, B. R. y col. (2000) J Control Release, 64: 217-28]. Así, pueden incluirse secuencias en secuencias de URP que mejoran la permeabilidad de la membrana. Muchas secuencias que mejoran la permeabilidad de la membrana son conocidas y ejemplos son 5 secuencias ricas en arginina [Takenobu, T. y col. (2002) Mol Cancer Ther, 1: 1043-9]. Así, pueden diseñarse secuencias de URP que mejoran la captación celular u oral de proteínas combinando dos funciones, una reducción en la degradación proteolítica de la proteína de interés, además de un aumento en la permeabilidad de la membrana del producto de fusión. Opcional, puede añadirse secuencia a la secuencia de URP que es sensible a una proteasa que está preferencialmente localizada en el tejido diana para el fármaco de interés, pero es estable a proteasas en el 10 tubo digestivo. Ejemplos de tales secuencias de URP son secuencias que contienen regiones largas de GRS, además de secuencias que son ricas en aminoácidos básicos, en particular arginina, y facilitan la transferencia de la membrana. URP puede utilizarse de un modo similar para mejorar la captación de proteínas mediante las vías intranasal, intrapulmonar, u otras vías de administración.

#### 15 **Ejemplos de productos específicos:**

**[0205]** *Agonista de DR4/DR5* – DR4 y DR5 son receptores de muerte que se expresan en muchas células tumorales. Estos receptores pueden desencadenarse por trimerización, que conduce a muerte celular y regresión tumoral. Los dominios de unión con especificidad por DR4 o DR5 pueden obtenerse por inmunopurificación de fago u otros procedimientos de expresión: Estos dominios específicos de unión a DR4 o DR5 pueden multimerizarse usando módulos de URP como ligadores como se ilustra en la Figura 12. Son de particular interés MURP que contienen tres o más módulos de unión con especificidad por DR4 o DR5 o ambos. Como se ilustra en la Figura 12, las MURP pueden contener módulos de unión adicionales con especificidad por antígenos de tumor que se expresan en exceso en tejidos tumorales. Esto permite construir MURP que se acumulan específicamente en tejido 20 tumoral y desencadenan muerte celular. Las MURP pueden contener módulos que se unen a o bien DR4 o a DR5. Son de particular interés MURP que contienen módulos de unión que se unen tanto a DR4 como a DR5.

**[0206]** *Interleucina 2 elegida como diana por tumor* – La interleucina 2 (IL2) es una citocina que puede potenciar la respuesta inmunitaria a tejido tumoral. Sin embargo, la terapia con IL2 sistémica se caracteriza por 30 efectos secundarios significativos. Las MURP pueden construirse de forma que combinen dominios de unión con especificidad por antígenos de tumor e IL2 como módulo efector como se ilustra en la Figura 13. Tales MURP pueden acumularse selectivamente en tejido tumoral y así provocar una respuesta inmunitaria selectiva para tumor, a la vez que se minimizan los efectos secundarios sistémicos de la terapia con citocinas. Tales MURP pueden elegir como diana una variedad de antígenos de tumor como EpCAM, Her2, CEA, EGFR, antígeno de Thomsen-Friedenreich. Son de particular utilidad MURP que se unen a antígenos de tumor que muestran lenta internalización. 35 Pueden diseñarse MURP similares usando otras citocinas o factor de necrosis tumoral-alfa como módulos efectores.

**[0207]** *Asparaginasa selectiva para tumor* – La asparaginasa se usa para tratar pacientes con leucemia aguda. Tanto la asparaginasa de *E. coli* como la asparaginasa de *Erwinia* se usan para tratamiento. Ambas enzimas 40 pueden conducir a inmunogenicidad y reacciones hipersensibles. Oncaspar es la versión PEGilada de la asparaginasa que tiene inmunogenicidad reducida. Sin embargo, la proteína es difícil de fabricar y se administra como una mezcla de isómeros. La adición de secuencias de URP a los extremos y/o a bucles internos permite la fabricación recombinante directa de una variante de asparaginasa que es homogénea y tiene baja inmunogenicidad. Pueden compararse diversas secuencias de URP y sitios de unión para determinar la posición óptima para la unión 45 a secuencia de URP. Varias otras enzimas que pueden degradar aminoácidos han informado actividad antitumoral. Ejemplos son arginasa, metioninasa, fenilalanina amoniaco liasa y triptofanasa. Es de particular interés la fenilalanina amoniaco liasa de *Streptomyces maritimus*, que tiene una alta actividad específica y no requiere un co-factor [Calabrese, J. C. y col. (2004) Biochemistry, 43: 11403-16]. La mayoría de estas enzimas son de origen bacteriano u otro origen no humano y es probable que provoquen reacciones inmunitarias. La inmunogenicidad de 50 estas enzimas puede reducirse añadiendo una o más secuencias de URP. Además, el índice terapéutico y las propiedades PK de estas enzimas pueden mejorarse aumentando su radio hidrodinámico como resultado de unión a secuencias de URP.

**[0208]** Las MURP de la invención pueden diseñarse para elegir como diana cualquier proteína celular. Una 55 lista no limitante se proporciona a continuación.

**[0209]** VEGF, VEGF-R1, VEGF-R2, VEGF-R3, Her-1, Her-2, Her-3, EGF-1, EGF-2, EGF-3, Alfa3, cMet, ICOS, CD40L, LFA-1, c-Met, ICOS, LFA-1, IL-6, B7,1, B7,2, OX40, IL-1b, TACI, IgE, BAFF o BLys, TPO-R, CD19, CD20, CD22, CD33, CD28, IL-1-R1, TNF $\alpha$ , TRAIL-R1, receptor 1 del complemento, FGFa, osteopontina, vitronectina, efrina

A1-A5, efrina B1-B3, alfa-2-macroglobulina, CCL1, CCL2, CCL3, CCL4, CCL5, CCL6, CCL7, CXCL8, CXCL9, CXCL10, CXCL11, CXCL12, CCL13, CCL14, CCL15, CXCL16, CCL16, CCL17, CCL18, CCL19, CCL20, CCL21, CCL22, PDGF, TGFb, GMCSF, SCF, p40 (IL12/IL23), IL1b, IL1a, IL1ra, IL2, IL3, IL4, IL5, IL6, IL8, IL10, IL12, IL15, IL23, Fas, FasL, ligando Flt3, 41BB, ACE, ACE-2, KGF, FGF-7, SCF, netrina 1, 2, IFNa, b, g, caspasa 2, 3, 7, 8, 10, 5 ADAM S1, S5, 8, 9, 15, TS1, TS5; adiponectina, ALCAM, ALK-1, APRIL, anexina V, angiogenina, anfiregulina, angiopoyetina 1, 2, 4, B7-1/CD80, B7-2/CD86, B7-H1, B7-H2, B7-H3, Bc1-2, BACE-1, BAK, BCAM, BDNF, bNGF, bECGF, BMP2, 3, 4, 5, 6, 7, 8; CRP, cadherina 6, 8, 11; catepsina A, B, C, D, E, L, S, V, X; CD11a/LFA-1, LFA-3, GP2b3a, receptor de GH, proteína RSV F, IL-23 (p40, p19), IL-12, CD80, CD86, CD28, CTLA-4,  $\alpha$ 4 $\beta$ 1,  $\alpha$ 4 $\beta$ 7, TNF/linfotóxina, IgE, CD3, CD20, IL-6, IL-6R, BLYS/BAFF, IL-2R, HER2, EGFR, CD33, CD52, digoxina, Rho (D), 10 varicela, hepatitis, CMV, tétanos, variolovacuna, antídoto, bolutinica, Trail-R1, Trail-R2, cMet, familia de TNF-R, tal como LA NGF-R, CD27, CD30, CD40, CD95, receptor de linfotóxina a/b, Wsl-1, TL1A/TNFSF15, BAFF, BAFF-R/TNFSF13C, TRAIL R2/TNFSF10B, TRAIL R2/TNFSF10B, Fas/TNFSF6 CD27/TNFSF7, DR3/TNFSF25, HVEM/TNFSF14, TROY/TNFSF19, ligando CD40/TNFSF5, BCMA/TNFSF17, CD30/TNFSF5, LIGHT/TNFSF14, 4-1BB/TNFSF9, CD40/TNFSF5, GITR/TNFSF18, osteoprotegerina/TNFSF11B, 15 RANK/TNFSF11A, TRAIL R3/TNFSF10C, TRAIL/TNFSF10, TRANCE/RANK L/TNFSF11, 4-1BB Ligando/TNFSF9, TWEAK/TNFSF12, ligando CD40/TNFSF5, ligando Fas/TNFSF6, RELT/TNFSF19L, APRIL/TNFSF13, DcR3/TNFSF6B, TNF RI/TNFSF1A, TRAIL R1/TNFSF10A, TRAIL R4/TNFSF10D, ligando CD30/TNFSF8, ligando GITR/TNFSF18, TNFSF18, TACI/TNFSF13B, NGF R/TNFSF16, ligando OX40/TNFSF4, TRAIL R2/TNFSF10B, TRAIL R3/TNFSF10C, TWEAK R/TNFSF12, BAFF/BLYS/TNFSF13, DR6/TNFSF21, 20 TNF-alfa/TNFSF1A, Pro-TNFalfa/TNFSF1A, R de linfotóxina beta/TNFSF3, quimera de R de linfotóxina beta (LTbR)/Fc, TNF RI/TNFSF1A, TNF-beta/TNFSF1B, PGRP-S, TNF RI/TNFSF1A, TNF RII/TNFSF1B, EDA-A2, TNF-alfa/TNFSF1A, EDAR, XEDAR, TNF RI/TNFSF1A.

**[0210]** Son de particular interés proteínas diana humana que están comercialmente disponibles en forma 25 purificada. Ejemplos son: 4EBP1, 14-3-3 zeta, 53BP1, 2B4/SLAMF4, CCL21/6Ccina, 4-1BB/TNFSF9, 8D6A, ligando 4-1BB/TNFSF9, 8-oxo-dG, 4-amino-1,8-naftalimida, A2B5, aminopeptidasa LRAP/ERAP2, A33, aminopeptidasa N/ANPEP, Aag, aminopeptidasa P2/YPNPEP2, ABCG2, aminopeptidasa P1/XPNPEP1, ACE, aminopeptidasa PILS/ARTS1, ACE-2, Amnionless, actina, anfiregulina, beta-actina, AMPK alfa 1/2, activina A, AMPK alfa 1, activina AB, AMPK alfa 2, activina B, AMPK beta 1, activina C, AMPK beta 2, activina RIA/ALK-2, R de 30 andrógenos/NR3C4, activina RIB/ALK-4, angiogenina, activina RIIA, angiopoyetina-1, activina RIIIB, angiopoyetina-2, ADAM8, angiopoyetina-3, ADAM9, angiopoyetina-4, ADAM10, 1 similar a angiopoyetina, ADAM12, 2 similar a angiopoyetina, ADAM15, 3 similar a angiopoyetina, TACE/ADAM17, 4 similar a angiopoyetina, ADAM19, 7 similar a angiopoyetina/CDT6, ADAM33, angiostatina, ADAMTS4, anexina A1/anexina I, ADAMTS5, anexina A7, ADAMTS1, anexina A10, ADAMTSL-1/punctina, anexina V, adiponectina/Acrp30, ANP, AEBSF, sitio AP, agregano, APAJF-1, 35 agrina, APC, AgRP, APE, AGTR-2, APJ, AIF, APLP-1, Akt, APLP-2, Akt1, apolipoproteína AI, Akt2, apolipoproteína B, Akt3, APP, albúmina de suero, APRIL/TNFSF13, ALCAM, ARC, ALK-1, artemina, ALK-7, arilsulfatasa A/ARSA, fosfatasa alcalina, ASAH2/N-acilesfingosina amidohidrolasa-2, alfa 2u-globulina, ASC, glicoproteína alfa-1-ácida, ASGR1, alfa-fetoproteína, ASK1, ALS, ATM, ameloblastina, ATRIP, AMICA/JAML, Aurora A, AMIGO, Aurora B, AMIGO2, axina-1, AMIGO3, Axl, aminoacilasa/ACY1, azurocidina/CAP37/HBP, aminopeptidasa A/ENPEP, B4 40 gALT1, BIM, B7-1/CD80, 6-biotina-17-NAD, B7-2/CD86, BLAME/SLAMF8, B7-H1/PD-L1, CXCL13/BLC/BCA-1, B7-H2, BLIMP1, B7-H3, Blk, B7-H4, BMI-1, BACE-1, BMP-1/PCP, BACE-2, BMP-2, Bad, BMP-3, BAFF/TNFSF13B, BMP-3b/GDF-10, BAFF R/TNFSF13C, BMP-4, Bag-1, BMP-5, BAK, BMP-6, BAMBI/NMA, BMP-7, BARD1, BMP-8, Bax, BMP-9, BCAM, BMP-10, Bcl-10, BMP-15/GDF-9B, Bcl-2, BMPR-IA/ALK-3, proteína A1 relacionada con Bcl-2, BMPR-IB/ALK-6, Bcl-w, BMPR-II, Bcl-x, BNIP3L, Bcl-xL, BOC, BCMA/TNFSF17, BOK, BDNF, BPDE, benzamida, 45 Brachyury, cadena beta común, B-Raf, beta IG-H3, CXCL14/BRAK, betacelulina, BRCA1, beta-defensina 2, BRCA2, BID, BTLA, biglicano, Bub-1, proteína asesina similar a Bik, c-jun, CD90/Thy1, c-Rel, CD94, CCL6/C10, CD97, Clq R1/CD93, CD151, ClqTNF1, CD160, ClqTNF4, CD163, ClqTNF5, CD164, componente del complemento C1r, CD200, componente del complemento C1s, CD200 R1, componente del complemento C2, CD229/SLAMF3, componente del complemento C3a, CD23/Fc épsilon RII, componente del complemento C3d, CD2F-10/SLAMF9, 50 componente del complemento C5a, CD5L, cadherina-4/R-cadherina, CD69, cadherina-6, CDC2, cadherina-8, CDC25A, cadherina-11, CDC25B, cadherina-12, CDCP1, cadherina-13, CDO, cadherina-17, CDX4, E-cadherina, CEACAM-1/CD66a, N-cadherina, CEACAM-6, P-cadherina, Cerberus 1, VE-cadherina, CFTR, calbindina D, cGMP, calcineurina A, Chem R23, calcineurina B, Chemerin, calreticulina-2, envases de muestreo de quimiocinas, CaM cinasa II, 1 similar a quitinasa 3, cAMP, quitotriosidasa/CHIT1, R1 de cannabinoide; Chk1, R2 de 55 cannabinoide/CB2/CNR2, Chk2, CAR/NR113, CHL-1/L1CAM-2, anhidrasa carbónica I, colina acetiltransferasa/ChAT, anhidrasa carbónica II, condrolectina, anhidrasa carbónica III, cordina, anhidrasa carbónica IV, 1 similar a cordina, anhidrasa carbónica VA, 2 similar a cordina, anhidrasa carbónica VB, CINC-1, anhidrasa carbónica VI, CINC-2, anhidrasa carbónica VII, CINC-3, anhidrasa carbónica VIII, claspina, anhidrasa carbónica IX, claudina-6, anhidrasa carbónica X, CLC, anhidrasa carbónica XII, CLEC-1, anhidrasa carbónica XIII, CLEC-2, anhidrasa carbónica XIV,

CLECSF13/CLEC4F, carboximetil lisina, CLECSFB, carboxipeptidasa A1/CPA1, CLF-1, carboxipeptidasa A2, CLP1/COLEC12, carboxipeptidasa A4, clusterina, carboxipeptidasa B1, 1 similar a clusterina, carboxipeptidasa E/CPE, CMG-2, carboxipeptidasa X1, CMV UL146, cardiotropina-1, CMV UL147, carnosina dipeptidasa 1, CNP, caronte, CNTF, CART, CNTF R alfa, caspasa, factor de coagulación II/trombina, caspasa-1, factor de coagulación III/factor tisular, caspasa-2, factor de coagulación VII, caspasa-3, factor de coagulación X, caspasa-4, factor de coagulación XI, caspasa-6, factor de coagulación XIV/proteína C, caspasa-7, COCO, caspasa-8, cohesina, caspasa-9, colágeno 1, caspasa-10, colágeno II, caspasa-12, colágeno IV, caspasa-13, cadena gamma común/IL-2 R gamma, inhibidores de péptidos de caspasa, COMP/trombospondina-5, catalasa, componente del complemento C1rLP, beta-catenina, componente del complemento ClqA, catepsina 1, componente del complemento C1qC, catepsina 3, factor D del complemento, catepsina 6, factor I del complemento, catepsina A, complemento MASP3, catepsina B, conexina 43, catepsina C/DPPI, contactina-1, catepsina D, contactina-2/TAG1, catepsina E, contactina-4, catepsina F, contactina-5, catepsina H, corina, catepsina L, comulina, catepsina O, CORS26/ClqTNF,3, catepsina S, citoblastos corticales de rata, catepsina V, cortisol, catepsina X/Z/P, COUP-TF I/NR2F1, CBP, COUP-TF II/NR2F2, CCl, COX-1, CCK-A R, COX-2, CCL28, CRACC/SLAMF7, CCR1, proteína C reactiva, CCR2, creatina cinasa, músculo/CKMM, CCR3, creatinina, CCR4, CREB, CCR5, CREG, CCR6, CRELD1, CCR7, CRELD2, CCR8, CRHBP, CCR9, CRHR-1, CCR10, CRIM1, CD155/PVR, Cripto, CD2, CRISP-2, CD3, CRISP-3, CD4, Crossveinless-2, CD4+/45RA-, CRTAM, CD4+/45RO-, CRTH-2, CD4+/CD62L-/CD44, CRY1, CD4+/CD62L+/CD44, críptico, CD5, CSB/ERCC6, CD6, CCL27/CTACK, CD8, CTGF/CCN2, CD8+/45RA-, CTLA-4, CD8+/45RO-, cubilina, CD9, CX3CR1, CD14, CXADP, CD27/TNFRSF7, CXCL16, ligando CD27/TNFSF7, CXCR3, CD28, CXCR4, CD30/TNFRSF8, CXCR5, ligando CD30/TNFSF8, CXCR6, CD31/PECAM-1, ciclofilina A, CD34, Cyr61/CCN1, CD36/SR-B3, cistatina A, CD38, cistatina B, CD40/TNFRSF5, cistatina C, ligando CD40/TNFSF5, cistatina D, CD43, cistatina E/M, CD44, cistatina F, CD45, cistatina H, CD46, cistatina H2, CD47, cistatina S, CD48/SLAMF2, cistatina SA, CD55/DAF, cistatina SN, CD58/LFA-3, citocromo c, CD59, apocitocromo c, CD68, holocitocromo c, CD72, citoqueratina 8, CD74, citoqueratina 14, CD83, citoqueratina 19, CD84/SLAMF5, citonina, D6, DISP1, DAN, Dkk-1, DANCE, Dkk-2, DARPP-32, Dkk-3, DAX1/NR0B1, Dkk-4, DCC, DLEC, DCIR/CLEC4A, DLL1, DCAR, DLL4, DcR3/TNFRSF6B, d-luciferina, DC-SIGN, ADN ligasa IV, DC-SIGNR/CD299, ADN polimerasa beta, DcTRAIL R1/TNFRSF23, ADNM-1, DcTRAIL R2/TNFRSF22, ADN-PKcs, DDR1, DNER, DDR2, dopa descarboxilasa/DDC, DEC-205, DPCR-1, decapentaplégico, DPP6, decorina, DPPA4, dectina-1/CLEC7A, DPPA5/ESG1, dectina-2/CLEC6A, DPPII/QPP/DPP7, DEP-1/CD148, DPPIV/CD26, Desert Hedgehog, DR3/TNFRSF25, desmina, DR6/TNFRSF21, desmogleína-1, DSCAM, desmogleína-2, DSCAM-L1, desmogleína-3, DSPG3, Dishevelled-1, Dtk, Dishevelled-3, dinamina, EAR2/NR2F6, EphA5, ECE-1, EphA6, ECE-2, EphA7, ECF-L/CHL3L3, EphA8, ECM-1, EphB1, ecotina, EphB2, EDA, EphB3, EDA-A2, EphB4, EDAR, EphB6, EDG-1, efrina, EDG-5, efrina-A1, EDG-8, efrina-A2, eEF-2, efrina-A3, EGF, efrina-A4, EGF R, efrina-A5, EGR1, efrina-B, EG-VEGF/PK1, efrina-B1, eIF2 alfa, efrina-B2, eIF4E, efrina-B3, Elk-1, epigen, EMAP-II, epimorfina/sintaxina 2, EMMPRIN/CD147, epieregulina, CXCL5/ENA, EPR-1/receptor Xa, endocano, ErbB2, endoglicina/CD105, ErbB3, endoglicano, ErbB4, endonucleasa III, ERCC1, endonucleasa IV, ERCC3, endonucleasa V, ERK1/ERK2, endonucleasa VIII, ERK1, endorepentina/perlecano, ERK2, endostatina, ERK3, endotelina-1, ERK5/BMK1, Engrailed-2, ERR alfa/NR3B1, EN-RAGE, ERR beta/NR3B2, enteropeptidasa/enterocinasa, ERR gamma/NR3B3, CCL11/eotaxina, eritropoyetina, CCL24/eotaxina-2, eritropoyetina R, CCL26/eotaxina-3, ESAM, EpCAM/TROP-1, ER alfa/NR3A1, EPCR, ER beta/NR3A2, Eph, exonucleasa III, EphA1, 2 similar a exostosina/EXTL2, EphA2, 3 similar a exostosina /EXTL3, EphA3, FABP1, FGF-BP, FABP2, FGF R1-4, FABP3, FGF R1, FABP4, FGF R2, FABP5, FGF R3, FABP7, FGF R4, FABP9, FGF R5, factor B del complemento, Fgr, FADD, FHR5, FAM3A, fibronectina, FAM3B, ficolina-2, FAM3C, ficolina-3, FAM3D, FITC, proteína alfa de activación de fibroblastos/FAP, FKBP38, Fas/TNFRSF6, Flap, ligando Fas /TNFSF6, FLIP, FATP1, FLRG, FATP4, FLRT1, FATP5, FLRT2, Fc gamma RI/CD64, FLRT3, Fc gamma RIIB/CD32b, Flt-3, Fc gamma RIIC/CD32c, ligando Flt-3, Fc gamma RIIA/CD32a, folistatina, Fc gamma RIIB/CD16, 1 similar a folistatina, FcRH1/IRTA5, FosB/G0S3, FcRH2/IRTA4, FoxD3, FcRH4/IRTA1, FoxJ1, FcRH5/IRTA2, FoxP3, 3 similar a receptor Fc/CD16-2, Fpg, FEN-1, FPR1, fetuína A, FPRL1, fetuína B, FPRL2, FGF ácido, CX3CL1/fractalcina, FGF básico, Frizzled-1, FGF-3, Frizzled-2, FGF-4, Frizzled-3, FGF-5, Frizzled-4, FGF-6, Frizzled-5, FGF-8, Frizzled-6, FGF-9, Frizzled-7, FGF-10, Frizzled-8, FGF-11, Frizzled-9, FGF-12, Frk, FGF-13, sFRP-1, FGF-16, sFRP-2, FGF-17, sFRP-3, FGF-19, sFRP-4, FGF-20, furina, FGF-21, FXR/NR1H4, FGF-22, Fyn, FGF-23, G9a/EHMT2, GFR alfa-3/GDNF R alfa-3, GABA-A-R alfa 1, GFR alfa-4/GDNF R alfa-4, GABA-A-R alfa 2, GITR/TNFRSF18, GABA-A-R alfa 4, ligando GITR/TNFSF18, GABA-A-R alfa 5, GLI-1, GABA-A-R alfa 6, GLI-2, GABA-A-R beta 1, GLP/EHMT1, GABA-A-R beta 2, GLP-1 R, GABA-A-R beta 3, glucagón, GABA-A-R gamma 2, glucosamina (N-acetil)-6-sulfatasa/GNS, GABA-B-R2, GluR1, GAD1/GAD67, GluR2/3, GAD2/GAD65, GluR2, GADD45 alfa, GluR3, GADD45 beta, Glut1, GADD45 gamma, Glut2, galectina-1, Glut3, galectina-2, Glut4, galectina-3, Glut5, galectina-3 BP, glutaredoxina 1, galectina-4, glicina R, galectina-7, glicoforina A, galectina-8, glipecano 2, galectina-9, glipecano 3, GalNAc4S-6ST, glipecano 5, GAP-43, glipecano 6, GAPDH, GM-CSF, Gas1, GM-CSF R alfa, Gas6, GMF-beta, GASP-1/WFIKKNRP, gp130, GASP-2/WFIKKN, glicógeno fosforilasa BB/GPBB, GATA-1, GPR15, GATA-2, GPR39, GATA-3, GPVI, GATA-4, GR/NR3C1, GATA-5, Gr-1/Ly-6G, GATA-6, granulisina, GBL, granzima A, GCNT/NR6A 1, granzima B, CXCL6/GCP-2, granzima D, G-

CSF, granzima G, G-CSF R, granzima H, GDF-1, GRASP, GDF-3 GRB2, GDF-5, gremlina, GDF-6, GRO, GDF-7, CXCL1/GRO alfa, GDF-8, CXCL2/GRO beta, GDF-9, CXCL3/GRO gamma, GDF-11, hormona de crecimiento, GDF-15, hormona de crecimiento R, GDNF, GRP75/HSPA9B, GFAP, GSK-3 alfa/beta, GFI-1, GSK-3 alfa, GFR alfa-1/GDNF R alfa-1, GSK-3 beta, GFR alfa-2/GDNF R alfa-2, EZFIT, H2AX, histidina, H60, HM74A, HAI-1, HMGA2, 5 HAI-2, HMGB1, HAI-2A, TCF-2/HNF-1 beta, HAI-2B, HNF-3 beta/FoxA2, HAND1, HNF-4 alfa/NR2A1, HAPLN1, HNF-4 gamma/NR2A2, proteasa similar a tripsina de las vías respiratorias/HAT, HO-1/HMOX1/HSP32, HB-EGF, HO-2/HMOX2, CCL14a/HCC-1, HPRG, CCL14b/HCC-3, Hrk, CCL16/HCC-4, HRP-1, alfa HCG, HS6ST2, Hck HSD-1, HCR/CRAM-A/B, HSD-2, HDGF, HSP10/EPF, hemoglobina, HSP27, Hepassocin, HSP60, HES-1, HSP70, HES-4, HSP90, HGF, HTRA/proteasa Do, activador de HGF, HTRA1/PRSS11, HGF R, HTRA2/Omi, HIF-1 alfa, 10 HVEM/TNFRSF14, HIF-2 alfa, hialuronano, HIN-1/secretoglobulina 3A1, 4-hidroxinonenal, Hip, CCL1/I-309/TCA-3, IL-10, cIAP (pan), IL-10 R alfa, cIAP-1/HiAP-2, IL-10 R beta, cIAP-2/HLAP-1, IL-11, IBSP/sialoproteína II, IL-11 R alfa, ICAM-1/CD54, IL-12, ICAM-2/CD102, IL-12/IL-23 p40, ICAM-3/CD50, IL-12 R beta 1, ICAM-5, IL-12 R beta 2, ICAT, IL-13, ICOS, IL-13 R alfa 1, iduronato 2-sulfatasa/IDS, IL-13 R alfa 2, IFN, IL-15, IFN-alfa, IL-15 R alfa, IFN-alfa 1, IL-16, IFN-alfa 2, IL-17, IFN-alfa 4b, IL-17 R, IFN-alfa A, IL-17 RC, IFN-alfa B2, IL-17 RD, IFN-alfa C, 1L-17B, 15 IFN-alfa D, IL-17B R, IFN-alfa F, IL-17C, IFN-alfa G, IL-17D, IFN-alfa H2, IL-17E, IFN-alfa I, IL-17F, IFN-alfa J1, IL-18/IL-1F4, IFN-alfa K, IL-18 BPa, IFN-alfa WA, IL-18 BPc, IFN-alfa/beta R1, IL-18 BPd, IFN-alfa/beta R2, IL-18 R alfa/IL-1 R5, IFN-beta, IL-18 R beta/IL-1 R7, IFN-gamma, IL-19, IFN-gamma R1, IL-20, IFN-gamma R2, IL-20 R alfa, IFN-omega, IL-20 R beta, IgE, IL-21, IGFBP-1, IL-21 R, IGFBP-2, IL-22, IGFBP-3, IL-22 R, IGFBP-4, IL-22BP, IGFBP-5, IL-23, IGFBP-6, IL-23 R, IGFBP-L1, IL-24, IGFBP-rp1/IGFBP-7, IL-26/AK155, IGFBP-rP10, IL-27, IGF-I, IL-28A, IGF-I R, IL-28B, IGF-II, IL-29/IFN-lambda 1, IGF-II R, IL-31, IgG, IL-31 RA, IgM, IL-32 alfa, IGSF2, IL-33, 20 IGSF4A/SynCAM, ILT2/CD85j, IGSF4B, ILT3/CD85k, IGSF8, ILT4/CD85d, IgY, ILT5/CD85a, Ikb-beta, ILT6/CD85e, IKK alfa, Hedgehog indio, IKK epsilon, INSR, IKK gamma, insulina, IL-1 alfa/IL-1F1, insulina R/CD220, IL-1 beta/IL-1F2, proinsulina, IL-1ra/IL-1F3, insulinas/IDE, 1L-1F5/FIL1 delta, integrina alfa 2/CD49b, IL-1F6/FIL1 epsilon, integrina alfa 3/CD49c, IL-1F7/FIL1 zeta, integrina alfa 3 beta 1/VLA-3, IL-1F8/FIL1 eta, integrina alfa 4/CD49d, IL-25 1F9/IL-1 H1, integrina alfa 5/CD49e, IL-1F10/IL-1HY2, integrina alfa 5 beta 1, IL-1 RI, integrina alfa 6/CD49f, IL-1 RII, integrina alfa 7, IL-1 R3/IL-1 R AcP, integrina alfa 9, IL-1 R4/ST2, integrina alfa E/CD103, IL-1 R6/IL-1 R rp2; integrina alfa L/CD11a, IL-1 R8, integrina alfa L beta 2, IL-1 R9, integrina alfa M/CD11b, IL-2, integrina alfa M beta 2, IL-2 R alfa, integrina alfa V/CD51, IL-2 R beta, integrina alfa V beta 5, IL-3, integrina alfa V beta 3, IL-3 R alfa, integrina alfa V beta 6, IL-3 R beta, integrina alfa X/CD11c, IL-4, integrina beta 1/CD29, IL-4 R, integrina beta 2/CD18, IL-5, integrina beta 3/CD61, IL-5 R alfa, integrina beta 5, IL-6, integrina beta 6, IL-6 R, integrina beta 7, IL-7, 30 CXCL10/IP-10/CRG-2, IL-7 R alfa/CD127, IRAK1, CXCR1/IL-8 RA, IRAK4, CXCR2/IL-8 RB, IRS-1, CXCL8/IL-8, Islet-1, IL-9, CXCL11/1-TAC, IL-9 R, Jagged 1, JAM-4/IGSF5, Jagged 2, JNK, JAM-A, JNK1/JNK2, JAM-B/VE-JAM, JNK1, JAM-C, JNK2, quininógeno, calcireína 3/PSA, quininostatina, calcireína 4, KIR/CD158, calcireína 5, KIR2DL1, calcireína 6/neuroquina, KIR2DL3, calcireína 7, KIR2DL4/CD158d, calcireína 8/neuroquina, KIR2DS4, calcireína 9, 35 KIR3DL1, calcireína plasmática/KLKB1, KIR3DL2, calcireína 10, Kirrel2, calcireína 11, KLF4, calcireína 12, KLF5, calcireína 13, KLF6, calcireína 14, Klotho, calcireína 15, Klotho beta, KC, KOR, Keap1, Kremen-1, Kell, Kremen-2, KGF/FGF-7, LAG-3, LINGO-2, LAIR1, lipina 2, LAIR2, lipocalina-1, laminina alfa 4, lipocalina-2/NGAL, laminina gamma 1, 5-lipoxigenasa, laminina I, LXR alfa/NR1H3, laminina S, LXR beta/NR1H2, laminina-1, livina, laminina-5, LIX, LAMP, LMIR1/CD300A, langerina, LMIR2/CD300c, LAR, LMIR3/CD300LF, latexina, LMIRS/CD300LB, layilina, 40 LMIR6/CD300LE, LBP, LM02, LDL R, LOM-1/SR-E1, LECT2, LRH-1/NR5A2, LEDGF, LRIG1, Lefty, LRIG3, Lefty-1, LRP-1, Lefty-A, LRP-6, legumina, LSECtin/CLEC4G, leptina, lumicina, leptina R, CXCL15/Lungkine, leucotrieno B4, XCL1/linfotactina, leucotrieno B4 R1, linfotoxina, LIF, linfotoxina beta/TNFSF3, LIF R alfa, linfotoxina beta R/TNFRSF3, LIGHT/TNFSF14, Lyn, limitina, Lyp, LIMPII/SR-B2, homólogo 2 de lisil oxidasa, LIN-28, LYVE-1, LINGO-1, alfa 2-macroglobulina, CXCL9/MIG, MAD2L1, mimecano, MAdCAM-1, mindina, MafB, R de 45 mineralocorticoide/NR3C2, Maff, CCL3L1/MIP-1 alfa isoforma LD78 beta, MafG, CCL3/MIP-1 alfa, MafK, CCL4L1/LAG-1, MAG/Siglec-4a, CCL4/MIP-1 beta, MANF, CCL15/MIP-1 delta, MAP2, CCL9/10/MIP-1 gamma, MAPK, MIP-2, marapsina/pancreasina, CCL19/MIP-3 beta, MARCKS, CCL20/MIP-3 alfa, MARCO, MIP-I, Mash1, MIP-II, matrilina-2, MIP-III, matrilina-3, MIS/AMH, matrilina-4, MIS RII, matriptasa/ST14, MIXL1, MBL, MKK3/MKK6, MBL-2, MKK3, melanocortina 3R/MC3R, MKK4, MCAM/CD146, MKK6, MCK-2, MKK7, Mcl-1, MKP-3, MCP-6, MLH-50 1, CCL2/MCP-1, MLK4 alfa, MCP-11, MMP, CCL8/MCP-2, MMP-1, CCL7/MCP-3/MARC, MMP-2, CCL13/MCP-4, MMP-3, CCL12/MCP-5, MMP-7, M-CSF, MMP-8, M-CSF R, MMP-9, MCV-tipo II, MMP-10, MD-1, MMP-11, MD-2, MMP-12, CCL22/MDC, MMP-13, MDL-1/CLEC5A, MMP-14, MDM2, MMP-15, MEA-1, MMP-16/MT3-MMP, MEK1/MEK2, MMP-24/MT5-MMP, MEK1, MMP-25/MT6-MMP, MEK2, MMP-26, melusina, MMR, MEPE, MOG, meprina alfa, CCL23/MPIF-1, meprina beta, M-Ras/R-Ras3, Mer, Mre11, mesotelina, MRP1, meteorina, 55 MSK1/MSK2, metionina aminopeptidasa 1, MSK1, metionina aminopeptidasa, MSK2, metionina aminopeptidasa 2, MSP, MFG-E8, MSP R/Ron, MFRP, Mug, MgcRacGAP, MULT-1, MGL2, Musashi-1, MGMT, Musashi-2, MIA, MuSK, MICA, MutY ADN glicosilasa, MICB, MyD88, M1CL/CLEC12A, mieloperoxidasa, beta 2 microglobulina, miocardina, Midkine, miocilina, MIF, mioglobina, NAIP NGFI-B gamma/NR4A3, Nanog, Ngr2/NgrH1, CXCL7/NAP-2, Ngr3/NgrH2, Nbsl, nidogen-1/entactina, NCAM-1/CD56, nidogen-2, NCAM-L1, óxido nítrico, nectina-1, nitrotirosina,

nectina-2/CD112, NKG2A, nectina-3, NKG2C, nectina-4, NKG2D, neogenina, NKp30, neprilisa/CD10, NKp44, neprilisa-2/MMEL1/MMEL2, NKp46/NCR1, nestina, NKp80/KLRF1, NET02, NKX2.5, netrina-1, NMDA R, subunidad de NR1, netrina-2, NMDA R, subunidad de NR2A, netrina-4, NMDA R, subunidad de NR2B, netrina-G1a, NMDA R, subunidad de NR2C, netrina-G2a, N-Me-6,7-diOH-TIQ, neuregulina-1/NRG1, nodal, neuregulina-3/NRG3, 5 noggina, neuritina, receptor de nogo, NeuroD1, Nogo-A, neurofascina, NOMO, neurogenina-1, nope, neurogenina-2, norrina, neurogenina-3, eNOS, neurolisina, iNOS, neurofisina II, nNOS, neurofilina-1, Notch-1, neurofilina-2, Notch-2, neuropoyetina, Notch-3, neurotrimina, Notch-4, neurturina, NOV/CCN3, NFAM1, NRAGE, NF-H, NrCAM, NFkB1, NRL, NFkB2, NT-3, NF-L, NT-4, NF-M, NTB-A/SLAMF6, NG2/MCSP, NTH1, NGF R/TNFRSF16, nucleostemina, beta-NGF, Nurr-1/NR4A2, NGFI-B alfa/NR4A1, OAS2, orexina B, OBCAM, OSCAR, OCAM, OSF-2/periostina, 10 OCIL/CLEC2d, oncostatina M/OSM, OCILRP2/CLEC2i, OSM R beta, Oct-3/4, osteoactivina/GPNMB, OGG1, osteoadherina, Olig 1, 2, 3, osteocalcina, Olig1, osteocrina, Olig2, osteopontina, Olig3, osteoprotegerina/TNFRSF11B, marcador de oligodendrocitos O1, Otx2, marcador de oligodendrocitos 04, OV-6, OMgp, OX40/TNFRSF4, opticina, ligando OX40/TNFSF4, orexina A, OAS2, orexina B, OBCAM, OSCAR, OCAM, OSF-2/periostina, OCIL/CLEC2d, oncostatina M/OSM, OCILRP2/CLEC2i, OSM R beta, Oct-3/4, 15 osteoactivina/GPNMB, OGG1, osteoadherina, Olig 1, 2, 3, osteocalcina, Olig1, osteocrina, Olig2, osteopontina, Olig3, osteoprotegerina/TNFRSF11B, marcador de oligodendrocitos O1, Otx2, marcador de oligodendrocitos 04, OV-6, OMgp, OX40/TNFRSF4, opticina, ligando OX40/TNFSF4, orexina A, RACK1, Ret, Rad1, REV-ERB alfa/NR1D1, Rad17, REV-ERB beta/NR1D2, Rad51, Rex-1, Rae-1, RGM-A, Rae-1 alfa, RGM-B, Rae-1 beta, RGM-C, Rae-1 delta, Rheb, Rae-1 épsilon, proteína ribosómica S6, Rae-1 gamma, RIP1, Raf-1, ROBO1, RAGE, ROBO2, 20 Ra1A/Ra1B, ROB03, Ra1A, ROBO4, Ra1B, ROR/NR1F1-3 (pan), RANK/TNFRSF11A, ROR alfa/NR1F1, CCL5/RANTES, ROR gamma/NR1F3, Rap1A/B, receptor 1 de orfano similar a RTK/ROR1, RAR alfa/NR1B1, receptor 2 de orfano similar a RTK/ROR2, RAR beta/NR1B2, RP105, RAR gamma/NR1B3, RPA2, Ras, RSK (pan), RBP4, RSK1/RSK2, RECK, RSK1, Reg 2/PAP, RSK2, Reg I, RSK3, Reg II, RSK4, Reg III, R-espondina 1, Reg IIIa, R-espondina 2, Reg IV, R-espondina 3, relaxina-1, RUNX1/CBFA2, relaxina-2, RUNX2/CBFA1, relaxina-3, 25 RUNX3/CBFA3, RELM alfa, RXR alfa/NR2B1, RELM beta, RXR beta/NR2B2, RELT/TNFRSF19L, RXR gamma/NR2B3, resistina, S100A10, SLITRK5, S100A8, SLPI, S100A9, SMAC/diablo, S100B, Smad1, S100P, Smad2, SALL1, Smad3, delta-sarcoglicano, Smad4, Sca-1/Ly6, Smad5, SCD-1, Smad7, SCF, Smad8, SCF R/c-kit, SMC1, SCGF, alfa-actina de músculo liso, SCL/Tal1, SMUG1, SCP3/SYCP3, Snail, CXCL12/SDF-1, intercambiador 1 de sodio y calcio, SDNSF/MCFD2, Soggy-1, alfa-secretasa, Sonic Hedgehog, gamma-secretasa, SorCS1, beta-secretasa, SorCS3, E-selectina, sortilina, L-selectina, SOST, P-selectina, SOX1, semaforina 3A, SOX2, semaforina 30 3C, SOX3, semaforina 3E, SOX7, semaforina 3F, SOX9, semaforina 6A, SOX10, semaforina 6B, SOX17, semaforina 6C, SOX21 semaforina 6D, SPARC, semaforina 7A, 1 similar a SPARC, separasa, SP-D, substrato I de serina/treonina fosfatasa, espinesina, serpina A1, F-espondina, serpina A3, SR-AI/MSR, serpina A4/calistatina, Src, serpina A5/inhibidor de proteína C, SREC-I/SR-F1, serpina A8/angiotensinógeno, SREC-II, serpina B5, SSEA-1, serpina 35 C1/antitrombina-III, SSEA-3, serpina D1/cofactor II de heparina, SSEA-4, serpina E1/PAI-1, ST7/LRP12, serpina E2, estabilina-1, serpina F1, estabilina-2, serpina F2, estaniocalcina 1, serpina G1/inhibidor de C1, estaniocalcina 2, serpina I2, STAT1, amiloide A1 del suero, STAT2, SF-1/NR5A1, STAT3, SGK, STAT4, SHBG, STAT5a/b, SHIP, STAT5a, SHP/NR0B2, STAT5b, SHP-1, STAT6, SHP-2, VE-estatina, SIGIRR, Stella/Dppa3, Siglec-2/CD22, STRO-1, Siglec-3/CD33, sustancia P, Siglec-5, sulfamidasa/SGSH, Siglec-6, factor 1 modificador de sulfatasa/SUMF1, 40 Siglec-7, factor 2 modificador de sulfatasa/SUMF2, Siglec-9, SUMO1, Siglec-10, SUMO2/3/4, Siglec-11, SUMO3, Siglec-F, superóxido dismutasa, SIGNR1/CD209, superóxido dismutasa-1/Cu-Zn SOD, SIGNR4, superóxido dismutasa-2/Mn-SOD, SIRP beta 1, superóxido dismutasa-3/EC-SOD, SKI, survivina, SLAM/CD150, sinapsina I, transposasa Sleeping Beauty, sindecano-1/CD138, Slit3, sindecano-2, SLITRK1, sindecano-3, SLITRK2, sindecano-4, SLITRK4, TACI/TNFRSF13B, TMEFF1/tomoregulina-1, TAO2, TMEFF2, TAPP1, TNF-alfa/TNFSF1A, 45 CCL17/TARC, TNF-beta/TNFSF1B, Tau, TNF RI/TNFRSF1A, TC21/R-Ras2, TNF RII/TNFRSF1B, TCAM-1, TOR, TCCR/WSX-1, TP-1, TC-PTP, TP63/TP73L, TDG, TR, CCL25/TECK, TR alfa/NR1A1, tenascina C, TR beta 1/NR1A2, tenascina R, TR2/NR2C1, TER-119, TR4/NR2C2, TERT, TRA-1-85, testicano 1/SPOCK1, TRADD, testicano 2/SPOCK2, TRAF-1, testicano 3/SPOCK3, TRAF-2, TFPI, TRAF-3, TFPI-2, TRAF-4, TGF-alfa, TRAF-6, TGF-beta, TRAIL/TNFSF10, TGF-beta 1, TRAIL R1/TNFRSF10A, LAP (TGF-beta 1), TRAIL R2/TNFRSF10B, TGF-beta 1 latente, TRAIL R3/TNFRSF10C, TGF-beta 1,2, TRAIL R4/TNFRSF10D, TGF-beta 2, TRANCE/TNFSF11, 50 TGF-beta 3, TfR (transferrina R), TGF-beta 5, apo-transferrina, TGF-beta latente pb 1, holo-transferrina, TGF-beta latente pb 2, trapina-2/elafina, TGF-beta latente pb4, TREM-1, TGF-beta RI/ALK-5, TREM-2, TGF-beta RII, TREM-3, TGF-beta RIIB, TREML1/TLT-1, TGF-beta RIII, TRF-1, termolisina, TRF-2, tiorredoxina-1, ectoenzima degradadora de TRH/TRHDE, tiorredoxina-2, TRIM5, tiorredoxina-80, tripeptidil-peptidasa 1, 5 similar a tiorredoxina /TRP14, TrkA, 55 THOP1, TrkB, trombospondulina/CD141, TrkC, trombosoyetina, TROP-2, trombosoyetina R, péptido 3 de troponina I, trombospondina-1, troponina T, trombospondina-2, TROY/TNFRSF19, trombospondina-4, tripsina 1, timopoyetina, tripsina 2/PRSS2, quimiocina-1 del timo, tripsina 3/PRSS3, Tie-1, triptasa-5/Prss32, Tie-2, triptasa alfa/TPS1, TIM-1/KIM-1/HAVCR, triptasa beta-1/MCPT-7, TIM-2, triptasa beta-2/TPSB2, TIM-3, triptasa épsilon/BSSP-4, TIM-4, triptasa gamma-1/TPSG1, TIM-5, triptófano hidroxilasa, TIM-6, TSC22, TIMP-1, TSG, TIMP-2, TSG-6, TIMP-3, TSK,

TIMP-4, TSLP, TL1A/TNFSF15, TSLP R, TLR1, TSP50, TLR2, beta-III tubulina, TLR3, TWEAK/TNFSF12, TLR4, TWEAK R/TNFSF12, TLR5, Tyk2, TLR6, fosfo-tirosina, TLR9, tirosina hidroxitasa, TLX/NR2E1, sustrato I de tirosina fosfatasa, ubiquitina, UNCSH3, Ugi, UNCSH4, UGRP1, UNG, ULBP-1, uPA, ULBP-2, uPAR, ULBP-3, URB, UNC5H1, UVDE, UNC5H2, vainilloide R1, VEGF R, VASA, VEGF R1/Flt-1, vasohibina, VEGF R2/KDR/Fik-1, 5 vasorina, VEGF R3/Flt-4, vasostatina, versicano, Vav-1, VG5Q, VCAM-1, VHR, VDR/NR111, vimentina, VEGF, vitronectina, VEGF-B, VLDLR, VEGF-C, vWF-A2, VEGF-D, sinucleína-alfa, Ku70, WASP, Wnt-7b, WIF-1, Wnt-8a WISP-1/CCN4, Wnt-8b, WNK1, Wnt-9a, Wnt-1, Wnt-9b, Wnt-3a, Wnt-10a, Wnt-4, Wnt-10b, Wnt-5a, Wnt-11, Wnt-5b, wnvNS3, Wnt7a, XCR1, XPE/DDB1, XEDAR, XPE/DDB2, Xg, XPF, XIAP, XPG, XPA, XPV, XPD, XRCC1, Sí, YY1, EphA4.

10

**[0211]** Numerosos canales de iones humanos son dianas de particular interés. Ejemplos no limitantes incluyen subunidad B del receptor 3 de 5-hidroxitriptamina, precursor del receptor 3 de 5-hidroxitriptamina, subunidad C del receptor 3 de 5-hidroxitriptamina, proteína 14 de AAD, proteína de receptor de acetilcolina, precursor de la subunidad alfa, proteína de receptor de acetilcolina, precursor de la subunidad beta, proteína de receptor de acetilcolina, precursor de la subunidad delta, proteína de receptor de acetilcolina, precursor de la subunidad épsilon, proteína de receptor de acetilcolina, precursor de la subunidad gamma, variante de corte y empalme b de canales de iones sensibles a ácidos 3, variante de corte y empalme c de canales de iones sensible a ácidos 3, canal de iones sensible a ácidos 4, ADP-ribosa pirofosfatasa, precursor mitocondrial, canal de calcio dependiente de voltaje alfa1A, canal de cationes sensible a amilorida 1, neuronal, canal de cationes sensible a amilorida 2, canal de cationes sensible a amilorida 4 neuronal, isoforma 2, canal de sodio sensible a amilorida, subunidad alfa de canales de sodio sensibles a amilorida, subunidad beta de canales de sodio sensibles a amilorida, subunidad delta de canales de sodio sensibles a amilorida, subunidad gamma de canales de sodio sensibles a amilorida, anexina A7, proteína similar a apical, canal de potasio de rectificación interna sensible a ATP 1, canal de potasio de rectificación interna sensible a ATP 10, canal de potasio de rectificación interna sensible a ATP 11, canal de potasio de rectificación interna sensible al ATP 14, canal de potasio de rectificación interna sensible al ATP 15, canal de potasio de rectificación interna sensible al ATP 8, subunidad alfa 12.2 de canales de calcio, subunidad alfa 12.2 de canales de calcio, subunidad alfa1E de canales de calcio, variante de corte y empalme delta19 delta40 delta46, subunidad 1 de canales de potasio activados por calcio alfa, subunidad 1 de canales de potasio activados por calcio beta, subunidad 2 de canales de potasio activados por calcio beta, subunidad 3 de canales de potasio activados por calcio beta, 30 canal de cloruro dependiente de calcio 1, canal de cationes TRPM4B, CADN FLJ90453 fis, clon NT2RP3001542, altamente similar a 6 que contiene dominio de tetramerización de canales de potasio, CADN FLJ90663 fis, clon PLACE1005031, altamente similar a proteína 5 de canales intracelulares de cloruro, subunidad beta de canales de cationes regulados por CGMP, proteína de canales de cloruro, proteína de canales de cloruro 2, proteína de canales de cloruro 3, proteína de canales de cloruro 4, proteína de canales de cloruro 5, proteína de canales de cloruro 6, 35 proteína de canales de cloruro C1C-Ka, proteína de canales de cloruro C1C-Kb, proteína de canales de cloruro, músculo esquelético, canal intracelular de cloruro 6, proteína 3 de canales intracelulares de cloruro, proteína 4 de canales intracelulares de cloruro, proteína 5 de canales intracelulares de cloruro, proteína CHRNA3, proteína Clcn3e, proteína CLCNKB, proteína CNGA4, culina-5, canal de potasio regulado por GMP cíclico, canal de cationes regulado por nucleótido cíclico 4, canal de cationes regulado por nucleótido cíclico alfa 3, canal de cationes regulado por nucleótido cíclico beta 3, canal olfativo regulado por nucleótidos cíclicos, regulador de la conductancia transmembrana de fibrosis quística, cadena pesada de citocromo B-245, tipo L sensible a dihidropiridina, precursor de subunidades alfa-2/delta de canales de calcio, precursor del regulador 3 del transporte de iones que contiene dominio de FXYP, precursor del regulador 5 del transporte de iones que contiene dominio de FXYP, precursor del regulador 6 del transporte de iones que contiene dominio de FXYP, regulador 7 del transporte de iones que contiene dominio de FXYP, precursor del regulador 8 del transporte de iones que contiene dominio de FXYP, canal de potasio de rectificación interna activado por proteína G 1, canal de potasio de rectificación interna activado por proteína G 2, canal de potasio de rectificación interna activado por proteína G 3, canal de potasio de rectificación interna activado por proteína G 4, precursor de la subunidad alfa-1 de receptor de ácido gamma-aminobutírico, precursor de la subunidad alfa-2 de receptor de ácido gamma-aminobutírico, precursor de la subunidad alfa-3 de receptor de ácido gamma-aminobutírico, precursor de la subunidad alfa-4 de receptor de ácido gamma-aminobutírico, precursor de la subunidad alfa-5 de receptor de ácido gamma-aminobutírico, precursor de la subunidad alfa-6 de receptor de ácido gamma-aminobutírico, precursor de la subunidad beta-1 de receptor de ácido gamma-aminobutírico, precursor de la subunidad beta-2 de receptor de ácido gamma-aminobutírico, precursor de la subunidad beta-3 de receptor de ácido gamma-aminobutírico, precursor de la subunidad delta de receptor de ácido gamma-aminobutírico, precursor de la subunidad épsilon de receptor de ácido gamma-aminobutírico, precursor de la subunidad gamma-1 de receptor de ácido gamma-aminobutírico, precursor de la subunidad gamma-3 de receptor de ácido gamma-aminobutírico, precursor de la subunidad pi de receptor de ácido gamma-aminobutírico, precursor de la subunidad rho-1 de receptor de ácido gamma-aminobutírico, precursor de la subunidad rho-2 de receptor de ácido gamma-aminobutírico, precursor de la subunidad theta de receptor de ácido gamma-aminobutírico, receptor de

kainato GluR6, precursor del receptor 1 de glutamato, precursor del receptor 2 de glutamato, precursor del receptor 3 de glutamato, precursor del receptor 4 de glutamato, receptor 7 de glutamato, receptor B de glutamato, precursor de la subunidad delta-1 del receptor de glutamato, receptor de glutamato, precursor de kainato 1 ionotrópico, receptor de glutamato, precursor de kainato 2 ionotrópico, receptor de glutamato, precursor de kainato 3 ionotrópico, receptor de glutamato, precursor de kainato 4 ionotrópico, receptor de glutamato, precursor de kainato 5 ionotrópico, precursor de la subunidad 3A del receptor de glutamato [NMDA], precursor de la subunidad 3B del receptor de glutamato [NMDA], precursor de la subunidad épsilon 1 del receptor de glutamato [NMDA], precursor de la subunidad épsilon 2 del receptor de glutamato [NMDA], precursor de la subunidad épsilon 4 del receptor de glutamato [NMDA], precursor de la subunidad zeta I del receptor de glutamato [NMDA], precursor de la cadena alfa-1 de receptor de glicina, precursor de la cadena alfa-2 de receptor de glicina, precursor de la cadena alfa-3 de receptor de glicina, precursor de la cadena beta de receptor de glicina, subunidad 1 del complejo de ribonucleoproteína H/ACA, subunidad beta del receptor épsilon de inmunoglobulina de alta afinidad, proteína hipotética DKFZp313I0334, proteína hipotética DKFZp761M1724, proteína hipotética FLJ12242, proteína hipotética FLJ14389, proteína hipotética FLJ14798, proteína hipotética FLJ14995, proteína hipotética FLJ16180, proteína hipotética FLJ16802, proteína hipotética FLJ32069, proteína hipotética FLJ37401, proteína hipotética FLJ38750, proteína hipotética FLJ40162, proteína hipotética FLJ41415, proteína hipotética FLJ90576, proteína hipotética FLJ90590, proteína hipotética FLJ90622, proteína hipotética KCTD15, proteína hipotética MGC15619, receptor tipo 1 de 1,4,5-trisfosfato de inositol, receptor tipo 2 de 1,4,5-trisfosfato de inositol, receptor tipo 3 de 1,4,5-trisfosfato de inositol, proteína 4 de canales de potasio activados por calcio de conductancia intermedia, canal de potasio de rectificación interna 13, canal de potasio de rectificación interna 16, canal de potasio de rectificación interna 4, regulador Kir2.2v negativo de canales de K(+) de rectificación interna, subunidad KA2a del receptor de kainato, proteína KCNH5, proteína KCTD 17, proteína KCTD2, proteína transmembrana asociada a queratinocitos 1, proteína 4 de interacción con canales Kv, melastatina 1, proteína MLC1 de membrana, proteína MGC15619, mucolipina-1, mucolipina-2, mucolipina-3, proteína 4 asociada a resistencia a múltiples fármacos, precursor de la subunidad 2C del receptor de N-metil-D-aspartato, homólogo de NADPH oxidasa 1, proteína de receptor de acetilcolina neuronal Nav1.5, precursor de la subunidad alfa-10, proteína del receptor de acetilcolina neuronal, precursor de la subunidad alfa-2, proteína del receptor de acetilcolina neuronal, precursor de la subunidad alfa-3, proteína del receptor de acetilcolina neuronal, precursor de la subunidad alfa-4, proteína del receptor de acetilcolina neuronal, precursor de la subunidad alfa-5, proteína del receptor de acetilcolina neuronal, precursor de la subunidad alfa-6, proteína del receptor de acetilcolina neuronal, precursor de la subunidad alfa-7, proteína del receptor de acetilcolina neuronal, precursor de la subunidad alfa-9, proteína del receptor de acetilcolina neuronal, precursor de la subunidad beta-2, proteína del receptor de acetilcolina neuronal, precursor de la subunidad beta-3, proteína del receptor de acetilcolina neuronal, precursor de la subunidad beta-4, subunidad alfa 2D de canales de calcio dependientes de voltaje neuronales, receptor purinérgico P2X 1, receptor purinérgico P2X 2, receptor purinérgico P2X 3, receptor purinérgico P2X 4, receptor purinérgico P2X 5, receptor purinérgico P2X 6, receptor purinérgico P2X 7, canal de potasio pancreático TALK-1b, canal de potasio pancreático TALK-1c, canal de potasio pancreático TALK-1d, precursor de Phospholemman, plasmolipina, proteína relacionada con 2 de enfermedad del riñón poliquístico, proteína 1 similar a 2 de enfermedad del riñón poliquístico, proteína 2 similar a 2 de enfermedad del riñón poliquístico, enfermedad del riñón poliquístico y receptor para precursor de la proteína relacionada con la gelatina del huevo, policistina-2, regulador de los canales de potasio, miembro 1 de la subfamilia K de canales de potasio, miembro 10 de la subfamilia K de canales de potasio, miembro 12 de la subfamilia K de canales de potasio, miembro 13 de la subfamilia K de canales de potasio, miembro 15 de la subfamilia K de canales de potasio, miembro 16 de la subfamilia K de canales de potasio, miembro 17 de la subfamilia K de canales de potasio, miembro 2 de la subfamilia K de canales de potasio, miembro 3 de la subfamilia K de canales de potasio, miembro 4 de la subfamilia K de canales de potasio, miembro 5 de la subfamilia K de canales de potasio, miembro 6 de la subfamilia K de canales de potasio, miembro 7 de la subfamilia K de canales de potasio, miembro 9 de la subfamilia K de canales de potasio, 3 que contiene dominios de tetramerización de canales de potasio, proteína 12 que contiene dominios de tetramerización de canales de potasio, proteína 14 que contiene dominios de tetramerización de canales de potasio, proteína 2 que contiene dominios de tetramerización de canales de potasio, proteína 4 que contiene dominios de tetramerización de canales de potasio, proteína 5 que contiene dominios de tetramerización de canales de potasio, 10 que contiene dominios de tetramerización de canales de potasio, proteína 13 que contiene dominios de tetramerización de canales de potasio, 1 que contiene dominios de tetramerización de canales de potasio, miembro 1 de la subfamilia A de canales regulados por voltaje de potasio, miembro 2 de la subfamilia A de canales regulados por voltaje de potasio, miembro 4 de la subfamilia A de canales regulados por voltaje de potasio, miembro 5 de la subfamilia A de canales regulados por voltaje de potasio, miembro 6 de la subfamilia A de canales regulados por voltaje de potasio, miembro 1 de la subfamilia B de canales regulados por voltaje de potasio, miembro 2 de la subfamilia B de canales regulados por voltaje de potasio, miembro 1 de la subfamilia C de canales regulados por voltaje de potasio, miembro 3 de la subfamilia C de canales regulados por voltaje de potasio, miembro 4 de la subfamilia C de canales regulados por voltaje de potasio, miembro 1 de la subfamilia D de canales regulados por voltaje de potasio, miembro 2 de la subfamilia D de canales regulados

por voltaje de potasio, miembro 3 de la subfamilia D de canales regulados por voltaje de potasio, miembro 1 de la subfamilia E de canales regulados por voltaje de potasio, miembro 2 de la subfamilia E de canales regulados por voltaje de potasio, miembro 3 de la subfamilia E de canales regulados por voltaje de potasio, miembro 4 de la subfamilia E de canales regulados por voltaje de potasio, miembro 1 de la subfamilia F de canales regulados por voltaje de potasio, miembro 1 de la subfamilia G de canales regulados por voltaje de potasio, miembro 2 de la subfamilia G de canales regulados por voltaje de potasio, miembro 3 de de la subfamilia G de canales regulados por voltaje de potasio, miembro 4 de la subfamilia G de canales regulados por voltaje de potasio, miembro 1 de la subfamilia H de canales regulados por voltaje de potasio, miembro 2 de la subfamilia H de canales regulados por voltaje de potasio, miembro 3 de la subfamilia H de canales regulados por voltaje de potasio, miembro 4 de la subfamilia H de canales regulados por voltaje de potasio, miembro 5 de la subfamilia H de canales regulados por voltaje de potasio, miembro 6 de la subfamilia H de canales regulados por voltaje de potasio, miembro 7 de la subfamilia H de canales regulados por voltaje de potasio, miembro 8 de la subfamilia H de canales regulados por voltaje de potasio, miembro 1 de la subfamilia KQT de canales regulados por voltaje de potasio, miembro 2 de la subfamilia KQT de canales regulados por voltaje de potasio, miembro 3 de la subfamilia KQT de canales regulados por voltaje de potasio, miembro 4 de la subfamilia KQT de canales regulados por voltaje de potasio, miembro 5 de la subfamilia KQT de canales regulados por voltaje de potasio, miembro 1 de la subfamilia S de canales regulados por voltaje de potasio, miembro 2 de la subfamilia S de canales regulados por voltaje de potasio, miembro 3 de la subfamilia S de canales regulados por voltaje de potasio, miembro 2 de la subfamilia V de canales regulados por voltaje de potasio, canales regulados por voltaje de potasio, subfamilia H, miembro 7, isoforma 2, canal regulado por nucleótidos cíclicos activados por hiperpolarización de potasio/sodio 1, canal regulado por nucleótidos cíclicos activados por hiperpolarización de potasio/sodio 2, canal regulado por nucleótidos cíclicos activados por hiperpolarización de potasio/sodio 3, canal regulado por nucleótidos cíclicos activados por hiperpolarización de potasio/sodio 4, homólogos de TOM40 de la subunidad de receptores de importación mitocondrial probable, receptor purinérgico, isoforma A, canal de iones seleccionados por voltaje de 4 repeticiones putativas, proteína 7 de canales de cloruro putativos, receptor de kainato GluR6 putativo, variante 1 de la proteína CATSPER2 de canales de iones putativos, variante 2 de la proteína CATSPER2 de canales de iones putativos, variante 3 de la proteína CATSPER2 de canales de iones putativos, regulador putativo de la variante 1 de proteína de canales de potasio, proteína tirosina fosfatasa TPTE putativa, receptor 1 de rianodina, receptor 2 de rianodina, receptor 3 de rianodina, proteína 1 de unión a SH3KBP1, canal 1 de potencial receptor transitorio corto, canal 4 de potencial receptor transitorio corto, canal 5 de potencial receptor transitorio corto, canal 6 de potencial receptor transitorio corto, canal 7 de potencial receptor transitorio corto, proteína 1 de canales de potasio activados por calcio de conductancia pequeña, proteína 2 de canales de potasio activados por calcio de conductancia pequeña, isoforma b, proteína 3 de canales de potasio activados por calcio de conductancia pequeña, isoforma b, canal de potasio activado por calcio de conductancia pequeña SK2, canal de potasio activado por calcio de conductancia pequeña SK3, canal de sodio, precursor de la subunidad beta-1 de canales de sodio, subunidad alfa de proteína tipo II de canales de sodio, subunidad alfa de proteína tipo III de canales de sodio, subunidad alfa de proteína tipo IV de canales de sodio, subunidad alfa de proteína tipo IX de canales de sodio, subunidad alfa de proteína tipo V de canales de sodio, subunidad alfa de proteína tipo VI de canales de sodio, subunidad alfa de proteína tipo VIII de canales de sodio, subunidad alfa de proteína tipo X de canales de sodio, subunidad alfa de proteína tipo XI de canales de sodio, canal de potasio sensible al ATP activado por sodio y cloruro, cadena gamma de ATPasa transportadora de sodio/potasio, canal de cationes asociado a espermatozoides 1, canal de cationes asociado a espermatozoides 2, isoforma 4, sintaxina-1 B1, miembro 1 de la subfamilia A de canales de cationes de potencial receptor transitorio, miembro 2 de la subfamilia M de canales de cationes de potencial receptor transitorio, miembro 3 de la subfamilia M de canales de cationes de potencial receptor transitorio, miembro 6 de la subfamilia M de canales de cationes de potencial receptor transitorio, miembro 7 de la subfamilia M de canales de cationes de potencial receptor transitorio, miembro 1 de la subfamilia V de canales de cationes de potencial receptor transitorio, miembro 2 de la subfamilia V de canales de cationes de potencial receptor transitorio, miembro 3 de la subfamilia V de canales de cationes de potencial receptor transitorio, miembro 4 de la subfamilia V de canales de cationes de potencial receptor transitorio, miembro 5 de la subfamilia V de canales de cationes de potencial receptor transitorio, miembro 6 de la subfamilia A de canales de cationes de potencial receptor transitorio, variante de corte y empalme épsilon de canales de potencial receptor transitorio 4, variante de corte y empalme zeta de canales de potencial receptor transitorio 4, variante de corte y empalme gamma de canales de potencial receptor transitorio 7, factor de necrosis tumoral, proteína 1 alfa-inducida, endotelial, proteína 2 de canales de calcio de dos poros, proteína VDAC4, canal de potasio regulado por voltaje Kv3.2b, subunidad beta1B de canales de sodio regulados por voltaje, canales de aniones dependientes de voltaje, canales de aniones dependientes de voltaje 2, proteína 1 de canales selectivos de aniones dependientes de voltaje, proteína 2 de canales selectivos de aniones dependientes de voltaje, proteína 3 de canales selectivos de aniones dependientes de voltaje, subunidad gamma-1 de canales de calcio dependientes de voltaje, subunidad gamma-2 de canales de calcio dependientes de voltaje, subunidad gamma-3 de canales de calcio dependientes de voltaje, subunidad gamma-4 de canales de calcio dependientes de voltaje, subunidad gamma-5 de canales de calcio

- dependientes de voltaje, subunidad gamma-6 de canales de calcio dependientes de voltaje, subunidad gamma-7 de canales de calcio dependientes de voltaje, subunidad gamma-8 de canales de calcio dependientes de voltaje, subunidad alfa-1C de canales de calcio tipo L dependientes de voltaje, subunidad alfa-1D de canales de calcio tipo L dependientes de voltaje, subunidad alfa-1S de canales de calcio tipo L dependientes de voltaje, subunidad beta-1 de canales de calcio tipo L dependientes de voltaje, subunidad beta-2 de canales de calcio tipo L dependientes de voltaje, subunidad beta-3 de canales de calcio tipo L dependientes de voltaje, subunidad beta-4 de canales de calcio tipo L dependientes de voltaje, subunidad alfa-1B de canales de calcio tipo N dependientes de voltaje, subunidad alfa-1A de canales de calcio tipo P/G dependientes de voltaje, subunidad alfa-1E de canales de calcio tipo R dependientes de voltaje, subunidad alfa-1G de canales de calcio tipo T dependientes de voltaje, subunidad alfa-1H de canales de calcio tipo L dependientes de voltaje, subunidad alfa-1I de canales de calcio tipo T dependientes de voltaje, subunidad alfa-1 de canales de calcio tipo L dependientes de voltaje, subunidad beta-1 de canales de potasio dependientes de voltaje, subunidad beta-2 de canales de potasio dependientes de voltaje, subunidad beta-3 de canales de potasio dependientes de voltaje, canal de potasio dependiente de voltaje KCNA7.
- 15 **[0212]** GPCR a modo de ejemplo incluyen, pero no se limitan a, rodopsina de clase A como receptores tales como acetilcolina muscular de vertebrado tipo 1, acetilcolina muscular de vertebrado tipo 2, acetilcolina muscular de vertebrado tipo 3, acetilcolina muscular de vertebrado tipo 4; receptores adrenérgicos (receptores adrenérgicos alfa tipo 1, receptores adrenérgicos alfa tipo 2, receptores adrenérgicos beta tipo 1, receptores adrenérgicos beta tipo 2, receptores adrenérgicos beta tipo 3, dopamina de vertebrado tipo 1, dopamina de vertebrado tipo 2, dopamina de vertebrado tipo 3, dopamina de vertebrado tipo 4, histamina tipo 1, histamina tipo 2, histamina tipo 3, histamina tipo 4, serotonina tipo 1, serotonina tipo 2, serotonina tipo 3, serotonina tipo 4, serotonina tipo 5, serotonina tipo 6, serotonina tipo 7, serotonina tipo 8, otros tipos de serotonina, amina Trace, angiotensina tipo 1, angiotensina tipo 2, bombesina, bradiquinina, anafilatoxina C5a, Fmet-leu-phe, similar a APJ, interleucina-8 tipo A, interleucina-8 tipo B, otros tipos de interleucina-8, quimiocina C-C tipo 1 a tipo 11 y otros tipos, quimiocina C-X-C (tipos 2 a 6 y otros), quimiocina C-X3-C, colecistoquinina CCK, CCK tipo A, CCK tipo B, otras CCK, endotelina, melanocortina (hormona estimulante de melanocitos, hormona adrenocorticotrópica, hormona melanocortina), antígeno Duffy, péptido liberador de prolactina (GPR10), neuropéptido Y (tipo 1 a 7), neuropéptido Y, otros neuropéptidos Y, neurotensina, opioide (tipo D, K, M, X), somatostatina (tipo 1 a 5), taquiquinina (sustancia P (NK1), sustancia K (NK2), neuromedina K (NK3), 1 similar a taquiquinina, 2 similar a taquiquinina, vasopresina / vasotocina (tipo 1 a 2), vasotocina, oxitocina / mesotocina, conopresina, similar a galanina, similar a activada con proteinasa, orexina y neuropéptidos FF, QRFP, similar a receptor de quimiocina, similar a neuromedina U (neuromedina U, PRXamida), proteína hormonal (hormona estimulante del folículo, hormona lutropina-coriogonadotrópica, tirotropina, gonadotropina tipo I, gonadotropina tipo II), (rod)opsina, rodopsina de vertebrado (tipos 1-5), rodopsina de vertebrado tipo 5, rodopsina de artrópodo, rodopsina de artrópodo tipo 1, rodopsina de artrópodo tipo 2, rodopsina de artrópodo tipo 3, rodopsina de molusco, rodopsina, olfativa (olfativa II fam 1 a 13), prostaglandina (prostaglandina E2 subtipo EP1, prostaglandina E2/D2 subtipo EP2, prostaglandina E2 subtipo EP3, prostaglandina E2 subtipo EP4, prostaglandina F2-alfa, prostaciclina, tromboxano, adenosina tipo 1 a 3, receptores purinérgicos, receptor purinérgico P2RY1-4, 6, 11 GPR91, receptor purinérgico P2RY5, 8, 9, 10 GPR35, 92, 174, receptor purinérgico P2RY12-14 GPR87 (UDP-Glucosa), cannabinoide, factor de activación de plaquetas, hormona liberadora de gonadotropina, hormona liberadora de gonadotropina tipo I, hormona liberadora de gonadotropina tipo II, similar a hormona adipocinética, corazonina, hormona liberadora de tirotropina y secretagogo, hormona liberadora de tirotropina, secretagogo de hormona de crecimiento, similar a secretagogo de hormona de crecimiento, hormona desencadenante de ecdisis (ETHR), melatonina, lisoesfingolípido y LPA (EDG), esfingosina 1-fosfato Edg-1, ácido lisofosfatídico Edg-2, esfingosina 1-fosfato Edg-3, ácido lisofosfatídico Edg-4, esfingosina 1-fosfato Edg-5, esfingosina 1-fosfato Edg-6, ácido lisofosfatídico Edg-7, esfingosina 1-fosfato Edg-8, Edg. Otro receptor de leucotrieno B4, receptor BLT1 de leucotrieno B4, receptor BLT2 de leucotrieno B4, orfano de clase A/otro, neurotransmisores putativos, SREB, proto-oncogén mas y relacionado con mas (MRG), similar a GPR45, cisteinil-leucotrieno, receptor de ácido biliar acoplado a proteína G, receptor de ácidos grasos libres (GP40, GP41, GP43), similar a secretina de clase B, calcitonina, factor liberador de corticotropina, péptido inhibidor gástrico, glucagón, hormona liberadora de hormona de crecimiento, hormona paratiroidea, PACAP, secretina, polipéptido intestinal vasoactivo, latrofilina, latrofilina tipo 1, latrofilina tipo 2, latrofilina tipo 3, receptores de ETL, inhibidor de angiogénesis específica para cerebro (BAI), proteínas similares a Methuselah (MTH), cadherina EGF LAG (CELSR), receptor acoplado a proteína G muy grande, glutamato metabotrópico de clase C/feromona, glutamato metabotrópico grupo I a III, similar a sensibilización al calcio, sensibilización al calcio extracelular, feromona, otra similar a sensibilización al calcio, receptores de feromonas putativas, GABA-B, GABA-B subtipo 1, GABA-B subtipo 2, similar a GABA-B, orfano GPRC5, orfano GPCR6, novia de proteínas sevenless (BOSS), receptores de Taste (T1R), feromonas fúngicas de clase D, similar a factor de feromonas fúngicas A (STE2, STE3), similar a feromona B fúngica (BAR, BBR, RCB, PRA), receptores de cAMP de clase E, proteínas del albinismo ocular, familia de Frizzled/Smoothened, frizzled grupo A (Fz 1 y 2 y 4 y 5 y 7-9), frizzled grupo B (Fz 3 y 6), frizzled grupo C (otros), receptores vomeronasales,

quimiorreceptores de nematodos, receptores olorosos de insectos y opsinas arqueales/bacterianas/fúngicas de clase Z.

5 **[0213]** Las MURP de la invención pueden diseñarse para elegir como diana cualquier proteína celular que incluye, pero no se limita a, proteína de la superficie celular, proteína secretada, proteína citosólica y proteína nuclear. Una diana de particular interés es un canal de iones.

10 **[0214]** Los canales de iones constituyen una superfamilia de proteínas, que incluyen la familia de canales de potasio (canales de K), la familia de canales de sodio (canales de Na), la familia de canales de calcio (canales de Ca), la familia de canales de cloro (canales de Cl) y la familia de canales de acetilcolina. Cada una de estas familias contiene subfamilias y cada subfamilia normalmente contiene canales específicos derivados de genes individuales. Por ejemplo, la familia de los canales de K contiene subfamilias de canales de potasio regulados por voltaje llamados Kv1.x y Kv3.x. La subfamilia Kv 1.x contiene los canales Kv1.1, Kv1.2 y Kv1.3, que se corresponden con los productos de genes individuales y así se llaman 'especies'. La clasificación se aplica a las familias de canales de 15 Na, Ca, Cl y a otras familias también.

20 **[0215]** Los canales de iones también pueden clasificarse según los mecanismos por los que operan los canales. Específicamente, los principales tipos de proteínas de canales de iones se caracterizan por el procedimiento empleado de abrir o cerrar la proteína del canal para tanto permitir como prevenir que iones específicos entren en la proteína del canal y atraviesen una membrana celular bicapa lipídica. Un tipo importante de proteína de canales es la proteína de canales regulados por voltaje que se abre o cierra (regula) en respuesta a cambios en el potencial eléctrico a través de la membrana celular. El canal de sodio regulado por voltaje 1.6 (Nav1.6) es de particular interés como diana terapéutica. Otro tipo de proteína de canales de iones es el canal mecánicamente regulado, para el cual una tensión mecánica sobre la proteína abre o cierra el canal. Todavía otro 25 tipo se llama canal regulado por ligando, que se abre o cierra dependiendo de si un ligando particular se une o no a la proteína. El ligando puede ser tanto un resto extracelular, tal como un neurotransmisor, como un resto intracelular, tal como un ión o nucleótido.

30 **[0216]** Los canales de iones generalmente permiten flujo pasivo de iones hacia un gradiente electroquímico, mientras que las bombas de iones usan ATP para transportar contra un gradiente. Los transportadores acoplados, tanto antiportadores como simportadores, permiten el movimiento de una especie de ión contra su gradiente, accionado por el movimiento descendente de otra especie de ión.

35 **[0217]** Uno de los tipos más comunes de proteínas de canales, encontradas en la membrana de casi todas las células animales, permite la permeación específica de iones potasio a través de una membrana celular. En particular, los iones potasio entran rápidamente a través de membranas celulares a proteínas de canales de K<sup>+</sup> (hasta 10<sup>8</sup> iones por segundo). Además, las proteínas de canales de potasio tienen la capacidad para distinguir entre iones potasio y otros iones alcalinos metálicos pequeños tales como Li<sup>+</sup> o Na<sup>+</sup> con gran fidelidad. En particular, los iones potasio son al menos diez mil veces más permanentes que los iones sodio. Las proteínas de los canales 40 de potasio normalmente comprenden cuatro subunidades (normalmente idénticas), así sus dianas de la superficie celular están presentes como tetrámeros, permitiendo la unión tetravalente de MURP. Un tipo de subunidad contiene seis segmentos hidrófobos de longitud (que puede atravesar la membrana), mientras que los otros tipos contienen dos segmentos hidrófobos.

45 **[0218]** Otra familia significativa de canales es el canal de calcio. Los canales de calcio se clasifican generalmente según sus propiedades electrofisiológicas como canales activados por bajo voltaje (LVA) o activados por alto voltaje (HVA). Los canales HVA comprenden al menos tres grupos de canales, conocidos como canales tipo L, N y P/Q. Estos canales se han diferenciado entre sí electrofisiológicamente, además de bioquímicamente, basándose en su farmacología y propiedades de unión a ligando. Por ejemplo, las dihidropiridinas, difenil- 50 alquilaminas y piperidinas se unen a la subunidad  $\alpha_1$  del canal de calcio tipo L y bloquean una proporción de corrientes de calcio HVA en tejido neuronal, que se llaman corrientes de calcio tipo L. Los canales de calcio tipo N son sensibles a conopeptidos omega, pero son relativamente insensibles a compuestos de dihidropiridina tales como nimodipina y nifedipina. Los canales tipo P/Q, por otra parte, son insensibles a dihidropiridinas, pero son sensibles a la toxina de la araña de tela en embudo Aga IIIA. Los canales de calcio tipo R, al igual que los canales tipo L, N, P y 55 Q, se activan por grandes despolarizaciones de membrana, y así se clasifican como canales activados por alto voltaje (HVA). Los canales de tipo R son generalmente insensibles a las dihidropiridinas y conopeptidos omega, pero, al igual que los canales P/Q, L y N, son sensibles a la toxina de la araña de tela en embudo AgaIVA. Estudios de tinción inmunocitoquímica indican que estos canales están localizados en todo el cerebro, particularmente en las estructuras de la línea media profunda (caudado-putamen, tálamo, hipotálamo, amígdala, cerebelo) y en los núcleos

del cerebro medio ventral y tronco encefálico. Los canales de calcio sensibles al voltaje neuronal normalmente consisten en una subunidad central  $\alpha_1$ , una subunidad  $\alpha_2/\delta$ , una subunidad  $\beta$  y una subunidad de 95 kD.

**[0219]** Ejemplos no limitantes adicionales incluyen Kir (un canal de potasio de rectificación interna), Kv (un canal de potasio regulado por voltaje), Nav (un canal de sodio regulado por voltaje), Cav (un canal de calcio regulado por voltaje), CNG (canal regulado por nucleótidos cíclicos), HCN (canal activado por hiperpolarización), TRP (un canal de potencial receptor transitorio), CIC (un canal de cloruro), CFTR (un regulador de la conductancia transmembrana de fibrosis quística, un canal de cloruro), IP3R (un receptor de trifosfato de inositol), RYR (un receptor de rianodina). Otros tipos de canales son canales de 2 poros, receptores de glutamato (AMPA, NMDA, KA), M2, conexinas y receptores de bucle Cys.

**[0220]** Un diseño común para proteínas de los canales de iones tales como Kv1.2, Kv3.1, Shaker, TRPC1 y TRPC5 es que tengan seis segmentos que atraviesan la membrana dispuestos del siguiente modo:

extremo N---S1---EI---S2---X1---S3---E2---S4---X2---S5---E3---S6---extremo C

**[0221]** En la que S1-6 son secuencias que atraviesan la membrana, EI-3 son bucles de la superficie extracelular y X1-2 son bucles de la superficie intracelular. El bucle E3 es generalmente el más largo de los tres bucles extracelulares y es hidrófilo de forma que es una buena diana para la unión de fármacos y MURP. La parte que forma los poros de la mayoría de los canales es un complejo multimérico (por ejemplo, tetramérico o raramente pentamérico) de hélices alfa que atraviesa la membrana. Generalmente hay un bucle de poros, que es una región de la proteína que retorna a la membrana para formar el filtro de selectividad que determina qué especie de ión puede entrar. Tales canales se llaman canales de 'poro-bucle'.

**[0222]** Los canales de iones son dianas valiosas para el diseño de fármacos debido a que participan en una amplia gama de procesos fisiológicos. En seres humanos existen aproximadamente más de trescientas proteínas de los canales de iones, muchas de las cuales participan en enfermedades genéticas. Por ejemplo, la expresión o función anómala de los canales de iones se ha mostrado que produce una amplia gama de enfermedades que incluyen enfermedades cardíacas, neuronales, musculares, respiratorias y metabólicas. Esta sección se basa en canales de iones, pero los mismos conceptos y enfoques son igualmente aplicables a todas las proteínas de la membrana, que incluyen 7TM, 1TM, proteínas G y receptores acoplados a proteína G (GPCR), etc. Algunos de los canales de iones son GPGR.

**[0223]** Los canales de iones normalmente forman grandes complejos macromoleculares que incluyen subunidades de proteínas accesorias estrechamente unidas y el uso combinatorio de tales subunidades contribuye a la diversidad de los canales de iones. Estas proteínas accesorias también puede ser las dianas de unión de las MURP de la invención, microproteínas y toxinas.

**[0224]** Las MURP de la invención pueden diseñarse para unirse a cualquiera de los canales conocidos en la técnica y a aquellos específicamente ejemplificados en el presente documento. Las MURP que presentan una capacidad de unión a canales de iones deseada (que engloba especificidad y avidéz) pueden seleccionarse por cualquier técnica recombinante y bioquímica (por ejemplo, expresión y expresión) conocida en la técnica. Por ejemplo, las MURP pueden expresarse por un paquete genético que incluye, pero no se limita a, fagos y esporas, y someterse a inmunopurificación contra membranas celulares intactas, o preferentemente células intactas tales como células de mamífero completas. Para eliminar el fago que se unen a las otras moléculas de la superficie celular no diana, el enfoque convencional era realizar inmunopurificación de sustracción contra líneas celulares similares que tenían un nivel bajo o no detectable del receptor diana. Sin embargo, Popkov y col. (J. Immunol. Methods 291:137-151 (2004)) mostraron que tipos de células relacionados no son ideales para la sustracción debido a que generalmente tienen un nivel reducido, pero todavía significativo, de la diana sobre su superficie, que reduce el número de clones de fago deseados. Este problema se produce incluso cuando se inmunopurifica sobre células que han sido transfectadas con el gen que codifica la diana, seguido de selección/sustracción negativa sobre la misma línea celular que no se transfectó, especialmente si el gen diana nativo no estaba inactivado. En su lugar, Popkov y col. mostraron que la selección o inmunopurificación por sustracción negativa funciona mucho mejor si se realiza con un exceso de las mismas células que se usan para inmunopurificación normal (selección positiva), excepto que la diana se ha bloqueado ahora con un inhibidor específico para diana de alta afinidad tal como una molécula pequeña, péptido o un anticuerpo para la diana, que hace que el sitio activo no esté disponible. Este procedimiento se llama "selección negativa con células enmascaradas por epítipo", que es particularmente útil en seleccionar MURP de la invención con una capacidad de unión a canales de iones deseada.

- [0225]** En una realización separada, la presente divulgación proporciona microproteínas, y particularmente microproteínas que presentan capacidad de unión hacia al menos una familia de canales de iones. La presente divulgación también proporciona un paquete genético que expresa tales microproteínas. Ejemplos no limitantes de canales de iones a los que las microproteínas objeto se unen son canales de sodio, potasio, calcio, acetilcolina y cloro. Son de particular interés aquellas microproteínas y los paquetes genéticos que expresan tales microproteínas, que presentan capacidad de unión hacia dianas nativas. Las dianas nativas son generalmente moléculas o fragmentos naturales, derivados de los mismos de los que se sabe que la microproteína se une, incluyendo normalmente aquellas dianas de unión conocidas de las que se ha informado en la bibliografía.
- 10 **[0226]** La presente divulgación también proporciona un paquete genético que expresa una microproteína de unión a canales de iones que ha sido modificada. La microproteína modificada puede (a) unirse a una familia diferente de canal con respecto a la microproteína sin modificar correspondiente; (b) unirse a una subfamilia diferente de la misma familia del canal con respecto a la microproteína sin modificar correspondiente; (c) unirse a una especie diferente de la misma subfamilia de canal con respecto a la microproteína sin modificar correspondiente;
- 15 (d) la microproteína se une a un sitio diferente sobre el mismo canal con respecto a la microproteína sin modificar correspondiente; y/o (e) unirse al mismo sitio del mismo canal, pero da un efecto biológico diferente con respecto a la microproteína sin modificar correspondiente.
- [0227]** La Figura 22 y 46 muestran cómo los dominios de microproteína o toxinas que cada uno se unen a diferentes sitios del mismo canal de iones pueden combinarse en una única proteína. Los dos sitios de unión en los que se unen estas dos microproteínas pueden estar sobre dos canales de familias diferentes, dos canales de la misma familia, pero una subfamilia diferente, dos canales de la misma subfamilia, pero una especie diferente (producto génico), o dos sitios de unión diferentes sobre el mismo canal (especie) o pueden unirse (simultáneamente o no) al mismo sitio de unión sobre el mismo canal (especie), ya que los canales son multiméricos. Los módulos y dominios de unión que se unen a sitios sobre los canales pueden ser dominios de microproteína (natural o no natural, que contienen 2 a 8 disulfuros), péptidos de un disulfuro, o péptidos lineales. Estos módulos pueden seleccionarse independientemente y combinarse, o pueden seleccionarse de una biblioteca para unirse en presencia de un módulo de unión activo fijo. En el último caso, la biblioteca de expresión expresaría múltiples módulos de que los que uno contendría una biblioteca de variantes. Un objetivo típico es seleccionar un dímero de esta biblioteca que tiene una mayor afinidad que el monómero activo que era el punto de partida.
- 20 **[0228]** En otra realización, la presente divulgación proporciona una proteína que comprende una pluralidad de dominios de unión a canales de iones, en la que los dominios individuales son dominios de microproteína que han sido modificados de forma que (a) los dominios de microproteína se unen a una familia diferente de canal con respecto a los dominios de microproteína sin modificar correspondientes; (b) los dominios de microproteína se unen a una subfamilia diferente de la misma familia del canal con respecto a los dominios de microproteína sin modificar correspondientes; (c) los dominios de microproteína se unen a una especie diferente de la misma subfamilia con respecto a los dominios de microproteína sin modificar correspondientes; (d) los dominios de microproteína se unen a un sitio diferente sobre el mismo canal con respecto a los dominios de microproteína sin modificar correspondientes; (e) los dominios de microproteína se unen al mismo sitio del mismo canal, pero dan un efecto biológico diferente con respecto a los dominios de microproteína sin modificar correspondientes; y/o (f) los dominios de microproteína se unen al mismo sitio del mismo canal y dan el mismo efecto biológico con respecto a los dominios de microproteína sin modificar correspondientes. Si se desea, los dominios de microproteína pueden comprender secuencias naturales o no naturales. Los dominios individuales pueden ligarse juntos mediante un ligador heterólogo. Los dominios individuales de microproteína pueden unirse a la misma familia de canales o diferente, la misma subfamilia de canales o diferente, la misma especie o diferente de la misma subfamilia, el mismo sitio o diferente en el mismo canal.
- 35 **[0229]** Las microproteínas objeto pueden ser una toxina. Preferentemente, la toxina retiene en parte o por completo su espectro de toxicidad. En particular, animales venenosos, tales como serpientes, tienen una gama de especies de presa e intrusas y las toxinas de veneno se diferencian en la actividad para los diferentes receptores de las diferentes especies. El veneno consiste en un gran número de toxinas relacionadas y sin relacionar, teniendo cada toxina un "espectro de actividad", que puede definirse como todos los receptores de todas las especies sobre las que esa toxina tiene actividad medible. Todas las dianas en el 'espectro de actividad' se consideran "dianas nativas" y esto incluye cualquier diana humana contra la que la toxina es activa. La(s) diana(s) nativa(s) de una microproteína o toxina incluyen todas las dianas de las que se informa en la bibliografía que la toxina inhibe. Cuanto mayor sea la afinidad o actividad sobre una diana, más probable será que esa diana sea la diana nativa natural, pero es común para toxinas que actúan sobre múltiples dianas dentro de la misma especie. La(s) diana(s) nativa(s) puede(n) ser receptor(es) humano(s) o no humano(s) contra los que la toxina es activa.
- 40 45 50 55

**[0230]** Para que la toxina retenga la capacidad para unirse a células después de la fusión con el vector de expresión, puede desearse probar tanto el extremo N como el extremo C para la fusión y probar una variedad de sitios de fusión (es decir, 0, 1, 2, 3, 4, 5, 6 aminoácidos antes de la primera cisteína o después de la última cisteína del dominio de toxina, si el dominio de toxina es un dominio que contiene cisteína) usando un enfoque de biblioteca de ADN sintético, que preferentemente codifica una biblioteca de ligadores ricos en glicina, que forman la cadena de aminoácidos más pequeña, no tienen carga y lo más probablemente es que sean compatibles con la unión de la toxina a la diana. Como el grupo amino del extremo N y los grupos carboxilo del extremo C pueden participar en la unión a diana, la biblioteca debe contener una lisina o una arginina para imitar el grupo amino positivamente cargado (o fusiones con el extremo N de la toxina) y un glutamato o un aspartato para imitar el grupo carboxilo negativamente cargado (para fusiones con el extremo C de la toxina).

**[0231]** El (Los) inhibidor(s) que se usan para bloquear la diana durante la selección negativa pueden ser moléculas pequeñas, péptidos o proteínas, y naturales o no naturales. Además de la simple sustracción, la elección de la mezcla de inhibidores es una herramienta valiosa para controlar la especificidad de los inhibidores de los canales de iones que están siendo diseñados. Debido a que hay más de trescientos canales de iones en total, con especificidades parcialmente solapantes y similitudes de secuencia, y múltiples sitios moduladores por canal, teniendo cada uno un efecto diferente, el requisito de especificidad puede ser complejo.

**[0232]** Si se modifica la actividad de una toxina, o si se combinan dos toxinas diferentes en una única proteína, las dos toxinas pueden unirse al mismo canal en el mismo sitio y tener el mismo efecto fisiológico, o las dos toxinas pueden unirse al mismo canal en el mismo sitio y tener un efecto fisiológico diferente, o las dos toxinas pueden unirse al mismo canal en un sitio diferente, o las dos toxinas pueden unirse a diferentes canales que pertenecen a la misma subfamilia (es decir, Kv1.3 y Kv1.2; que significa producto de un gen o 'especie' diferente), o las dos toxinas pueden unirse a diferentes canales que pertenecen a la misma familia (es decir, ambos son canales de K), o las dos toxinas pueden unirse a canales que pertenecen a familias diferentes (es decir, canales de K frente a canales de Na).

**[0233]** Los canales de iones tienen normalmente muchos segmentos transmembrana (24 para canales de sodio) y así ofrecen varios sitios de unión diferentes, no competitivos y no solapantes para que los moduladores alteren la actividad del canal de diferentes formas. Un enfoque es crear ligantes para un sitio sobre el mismo canal de iones de ligantes existentes para un sitio diferente, aunque estos sitios no estén relacionados. Para lograr esto, la toxina existente puede usarse como agente que elige diana para una biblioteca de proteínas de 1, 2, 3 ó 4 disulfuros que están separados de la toxina que elige diana por un ligador flexible de 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40 ó 50 aminoácidos. Es útil si la afinidad del agente que elige diana no es demasiado alta, de manera que la afinidad de la nueva biblioteca pueda tener una contribución significativa a la afinidad global. Otro enfoque es crear nuevos moduladores para canales de moduladores existentes para otros canales que están relacionados en secuencia o en estructura. La familia de las conotoxinas, por ejemplo, contiene moduladores relacionados con secuencia y relacionados con estructura para canales de Ca, K, Na y receptores de acetilcolina nicotínicos. Parece factible convertir un modulador de canales de K en un modulador de canales de Na usando una biblioteca de derivados de conotoxina, o viceversa. Por ejemplo, las conotoxinas kappa inhiben canales de K, las conotoxinas mu y las conotoxinas delta inhiben canales de Na, las conotoxinas omega inhiben canales de Ca y las conotoxinas alfa inhiben receptores de acetilcolina.

**[0234]** La proximidad de sitios de unión diferentes, cada uno con un efecto diferente sobre la actividad de los canales, del mismo canal de iones hace atractivo ligar los inhibidores usando ligadores flexibles, que crean un único inhibidor con dos dominios, cada unión en un sitio diferente, o una única proteína con dos dominios que se unen en copias diferentes del mismo sitio, dando una interacción de alta afinidad bivalente (avidez). Este enfoque no ha sido tomado por toxinas naturales, supuestamente debido a que deben actuar rápidamente y así permanecer pequeñas con el fin de tener máxima penetración en el tejido, pero para la farmacéutica la velocidad de acción es menos importante, haciendo esto un enfoque atractivo.

**[0235]** Así pueden crearse bibliotecas combinatorias de toxinas/moduladores diméricos, triméricos, tetraméricos o multiméricos, cada uno nativo o modificado, y cribar directamente estas bibliotecas al nivel de proteínas o inmunopurificar estas bibliotecas usando paquetes genéticos para afinidad mejorada (avidez, si se produce la unión simultáneamente en múltiples sitios) y luego caracterizar la especificidad y actividad de tales clones multiméricos por expresión de proteínas y purificación, seguido de ensayos de actividad basados en células, que incluyen ensayos de fijación de parches. Los módulos individuales pueden inmunopurificarse y seleccionarse por separado, en aislamiento entre sí, o pueden diseñarse en presencia de los otros, de forma que el nuevo dominio se

añade a un sistema de expresión como una biblioteca que también contiene una copia activa fija que sirve de elemento que elige diana para la biblioteca y solo clones que son significativamente mejores que el monómero activo fijado se seleccionan y se caracterizan.

5 **[0236]** Las Figs. 46 y 47 muestran algunos de los derivados monoméricos que pueden prepararse a partir de proteínas nativas (naturales), y algunos de los multímeros que pueden prepararse para unirse a múltiples sitios de unión diferentes de la diana. Los ligadores se muestran como PEGr rico en glicina, pero los ligadores podrían ser cualquier secuencia y también podrían optimizarse usando bibliotecas moleculares, seguido de inmunopurificación. Pueden crearse bibliotecas dentro de la propia toxina nativa activa usando una variedad de estrategias de mutagénesis como se ha descrito anteriormente, o pueden extender el área existente de contacto con la diana creando bibliotecas sobre el lado del extremo N o extremo C de las toxinas activas, esperando crear contactos adicionales con la diana. Tales bibliotecas pueden basarse en toxinas existentes con actividad conocida para ese sitio, o pueden ser o bibliotecas de 1, 2, 3, 4 disulfuros sin tratamiento previo basándose en andamiajes de microproteína sin relacionar. Estos elementos de contacto adicionales pueden añadirse a uno o ambos lados de los dominios activos, y pueden ser directamente adyacentes al dominio modulador existente o pueden separarse del mismo por ligadores flexibles. El multímero inicial o el multímero mejorado final puede ser un homomultímero o un heteromultímero, basándose en similitud de secuencias de los dominios o basándose en la especificidad diana de los dominios del multímero. Así, los monómeros que comprenden el multímero pueden unirse a los mismos sitios diana, pero tener las mismas secuencias o diferentes. Con 10-100 toxinas nativas diferentes que son conocidas por unirse a cada familia de canales, y con 2, 3, 4, 5 ó 6 dominios por clon, pueden crearse las bibliotecas de expresión con una enorme diversidad combinatoria, aunque solo una usa secuencias de toxina nativa. La mutagénesis sintética de bajo nivel basada en similitud de aminoácidos o en tasas de sustitución filogenética dentro de la familia puede usarse para crear bibliotecas de alta calidad de mutantes, de las cuales se espera que una fracción muy alta retenga la función, con una alta probabilidad de función potenciada en algunas de las propiedades de interés.

25 **[0237]** La capacidad de unión de las MURP de la invención, microproteínas o toxinas a un canal de iones dado puede medirse en términos del coeficiente de Hill. El coeficiente de Hill indica la estequiometría de la interacción de unión. Un coeficiente de Hill de 2 indica que 2 inhibidores se unen a cada canal. También puede evaluarse la modulación alostérica, que es la modulación de la actividad en un sitio producida por la unión en un sitio remoto.

**[0238]** La actividad biológica o efecto de un canal de iones y la capacidad de las MURP de la invención, microproteínas o toxinas para regular una actividad de canal de iones puede evaluarse usando una variedad de ensayos *in vitro* e *in vivo*. Por ejemplo, en la técnica están disponibles procedimientos para medir voltaje, medir corriente, medir potencial de membrana, medir flujo de iones, por ejemplo, potasio o rubidio, medir concentración de iones, medir regulación, medir segundos mensajeros y niveles de transcripción, y usar, por ejemplo, colorantes sensibles al voltaje, trazadores radiactivos y electrofisiología por fijación de parches. En particular, tales ensayos pueden usarse para probar microproteínas y toxinas que pueden inhibir o activar un canal de iones de interés.

40 **[0239]** Específicamente, pueden probarse inhibidores o activadores de canales de potencial en comparación con un control adecuado para examinar el grado de modulación. Las muestras de control también puede ser muestras sin tratar con los activadores o inhibidores candidatos. La inhibición está presente cuando un valor de actividad de canales de iones dado con respecto al control es aproximadamente el 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, o incluso menos. La  $CI_{50}$  es una unidad comúnmente usada (la concentración de inhibidor que reduce la actividad del canal de iones al 50%) para determinar el efecto inhibidor. Similar a  $CI_{90}$ . La activación de canales se logra cuando un valor de actividad de canal de iones dado con respecto al control aumenta el 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 200%, 500%, o más.

50 **[0240]** Pueden evaluarse cambios en el flujo de iones determinando cambios en la polarización (es decir, potencial eléctrico) de la célula o membrana que expresa el canal de interés. Por ejemplo, un procedimiento es determinar cambios en la polarización celular midiendo cambios en la corriente (midiendo así cambios en la polarización) con técnicas de fijación de voltaje y fijación de parches, por ejemplo, el modo "unido a célula", el modo "al revés" y el modo "célula completa" (véase, por ejemplo, Ackerman y col., *New Engl. J. Med.* 336:1575-1595 (1997)). Las corrientes de célula completa se determinan convenientemente usando la metodología convencional (véase, por ejemplo, Hamil y col., *Pflugers. Archiv.* 391:85 (1981)). Otros ensayos conocidos incluyen: ensayos de flujo de rubidio radiomacado y ensayos de fluorescencia usando colorantes sensibles al voltaje (véase, por ejemplo, Vestergarrd-Bogind y col., *J. Membrane Biol.* 88:67-75 (1988); Daniel y col., *J. Pharmacol. Meth.* 25:185-193 (1991); Holevinsky y col., *J. Membrane Biology* 137:59-70 (1994)).

**[0241]** Los efectos de las MURP candidatas, microproteínas o toxinas sobre la función de un canal de interés pueden medirse por cambios en las corrientes eléctricas o flujo iónico o por las consecuencias de los cambios en las corrientes y flujo. Puede variarse el efecto aguas abajo de las proteínas candidatas sobre el flujo de iones. Por consiguiente, cualquier cambio fisiológico adecuado puede usarse para evaluar la influencia de una proteína candidata sobre los canales de prueba. Los efectos de la proteína candidata pueden medirse por un ensayo de unión a toxina. Si las consecuencias funcionales se determinan usando células o animales intactos, también puede medirse una variedad de efectos tales como liberación de transmisor (por ejemplo, dopamina), liberación de hormona (por ejemplo, insulina), cambios transcripcionales a tanto marcadores genéticos conocidos como no caracterizados (por ejemplo, transferencias Northern), cambios del volumen de células (por ejemplo, en glóbulos rojos), inmunorrespuestas (por ejemplo, activación de linfocitos T), cambios en el metabolismo de las células tal como crecimiento celular o cambios de pH, y cambios en los segundos mensajeros intracelulares tales como  $Ca^{2+}$ .

**[0242]** Otras actividades biológicas clave de los canales de iones son selectividad por iones y regulación. La selectividad es la capacidad de algunos canales para discriminar entre especies de iones, permitiendo que algunos pasen al poro mientras que excluyen otros. La regulación es la transición entre estados abiertos y cerrados. Puede evaluarse por cualquiera de los procedimientos conocidos en la técnica o desvelados en el presente documento

**[0243]** Todavía otra propiedad biológica por la que la MURP de la invención, microproteína o toxina puede seleccionarse es la frecuencia de apertura y cierre de los canales diana, llamado frecuencia de regulación. La frecuencia de regulación está influida por el voltaje (en canales regulados por voltaje, que se abren o se cierran por cambios en el voltaje de la membrana) y la unión a ligando. La tasa de transición entre estados abiertos y cerrados es normalmente <10 microsegundos, pero puede aumentarse o disminuirse por otras moléculas. La tasa de flujo (corriente) en el poro cuando se abre es del orden de  $10e^7$  iones por segundo para canales de iones y mucho menos para intercambiadores acoplados. Tras la apertura, algunos canales regulados por voltaje entran en un estado no conductor inactivado en el que son resistentes a la despolarización.

## EJEMPLOS

### Ejemplo: Diseño de un oligómero de glicina-serina basado en secuencias humanas

**[0244]** Se buscó en la base de datos del genoma humano secuencias que eran ricas en glicina. Se identificaron tres secuencias como secuencias de donante adecuadas como se muestra en la Tabla X.

Tabla X: Secuencias de donante para el diseño A de GRS.

Acceso	Secuencias	Aminoácido	Proteína
NP_009060	GGGSGGGSGSGGGG	486-499	proteína de dedo de cinc
Q9Y2X9	GSGSGGGGSGG	19-31	proteína de dedo de cinc
CAG38801	SGGGSGGGSGSG	7-19	MAP2K4

**[0245]** Basándose en las secuencias en la Tabla X, los presentes inventores diseñaron una secuencia rica en glicina que contenía múltiples repeticiones del péptido A con la secuencia GGGSGSGGGGS. El péptido A puede oligomerizarse para formar estructuras con la fórmula  $(GGGSGSGGGGS)_n$  en la que n está entre 2 y 40. La Figura 5 muestra que todas las posibles subsecuencias 9-meras en oligómeros del péptido A están contenidas en al menos una de las proteínas enumeradas en la Tabla 3. Así, los oligómeros del péptido A no contienen epítomos de linfocitos T humanos. La inspección de la Figura 5 revela que GRS basada en oligómeros de péptido A puede empezar y terminar en cualquiera de las posiciones del péptido A.

### 45 Ejemplo: Diseño de un oligómero de glicina-prolina basado en secuencias humanas

**[0246]** Se diseñaron secuencias ricas en glicina basándose en la secuencia GPGGGGPGGGGGPGGGGPGGGGGGPGGGGGGPGGGGGGPGGG que representa los aminoácidos 146-182 del dominio POU de clase 4 humano con número de acceso NP\_006228. La Figura 6 ilustra que los oligómeros del péptido B con la secuencia GGGGGPGGGGP pueden utilizarse como GRS. Todas las subsecuencias 9-meras que están contenidas en péptidos con la secuencia  $(GGGGPGGGGP)_n$  también están contenidas en la secuencia del dominio POU. Así, tales secuencias oligoméricas no contienen epítomos de linfocitos T.

### Ejemplo: Diseño de oligómero de glicina-ácido glutámico

55

**[0247]** Pueden diseñarse secuencias ricas en glicina basándose en la subsecuencia GAGGEGGGEGGGPGG que es parte de la proteína S6 cinasa ribosómica (número de acceso BAD92170). Por ejemplo, los oligómeros del péptido C con la secuencia GGGGE formarán secuencias en las que la mayoría de las subsecuencias 9-meras estarán contenidas en la secuencia de la proteína S6 cinasa ribosómica. Así, las GRS 5 oligoméricas de la estructura general (GGGGE)<sub>n</sub> tienen un riesgo muy bajo de contener epítomos de linfocitos T.

**Ejemplo: Identificación de secuencias ricas en glicina hidrófilas humanas**

**[0248]** Se buscó en una base de datos de proteínas humanas subsecuencias que eran ricas en residuos de glicina. Estas subsecuencias contuvieron al menos el 50% de glicina. Solo se dejó que se produjeran los siguientes 10 residuos de no glicina en la GRS: ADEHKPRST. Se identificaron 70 subsecuencias que tenían una longitud mínima de 20 aminoácidos. Estas subsecuencias se enumeran en el Apéndice A. Pueden utilizarse para construir GRS con bajo potencial inmunogénico en seres humanos.

**15 Ejemplo: Construcción de rPEG J288**

**[0249]** El siguiente ejemplo describe la construcción de un gen de codones optimizados que codifica una secuencia de URP con 288 aminoácidos y la secuencia (GSGGEG)<sub>48</sub>. En primer lugar, los presentes inventores 20 construyeron un vector de relleno pCW0051 como se ilustra en la figura 40. La secuencia del casete de expresión en pCW0051 se muestra en la figura 42. El vector de relleno se basó en un vector pET e incluyó un promotor T7. El vector codifica una secuencia Flag, seguida de una secuencia de relleno que está flanqueada por sitios Bsal, BbsI y KpnI. Los sitios Bsal y BbsI se insertaron de forma que generaron residuos protuberantes compatibles después de la digestión tal como se ilustra en la figura 42. La secuencia de relleno fue seguida de una etiqueta de His<sub>6</sub> y el gen de proteína verde fluorescente (GFP). La secuencia de relleno contiene codones de terminación y así células de *E. coli* 25 que llevan el plásmido de relleno pCW0051 formaron colonias no fluorescentes. El vector de relleno pCW0051 se digirió con Bsal y KpnI. Se construyó una biblioteca de codones que codifica secuencias de URP de 36 aminoácidos de longitud como se muestra en la figura 41. La secuencia de URP se designó como rPEG\_136 y tenía la secuencia de aminoácidos (GSGGEG)<sub>6</sub>. El inserto se obtuvo hibridando pares de oligonucleótidos sintéticos que codifican la secuencia de aminoácidos GSGGEGGSGGEG, además de una pareja de oligonucleótidos que codifican un 30 adaptador con el sitio KpnI. Se usaron los siguientes oligonucleótidos: pr\_LCW0057for: AGGTAGTGGWGGWGARGGGWGGWTCYGGWGGAGAAGG, pr\_LCW0057rev:

**[0250]** ACCTCCTTCTCCWCCRGAWCCWCCYTCWCCWCCACT, pr\_3KpnIstopperFor: 35 AGGTTTCGTTCTTCACTCGAGGGTAC, pr\_3KpnIstopperRev: CCTCGAGTGAAGACGA. Los pares de oligonucleótidos hibridados se ligaron, que produjo una mezcla de productos con longitud variable que representa el número variable de repeticiones de rPEG\_J12. El producto correspondiente a la longitud de rPEG\_J36 se aisló de la mezcla por electroforesis en gel de agarosa y se ligó al vector de relleno digerido con Bsal/KpnI pCW0051. La mayoría de los clones en la biblioteca resultante designada LCW0057 mostró fluorescencia verde después de la inducción, que muestra que la secuencia de rPEG\_J36 se había ligado en marco con el gen GFP. El procedimiento 40 de cribado y multimerización iterativa de secuencias de rPEG\_J36 se ilustra en la Fig. 14. Los presentes inventores cribaron 288 cepas aisladas de la biblioteca LCW0057 para alto nivel de fluorescencia. Las 48 cepas aisladas con fuerte fluorescencia se analizaron por PCR para verificar la longitud del segmento de rPEG\_J y se identificó que 16 clones tenían la longitud esperada de rPEG\_J36. Este procedimiento produjo una colección de 16 cepas aisladas de rPEG\_J36, que muestran alta expresión y que se diferencian en su uso de codones. Las cepas aisladas se 45 reunieron y se dimerizaron usando un procedimiento explicado resumidamente en la Fig. 40. Una mezcla de plásmidos se digirió con Bsal/NcoI y un fragmento que comprendía la secuencia rPEG\_J36 y una parte de GFP se aisló. La misma mezcla de plásmidos también se digirió con BbsI/NcoI y el fragmento de vector que comprende rPEG\_J36, la mayor parte del vector de plásmido, y el resto del gen GFP se aisló. Ambos fragmentos se mezclaron, se ligaron y se transformaron en BL21 y se cribaron cepas aisladas para fluorescencia. Este procedimiento 50 de dimerización se repitió dos rondas más como se ha explicado resumidamente en la Fig. 14. Durante cada ronda, los presentes inventores duplicaron la longitud del gen rPEG\_J y por último lugar obtuvieron una colección de genes que codificaban rPEG\_J288. La secuencia de aminoácidos y de nucleótidos de rPEG\_J288 se muestra en la Fig. 15. Puede observarse que el módulo de rPEG\_J288 contiene segmentos de rPEG\_J36 que se diferencian en su secuencia de nucleótidos a pesar de tener secuencia de aminoácidos idéntica. Así, los presentes inventores 55 minimizaron la homología interna en el gen y como resultado los presentes inventores redujeron el riesgo de recombinación espontánea. Los presentes inventores cultivaron BL21 de *E. coli* que alojaba plásmidos que codificaban rPEG\_J288 para al menos 20 duplicados y no se observó recombinación espontánea.

**Ejemplo: Construcción de rPEG H288**

**[0251]** Se construyó una biblioteca de genes que codifica un URP de 288 aminoácidos llamado rPEG\_H288 usando el mismo procedimiento que se usó para construir rPEG\_J288. rPEG\_H288 tiene la secuencia de aminoácidos (GSGGEGGSGGSG)<sub>24</sub>. El diagrama de flujo del procedimiento de construcción se muestra en la Fig. 14. La secuencia de aminoácidos completa, además de la secuencia de nucleótidos de una cepa aislada de rPEG\_H288, se facilitan en la Fig. 16.

**Ejemplo: Estabilidad en suero de rPEG J288**

10 **[0252]** Una proteína de fusión que contiene una marca Flag en el extremo N y la secuencia de URP rPEG\_J288 fusionada con el extremo N de la proteína verde fluorescente se incubó en 50% de suero de ratón a 37°C durante 3 días. Las muestras se extrajeron en diversos momentos de tiempo y se analizaron por SDS-PAGE, seguido de detección usando análisis Western. Se usó un anticuerpo contra la marca Flag del extremo N para detección por Western. Los resultados se muestran en la Fig. 28, que indica que una secuencia de URP de 288  
15 aminoácidos puede ser completamente estable en suero durante al menos tres días.

**Ejemplo: Ausencia de anticuerpos preexistentes para rPEG J288 en suero**

20 **[0253]** La existencia de anticuerpos contra URP sería una indicación de una posible respuesta inmunogénica a esta secuencia rica en glicina. Para probar la presencia de anticuerpos existentes en suero, una fusión URP-GFP se sometió a un ELISA inmovilizando URP-GFP sobre un soporte y posteriormente incubando con 30% de suero. La presencia de anticuerpos unidos a URP-GFP se detectó usando un anticuerpo anti-IgG-peroxidasa de rábano picante y sustrato. Los datos se muestran en la Fig. 29. Los datos muestran que la proteína de fusión puede detectarse por anticuerpos contra GFP o Flag, pero no por suero murino. Esto indica que el suero murino no  
25 contiene anticuerpos que contienen la secuencia de URP.

**Ejemplo: Purificación de una proteína de fusión que contiene rPEG J288**

30 **[0254]** Los presentes inventores purificaron una proteína con la arquitectura Flag-rPEG\_J288-H6-GFP. La proteína se expresó en BL21 de *E. coli* en medio SB. Los cultivos se indujeron con IPTG 0,5 mM durante la noche a 18 °C. Las células se recogieron por centrifugación. El sedimento se resuspendió en tampón TBS que contenía benzonasa y una mezcla de inhibidor de proteasa comercial. La suspensión se calentó durante 10 min a 75 °C en un baño de agua para lisar las células. El material insoluble se eliminó por centrifugación. El sobrenadante se purificó usando especificidad por ión metálico inmovilizado (IMAC), seguido de una columna con anticuerpo anti-Flag  
35 inmovilizado. La Fig. 43 muestra el análisis de PAGE del procedimiento de purificación. El procedimiento dio proteína con al menos el 90% de pureza.

**Ejemplo: Construcción de proteína de fusión entre rPEG J288 e interferón-alfa**

40 **[0255]** Se diseñó un gen que codifica interferón alfa humano usando la optimización de codones para expresión de *E. coli*. El gen sintético se fusionó con un gen que codificaba rPEG\_J288. Se dispuso una marca de His6 en el extremo N para facilitar la detección y purificación de la proteína de fusión. La secuencia de aminoácidos de la proteína de fusión se facilita en la Fig. 44.

45 **Ejemplo: Construcción de la fusión de rPEG J288-G-CSF**

**[0256]** Se diseñó un gen que codifica G-CSF humano usando la optimización de codones para expresión de *E. coli*. El gen sintético se fusionó con un gen que codificaba rPEG\_J288. Se dispuso una marca de His6 en el extremo N para facilitar la detección y purificación de la proteína de fusión. La secuencia de aminoácidos de la  
50 proteína de fusión se facilita en la Fig. 44.

**Ejemplo: Construcción de la fusión de rPEG J288-hGH**

**[0257]** Se diseñó un gen que codifica hormona de crecimiento humana usando la optimización de codones para expresión de *E. coli*. El gen sintético se fusionó con un gen que codificaba rPEG\_J288. Se dispuso una marca de His6 en el extremo N para facilitar la detección y purificación de la proteína de fusión. La secuencia de aminoácidos de la proteína de fusión se facilita en la Fig. 44.

**Ejemplo: Expresión de proteínas de fusión entre rPEG J288 y proteínas humanas**

**[0258]** Las proteínas de fusión entre rPEG\_J288 y dos proteínas humanas, interferón-alfa y hormona de crecimiento humana, se clonaron en un vector de expresión T7 y se transformaron en BL21 de *E. coli*. Las células se cultivaron a 37 °C a una densidad óptica de DO 0,5. Posteriormente, las células se cultivaron a 18 °C durante 30 min. Luego se añadió IPTG 0,5 mM y los cultivos se incubaron en una estufa de incubación con agitación a 18 °C durante la noche. Las células se recogieron por centrifugación y la proteína soluble se liberó usando BugBuster (Novagen). Ambas fracciones de proteína, insoluble e soluble, se separaron por SDS-PAGE y las proteínas de fusión se detectaron por Western usando un anticuerpo contra la marca His6 del extremo N para la detección. La Fig. 45 muestra el análisis Western de las dos proteínas de fusión, además de rPEG\_J288-GFP como control. Todas las proteínas de fusión se expresaron y la mayoría de la proteína estuvo en la fracción soluble. Esto es una prueba de la alta solubilidad de rPEG\_J288 debido a que la mayoría de los intentos en la expresión del interferón-alfa y hormona de crecimiento humana en el citosol de *E. coli*, que se han informado en la bibliografía, produjeron la formación de cuerpos de inclusión insolubles. La Fig. 45 muestra que la mayoría de las proteínas de fusión se expresan como proteínas de longitud completa, es decir, no se detectó ningún fragmento que sugiriera síntesis incompleta o degradación de la proteína parcial.

**Ejemplo: Construcción y unión de un multímero de VEGF**

**[0259]** Se construyeron bibliotecas de péptidos restringidos a cisteína como se ha publicado [Scholle, M. D. y col. (2005) Comb Chem High Throughput Screen, 8: 545-51]. Estas bibliotecas se inmunopurificaron contra VEGF humano y se identificaron dos módulos de unión constituidos por secuencias de aminoácidos FTCTNHWCP5 o FQCTRHWCPI. Los oligonucleótidos que codifican la secuencia de aminoácidos FTCTNHWCP5 se ligaron a una secuencia de nucleótidos que codifica la secuencia de URP rPEG\_A36 con la secuencia (GGG)<sub>12</sub>. Posteriormente, la secuencia de fusión se dimerizó usando enzimas de restricción y etapas de ligación para construir una molécula que contenía 4 copias del módulo de unión de VEGF separado por rPEG\_A36 fusionado con GFP. La afinidad de unión por VEGF de proteínas de fusión que contienen entre cero y cuatro unidades de unión a VEGF se comparó en la Fig. 30. Una proteína de fusión que contiene solo rPEG\_A36 fusionado con GFP no muestra afinidad por VEGF. La adición de números crecientes de módulos de unión a VEGF aumenta la afinidad de las proteínas de fusión resultantes.

**Ejemplo: Descubrimiento de módulos de unión 1SS contra dianas terapéuticas**

**[0260]** Se generaron bibliotecas de péptidos aleatorias según Scholle y col. [Scholle, M. D. y col. (2005) Comb Chem High Throughput Screen, 8: 545-51]. Las bibliotecas de péptidos sin tratamiento previo expresaron péptidos restringidos a cisteína separados 4 a 10 residuos aleatorios. El diseño de la biblioteca se ilustra en la tabla:

**Tabla X:** Bibliotecas de 1SS sin tratamiento previo:

LNG0001	XXXCXXCXXX	X <sub>3</sub> CX <sub>2</sub> CX <sub>3</sub>	NNS NNS NNS TGC NNS NNS TGT NNS NNS NNS
LNG0002	XXCXXXCXXX	X <sub>2</sub> CX <sub>3</sub> CX <sub>3</sub>	NNS NNS TGC NNS NNS NNS TGT NNS NNS NNS
LNG0003	XXCXXXXCXX	X <sub>2</sub> CX <sub>4</sub> CX <sub>2</sub>	NNS NNS TGC NNS NNS NNS NNS TGT NNS NNS.
LNG0004	XCXXXXXCXX	X <sub>1</sub> CX <sub>5</sub> CX <sub>2</sub>	NNS TGC NNS NNS NNS NNS NNS TGT NNS NNS
LNG0005	XCXXXXXCX	X <sub>1</sub> CX <sub>6</sub> CX <sub>1</sub>	NNS TGC NNS NNS NNS NNS NNS NNS TGT NNS
LNG0006	CXXXXXCX	CX <sub>7</sub> CX <sub>1</sub>	TGC NNS NNS NNS NNS NNS NNS NNS TGT NNS
LNG0007	CXXXXXC	CX <sub>8</sub> C	TGC NNS NNS NNS NNS NNS NNS NNS NNS TGT
LNG0008	CXXXXXC	CX <sub>9</sub> C	TGC NNS NNS NNS NNS NNS NNS NNS NNS NNS
LNG0009	CXXXXXC	CX <sub>10</sub> C	TG TGC NNS NNS NNS NNS NNS NNS NNS NNS NN TGT
LNG0010	XXXXXCXXCXXXXX	X <sub>6</sub> CX <sub>2</sub> CX <sub>6</sub>	NNS NNS NNS NNS NNS NNS TGC NNS NNS TGT NN NNS NNS NNS NNS NNS
LNG0011	XXXXCXXCXXXXX	X <sub>5</sub> CX <sub>3</sub> CX <sub>6</sub>	NNS NNS NNS NNS NNS TGC NNS NNS NNS TGT NN: NNS NNS NNS NNS NNS
LNG0012	XXXXCXXCXXXXX	X <sub>5</sub> CX <sub>4</sub> CX <sub>5</sub>	NNS NNS NNS NNS NNS TGC NNS NNS NNS NNS TG NNS NNS NNS NNS NNS
LNG0013	XXXXCXXCXXXXX	X <sub>4</sub> CX <sub>5</sub> CX <sub>5</sub>	NNS NNS NNS NNS TGC NNS NNS NNS NNS NNS TG: NNS NNS NNS NNS NNS
LNG0014	XXXXCXXCXXXXX	X <sub>4</sub> CX <sub>6</sub> CX <sub>4</sub>	NNS NNS NNS NNS TGC NNS NNS NNS NNS NNS NN: TGT NNS NNS NNS NNS
LNG0015	XXXCXXXXXCXXXX	X <sub>3</sub> CX <sub>7</sub> CX <sub>4</sub>	NNS NNS NNS TGC NNS NNS NNS NNS NNS NN: TGT NNS NNS NNS NNS

LNG0016	XXXCXXXXXXXXXCXXX	X <sub>3</sub> CX <sub>8</sub> CX <sub>3</sub>	NNS NNS NNS TGC NNS NNS NNS NNS NNS NNS NN: NNS TGT NNS NNS NNS
LNG0017	XXCXXXXXXXXXCXXX	X <sub>2</sub> CX <sub>9</sub> CX <sub>3</sub>	NNS NNS TGC NNS NNS NNS NNS NNS NNS NNS NN: NNS TGT NNS NNS NNS
LNG0018	XXCXXXXXXXXXCXXX	X <sub>2</sub> CX <sub>10</sub> CX <sub>2</sub>	NNS NNS TGC NNS NNS NNS NNS NNS NNS NNS NN: NNS NNS TGT NNS NNS

**[0261]** Las bibliotecas se inmunopurificaron de nuevo contra una serie de dianas terapéuticamente relevantes usando el siguiente protocolo: pocillos sobre placas de ELISA inmunosorbentes se recubrieron con 5 µg/ml de antígeno diana en PBS durante la noche a 4 °C. Las placas recubiertas se lavaron con PBS, y los sitios no específicos se bloquearon con tampón de bloqueo (PBS que contenía tanto 0,5% de BSA como 0,5% de albúmina de huevo) durante 2 h a temperatura ambiente. Las placas se lavaron entonces con PBST (PBS que contiene 0,05% de Tween 20), y partículas de fago a 1-5x10<sup>12</sup>/ml en tampón de unión (tampón de bloqueo que contiene 0,05% de Tween 20) se añadieron a los pocillos y se incubaron con agitación durante 2 h a temperatura ambiente. Entonces, los pocillos se vaciaron y se lavaron con PBST. Las partículas de fago unidas se eluyeron de los pocillos por incubación con HCl 100 mM durante 10 min a temperatura ambiente, se transfirieron a tubos estériles y se neutralizaron con base TRIS 1 M. Para la infección, SS320 de *E. coli* de fase logarítmica que se cultivan en Super Broth complementado con 5 µg/ml de tetraciclina se añadieron al eluato de fago neutralizado, y el cultivo se incubó con agitación durante 30 min a 37 °C. Entonces, los cultivos infectados se transfirieron a tubos mayores que contenían Super Broth con 5 µg/ml de tetraciclina y los cultivos se incubaron con agitación durante la noche a 37 °C. Los cultivos durante la noche se limpiaron de *E. coli* por centrifugación, y el fago precipitó en el sobrenadante tras la adición de una disolución de 20% de PEG y NaCl 2,5 M a una concentración de PEG final del 4%. El fago precipitado se recogió por centrifugación, y el sedimento de fago se resuspendió en 1 ml de PBS, se limpió de *E. coli* residual por centrifugación y se transfirió a un tubo nuevo. Las concentraciones de fago se estimaron espectrofotométricamente y el fago se utilizó para la siguiente ronda de selección. Los clones individuales se cribaron para afinidad de unión a diana después de 3 ó 4 rondas de inmunopurificación en fago. Las placas individuales de los clones de fago seleccionadas durante la inmunopurificación se recogieron en Super Broth que contenía 5 µg/ml de tetraciclina y se cultivaron durante la noche con agitación a 37 °C. Se prepararon placas de ELISA por recubrimiento de antígeno y proteínas de control (BSA, albúmina de huevo, IgG) a 3 µg/ml en PBS durante la noche a 4 °C. Las placas se lavaron con PBS y se bloquearon con tampón de bloqueo (PBS que contiene 0,5% de BSA) durante 2 h a temperatura ambiente. Los cultivos durante la noche se limpiaron de *E. coli* por centrifugación y el sobrenadante se diluyó 1:10 en tampón de unión (tampón de bloqueo que contiene 0,05% de Tween 20) y se transfirió a placas de ELISA después de lavar con PBST (PBS que contenía 0,05% de Tween 20). Las placas se incubaron con agitación durante 2 h a temperatura ambiente. Tras el lavado con PBST, anti-M13-HRP (Pharmacia), dilución 1:5000 en PBS, se añadió a pocillos. Las placas se incubaron con agitación durante 30 min a temperatura ambiente y se lavaron con PBST, seguido de PBS. Una disolución de sustrato que contiene 0,4 mg/ml de ABTS y 0,001% de H<sub>2</sub>O<sub>2</sub> en tampón fosfato-citrato 50 mM se añadió a los pocillos, y se dejó que se revelaran durante 40 min después de que las placas se leyeron en un lector de placas a 405 nm. Estas lecturas de ELISA permitieron la determinación de la especificidad por clon, y clones específicos para antígeno se secuenciaron comercialmente mediante procedimientos establecidos.

35 **Tabla X:** Secuencias de módulos de unión específica para EpCAM

			S	Y	I	C	H	N	C	L	L	S					sNG0017S3.021
			L	R	C	W	G	M	L	C	Y	A					sNG0017S3.017
			L	R	C	I	G	Q	I	C	W	R					sNG0017S3.022
			L	K	C	L	Y	N	I	C	W	V					sNG0017S3.024
R	P	G	M	A	C	S	G	Q	L	C	W	L	N	S	P		sNG0018S3.015
P	H	A	L	Q	C	Y	G	S	L	C	W	P	S	H	L		sNG0018S3.018
R	A	G	I	T	C	H	G	H	L	C	W	P	I	T	D		sNG0018S3.019
R	P	A	L	K	C	I	G	T	L	C	S	L	A	N	P		sNG0015S3.014
P	H	G	L	W	C	H	G	S	L	C	H	Y	P	L	A		sNG0018S3.012
P	H	G	L	I	C	A	G	S	I	C	F	W	P	P	P		sNG0018S3.007
P	R	N	L	T	C	Y	G	Q	I	C	F	Q	S	Q	H		sNG0018S3.011
P	H	N	L	A	C	Q	N	S	I	C	V	R	L	P	R		sNG0018S3.021
P	H	G	L	T	C	T	N	Q	I	C	F	Y	G	N	T		sNG0018S3.006
			L	F	C		G	N	V	C	H	F					sNG0017S3.006
			L	T	C	W	G	Q	V	C	F	R					sNC0017S3.009
			R	C	P	S	R	V	P	W	C	V					sNG0017S3.011
Q	L	V	C	G	F	S	D	S	S	R		C	Y	M	R		sNG0018S3.009

L	L	C	Y	I	T	S	P	G	N	R	L	C	S	P	Y	sNG0018S3.022
<b>Tabla X:</b> Secuencias de módulos de unión específica para VEGF																
A	P	F	W	E	C	T	Q	H	W	C	P	S	L	Q	T	sNG0025S3.021
T	P	Y	F	S	C	S	F	G	F	C	R	D	S	F	S	sNG0026S3.035
N	P	F	F	Y	C	V	A	G	K	C	V	D	A	P	L	sNG0026S3.045
D	M	R	F	L	C	R	H	G	K	C	H	D	L	P	L	sNG0026S3.029
P	P	F	F	V	C	S	L	G	K	C	R	D	A	H	L	sNG0026S3.034
P	P	Q	F	Q	C	V	R	G	K	C	F	D	L	T	F	sNG0026S3.043
I	S	T	F	F	C	S	N	G	S	C	V	D	V	P	A	sNG0026S3.053
P	P	H	F	R	C	F	N	G	S	C	V	D	L	S	R	sNG0026S3.006
N	V	H	F	W	C	H	N	H	K	C	H	D	L	V	S	sNG0026S3.051
L	F	F	K	C	D	V	G	H	G	C	Y	D	I	K	H	sNG0026S3.040
L	Y	F	Q	C	F	P	N	R	G	C	S	T	L	Q	P	sNG0026S3.038
P	S	F	F	C	S	P	L	L	G	C	R	D	S	L	S	sNG0026S3.002
G	T	P	R	C	N	P	F	R	Q	F	C	A	I	P	S	sNG0026S3.052
			L	C	L	P	L	G	R	W	C	P				sNG0026S3.032
T	S	P	A	C	N	P	F	R	H	F	C	T	L	P	T	sNG0025S3.016
Q	P	P	I	C	N	P	F	R	Q	L	C	G	I	P	L	sNG0026S3.058
V	H	T	F	C	N	P	F	R	Q	M	C	S	L	P	M	sNG0026S3.046
R	M	V	N	C	N	P	F	N	S	W	C	S	L	P	S	sNG0026S3.027
S	K	H	M	C	N	P	F	H	S	W	C	G	V	P	L	sNG0026S3.001
R	W	P	V	C	N	P	F	L	G	Y	C	G	I	P	N	sNG0026S3.047
S	K	P	T	C	N	V	F	N	S	W	C	S	V	P	L	sNG0026S3.056
R	P	P	A	C	N	L	F	L	S	W	C	S	Y	D	S	sNG0026S3.059
G	R	S	V	C	N	P	Y	K	S	W	C	P	V	R	Q	sNG0026S3.004
A	S	S	C	K	D	S	P	H	F	R	C	L	F	P	L	sNG0026S3.011
L	A	N	C	P	N	S	P	G	F	L	C	L	H	A	V	sNG0026S3.055
P	F	A	C	P	H	S	S	G	F	R	C	L	Y	N	I	sNG0026S3.024
S	F	T	C	S	L	F	P	S	P	H	C	T	T	L	R	sNG0026S3.005
L	R	L	C	T	Y	G	G	G	K	Y	D	C	S	S	T	sNG0026S3.054
G	S	Y	C	Q	Y	R	P	F	S	S	F	C	N	R	S	sNG0026S3.050
			C	Y	N	Q	V	L	G	R	A	C				sNG0026S3.048
P	H	C	R	Q	H	P	L	D	R	W	M	C	S	P	S	sNG0025S3.001
S	L	C	S	M	F	G	D	T	P	H	W	N	C	V	P	sNG0026S3.057
S	S	C	S	L	F	N	N	T	R	H	W	S	C	T	D	sNG0026S3.007
				L	F	N	N	T	R	H	W	S	C	T	D	sNG0026S3.008
<b>Tabla X:</b> Secuencias de módulos de unión específica para CD28																
T	T	A	Y	P	D	C	F	W	C	S	L	F	G	P	P	sNG0028S3.085
M	L	D	T	T	I	C	P	W	C	S	L	F	G	P	V	sNG0028S3.081
E	L	X	T	T	I	C	P	W	C	S	L	F	G	P	V	sNG0028S3.018
S	L	S	Q	Q	S	C	D	W	C	S	L	F	G	P	P	sNG0028S3.086
K	R	L	L	E	C	G	A	L	C	A	L	F	G	P	P	sNG0028S3.060
H	T	I	L	T	C	D	S	G	F	C	T	L	F	G	P	sNG0028S3.008
N	L	W	H	V	C	H	T	S	L	C	H	S	R	L	A	sNG0028S3.012
N	S	F	Y	L	C	H	S	S	V	C	G	Q	L	P	S	sNG0028S3.092
A	G	F	S	C	E	N	Y	F	F	C	P	P	K	N	L	sNG0028S3.082
S	W	C	T	V	F	G	N	H	D	P	S	C	N	S	R	sNG0028S3.016
			C	S	S	N	G	R	W	K	A	H	C			sNG0028S3.004
L	P	N	M	W	R	V	V	V	P	D	V	Y	D	R	R	sNG0028S3.076
				W	R	V	V	V	P	D	V	Y	D	R	R	sNG0028S3.068
<b>Tabla X:</b> Secuencias de módulos de unión específica para CD28																
K	H	Y	C	F	G	P	K	S	W	T	T	C	A	R	G	sNG0030S3.096
P	W	C	H	L	C	P	G	S	P	S	R	C	C	Q	P	sNG0030S3.091
P	E	S	K	L	I		E	E	D	L	N	G	D	V	S	sNG0030S3.042
<b>Tabla X:</b> Secuencias de módulos de unión específica para Tiel																
I	W	D	R	V	C	R	M	N	T	C	H	Q	H	S	H	sNG0032S3.096
P	Y	T	I	F	C	L	H	S	S	C	R	s	S	S	S	sNG0032S3.087
D	W	C	L	T	G	P	N	T	L	S	F	c	P	R	R	sNG0032S3.031
<b>Tabla X:</b> Secuencias de módulos de unión específica para DR4																

L	S	T	W	C	L	H	D	V	C	W	P	P	L	K	sNG0033S3.072	
<b>Tabla X:</b> Secuencias de módulos de unión específica para DR5																
V	Y	L	T	Q	C	G	A	Q	L	C	L	K	R	T	N	sNG0034S3.039
P	Y	L	T	S	C	G	D	R	V	C	L	K	R	P	P	sNG0034S3.001
P	V	L	S	R	C	G	G	R	I	C	M	H	D	R	L	sNG0034S3.026
L	K	L	T	P	C	S	H	G	V	C	M	H	R	L	R	sNG0034S3.087
Y	Y	L	T	N	C	P	K	G	H	C	L	R	R	V	D	sNG0034S3.080

<b>Tabla X:</b> Secuencias de módulos de unión específica para DR5																
L	Y	L	H	S	C	S	R	G	I	C	L	S	P	R	V	sNG0034S3.082
F	S	C	Q	S	S	F	P	G	R	R	M	C	E	L	R	sNG0034S3.040
H	R	C	S	A	H	G	S	S	S	S	F	C	P	G	S	sNG0034S3.029

<b>Tabla X:</b> Secuencias de módulos de unión específica para TrkA																
K	T	W	D	C	N	S	G	H	C	V	J	T	F	K	sNG0035S3.074	
A	T	W	D	C	R	D	H	N	F	S	C	V	R	L	S	sNG0035S3.089

**Ejemplo: Conjugados de fármaco de aEpCAM**

- 5 **[0262]** Se aislaron péptidos anti-EpCAM de bibliotecas de péptidos aleatorias que se generaron según Scholle y col. [Scholle, M. D. y col. (2005) Comb Chem High Throughput Screen, 8: 545-51]. Las bibliotecas de péptidos sin tratamiento previo expresaron péptidos restringidos a cisteína con cisteínas separadas 4 a 10 residuos aleatorios. Después de tres rondas de selección por afinidad con las bibliotecas anteriores se aislaron varios ligandos de péptidos específicos para EpCAM (EpCam1) (Tabla X). Las cepas aisladas de EpCam1 tienen una separación de
- 10 cisteínas conservadas de cuatro aminoácidos (CXXXXC). Los ligando de péptido EpCam1 se aleatorizaron entonces suavemente (excepto posiciones de cisteína) con codones que codificaban 3-9 residuos y se movieron en un vector de fagémido. Las bibliotecas de fagémido se seleccionaron posteriormente por afinidad contra EpCAM para aislar ligandos de péptido optimizados para la unión (Tabla X, EpCam2). Los ligandos EpCam2 contienen la separación de cisteínas CXXXXC conservada. Además, la mayoría de las secuencias anti-EpCam no contienen un residuo de
- 15 lisina, que permite la conjugación con grupos amina libres fuera de las secuencias de unión. Además, ligandos de péptido anti-EpCam pueden fusionarse genéticamente con secuencias de URP (de cualquier longitud) y multimerizarse usando dimerización iterativa. Las MURP anti-EpCAM resultantes pueden usarse para elegir específicamente como diana EpCAM con afinidad elevada con respecto a secuencias de monómero. Un ejemplo de una secuencia de aminoácidos de EpCAM-URP de tetrámero se muestra en la Fig. 31. Esta secuencia solo contiene
- 20 dos residuos de lisina que están localizados en la marca Flag del extremo N. Las cadenas laterales de estos residuos de lisina son particularmente adecuadas para la conjugación de fármacos.

**Tabla X.** Secuencias anti-EpCam

Nombre	Secuencia
EpCam 1	LRCWGMLCYA
	LRCIGQICWR
	LKCLYNICWV
	LFCWGNVCHF
	LTCWGQVCFR
	RPGMACSGQLCWLNSP
	PHALQCYGSLCWPSHL
	RAGITCHGHLCWPITD
	RPALKCIGTLCSLANP
	PHGLWCHGSLCHYPLA
	PHGLICAGSICFWPPP
	PRNLTCYGQICFQSQH
	PHNLACQNSICVRLPR
	PHGLTCTNQICFYGNT
EpCam 2	<b>HSLTCYGQICWVSNL</b>
	<b>PTLTTCYNQVCWVNRT</b>
	<b>PALRCLGQLCWVTPT</b>
	<b>PGLRCLGTLCWVPCR</b>
	<b>RNLTCWNTVCYAYPN</b>
<b>RGLKCLGQLCWVSSN</b>	

**PTLKCSGQICWVPPP**  
 RNLECLGNVCSLLNQ  
**PTLTCLNNLCWVPPQ**  
**RGLKCSGHLWCWVTPQ**  
**HGLTCHNTVCWVHHP**  
**HTLECLGNICWVINQ**  
 HGLTCYNQICWAPRP  
 HGLACYNQLCWVNP

RGLACQGNICWRLNP  
 RAITCLGTLCWPTSP  
 LTLECIIGNICYVPHH

**Ejemplo: Adición de secuencias aleatorias**

5 **[0263]** Los módulos de unión pueden madurarse por afinidad, o alargarse, mediante la adición de ligadores similares a URP y secuencia aleatoria para el extremo N, extremo C, o tanto el extremo N como C, de la secuencia de unión. La figura 32 muestra la adición de secuencias restringidas a cisteína sin tratamiento previo a un módulo de unión anti-EpCAM. Pueden generarse bibliotecas de adiciones de secuencias aleatorias usando un enfoque de clonación de ADN monocatenario o bicatenario. Una vez generadas, las bibliotecas pueden seleccionarse por  
 10 afinidad contra la proteína diana inicial o una segunda proteína. Por ejemplo, puede usarse una biblioteca de adición que contiene un módulo de unión anti-EpCAM para seleccionar secuencias que contienen 2 o más sitios de unión a la proteína diana.

**Ejemplo: Construcción de una biblioteca de formación 2SS**

15 **[0264]** Se diseñó una serie de oligonucleótidos para construir una biblioteca basada en el péptido I SS de unión a VEGF FTCTNHWCP. Los oligonucleótidos incorporan variaciones en los patrones de distancia de cisteína de las secuencias flanqueantes, mientras que la secuencia de péptidos de unión a VEGF se mantuvo fija.

20 Oligonucleótidos directos:

**[0265]** LMS70-1 CAGGCAGCGGGCCCGTCTGGCCCGT**GY**TTTACTTGTACGAATCATTGGTGTCC

**[0266]** LMS70-2 CAGGCAGCGGGCCCGTGTGGCCCGT**GYNNK**TTTAACTGTACGAATCATTGGTGTCC

25

**[0267]** LMS70-3  
 CAGGCAGCGGGCCCGTCTGGCCCGT**GYNNKNNK**TTTACTTGTACGAATCATTGGTGTCC

**[0268]** LMS70-4  
 30 CAGGCAGCGGGCCCGTCTGGCCCGT**GYNHTNHTNHT**TTTACTTGTACGAATCATTGGTGTCC

**[0269]** LMS70-5

**CAGGCAGCGGGCCCGTCTGGCCCGTGYNHTNHTNHTNHTTTTACTTGTACGAATCATTGGTGTCC**

35

**[0270]** LMS70-6

**CAGGCAGCGGGCCCGTCTGGCCCGTGYKMTKMTKMTKMTKMTTTTACTTGTACGAATCAT**  
**TGGTGTCC**

40 Oligonucleótidos inversos (complementados inversos):

**[0271]** LMS70-1R ACCGGAACCACCAGACTGGCC**RCACGAAGGACACCAATGATTCGTACAA**

[0272] LMS70-2R ACCGGAACCACCAGACTGGCCRC**AMNN**CGAAGGACACCAATGATTTCGTACAA

[0273] LMS70-3R

5 ACCGGAACCACCAGACTGGCCRC**AMNNMNN**CGAAGGACACCAATGATTTCGTACAA

[0274] LMS70-4R

ACCGGAACCACCAGACTGGCCRC**AAADNADNDN**CGAAGGACACCAATGATTTCGTACAA

10 [0275] LMS70-5R

ACCGGAACCACCAGACTGGCCRC**AAADNADNADNDN**CGAAGGACACCAATGATTTCGTACAA

[0276] LMS70-6R

**ACCGGAACCACCAGACTGGCCRCAAKMAKMAKMAKMAKMC**  
**GAAGGACACCAATGATTTCGTACAA**

15

Diluciones de oligonucleótidos

[0277] Mezcla 1 (de mezclas madre 100 µM): 100 µl de 70-6, 33 µl de 70-5, 11 µl de 70-4, 3,66 µl de 70-3, 1,2 µl de 70-2, 0,4 µl de 70-1. Mezcla 2 (de mezclas madre 100 µM): 100 µl de 70-6R, 33 µl de 70-5R, 11 µl de 70-4R, 3,66 µl de 70-3R, 1,2 µl de 70-2R, 0,4 µl de 70-1R

Ensamblaje por PCR

25 [0278] 10,0 µl de oligonucleótido molde (5 µM), 10,0 µl de 10 X tampón, 2,0 dNTP (10 mM), 1,0 µl de ADNc polimerasa (Clontech), 77 µl de H<sub>2</sub>O destilada. Programa de PCR: 95°C 1 min, (9,5°C 15 s, 54 °C 30 s, 68 °C 15 s) x5, 68°C 1 min

Amplificación por PCR

30

[0279] Cebadores, 10,0 µl de mezcla ensamblada, 10,0 µl de 10 X tampón, 2,0 dNTP (10 mM), 10,0 µl de LIBPTF (5 µM), 10,0 µl de LIBPTR (5 µM), 1,0 µl de ADNc polimerasa (Clontech), 57 µl de H<sub>2</sub>O destilada. Programa de PCR: 95°C 1 min, (95°C 15 s, 54°C 30 s, 68°C 15 s) x25, 68°C 1 min. El producto se purificó mediante columna Y10 de Amicon. El producto ensamblado se digirió con Sfil y BstXI y se ligó en el vector de fagémido pMP003. La ligación se realizó durante la noche a 16 °C en una máquina de PCR MJ. La ligación se purificó entonces por precipitación con EtOH. Transformación en células ER2738 competentes frescas por electroporación.

[0280] La biblioteca resultante se inmunopurificó contra VEGF como se describe a continuación. Se identificaron varias cepas aisladas que mostraron unión mejorada a VEGF con respecto a la secuencia de partida de 1SS. Los datos de unión y expresión se muestran en la Fig. 38. Las secuencias y resultados del análisis Western de clones de formación se muestra en la Fig. 39.

Ejemplo: Inmunopurificación en fago de bibliotecas de formación

45 [0281] Primera ronda de inmunopurificación:

[0282] 1) Primera ronda, recubrir 4 pocillos por biblioteca a cribar. Recubrir el pocillo de una placa de ELISA de 96 pocillos Costar con 0,25 µg de antígeno VEGF<sub>121</sub> en 25 µl de PBS. Cubrir la placa con un sellante de placas. El recubrimiento puede realizarse durante la noche a 4 °C o durante 1 h a 37 °C.

50

[0283] 2) Después de sacudir la disolución de recubrimiento, bloquear el pocillo añadiendo 150 µl de PBS/1% DE BSA. Sellar e incubar durante 1 h a 37 °C.

[0284] 3) Después de sacudir la disolución de bloqueo, añadir 50 µl de fago recientemente preparado (véase el protocolo de reamplificación de bibliotecas) al pocillo. Solo para la primera ronda, añadir también 5 µl de 5% de Tween. Sellar la placa e incubar durante 2 h a 37 °C.

- [0285]** Mientras tanto, inocular 2 ml de medio SB más 2 µl de 5 mg/ml de tetraciclina con 2 µl de una preparación de células ER 2738 y permitir el crecimiento a 250 rpm y 37 °C durante 2,5 h. Cultivar 1 cultivo para cada biblioteca que se criba que incluye selecciones negativas. Tomar todas las precauciones para evitar una contaminación del cultivo con fago.
- 5
- [0286]** 4) Sacudir la disolución de fago, añadir 150 µl de PBS/0,5% de Tween al pocillo y pipetear 5 veces vigorosamente arriba y abajo. Esperar 5 min, sacudir y repetir esta etapa de lavado. En la primera ronda, lavar de este modo 5 veces, en la segunda ronda 10 veces, y en la tercera, cuarta y quinta ronda 15 veces.
- 10 **[0287]** 5) Después de sacudir la disolución de lavado final, añadir 50 µl de 10 mg/ml de tripsina en PBS recientemente preparada, sellar e incubar durante 30 min a 37 °C. Pipetear 10 veces vigorosamente arriba y abajo y transferir el eluato (4 x 50 µl en la primera ronda, 2 x 50 ml en la segunda ronda, 1 x 50 µl en las rondas posteriores) al cultivo de 2 ml de *E. coli* preparado e incubar a temperatura ambiente durante 15 min.
- 15 **[0288]** 6) Añadir 6 ml de medio SB precalentado, 1,6 µl de carbenicilina y 6 µl de 5 mg/ml de tetraciclina. Transferir el cultivo a un tubo de polipropileno de 50 ml.
- [0289]** 7) Agitar el cultivo de 8 ml a 250 rpm y 37 °C durante 1 h, añadir 2,4 µl de 100 mg/ml de carbenicilina y agitar durante una hora adicional a 250 rpm y 37 °C.
- 20 **[0290]** 8) Añadir 1 ml de fago auxiliar VCSM 13 y transferir a una botella de centrifuga de polipropileno de 500 ml. Añadir 91 ml de medio SB precalentado (37 °C) y 46 µl de 100 mg/ml de carbenicilina y 92 µl de 5 mg/ml de tetraciclina. Agitar el cultivo de 100 ml a 300 rpm y 37 °C durante 1 1/2 a 2 h.
- 25 **[0291]** 9) Añadir 140 µl de 50 mg/ml de kanamicina y continuar la agitación a 300 rpm y 37 °C durante la noche.
- [0292]** 10) Centrifugar a 4000 rpm durante 15 min a 4 °C. Transferir el sobrenadante a una botella de centrifuga de 500 ml limpia y añadir 25 ml de 20% de PEG-8000/NaCl 2,5 M. Guardar sobre hielo durante 30 min.
- 30 **[0293]** 11) Centrifugar a 9000 rpm durante 15 min a 4 °C. Desechar el sobrenadante, drenar invertido sobre una toalla de papel durante al menos 10 min y limpiar el líquido restante de la parte superior de la botella de centrifuga con una toalla de papel.
- 35 **[0294]** 12) Resuspender el sedimento de fago en 2 ml de tampón PBS/0,5% de BSA/0,5% de Tween pipeteando arriba y abajo a lo largo del lado de la botella de centrifuga y transferir a un tubo de microcentrifuga de 2 ml. Resuspender adicionalmente pipeteando arriba y abajo usando una punta de pipeta de 1 ml, centrifugar a velocidad completa en una microcentrifuga durante 1 min a 4 °C y pasar el sobrenadante a un filtro de 0,2 µm en un tubo de microcentrifuga de 2 ml estéril.
- 40 **[0295]** 13) Continuar la etapa 3) durante la siguiente ronda o guardar la preparación de fago a 4 °C. Puede añadirse azida de sodio al 0,02% (peso/volumen) para almacenamiento a largo plazo. Solo debe usarse fago recientemente preparado para cada ronda.
- 45 **[0296]** Segunda ronda de inmunopurificación
- [0297]** Segunda ronda, recubrir 2 pocillos por biblioteca a cribar. Recubrir el pocillo de una placa de ELISA de 96 pocillos Costar con 0,25 µg de antígeno VEGF<sub>121</sub> en 25 µl de PBS. Cubrir la placa con un sellante de placas. El recubrimiento puede realizarse durante la noche a 4 °C o durante 1 h a 37 °C.
- 50 **[0298]** Bloquear también 2 pocillos sin recubrir para cada biblioteca que va a usarse como control negativo para el cálculo de la relación de enriquecimiento.
- [0299]** Tercera ronda de inmunopurificación
- 55 **[0300]** Tercera ronda, recubrir 1 pocillo por biblioteca a cribar. Recubrir el pocillo de una placa de ELISA de 96 pocillos Costar con 0,25 µg de antígeno VEGF<sub>121</sub> en 25 µl de PBS. Cubrir la placa con un sellante de placas. El recubrimiento puede realizarse durante la noche a 4 °C o durante 1 h a 37 °C.

**[0301]** Bloquear también 1 pocillos sin recubrir para cada biblioteca que va a usarse como control negativo para el cálculo de la relación de enriquecimiento.

**Ejemplo: Inmunopurificación basada en disolución:**

5

**[0302]** 1. Biotinilar la proteína diana según el fabricante.

**[0303]** 2. Recubrir un total de 8 pocillos (por selección) con 1,0 µg de neutravidina (Pierce) en PBS e incubar durante la noche a 4°C.

10

**[0304]** 3. Bloquear los pocillos con SuperBlock (Pierce) durante 1 h a temp ambiente. Guardar la placa con tampón de bloqueo hasta que se necesite (en la etapa 6).

15

**[0305]** 4. Usar 100 nM de proteína diana biotinilada y añadir 10<sup>12</sup> fagos/ml (en PBST) para un volumen total de 100-200 µl usando SuperBlock más 0,05% de Tween 20.

**[0306]** 5. Voltear la mezcla de fago-diana a temp ambiente durante al menos 1 h.

20

**[0307]** 6. Diluir 100 µl de mezcla de fago-diana con 700 µl de SuperBlock, mezclar y añadir 100 µl a cada uno de 8 pocillos recubiertos con neutravidina (de la etapa 3).

**[0308]** 7. Incubar durante 5 min a temp ambiente.

25

**[0309]** 8. Lavar 8X con PBST.

**[0310]** 9. Eluir el fago con 100 µl de HCl 100 mM durante 10 min.

**[0311]** 10. Neutralizar añadiendo 10 µl de TRIS 1 M a pH=8,0.

30

**[0312]** 11. Infectar las células para siembra o amplificar el fago para una ronda posterior de inmunopurificación en disolución.

**Ejemplo: Cribado por ELISA de fago para clones positivos para VEGF**

35

**[0313]** 1) Añadir 0,5 ml de SB que contiene 50 µg/ml de carbenicilina a placa de 96 pocillos profunda. Recoger una colonia e inocular pocillos.

**[0314]** 2) Agitar la placa que contiene los cultivos bacterianos a 300 rpm durante la noche a 37 °C.

40

**[0315]** 3) Preparar 4 ng/µl de disolución de proteína diana en PBS. Añadir 25 µl (100 ng) de proteína a cada pocillo e incubar durante la noche a 4 °C.

**[0316]** 4) Sacudir las placas de ELISA recubiertas y lavar 2x con PBS. Añadir 150 µl/pocillo de PBS + 0,5% de BSA (tampón de bloqueo). Bloquear durante 1 h a TA.

45

**[0317]** 5) Centrifugar las gradillas de microtubos (3000 rpm; 20 min).

**[0318]** 6) Preparar tampón de unión (tampón de bloqueo + 0,5% de Tween 20). Tomar alícuotas de 135 µl de tampón de unión por pocillo en placa de 96 pocillos de baja unión a proteína.

50

**[0319]** 7) Sacudir los pocillos sobre placas de ELISA y lavar 2 veces con PBST (PBS + 0,5% de Tween 20).

**[0320]** 8) Diluir 15 µl de fago de cultivos durante la noche 1:10 en PBST, mezclar por pipeteado y transferir 30 µl a cada pocillo recubierto de proteína. Incubar 2 h a TA con agitación suave.

55

**[0321]** 9) Lavar las placas 6 veces con PBST.

**[0322]** 10) Añadir 50 µl de anti-M13-HRP 1:5000 en tampón de unión a los pocillos. Incubar 30 min con agitación suave a TA.

[0323] 11) Lavar las placas 4 veces con PBST, seguido de 2 veces con H<sub>2</sub>O.

[0324] 12) Preparar 6 ml de disolución de ABTS (5,88 ml de tampón citrato más 120 µl de ABTS y 2 µl de H<sub>2</sub>O<sub>2</sub>). Tomar alícuotas de 50 µl por pocillo sobre cada placa de ELISA.

[0325] 13) Incubar a TA y leer la DO a 405 nm usando un lector de placas de ELISA en momentos de tiempo apropiados dependiendo de la señal (hasta 1 h)

#### 10 **Ejemplo: Dimerización de módulos de unión**

[0326] Se crearon bibliotecas expresadas en fago de 10e<sup>9</sup> a 10e<sup>11</sup> péptidos cíclicos con 4, 5, 6, 7, 8, 9, 10, 11 y 12 aminoácidos aleatorizados o parcialmente aleatorizados entre las cistinas unidas a disulfuro, y en algunos casos aminoácidos aleatorizados adicionales sobre el exterior del par de cisteínas mediante procedimientos convencionales. La inmunopurificación de estas bibliotecas de péptidos cíclicos contra varias dianas, que incluyen VEGF humano, dio fidedignamente péptidos que se unieron específicamente a VEGFh y no a BSA, albúmina de huevo o IgG.

#### **Ejemplo: Construcción e inmunopurificación de una biblioteca basada en plexina**

20 [0327] Se diseñaron dos bibliotecas basándose en el andamiaje de plexina. La base de datos de proteínas Pfam se usó para alineamiento filogenético de dominios de plexina que se producen naturalmente como se muestra en la Fig. 35. La parte central del andamiaje de plexina (Cys24-Gly25-Trp26-Cys27) se conserva en ambos diseños de biblioteca y sirvió de región cruzada para la generación de bibliotecas de N y C. Los esquemas de aleatorización de ambas bibliotecas de plexina se muestran en la Fig. 36. Las dos bibliotecas se generaron solapando dos oligonucleótidos que codifican bibliotecas en la región cruzada y usando PCR de estiramiento a través ("pull-through"), seguida de clonación por restricción (Sfil/BstXI) y clonación en el vector de fagémido pMP003. Las bibliotecas de plexina resultantes se designaron LMP031 (biblioteca del extremo N) y LMP032 (biblioteca del extremo C) y cada una se representó por una complejidad de aproximadamente 5 x 10<sup>8</sup> transformantes independientes. Para la validación, aproximadamente 24 clones resistentes a Carb de cada biblioteca sin seleccionar se analizaron por PCR. Los clones que dieron un fragmento de tamaño correcto (375 pb) se analizaron adicionalmente por secuenciación de ADN. Las secuencias de plexina de longitud completa correctas se obtuvieron para el 50% y 67% de clones derivados de las bibliotecas LMP031 y LMP032, respectivamente.

35 [0328] Las dos bibliotecas se mezclaron juntas a la relación 50/50 y se inmunopurificaron en paralelo contra VEGF, receptor de muerte Dr4, ErbB2 y HGFR inmovilizados sobre placas de ELISA de 96 pocillos. Se llevaron a cabo cuatro rondas de inmunopurificación usando 1000 ng de proteína diana en la primera ronda, 500 ng en la segunda ronda, 250 ng en la tercera ronda y 100 ng en la cuarta ronda. Después de la ronda final de inmunopurificación, 192 clones resistentes a Carb de cada selección se analizaron para unirse a 100 ng de proteína diana inmovilizada, IgG humana, albúmina de huevo y BSA por ELISA en fago usando Ab anti-M13 policlonal conjugado con peroxidasa de rábano picante para la detección. El mayor porcentaje de clones positivos se obtuvo para DR4 diana (69%), seguido de ErbB2 diana (53%), HGFR (13%) y BoNT diana (1%). Los clones positivos se analizaron adicionalmente por PCR y por secuenciación de ADN. Todos los clones revelaron secuencias únicas y todos, excepto uno (contra DR4), se derivaron de LMP032 (biblioteca del extremo C). Las secuencias de algunas de las cepas aisladas selectivas para diana identificadas se muestran en la Fig. 37.

[0329] Para análisis adicional, una selección de ligantes específicos para diana seleccionada se subclona primero en el vector de expresión de proteína pVS001, luego se produjeron como microproteínas solubles, y finalmente se purificaron por lisis térmica. Las microproteínas específicas para diana purificadas se analizan por ELISA de proteínas para confirmar el reconocimiento de diana, por SDS-PAGE para confirmar la formación de monómeros y por resonancia de plasmones superficiales para medir sus afinidades por la diana. Los mejores clones se usan en la siguiente ronda de generación de bibliotecas para mejorar adicionalmente sus propiedades.

#### **Ejemplo: Construcción de una biblioteca basada en veneno de serpiente**

55 [0330] Se crearon bibliotecas expresadas en fago de 10e<sup>8</sup> a 10e<sup>10</sup> del andamiaje de la toxina de 3 dedos (3FT) con aminoácidos parcialmente aleatorizados de la punta del dedo 1 y la parte descendente del dedo 2 o la punta del dedo 3 y la parte ascendente del dedo 2 mediante procedimientos convencionales.

**[0331]** Se usaron dos andamiajes de 3FT como molde para la generación de la biblioteca de 3FT (configuración de dedos 1 y 2). La estructura de un andamiaje de 3FT y un alineamiento de múltiples secuencias de secuencias relacionadas se muestra en la Fig. 33. Se diseñó una biblioteca de forma que dos bucles de superficie de la toxina están aleatorizados como se ilustra en la Fig. 34. La biblioteca de andamiaje de 3FT parcialmente aleatorizada se generó solapando cuatro oligonucleótidos que codifican la biblioteca en las regiones de hibridación y usando PCR de estiramiento a través ("pull-through"), seguido de clonación por restricción (SfiI/BstXI) en el vector de fagémido pMP003. La biblioteca de 3FT resultante se designó LMP041.

**Ejemplo: Injerto de péptidos de unión en andamiajes de microproteína – aleatorización asistida por péptidos específicos para diana**

**[0332]** El objetivo aquí es usar los péptidos que se han identificado que son específicos para diana de interés con el fin de generar 3SS más ligantes específicos para diana. Esta estrategia se ilustra usando transferencia de péptidos específica para VEGF en la punta del dedo 1 de andamiaje de 3FT y modificando los residuos de AA del dedo 2, que están en estrecha proximidad de la secuencia específica para diana para generar ligantes de VEGF de alta afinidad. Las bibliotecas expresadas de fago de  $10e^8$  a  $10e^{10}$  del andamiaje de la toxina de 3 dedos (3FT) con secuencia específica para VEGF de la punta de dedo 1 y parte descendente parcialmente aleatorizada del dedo 2 se creó mediante procedimientos convencionales como se describe en el ejemplo anterior, excepto que 2 cebadores directos del dedo 1 aleatorios se sustituyeron por cebador directo específico para F1-VEGF que codifica la siguiente secuencia: PSGPSCHTTNHWPI SAVT CPP.

**[0333]** La biblioteca del andamiaje de 3FT (específica para VEGF) centrada con el dedo 2 parcialmente aleatorizado se generó solapando cuatro oligonucleótidos que codifican la biblioteca en las regiones de hibridación y usando PCR de estiramiento a través ("pull-through"), seguido de clonación por restricción (SfiI/BstXI) en el vector de fagémido pMP003. La biblioteca de 3FT resultante se designó LMP042.

**Ejemplo: Semivida en plasma de una MURP**

**[0334]** La semivida en plasma de MURP pueden medirse después de inyección i.v. o i.p. de la MURP en ratas cateterizadas esencialmente como se describe por [Pepinsky, R. B. y col. (2001) J Pharmacol Exp Ther, 297: 1059-66]. Pueden extraerse muestras de sangre en diversos momentos de tiempo (5 min, 15 min, 30 min, 1 h, 3 h, 5 h, 1 d, 2 d, 3 d) y la concentración en plasma de la MURP puede medirse usando ELISA. Los parámetros farmacocinéticos pueden calcularse usando WinNonlin versión 2.0 (Scientific Consulting Inc., Apex, NC). Para analizar el efecto del módulo de URP puede compararse la semivida en plasma de una proteína que contiene el módulo de URP con la semivida en plasma de la misma proteína que carece del módulo de URP.

**Ejemplo: Prueba de solubilidad de una MURP**

**[0335]** La solubilidad de MURP puede determinarse concentrando muestras purificadas de MURP en tampones fisiológicos como solución salina tamponada con fosfato a concentraciones diversas en el intervalo de 0,01 mg/ml a 10 mg/ml. Las muestras pueden incubarse para hasta varias semanas. Muestras en las que la concentración supera la solubilidad de la MURP muestran precipitación como se indica por turbidez, que puede medirse en un lector de absorbancia. Puede eliminarse material precipitado por centrifugación o filtración y medir la concentración de proteína restante en el sobrenadante usando un ensayo de proteína como el ensayo de Bradford midiendo la absorbancia a 280 nm. Los estudios de solubilidad pueden acelerarse congelando las muestras a  $-20\text{ }^{\circ}\text{C}$  y posterior descongelación. Este procedimiento frecuentemente conduce a la precipitación de proteínas escasamente solubles.

**Ejemplo: Actividad de unión a suero de MURP**

**[0336]** Pueden recubrirse MURP de interés en placas de microtitulación y proteínas de control en otros pocillos de la placa. Posteriormente, pueden añadirse muestras de suero de interés a los pocillos durante 1 hora. Posteriormente, los pocillos pueden lavarse con un lavador de placas. Las proteínas del suero unidas pueden detectarse añadiendo anticuerpos contra las proteínas del suero que se han conjugado con enzimas como peroxidasa de rábano picante o fosfatasa alcalina para la detección. Otra forma de detectar la unión del suero a MURP es añadir la MURP de interés a suero durante aproximadamente 1 hora para permitir la unión. Posteriormente, puede inmunoprecipitarse la MURP usando un anticuerpo contra un epítipo en la secuencia de MURP. Las muestras precipitadas pueden analizarse por PAGE y opcionalmente por Western para detectar cualquier proteína que coprecipitó con la MURP. Pueden identificarse proteínas del suero que muestran

coprecipitación por espectrometría de masas.

Los siguientes párrafos numerados contienen afirmaciones de amplias combinaciones de las características de la invención aquí descritas.

1. Un polímero recombinante no estructurado (URP) que comprende al menos 40 aminoácidos contiguos, en el que
  - 5 (a) la suma de los residuos de glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) contenidos en el URP constituye más de aproximadamente el 80% de los aminoácidos totales del URP; y/o
  - (b) al menos el 50% de los aminoácidos carecen de estructura secundaria tal como se ha determinado por el algoritmo de Chou-Fasman.
- 10 2. Un polímero recombinante no estructurado (URP) que comprende al menos 40 aminoácidos contiguos, en el que dicho URP tiene una semivida de degradación en suero *in vitro* superior a aproximadamente 24 horas, y en el que
  - (a) la suma de los residuos de glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) contenidos en el URP constituye más de aproximadamente el 80% de los aminoácidos totales del URP; y/o
  - (b) al menos el 50% de los aminoácidos carecen de estructura secundaria tal como se ha determinado por el
- 15 algoritmo de Chou-Fasman.
3. El URP del párrafo 1 ó 2, en el que el URP comprende una secuencia de aminoácidos no naturales.
4. El URP del párrafo 1 ó 2, en el que el URP se selecciona para la incorporación en una proteína heteróloga, y en el que tras la incorporación del URP en una proteína heteróloga, dicha proteína heteróloga presenta una semivida en suero más larga y/o mayor solubilidad con respecto a la proteína correspondiente que es deficiente en dicho URP.
- 20 5. El URP del párrafo 1 ó 2, en el que tras la incorporación del URP en una proteína heteróloga, dicha proteína heteróloga presenta una semivida de secreción en suero que es al menos dos veces más larga en comparación con la proteína correspondiente que es deficiente en dicho URP.
6. El URP del párrafo 1 ó 2, en el que la incorporación del URP en una proteína heteróloga da lugar a al menos un incremento de dos veces en el peso molecular aparente de la proteína tal como se obtiene por aproximación
- 25 mediante cromatografía de exclusión por tamaño.
7. El URP del párrafo 1 ó 2, en el que el URP tiene una puntuación de epítipo T inferior a -4.
8. El URP del párrafo 1 ó 2, en el que los aminoácidos son predominantemente residuos hidrófilos.
9. El URP del párrafo 1 ó 2, en el que al menos el 50% de los aminoácidos del URP carecen de estructura secundaria tal como se ha determinado por el algoritmo de Chou-Fasman.
- 30 10. El URP del párrafo 1 ó 2, en el que los residuos de glicina contenidos en el URP constituyen al menos aproximadamente el 50% de los aminoácidos totales del URP.
11. El URP del párrafo 1 ó 2, en el que cualquier tipo de los aminoácidos solos seleccionados del grupo que consiste en glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) constituye más de aproximadamente el 20% de los aminoácidos totales del URP.
- 35 12. El URP del párrafo 1 ó 2, en el que cualquier tipo de los aminoácidos solos seleccionados del grupo que consiste en glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) constituye más de aproximadamente el 40% de los aminoácidos totales del URP.
13. El URP del párrafo 1 ó 2, en el que el URP comprende más de aproximadamente 100 aminoácidos contiguos.
14. El URP del párrafo 1 ó 2, en el que el URP comprende más de aproximadamente 200 aminoácidos contiguos.
- 40 15. El URP del párrafo 1 ó 2, que comprende secuencias repetidas.
16. El URP del párrafo 1 ó 2, en el que un tipo de los aminoácidos solos seleccionados del grupo que consiste en glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) constituye más de aproximadamente el 50% de los aminoácidos totales del URP.
17. Una proteína que comprende uno o más de URP del párrafo 1 ó 2, en el que dicho uno o más URP son
- 45 heterólogos con respecto a la proteína.
18. La proteína del párrafo 17, que comprende un módulo efector.
19. La proteína del párrafo 17, que comprende un módulo de unión.
20. La proteína del párrafo 17, que comprende un módulo efector y un módulo de unión.
21. La proteína del párrafo 17, que comprende una pluralidad de módulos de unión, en la que los módulos de unión
- 50 individuales muestran especificidades de unión a la misma o diferentes dianas.
22. La proteína del párrafo 17, en la que la longitud total de los URP en agregación supera aproximadamente los 150 aminoácidos.
23. La proteína del párrafo 17, que comprende uno o más módulos de unión, en la que el módulo de unión comprende un andamiaje que contiene disulfuro formado por apareamiento de cisteínas entre andamiajes.
- 55 24. La proteína del párrafo 17, que comprende un módulo efector que es citotóxico.
25. La proteína del párrafo 17, que comprende un módulo de unión específico para una molécula diana, en la que la diana se selecciona del grupo que consiste en proteína de la superficie celular, proteína secretada, proteína citosólica y proteína nuclear.
26. La proteína del párrafo 17, que comprende un módulo de unión específico para molécula diana, en la que la

diana es un canal de iones.

27. La proteína del párrafo 17, que muestra una semivida de secreción en suero prolongada en al menos 2 veces con respecto a la proteína correspondiente que es deficiente en dicho URP.
- 5 28. Una proteína no natural que comprende al menos 3 unidades de repetición de secuencias de aminoácidos, comprendiendo cada una de las unidades de repetición al menos 6 aminoácidos, en la que la mayoría de los segmentos que comprenden de aproximadamente 6 a aproximadamente 15 aminoácidos contiguos de dichas al menos 3 unidades de repetición están presentes en una o más proteínas humanas nativas.
- 10 29. La proteína del párrafo 28, en la que la mayoría de segmentos que comprende de aproximadamente 9 a aproximadamente 15 aminoácidos contiguos dentro de las unidades de repetición están presentes en una o más proteínas humanas nativas.
30. La proteína del párrafo 28, en la que cada segmento que comprende de aproximadamente 6 a aproximadamente 15 aminoácidos contiguos dentro de la proteína está presente en al menos una proteína humana nativa.
31. La proteína del párrafo 28, en la que cada segmento que comprende de aproximadamente 9 a aproximadamente 15 aminoácidos contiguos dentro de la proteína está presente en al menos una proteína humana nativa.
- 15 32. La proteína del párrafo 28, en la que al menos 3 unidades de repetición comparten una identidad de secuencia superior a aproximadamente el 80%.
33. La proteína del párrafo 28, en la que cada una de las unidades de repetición comprende de aproximadamente 6 a aproximadamente 15 aminoácidos contiguos.
- 20 34. La proteína del párrafo 28, en la que cada una de las unidades de repetición comprende aproximadamente 9 aminoácidos contiguos.
35. La proteína del párrafo 28, que comprende uno o más módulos seleccionados del grupo que consiste en módulos de unión, módulos efectores, módulos de multimerización, módulos C-terminales y módulos N-terminales.
36. La proteína del párrafo 28, en la que una unidad de repetición individual comprende un polímero recombinante no estructurado (URP).
- 25 37. La proteína del párrafo 36, en la que el URP comprende al menos 40 aminoácidos contiguos, en el que dicho URP es sustancialmente incapaz de unirse no específicamente a una proteína del suero, y en el que
- (a) la suma de los residuos de glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) contenidos en el URP constituye más de aproximadamente el 80% de los aminoácidos totales del URP; y/o
- (b) al menos el 50% de los aminoácidos carecen de estructura secundaria tal como se ha determinado por el
- 30 algoritmo de Chou-Fasman.
38. La proteína del párrafo 36, en la que el URP comprende al menos 40 aminoácidos contiguos, en el que dicho URP tiene una semivida de degradación en suero *in vitro* superior a aproximadamente 24 horas, y en el que
- (a) la suma de los residuos de glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) contenidos en el URP constituye más de aproximadamente el 80% de los aminoácidos totales del URP; y/o
- 35 (b) al menos el 50% de los aminoácidos carecen de estructura secundaria tal como se ha determinado por el algoritmo de Chou-Fasman.
39. Un polinucleótido recombinante que comprende una secuencia codificante que codifica el URP del párrafo 1 ó 2.
40. Un polinucleótido recombinante que comprende una secuencia codificante que codifica la proteína del párrafo 17.
- 40 41. Un polinucleótido recombinante que comprende una secuencia codificante que codifica la proteína del párrafo 28.
42. Una célula huésped que comprende los polinucleótidos recombinantes del párrafo 40.
43. Una célula huésped que comprende los polinucleótidos recombinantes del párrafo 41.
44. Un vector que comprende el polinucleótido recombinante del párrafo 40.
- 45 45. Un vector que comprende el polinucleótido recombinante del párrafo 41.
46. Una biblioteca seleccionable de vectores de expresión que comprenden más de un vector del párrafo 44.
47. Una biblioteca seleccionable de vectores de expresión que comprenden más de un vector del párrafo 45.
48. Un paquete genético que expresa la biblioteca del párrafo 46.
49. Un paquete genético que expresa la biblioteca del párrafo 47.
- 50 50. Un procedimiento de producción de una proteína que comprende un polímero recombinante no estructurado (URP), que comprende:
- (i) proporcionar una célula huésped que comprende un polinucleótido recombinante que codifica la proteína, comprendiendo dicha proteína uno o más URP, comprendiendo dicho URP al menos 40 aminoácidos contiguos, en el que dicho URP es sustancialmente incapaz de unirse no específicamente a una proteína del suero, y en el que
- 55 (a) la suma de los residuos de glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) contenidos en el URP constituye más de aproximadamente el 80% de los aminoácidos totales del URP; y/o
- (b) al menos el 50% de los aminoácidos carecen de estructura secundaria tal como se ha determinado por el algoritmo de Chou-Fasman;
- (ii) cultivar dicha célula huésped en un medio de cultivo adecuado en condiciones para efectuar la expresión de

dicha proteína de dicho polinucleótido.

51. El procedimiento del párrafo 50, en el que el URP tiene una semivida de degradación en suero *in vitro* superior a aproximadamente 24 horas.
52. El procedimiento del párrafo 50, en el que la célula huésped es una célula eucariota.
53. El procedimiento del párrafo 50, en el que la célula huésped es una célula CHO.
54. El procedimiento del párrafo 50, en el que la célula huésped es una célula procariota.
55. Un procedimiento de aumento de la semivida de secreción en suero de una proteína, que comprende: fusionar dicha proteína con uno o más polímeros recombinantes no estructurados (URP), en el que el URP comprende al menos aproximadamente 40 aminoácidos contiguos, y en el que
- 10 (a) la suma de los residuos de glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) contenidos en el URP constituye más de aproximadamente el 80% de los aminoácidos totales del URP; y/o
- (b) al menos el 50% de los aminoácidos del URP carecen de estructura secundaria como se ha determinado por el algoritmo de Chou-Fasman; y en el que dicho URP es sustancialmente incapaz de unirse no específicamente a una proteína del suero.
- 15 56. El procedimiento del párrafo 55, en el que la semivida de secreción en suero de la proteína se prolonga en al menos 2 veces.
57. Un procedimiento de detección de la presencia o ausencia de una interacción específica entre una diana y una proteína exógena que se expresa en un paquete genético, en el que dicha proteína comprende uno o más polímeros recombinantes no estructurados (URP), comprendiendo el procedimiento:
- 20 (a) proporcionar un paquete genético que expresa una proteína que comprende uno o más polímeros recombinantes no estructurados (URP);
- (b) poner en contacto el paquete genético con la diana en condiciones adecuadas para producir un complejo de proteína-diana estable; y
- (c) detectar la formación del complejo de proteína-diana estable en el paquete genético, detectando así la presencia
- 25 de una interacción específica.
58. El procedimiento del párrafo 57, que comprende además obtener una secuencia de nucleótidos del paquete genético que codifica la proteína exógena.
59. El procedimiento del párrafo 57, en el que la presencia o ausencia de una interacción específica es entre el URP y una diana que comprende una proteína de suero.
- 30 60. El procedimiento del párrafo 57, en el que la presencia o ausencia de una interacción específica es entre el URP y una diana que comprende una proteasa de suero.
61. El procedimiento del párrafo 57, en el que la diana o la proteína se selecciona del grupo que consiste en proteína de la superficie celular, proteína secretada, proteína citosólica y proteína nuclear.
62. El procedimiento del párrafo 57, en el que el paquete genético es un fago.
- 35 63. El procedimiento del párrafo 57, en el que el URP comprende al menos aproximadamente 40 aminoácidos contiguos, y en el que el URP es sustancialmente incapaz de unirse no específicamente a una proteína del suero, y en el que
- (a) la suma de los residuos de glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) contenidos en el URP constituye más de aproximadamente el 80% de los aminoácidos totales del URP; y/o
- 40 (b) al menos el 50% de los aminoácidos del URP carecen de estructura secundaria tal como se ha determinado por el algoritmo de Chou-Fasman.
64. El procedimiento del párrafo 57, en el que el URP comprende al menos aproximadamente 40 aminoácidos contiguos, y en el que el URP tiene una semivida de degradación en suero *in vitro* superior a aproximadamente 24 horas.
- 45 65. El procedimiento del párrafo 50, 55 ó 57, en el que el URP comprende una secuencia de aminoácidos no naturales.
66. El procedimiento del párrafo 50, 55 ó 57, en el que el URP tiene una puntuación de epítipo T igual o inferior a -4.
67. El procedimiento del párrafo 50, 55 ó 57, en el que el URP carece de estructura secundaria tal como se ha determinado por el algoritmo de Chou-Fasman.
- 50 68. El procedimiento del párrafo 50, 55 ó 57, en el que los residuos de glicina contenidos en el URP constituyen al menos aproximadamente el 50% de los aminoácidos totales del URP.
69. El procedimiento del párrafo 50, 55 ó 57, en el que la suma de los residuos de glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) contenidos en el URP constituye más de aproximadamente el 60% de los aminoácidos totales del URP.
- 55 70. El URP del párrafo 50 ó 51, en el un tipo de los aminoácidos solos seleccionados del grupo que consiste en glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) constituye más de aproximadamente el 50% de los aminoácidos totales del URP.
71. El procedimiento del párrafo 50, 55 ó 57, en el que el URP comprende más de 100 aminoácidos contiguos.
72. El procedimiento del párrafo 50, 55 ó 57, en el que el URP comprende secuencias de repetición.

73. El procedimiento del párrafo 50, 55 ó 57, en el que la proteína es una proteína terapéutica.

74. El procedimiento del párrafo 50, 55 ó 57, en el que la proteína comprende uno o más módulos seleccionados del grupo que consiste en módulos de unión, módulos efectores, módulos de multimerización, módulos C-terminales y módulos N-terminales.

5

10

15

20

25

## REIVINDICACIONES

1. Proteína de fusión que incorpora:  
 (I) un polímero recombinante no estructurado (URP) que comprende una secuencia contigua de al menos 40 aminoácidos, en el que dicho URP se **caracteriza porque**:  
 5 (a) contiene sólo 3, 4, 5, ó 6 tipos diferentes de aminoácidos;  
 (b) la suma de los residuos de glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) constituye más de aproximadamente el 80% de los aminoácidos totales;  
 (c) al menos el 50% de los aminoácidos no están presentes en la estructura secundaria tal como se ha  
 10 determinado por el algoritmo de Chou-Fasman; y  
 (d) tiene una puntuación de epítipo T inferior a -4;  
 comprendiendo además la proteína de fusión:  
 (II) una proteína heteróloga,  
 en la que dicha proteína de fusión muestra una semivida en suero al menos dos veces más larga en comparación  
 15 con la correspondiente proteína heteróloga que carece de dicho URP,  
 en la que dicha proteína heteróloga es una proteína farmacéuticamente activa seleccionada del grupo que consiste en citoquinas, factores de crecimiento, hormonas, eritropoyetina, adenosina desaminasa, asparaginasa, arginasa, interferón, hormona de crecimiento, hormona liberadora de hormona de crecimiento, G-CSF, GM-CSM, insulina, hirudina, receptor de TNF, uricasa, rasburicasa, axoquina, ARNasa, ADNasa, fosfatasa, exotoxina de *Pseudomonas*,  
 20 ricina, gelonina, desmoteplasa, laronidasa, trombina, enzima de coagulación de la sangre, VEGF, protropina, somatropina, alteplasa, interleucina, factor VII, factor VIII, factor X, factor IX, dornasa, glucocerebrosidasa, folitropina, glucagón, tirotropina, nesiritida, alteplasa, teriparatida, agalsidasa, laronidasa, metioninasa.
2. Proteína, según la reivindicación 1, en la que la proteína de fusión tiene un incremento de al menos 2 veces en el  
 25 peso molecular aparente en comparación con la proteína heteróloga que carece de dicho URP, tal como se obtiene por aproximación mediante cromatografía por exclusión de tamaño.
3. Proteína, según la reivindicación 1, en la que la proteína farmacéuticamente activa es una hormona del  
 30 crecimiento.
4. Proteína, según la reivindicación 1, en la que la proteína farmacéuticamente activa es la insulina.
5. Proteína, según la reivindicación 1, en la que la proteína farmacéuticamente activa es el factor VIII.
- 35 6. Proteína, según la reivindicación 1, en la que la proteína farmacéuticamente activa es el factor VII.
7. Proteína, según la reivindicación 1, en la que la proteína farmacéuticamente activa es el factor IX.
8. Proteína, según la reivindicación 1, en la que dicho URP tiene más de un 5% de glutamato (E) y comprende  
 40 menos de un 2% de arginina (R) o lisina (K).
9. Proteína, según la reivindicación 1, en la que cualquier tipo de los aminoácidos solos seleccionados del grupo que  
 consiste en glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) constituye más  
 45 de aproximadamente el 20% de los aminoácidos totales del URP.
10. Proteína, según la reivindicación 9, en la que cualquier tipo de los aminoácidos solos seleccionados del grupo  
 que consiste en glicina (G), aspartato (D), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) constituye  
 más de aproximadamente el 30% de los aminoácidos totales del URP.
- 50 11. Proteína, según la reivindicación 9, en la que los residuos de glicina contenidos en el URP constituyen al menos  
 aproximadamente el 20% de los aminoácidos totales de la URP.
12. Proteína, según la reivindicación 1, en la que el URP comprende más de aproximadamente 100 aminoácidos  
 55 contiguos.
13. Proteína, según la reivindicación 1, en la que el URP comprende más de aproximadamente 200 aminoácidos  
 contiguos.
14. Polinucleótido recombinante que comprende una secuencia codificante que codifica la proteína de fusión según

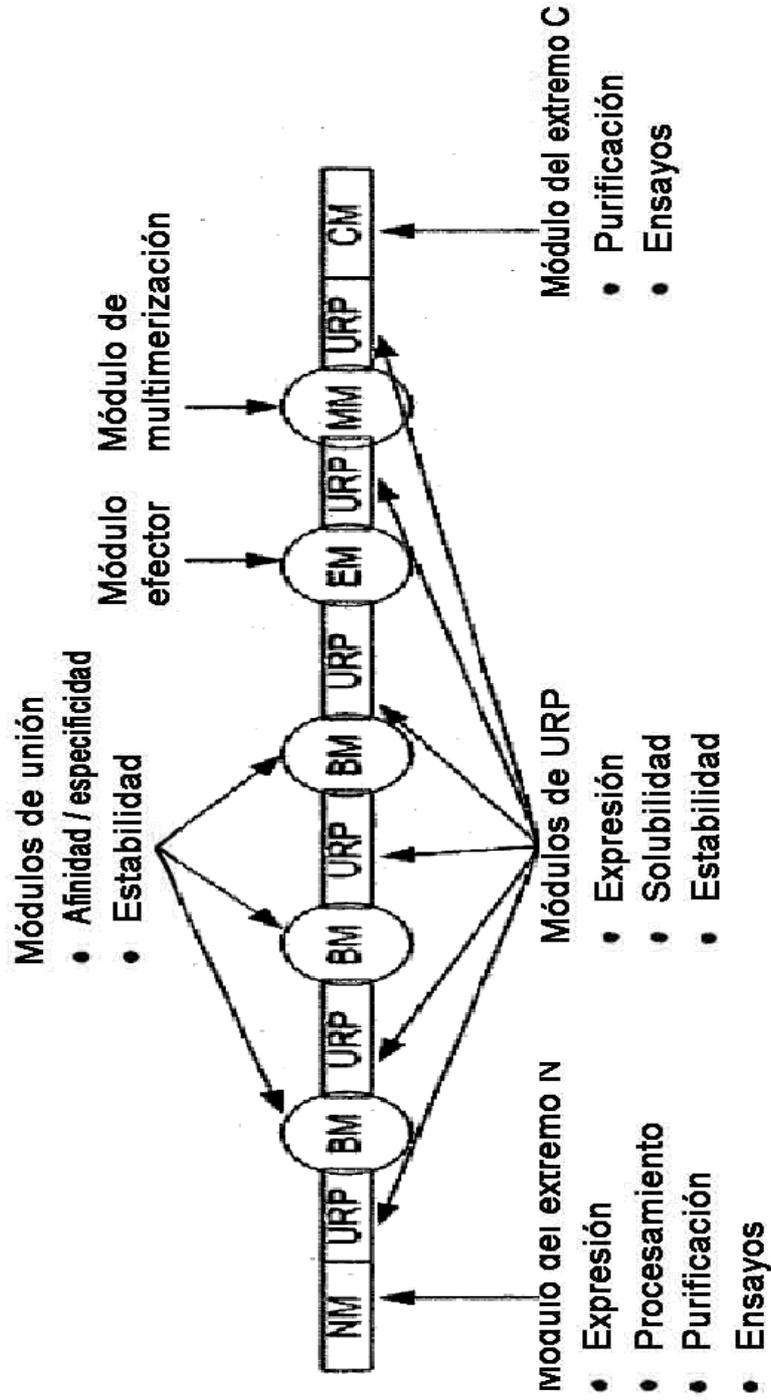
la reivindicación 1.

15. Célula huésped que comprende el polinucleótido recombinante según la reivindicación 14.

5

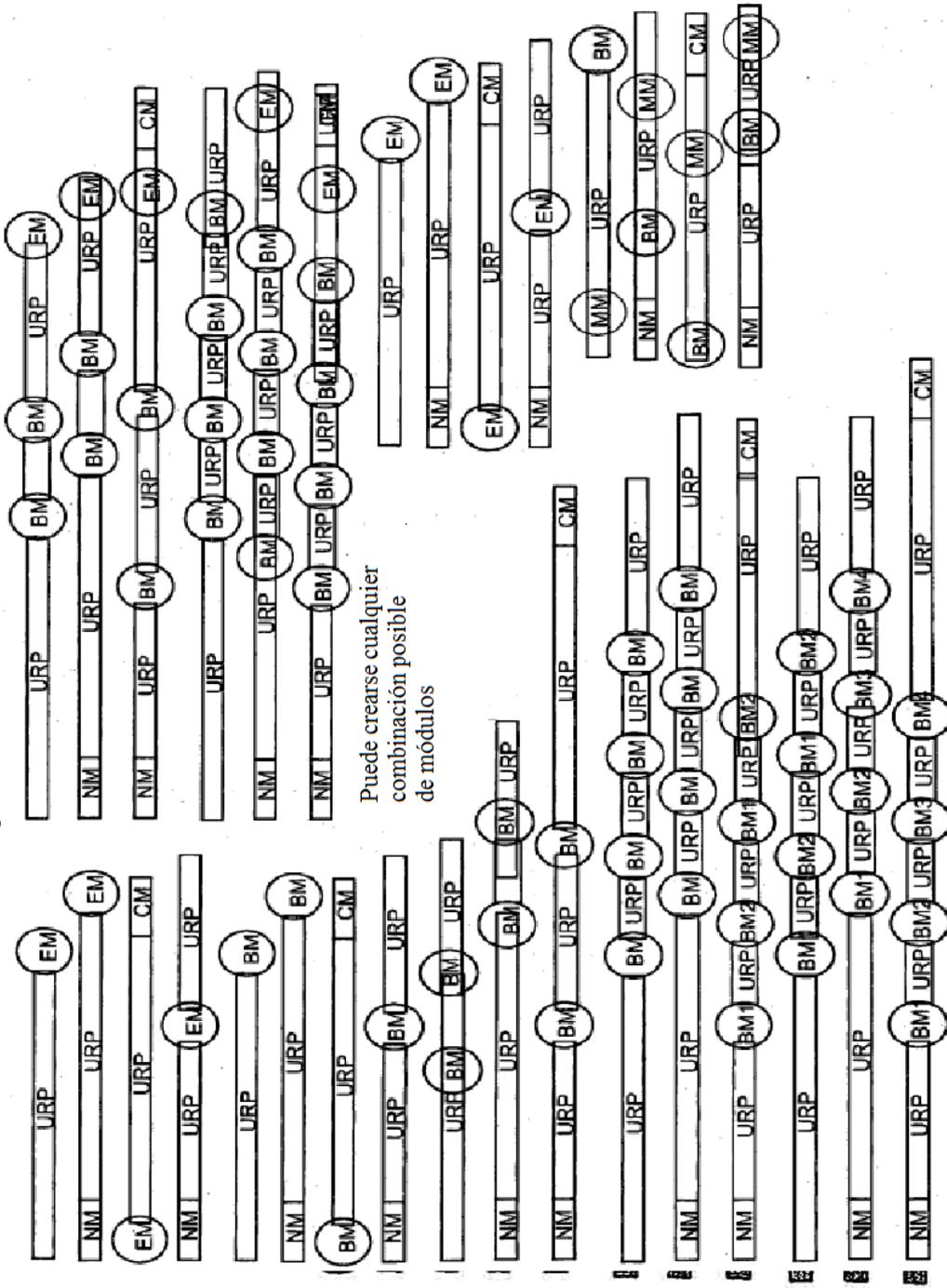
10

**Fig. 1: Proteínas recombinantes no estructuradas multifuncionales (MURP)**



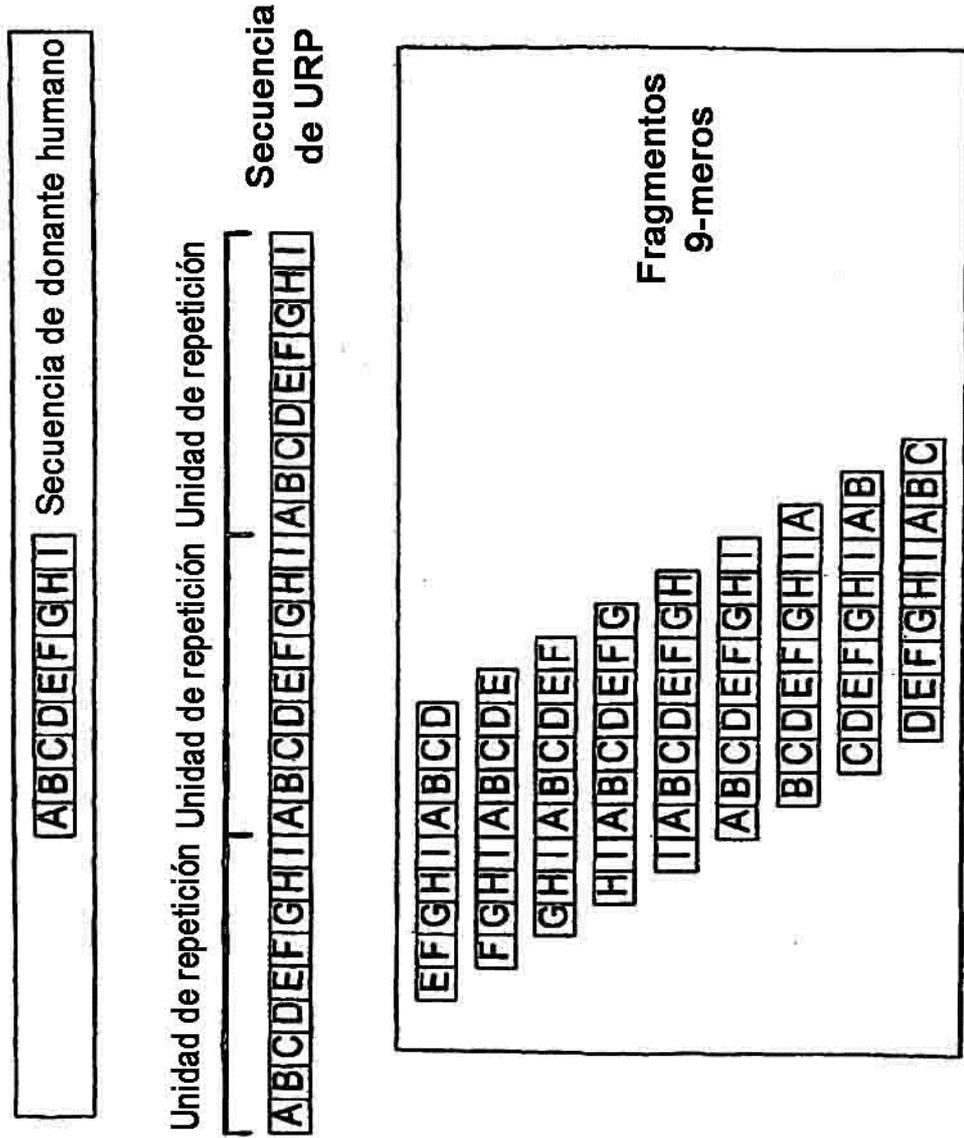
Los módulos de unión pueden tener la misma especificidad o diferente

Fig. 2: Ejemplos de MURP con pocos dominios

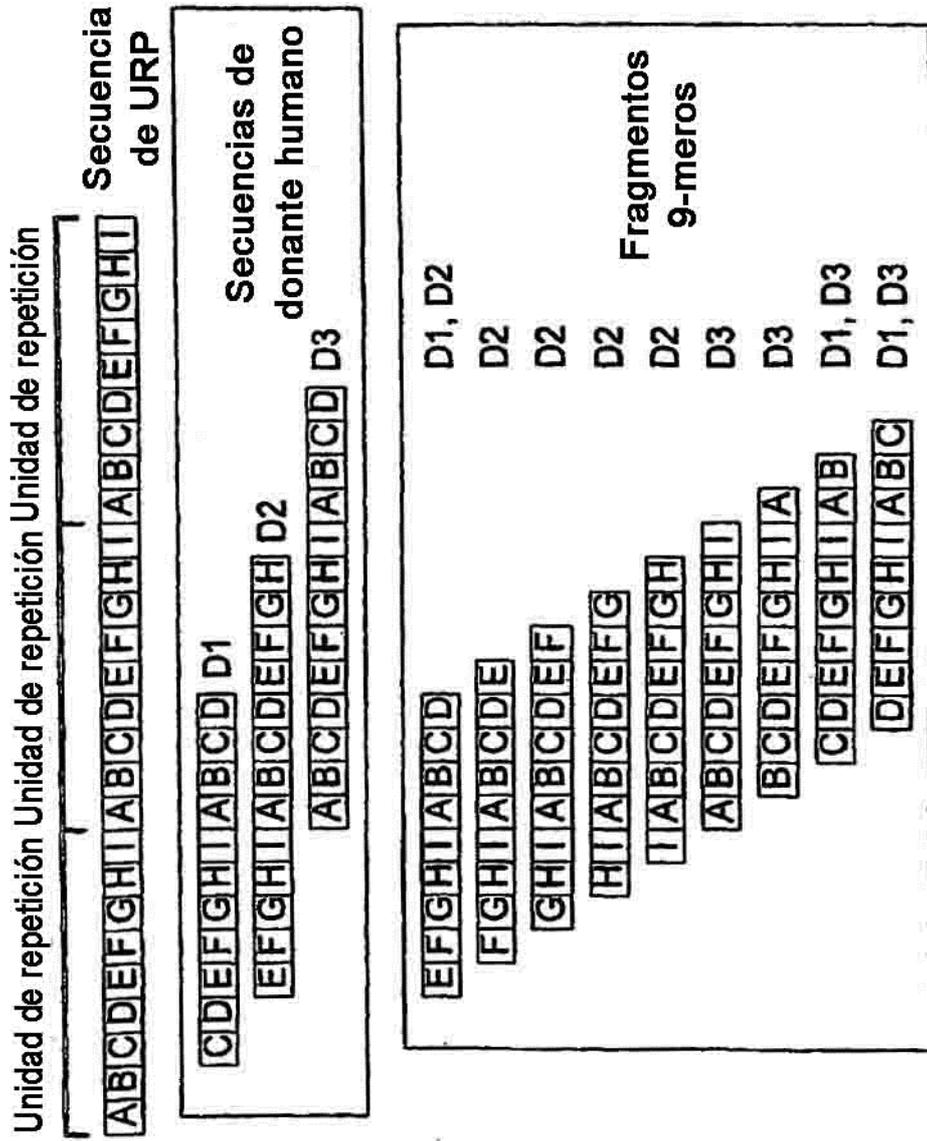


BM1, BM2, BM3 y BM4 son módulos de unión con 4 especificidades por diana diferentes

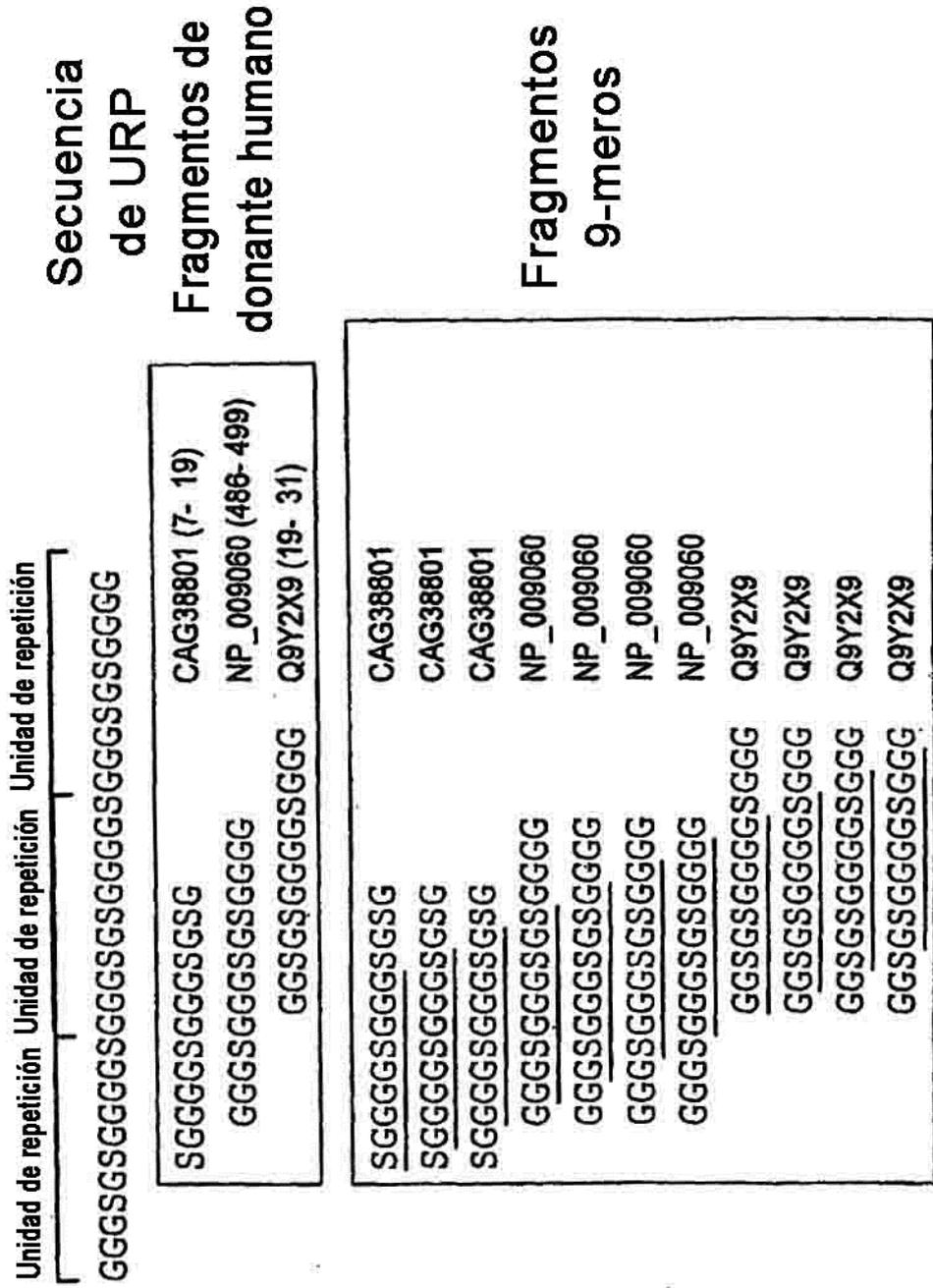
**Fig. 3: Secuencias de URP como oligómeros de secuencias humanas**



**Fig. 4: Secuencias de URP como oligómeros de secuencias humanas solapantes**

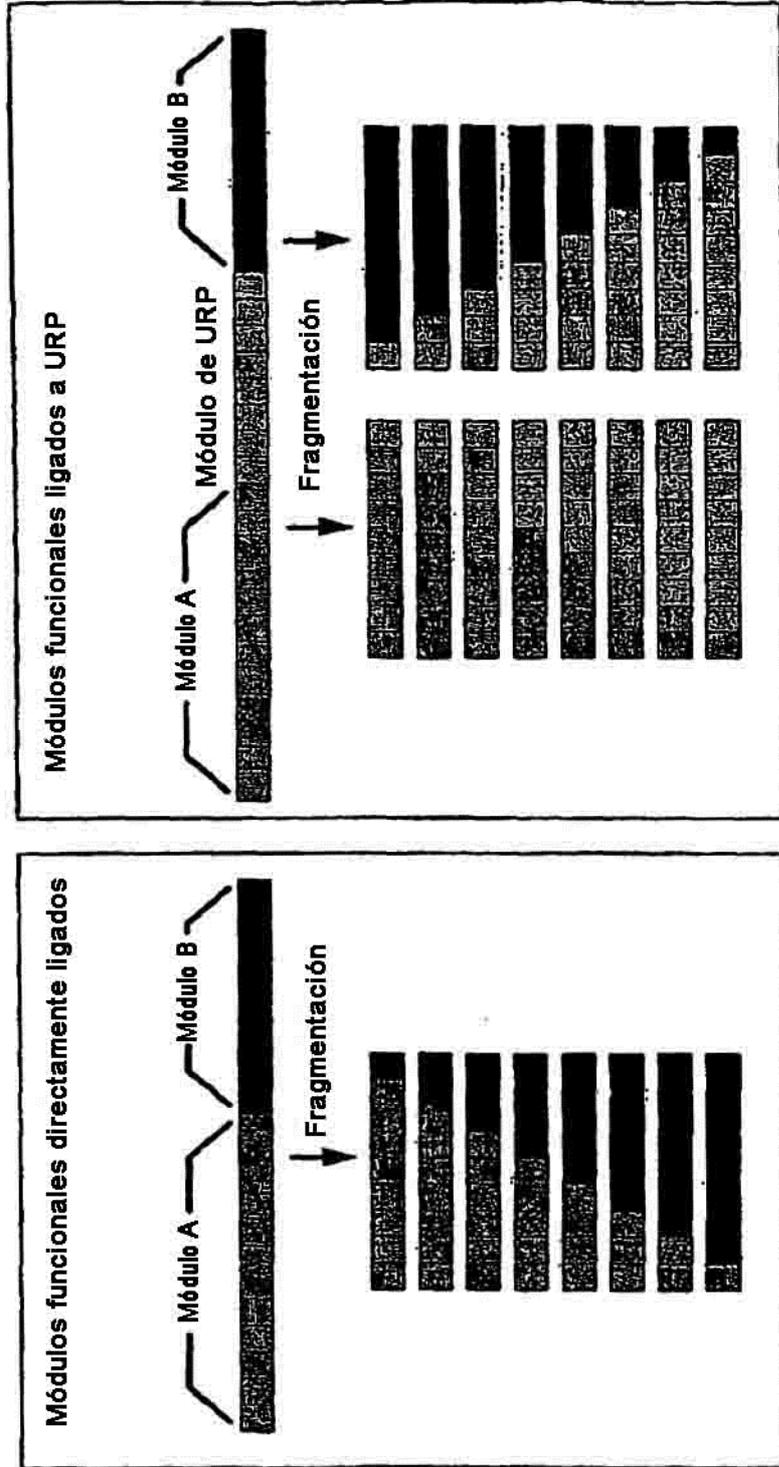


**Fig. 5: Secuencias de URP ricas en glicina/serina basadas en tres secuencias humanas solapantes**

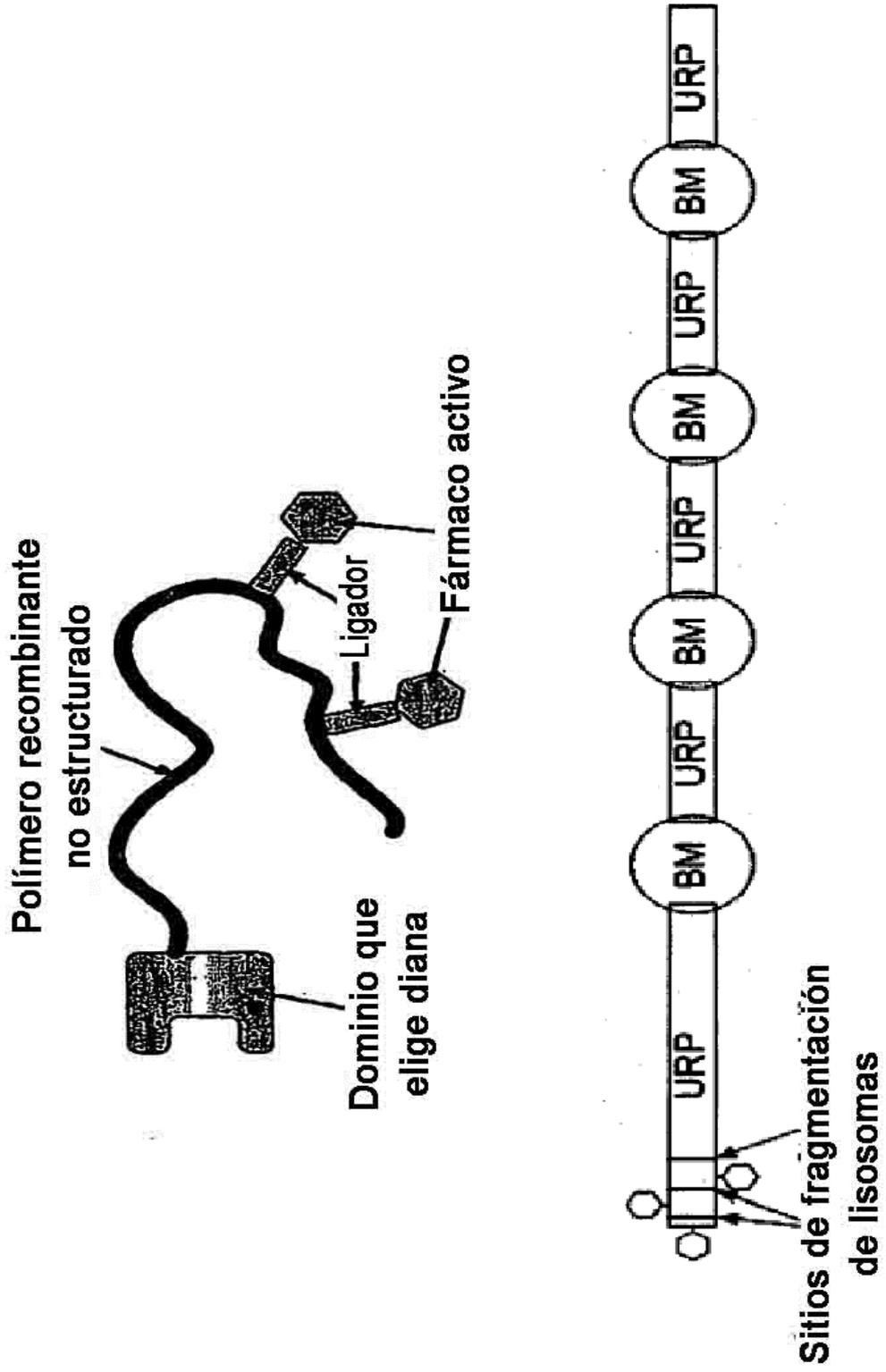




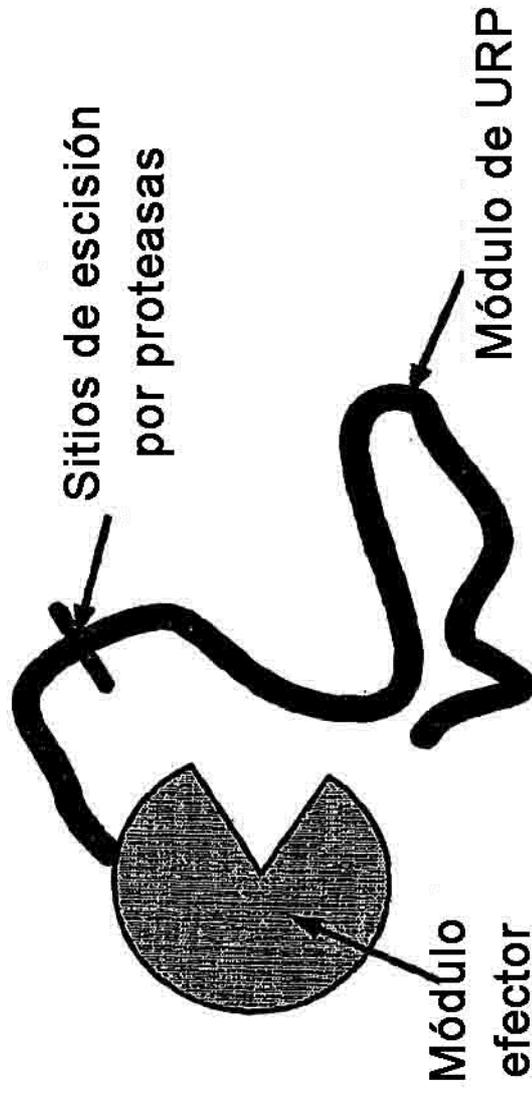
**Fig. 7: Interrupción de epítopes de linfocitos T que atraviesan módulos por módulos de URP**



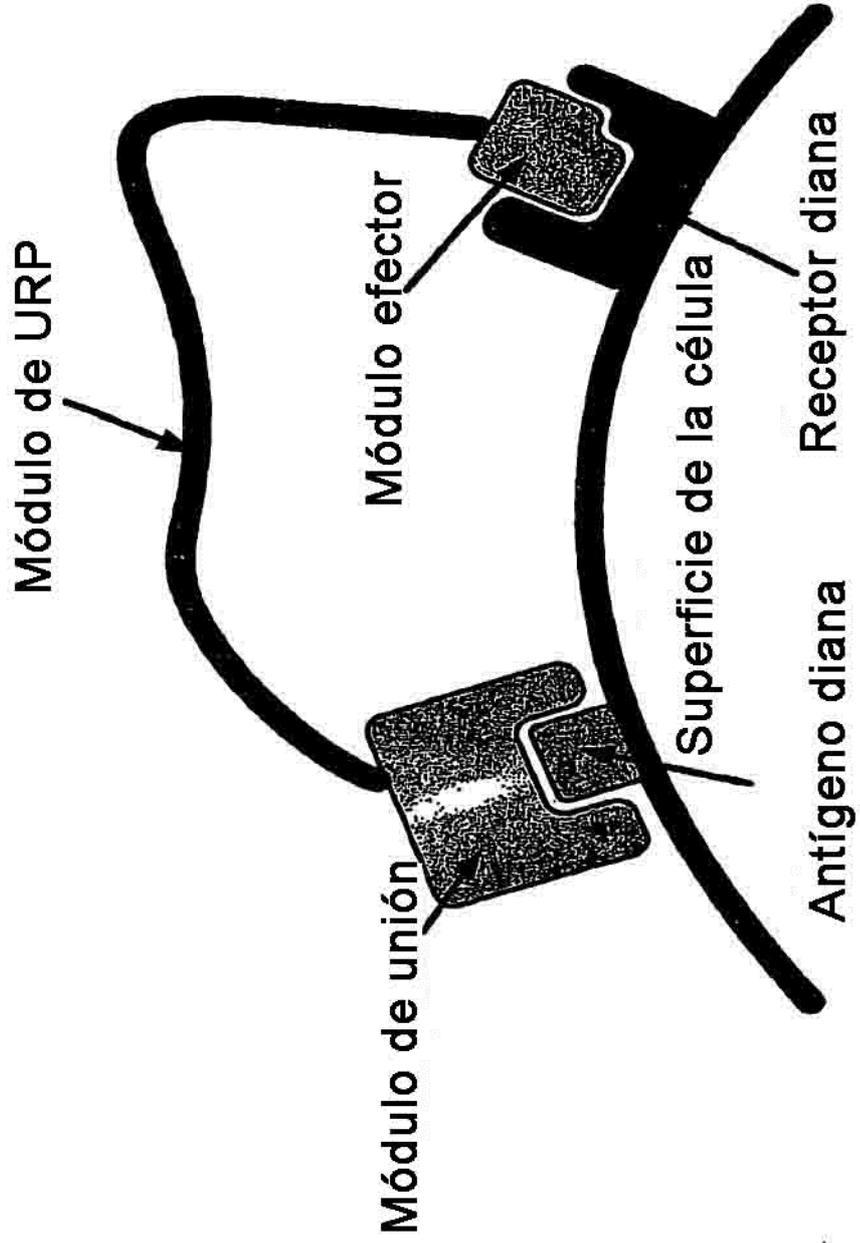
**Fig. 8: Construcciones de administración de fármaco**



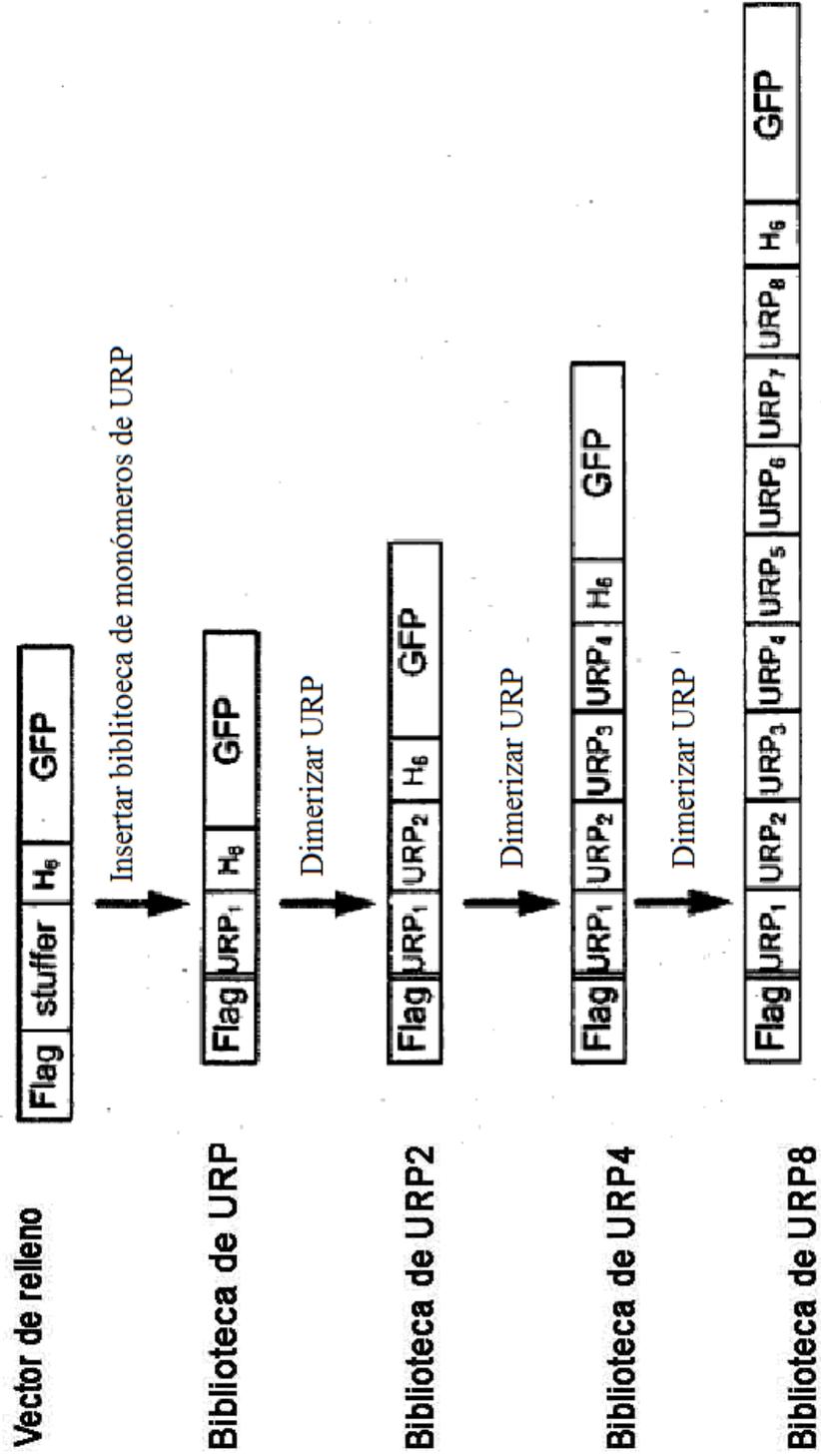
**Fig. 9: Ejemplo de una MURP que contiene un sitio sensible a proteasas**



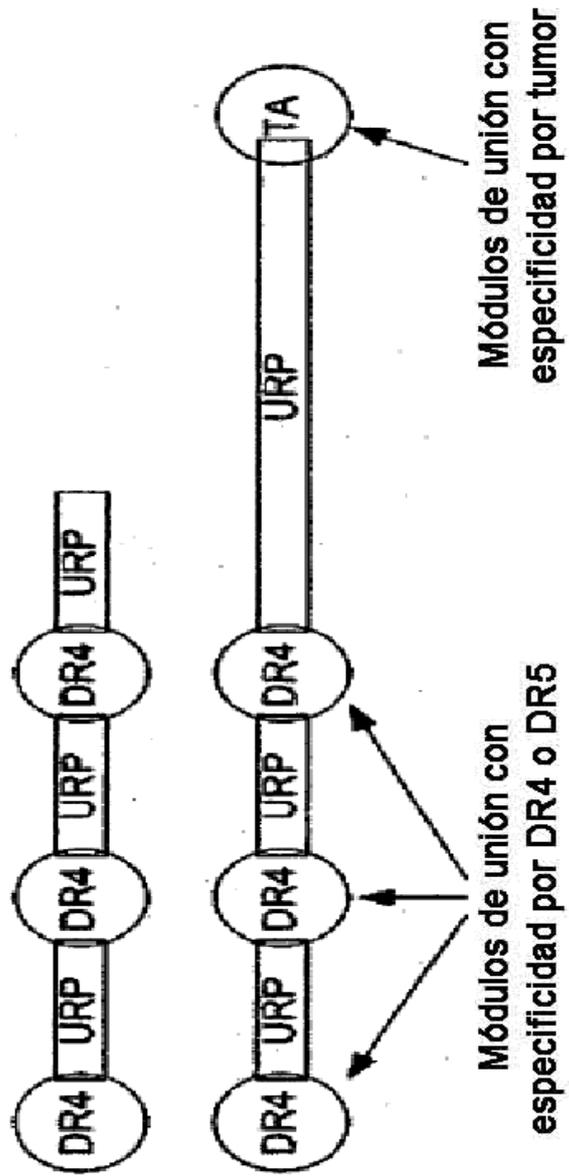
**Fig. 10. Aumento de la concentración local de un módulo efector**



**Fig. 11 Construcción de módulos de URP por dimerización iterativa**



**Fig. 12: Ejemplos de MURP con especificidad por receptores de muerte**



**Fig. 13 IL2 que elige antígeno tumoral  
como diana**

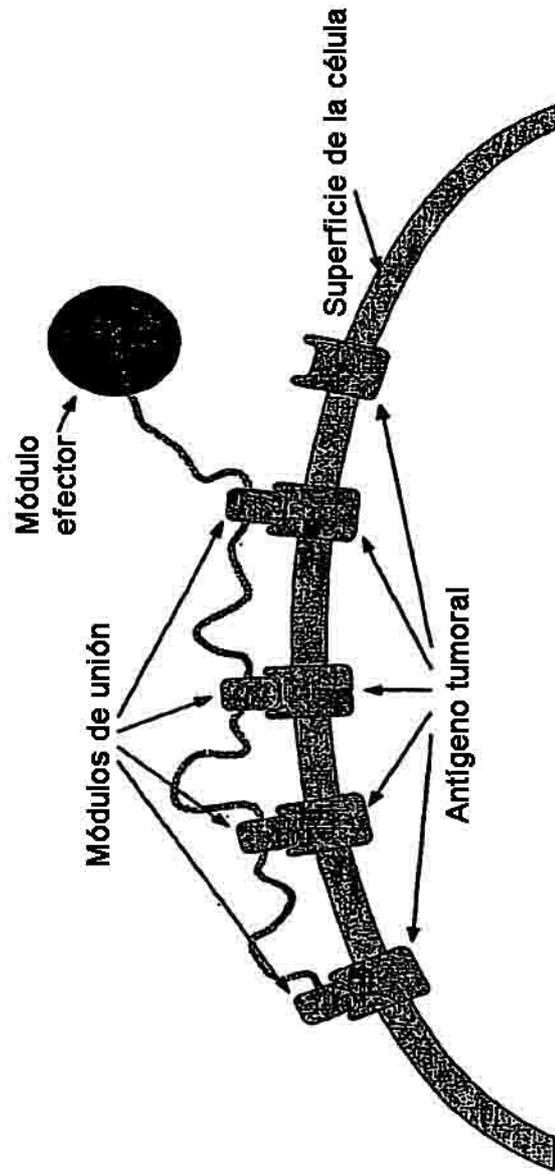


Fig. 14 Construcción de rPEG\_H288 y rPEG\_J288

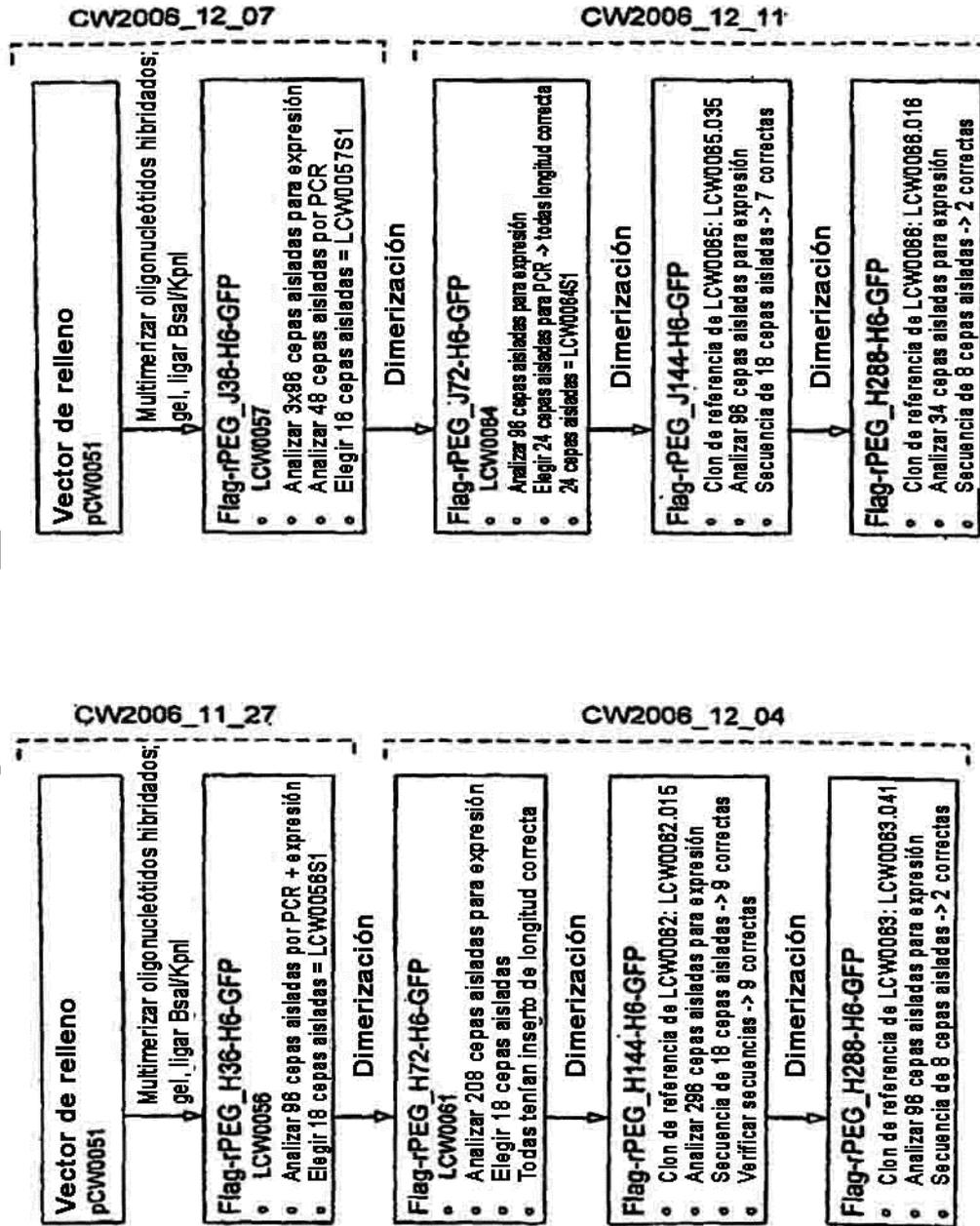


Fig. 15 Secuencia de rPEG\_J288

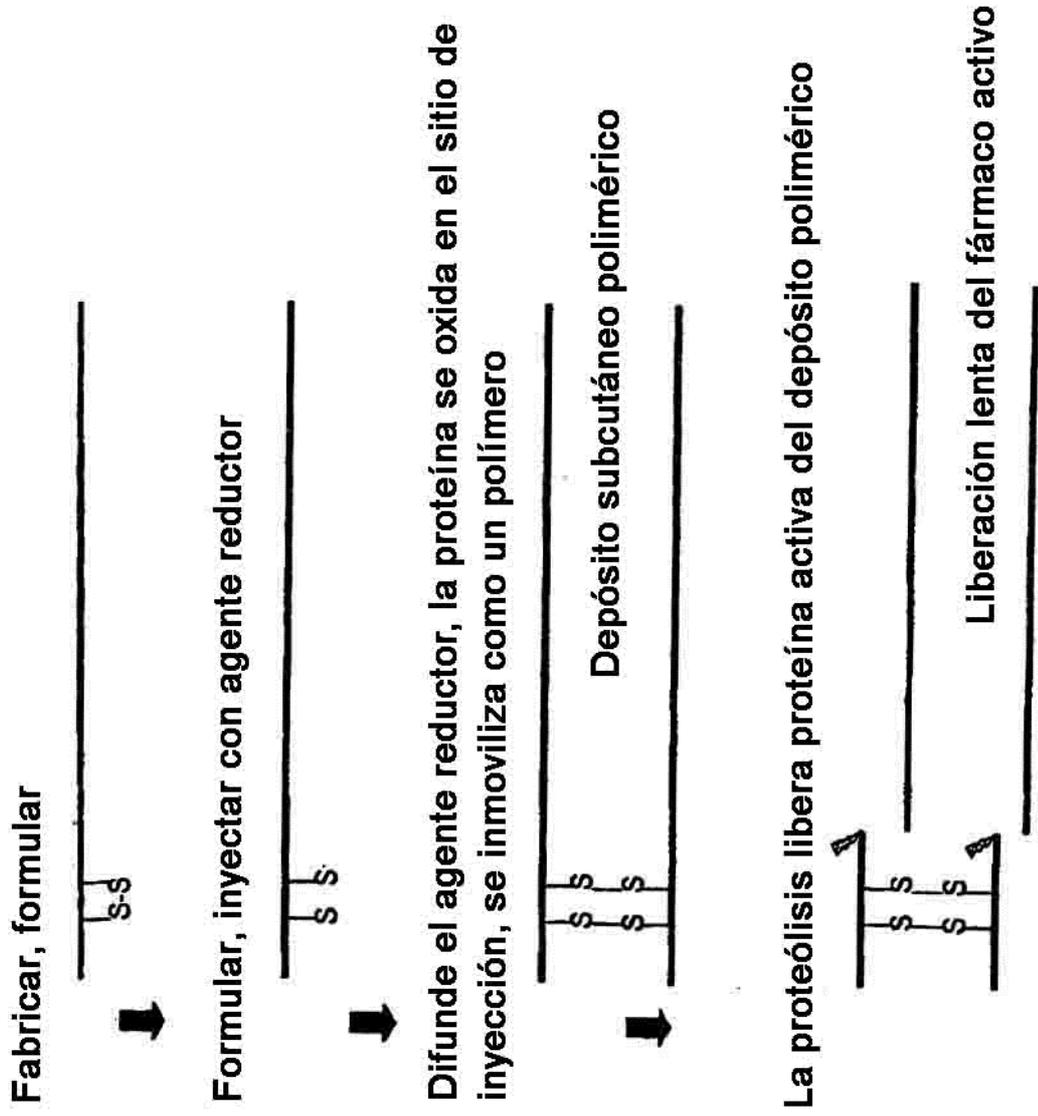
G S G G E G G S G G E G G S G G E G S G G S G G E G G S  
 GGTAGTGGTGAAGGAGGTTCTGGTGGAGAGGAGGTAGTGGAGGTGAAGGTGGATCCGGAGGAGAAGGAGGTAGT  
 G G E G G S G G E G G S G G E G G S G G E G G S G G E G G S G G  
 GGAGGTGAAGGAGGATCCGGAGGAGAAGGAGGTAGTGGTGAAGGAGGTTCTGGTGGAGAAGGAGGTAGTGGAGGT  
 E G G S G G E G G S G G E G G S G G E G G S G G E G G S G G E G  
 GAAGGTGGATCCGGAGGAGAAGGAGGTAGTGGAGGTGAAGGAGGTCCGGAGGAGAAGGAGGTAGTGGAGGTGAAGGT  
 G S G G E G G S G G E G G S G G E G G S G G E G G S G G E G G S  
 GGATCCGGTGGAGAAGGAGGTAGTGGAGGTGAAGGAGGTTCCGGTGGAGAAGGAGGTAGTGGAGGAGAGGTTGGATCT  
 G G E G G S G G E G G S G G E G G S G G E G G S G G E G G S G G  
 GGAGGAGAAGGAGGTAGTGGAGGAGAGGGTGGTCTGGAGGAGAAGGAGGTAGTGGTGGAGAGGTTGGATCTGGTGG  
 E G G S G G E G G S G G E G G S G G E G G S G G E G G S G G E G  
 GAAGGAGGTAGTGGAGGAGAAGGTTCTGGAGGAGAAGGAGGTAGTGGTGAAGGAGGTTCTGGTGGAGAAGGA  
 G S G G E G G S G G E G G S G G E G G S G G E G G S G G E G G S  
 GGTAGTGGAGGTGAAGGTGGATCCGGAGGAGAAGGAGGTAGTGGAGGTGAAGGAGGTCCGGAGGAGAAGGAGGTAGT  
 G G E G G S G G E G G S G G E G G S G G E G G S G G E G G S G G  
 GGTGGTGAAGGAGGTTCTGGTGGAGAAGGAGGTAGTGGAGGTGAAGGTGGATCCGGAGGAGAAGGAGGTAGTGGAGGT  
 E G G S G G E G G S G G E G G S G G E G G S G G E G G S G G E G  
 GAAGGAGGTCCGGAGGAGAAGGAGGTAGTGGAGGTGAAGGAGGTCCGGTGGAGAAGGAGGTAGTGGAGGTGAAGGA  
 G S G G E G G S G G E G G S G G E G G S G G E G G S G G E G G S  
 GGTCCGGTGGAGAAGGAGGTAGTGGAGGAGAGGGTGGATCTGGAGGAGAAGGAGGTAGTGGAGGAGAGGTTGGTCT  
 G G E G G S G G E G G S G G E G G S G G E G G S G G E G G S G G  
 GGAGGAGAAGGAGGTAGTGGAGGAGGTTGGATCTGGTGGAGAAGGAGGTAGTGGAGGAGAAGGAGGTCTGGAGGA  
 E G  
 GAAGGA

Fig. 16 Secuencia de rPEG\_H288

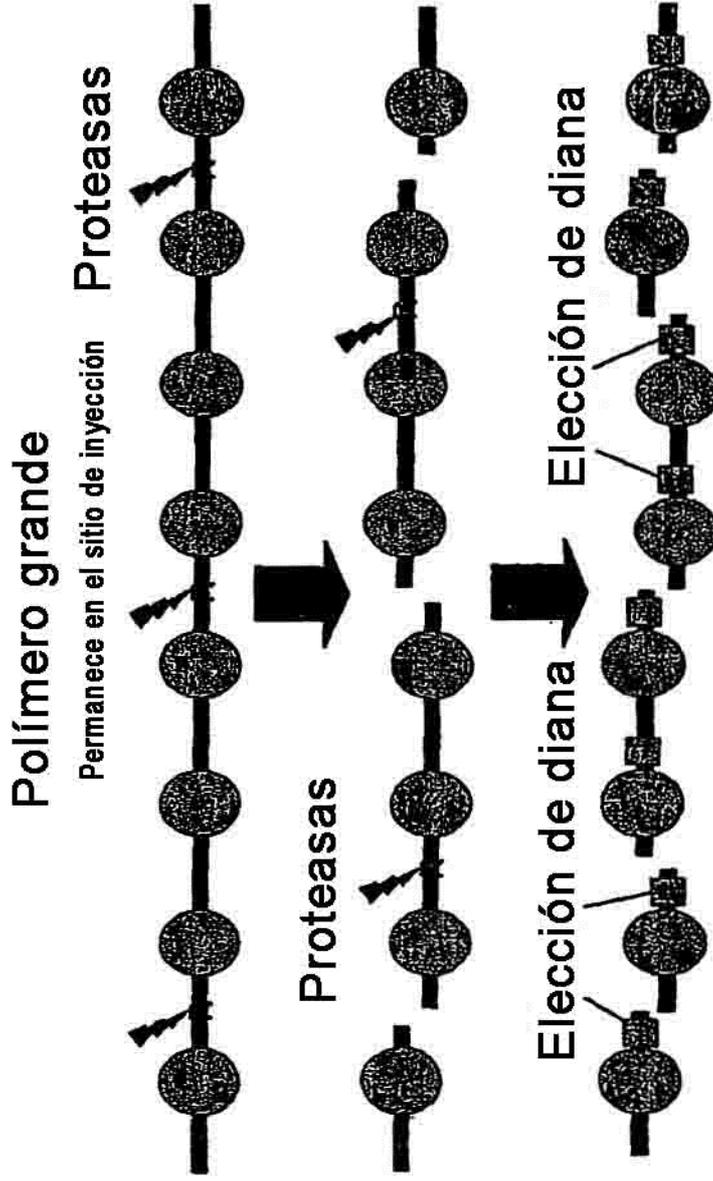
G S G G E G G S G G S G G S G G E G G S G G S G G S  
 GGTAGTGGTGGTGGATCCGGAGGAAGTGGAGGTAGTGGTGGACAGGAGGATCTGGTCCAAGTGCAGGTAGT  
 G G E G G S G G S G G S G G E G G S G G S G G S G G S G G  
 GGAGGTGAGGCAGGATCTGGTGAAGTGGAGGTAGTGGTGGTGGAGGTGGTCCGGAGGAAGTGGAGGTAGTGGAGGA  
 E G G S G G S G G S G G E G G S G G S G G S G G S G G E G  
 GAAGGTGGTCCGGTGGAAAGTGGAGGTAGTGGTGGAGAGGTGGATCTGGAGGAAGTGGAGGTAGTGGTGGAGGGT  
 G S G G S G G S G G E G G S G G S G G S G G S G G S G G S  
 GGTCCGGAGGAAGTGGAGGTAGTGGAGGAGAAAGGTGGTCCGGTGGAAAGTGGAGGTAGTGGTGGAGAGGTGGATCT  
 G G S G G S G G E G G S G G S G G S G G S G G E G G S G G  
 GGAGGAAGTGGAGGTAGTGGAGGAGGAGGATCTGGAGGAAGTGGAGGTAGTGGTGGTGAAGSAGGTTCCGGTGG  
 S G G S G G E G G S G G S G G S G G E G G S G G S G  
 AGTGCAGGTAGTGGTGGAGAGGAGGTCCCGAGGAAGTGGAGGTAGTGGTGGTGGAGGATCTGGTGGAGTGG  
 G S G G E G G S G G S G G S G G E G G S G G S G G S  
 GGTAGTGGAGGAGGGAGGTTCTGGTGAAGTGGAGGTAGTGGTGGTGGTCCGGTGGAACTGGAGGTAGT  
 G G E G G S G G S G G S G G E G G S G G S G G S G G  
 GGTGGTGAAGGAGGTTCTGGAGGAAGTGGAGGTAGTGGTGGAGAGGTGGTCCGGTGGAAAGTGGAGGTAGTGGAGGA  
 E G G S G G S G G S G G E G G S G G S G G S G G E G  
 GAAGGAGATCTGGAGGAAGTGGAGGTAGTGGTGGTGGAGGTTCCCGAGGAAGTGGAGGTAGTGGAGGAGAAAGGT  
 G S G G S G G S G G E G G S G G S G G S G G E G G S  
 GGTCCGGTGGAAAGTGGAGGTAGTGGTGGAGAGGTGGATCTGGAGGAAGTGGAGGTAGTGGTGGTGAAGGAGGTCC  
 G G S G G S G G E G G S G G S G G S G G E G G S G G  
 GGTGGAAGTGGAGGTAGTGGAGGTGAAGGTGGATCTGGTGAAGTGGAGGTAGTGGAGGTGAGGTTCCGGAGGA  
 S G  
 AGTGGGA



**Fig. 18 Derivados de liberación prolongada de MURP**



**Fig. 19 MURP manipulada para liberación lenta**



Liberación gradual de fragmentos del depósito del sitio de inyección en la sangre

**Fig. 20 Derivados de liberación prolongada de MURP basados en secuencias ricas en His**

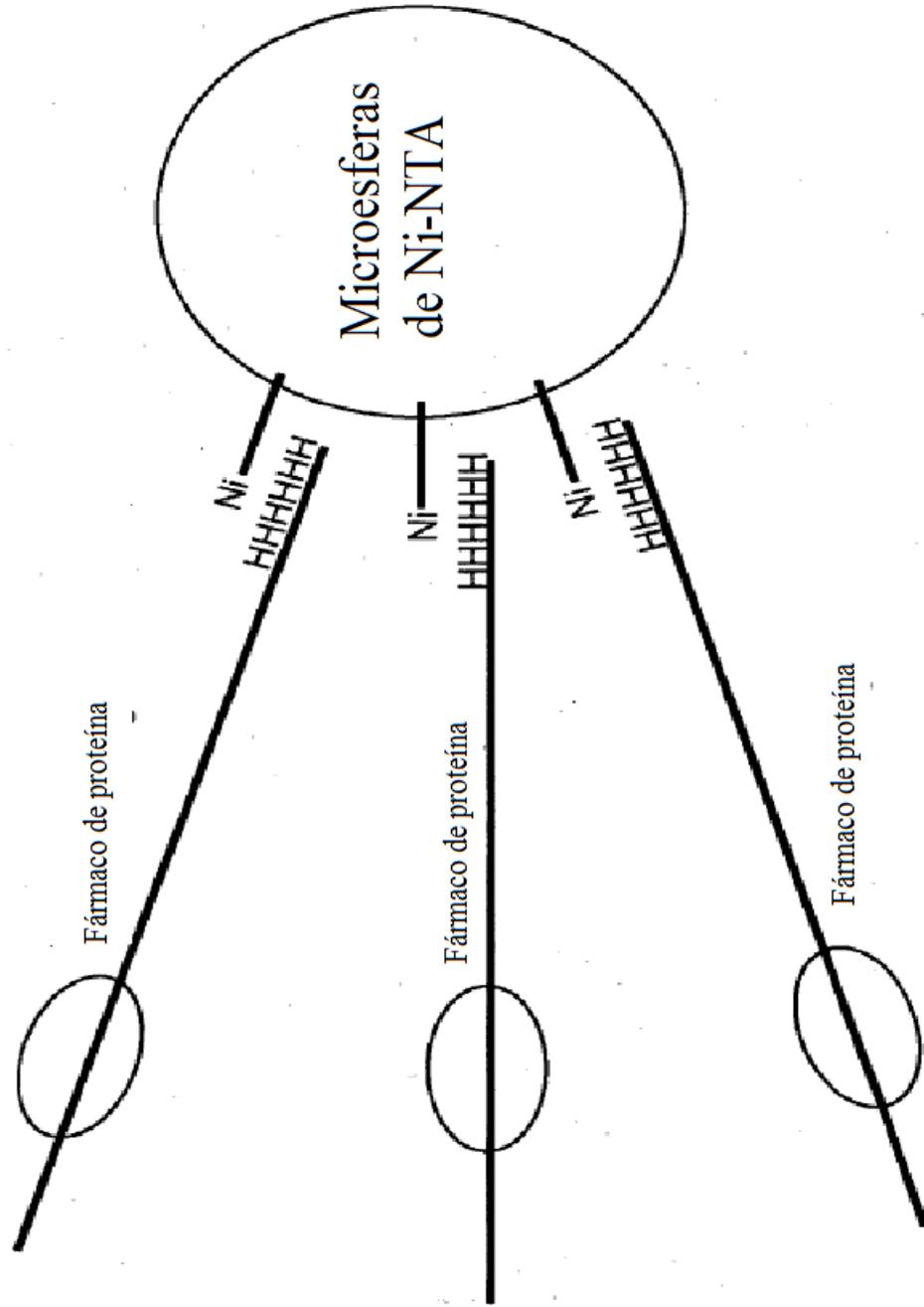
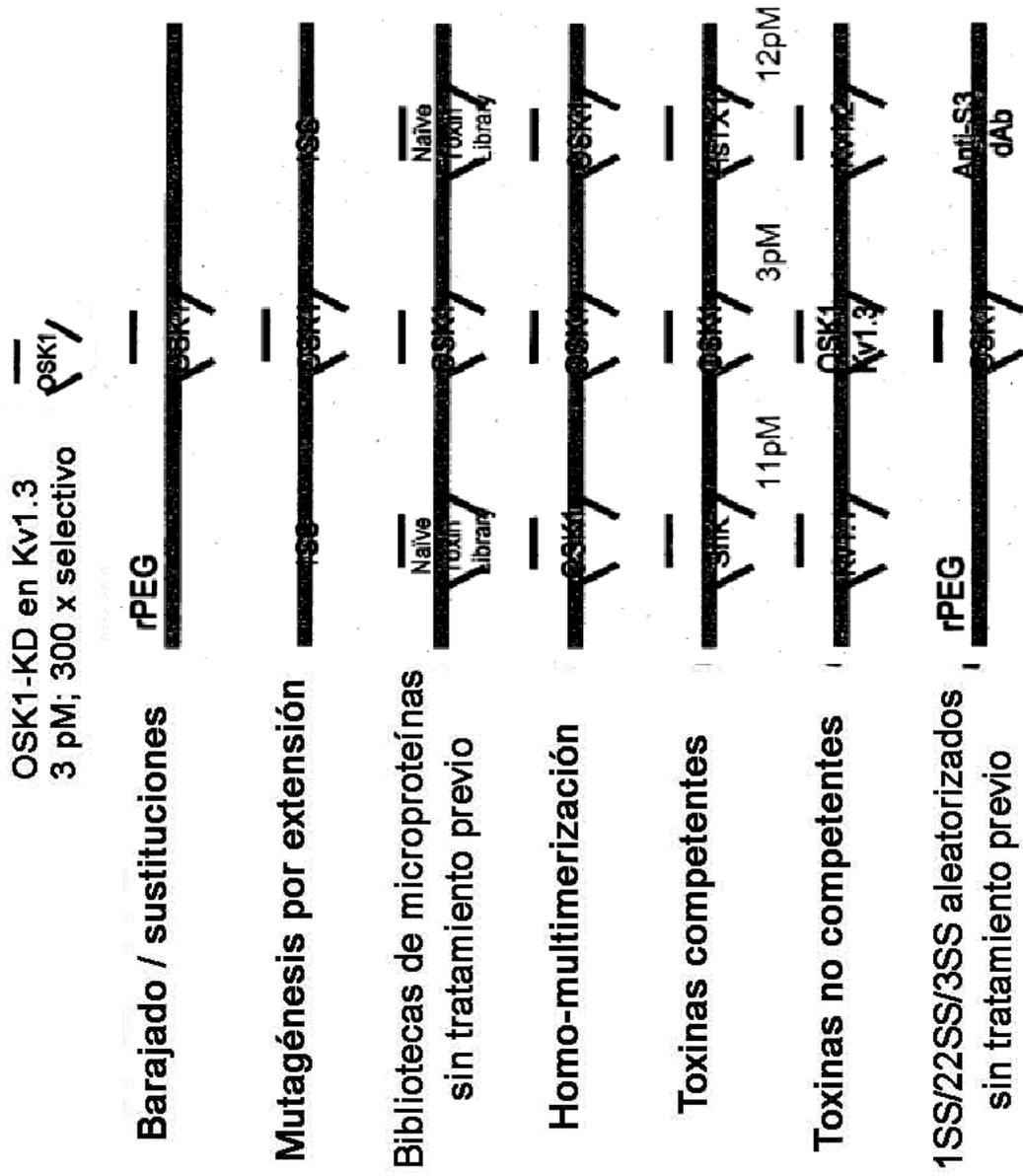
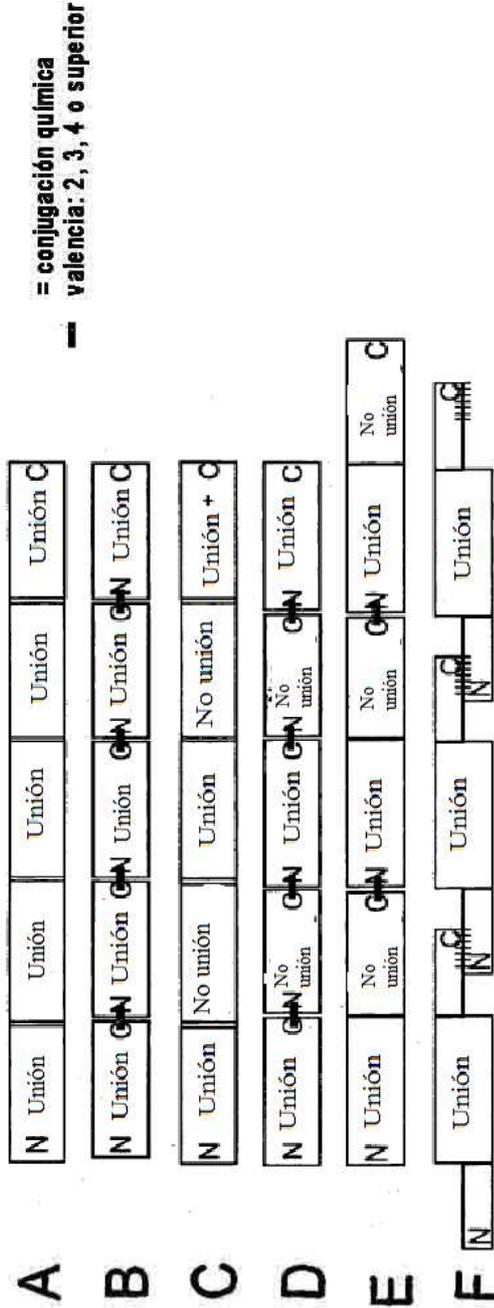




Fig. 22. Optimización de bloqueantes de canales de iones



**Fig. 23 Conceptos de semivida adicionales**



- A** Múltiples copias del motivo de unión en una única cadena de proteína. Las copias pueden ser iguales o diferentes.
- B** Las copias de un sitio de unión se expresan como proteínas separadas y se multimerizan por acoplamiento químico. Pueden usarse varios procedimientos de acoplamiento químico (añadir lista de Pierce de agentes de acoplamiento). Las copias pueden ser iguales o diferentes.
- C** Múltiples copias de un sitio de unión en una única cadena de proteína, pero separadas por ligadores de no unión.
- D** Copias de un sitio de unión y ligador de no unión se expresan como proteínas separadas y se multimerizan por acoplamiento químico. Pueden usarse varios procedimientos de acoplamiento químico (añadir lista de Pierce de agentes de acoplamiento). Las copias pueden ser iguales o diferentes.
- E** Copias de un sitio de unión y copias de un ligador de no unión se expresan cada una como proteínas separadas y se multimerizan por acoplamiento químico. Pueden usarse varios procedimientos de acoplamiento químico (añadir lista de Pierce de agentes de acoplamiento). Las copias pueden ser iguales o diferentes.
- F** Copias de un sitio de unión y ligador de no unión que contiene un sitio de auto-unión se expresan como proteínas separadas y se multimerizan por auto unión de las secuencias de péptidos. Pueden usarse varias secuencias de péptidos (añadir lista de referencias de péptidos). Las copias pueden ser iguales o diferentes.

Fig. 24 MURP universal

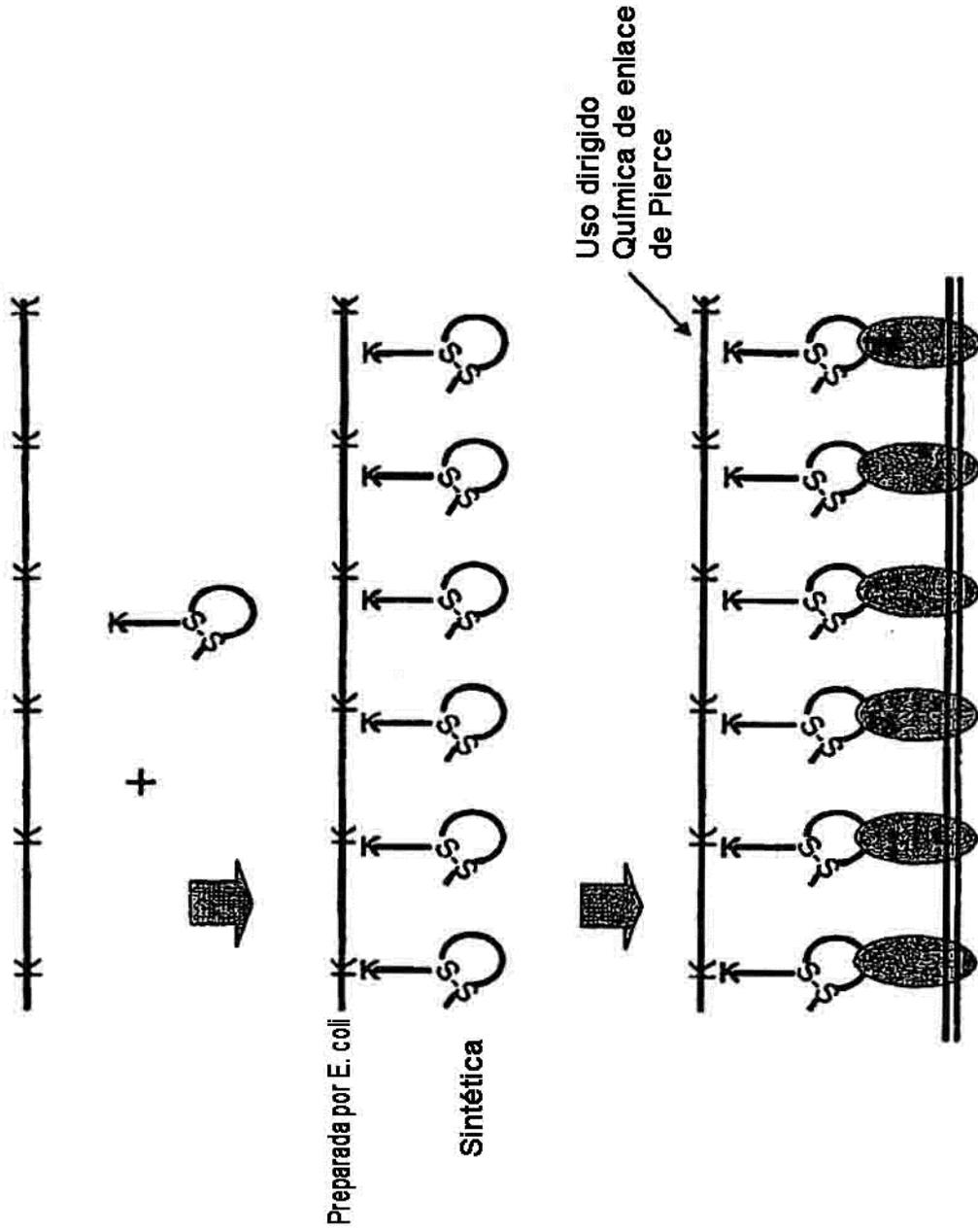
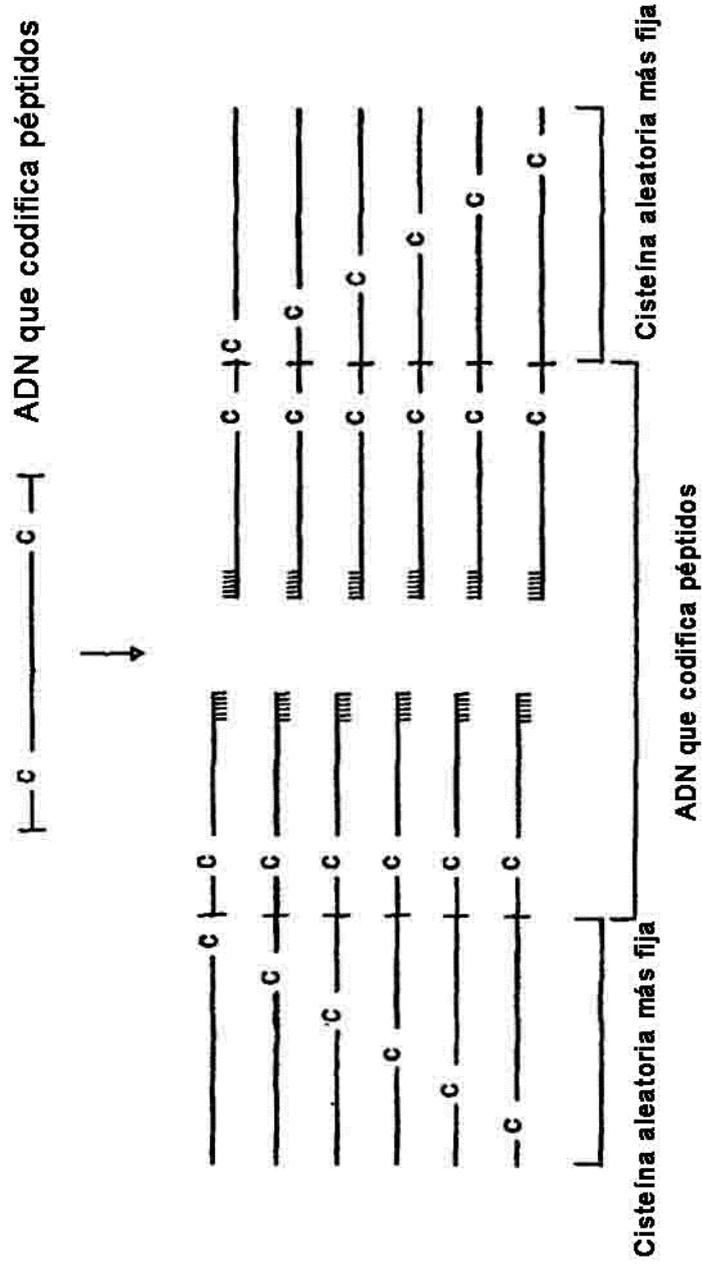


Fig. 25. Formación de módulos de unión 2SS



Reunir, hibridar y amplificar por PCR

Fig. 26. Formación de módulos de unión 2SS  
A

Péptido seleccionado y clasificado

```

C X X X X X C
L L L C C W W W W W C X X X X C L L L L
W W W C X X X X C L L L
W W W C X X X X C L L L
W W W C X X X X C L L L
W W W C X X X X C L L L
W W W C X X X X C L L L
    
```

**B**

Péptido seleccionado y clasificado

```

C X X X X X C
L L L C X X X C W W W C C L L L
L L L C X X X C W W W
L L L C X X X C W W W
L L L C X X X C W W W
L L L C X X X C W W W
    
```

**C**

Péptido seleccionado y clasificado

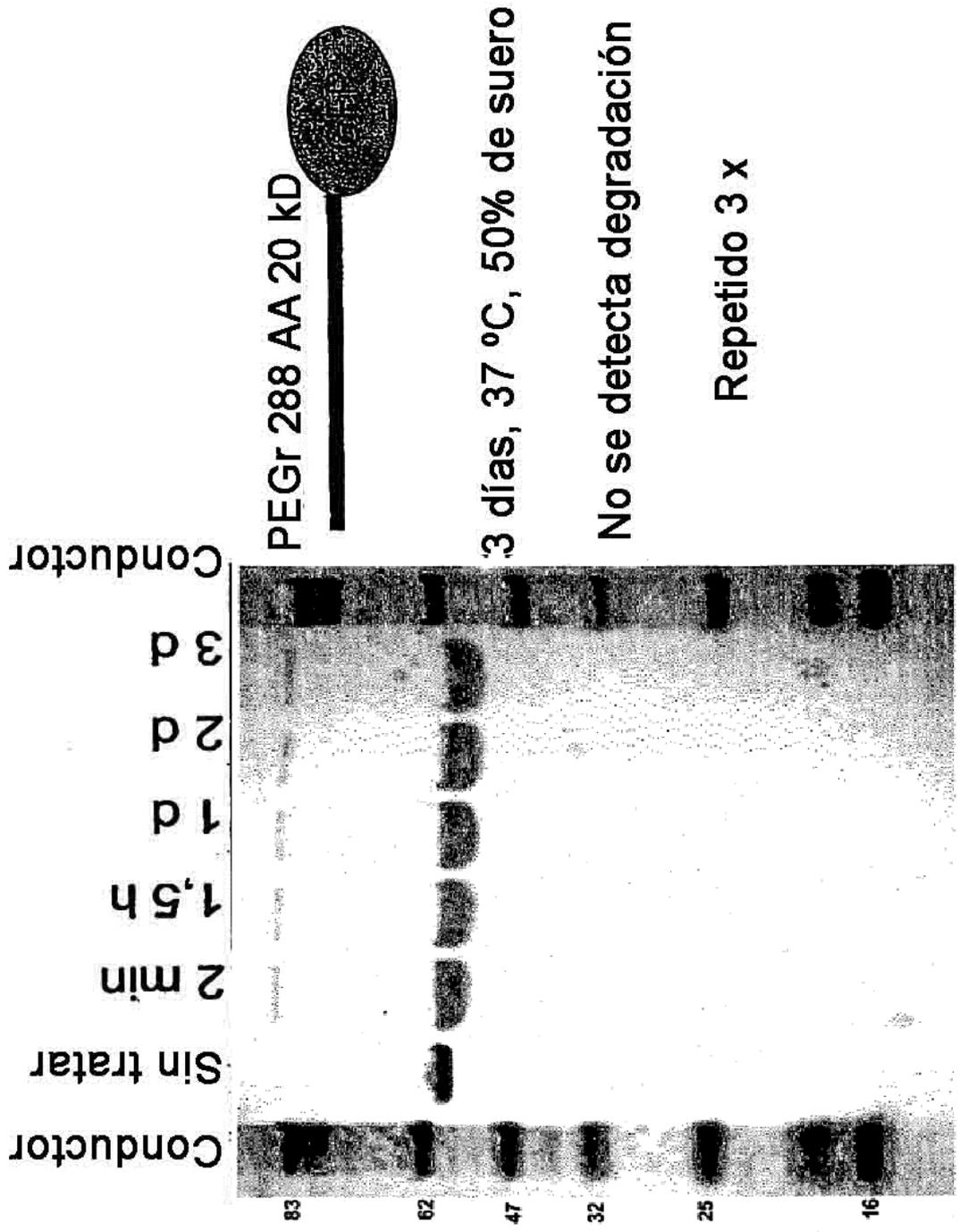
```

C X X X X X C
L L L C X C W W W W C X C L L L
L L L C X X C W W W W C X X C L L L
L L L C X X C W W W W C X X C L L L
L L L C X X C W W W W C X X C L L L
L L L C X X C W W W W C X X C L L L
L L L C X X C W W W W C X X C L L L
L L L C X X C W W W W C X X C L L L
L L L C X X C W W W W C X X C L L L
L L L C X X C W W W W C X X C L L L
    
```

W = base de Wobble para acomodar diversidad en el conjunto    L = Ligador



**Fig. 28. Estabilidad en plasma de rPEG\_J288**



**Fig. 29. Ausencia de anticuerpos preexistentes contra URP**

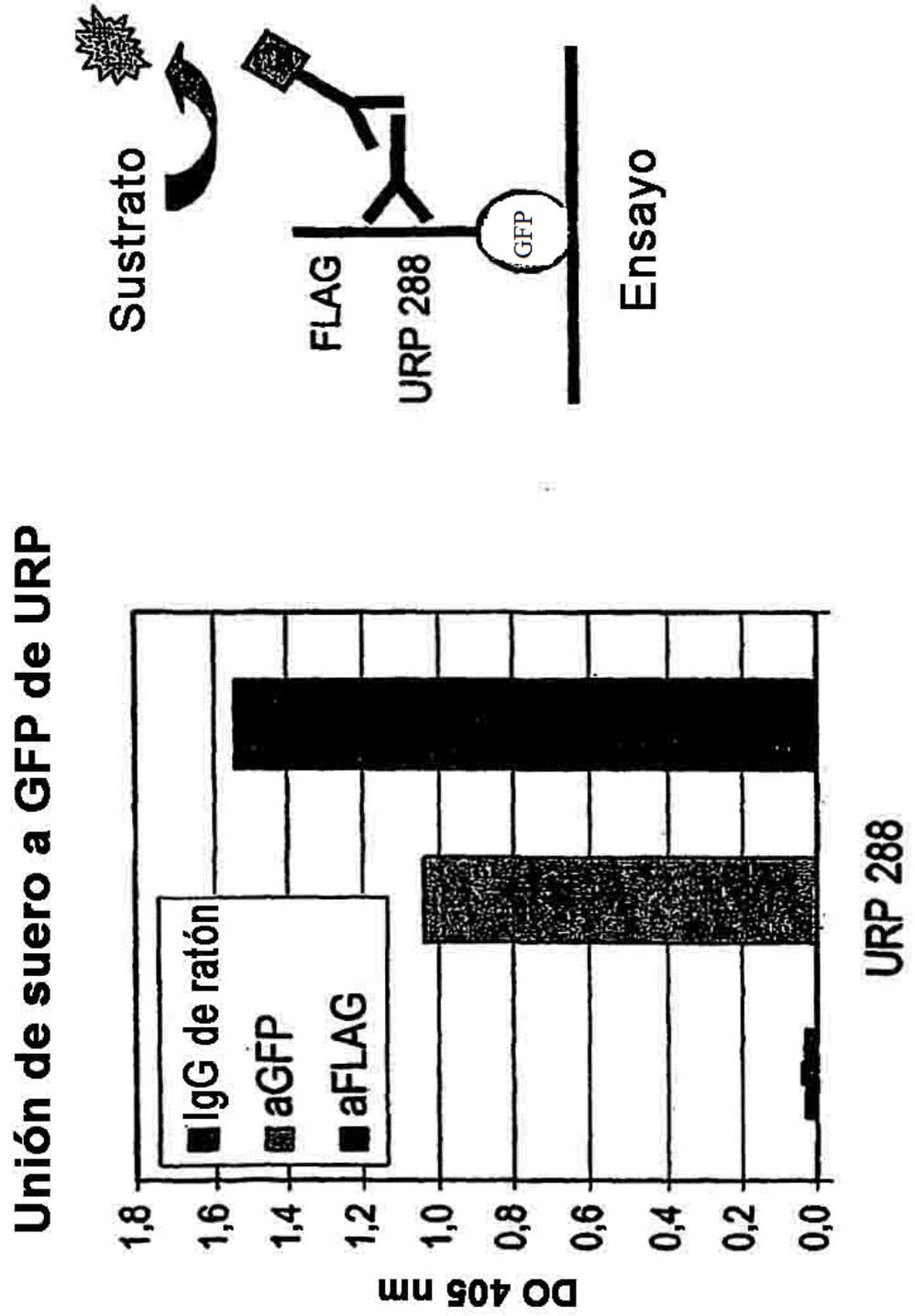
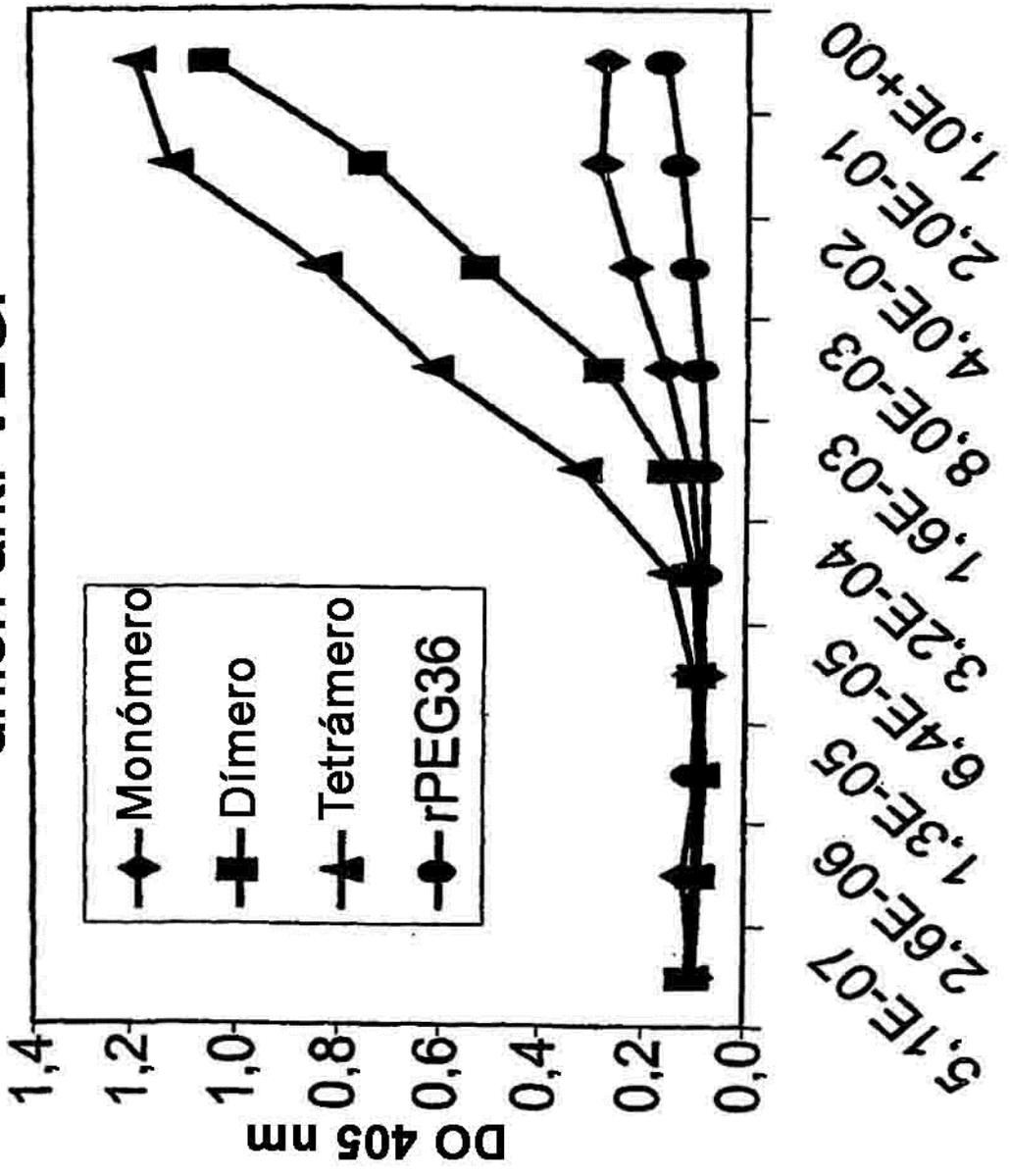


Fig. 30. Afinidad de unión de módulos de unión anti-VEGF





**Fig. 32. Formación de módulos de unión por  
adición de secuencias aleatorias**

**Monómero**

GGESGGESHTLECLGNICWVINOSSGSGGGSG  
 Módulo de unión

**Adición al extremo N**

Secuencia aleatoria SXXXXXXXXXX  
 Módulo de unión HTLECLGNICWVINOSSGSGGGSG

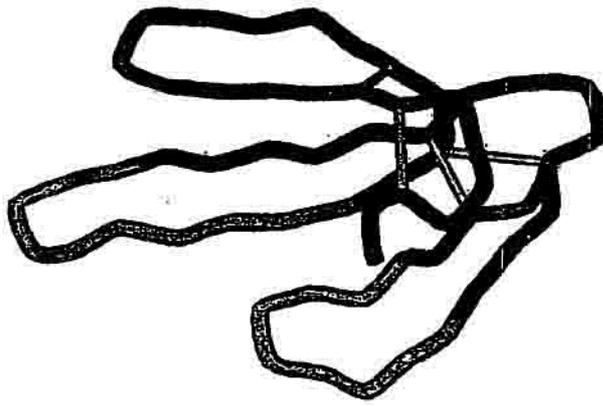
**Adición al extremo C**

GGESGGESHTLECLGNICWVINOSSGSGGGSGS  
 Módulo de unión  
 Secuencia aleatoria XXXXXXXXXXXXX

**Adición al extremo N más C**

Secuencia aleatoria SXXXXXXXXXX  
 Módulo de unión HTLECLGNICWVINOSSGS-  
 Secuencia aleatoria XXXXXXXXXXXXX

**Fig. 33 Estructura de toxina de tres dedos (3FT)**



Andamiaje de 3FT corto (NXSA\_LATSE/22-83):  
 RICFNHQSSQPQTTKTCSPESSCYNKQWSDFRGTIERGGCGCP  
 TVKPGIKLSCCESEVCNN

CFNHQSQPQTTKTCSPESSCYNKQWSD---F-RGTIIERGG--CPTVKPGIKLSCCESEVC  
 CHNQQSQPPTIKTCEGQ--CYKKTWRD---H-RGTISERGG--CPTVKPGIHI SCCASDKC  
 CYKQRSQFPITTVCPGEKN-CYKQWSG---H-RGTIIERGG--CPSVKKGIEINCCTTDKC  
 CHNQQSQPPTTKSCPGDTN-CYNKRWRD---H-RGTIIERGG--CPTVKPGINLKCCTTDRC  
 CYNQPSQHPPTTKACPEKN-CYRKQWSD---H-RGTIIERGG--CPTVKPGVKLHCCITEKC  
 CHNQQSQPPTTTCGGGETNCKKRWRD---H-RGYRTERGG--CPIVKNGIESNCCTTDRC  
 CHNQMSQPPTTRCSRWETNCKKRWRD---H-RGYKTERGG--CPTVKKGIQLHCCTSDNC  
 CFNQQSQPKTKSCPPGENSCYNKQWRD---H-RGSI TERGG--CPKVKPGIKLRCESEDC  
 CYNHQSTRATTKSC--EENSCYKKYWRD---H-RGTIIERGG--CPKVKPGVGIHCCQSDKC  
 CYNHQSTPATTKSC--VENSCYKSIWAD---H-RGTIIKRGCG--CPRVKS--KIKCCKSDNC  
 CYNQQSEAKTTTCSGGVSSCYKKTWSD---G-RGTIIERGG--CPSVKKGIERICCRTDKC  
 CLKQEPQPETTTCPEGEDACYNLFWSD---H-SEIKIEMGG--CPKTEPYTNLYCCKIDSC  
 CYSHKLQAKTKTC--EENSCY-KRSLP---KIPLIIIGRGCG--CPLTLPLFLRIKCCTSDKC  
 CYSHKTOPSATITC--EEKTCY-KKSVR---KLPAIVAGRGCG--CPSKEMLVAIHCCRSDKC  
 CUIHKALPRATKTC--VENTCY-KMFIR---TOREYISERGG--CPTAMWPHYQTECCCKGDRC  
 CYPHKSQAHTKSC--EGNTCY-KMFIR---TSREYISERGG--CPTAMWPHYQTECCCKGDRC

**Fig. 34. Diseño de una biblioteca de 3FT**

SCHX<sub>1</sub>X<sub>2</sub>X<sub>3</sub>X<sub>4</sub>X<sub>5</sub>X<sub>6</sub>X<sub>7</sub>X<sub>8</sub>AVTCTPPGENLCYRKMWX<sub>9</sub>X<sub>10</sub>X<sub>11</sub>X<sub>12</sub>X<sub>13</sub>X<sub>14</sub>X<sub>15</sub>X<sub>16</sub>X<sub>17</sub>X<sub>18</sub>X<sub>19</sub>GCAATCPSVKPYEEVTCSTDKC  
**G**

- X<sub>1</sub> – MNW (HIKLNQQRST)
- X<sub>2</sub> – VVK (AEGHKNPQRST)
- X<sub>3</sub> – 50% de MVA (KPQRT) y 50% de KMT (ADSY)
- X<sub>4</sub> – RVT (ADGNST)
- X<sub>5</sub> – MVW (HKNPQRST)
- X<sub>6</sub> – SCT (AP)
- X<sub>7</sub> – MNA (IKLPQRT)
- X<sub>8</sub> – RNM (ADEGIKNRSTV)
- X<sub>9</sub> – VKT (GILRSV)
- X<sub>10</sub> – SVT (ADGHPR)
- X<sub>11</sub> – RST((AGST)
- X<sub>12</sub> – WYA (ILST)
- X<sub>13</sub> – CST (PR)
- X<sub>14</sub> – GVG (AEG)
- X<sub>15</sub> – 50% de RRG (EGKR) y 50% de WHT (FINSTY)
- X<sub>16</sub> – RDA (EGIKRV)
- X<sub>17</sub> – DYA (AILSTV)
- X<sub>18</sub> – RDM (DEGIKNRSV)
- X<sub>19</sub> – MKG (LMR)

**Fig. 35. Alineamiento de secuencias de plexina**

CRHFQSCSQCLSA	PPFVQCGWCHDKCVRSEEC	LSGIWTQQIC	MET_HUMAN/519-562
CEHFQSCSQCLSA	PPFVQCGWCHDKCVRSEEC	PSGVWTQDVC	Q2IBC0_RHIFE/519-562
CEHFQSCSQCLSA	PPFVQCGWCHDKCVRLE	TCPSGAWTQEIC	Q2QLD2_CARPS/119-162
CEHFQSCSQCLSA	PPFVQCGWCHDKCVRSEEC	PSGSWTQETC	MET_OTOGA/520-563
CEHFQSCSQCLSA	PPFVQCGWCHDKCVRSEEC	PSGIWTQEIC	Q2VHX7_PIG/519-562
CEHFQSCSQCLSA	PPFVQCGWCHDRVHLEEC	PTGAWTQEVC	MET_CANFA/520-563
CGHFQSCSQCLSA	PPYFIQCGWCHNRCVHSNEC	PSGIWTQEIC	MET_RAT/520-563
CHHFQSCSQCLLA	PAPFMRGCGGQQCLRAP	ECNGGTQETC	Q90975_CHICK/519-562
CDHLTTCTSLVSS	RVTECGWCEGRCTRANQC	PPSVWTQEYC	Q9YGM7_FUGRU/528-571
CQHFLTCAVCLTAP	KFVCGWCSGVC	SWESDCDHHW-RNDSC	Q9YGN0_FUGRU/512-554
CQHFLTCAVCLTAP	QFMGCGWCSGVC	SWENQCDDRW-RNESC	Q4SR43_TETNG/490-532
CAHFRTCSMCLMAP	RFMNCGWCSGVC	SRQHECTSQW-TSASC	Q4RXB0_TETNG/431-473
CAHFRTCSMCLMAP	RFMNCGWCSGVC	SRQHCDMQW-EKDSC	Q9YGM5_FUGRU/509-551
CRHFLTQGRCLRAWH	FMGCGWCGNMCGQK	KECIPGSW-QQDHC	RON_HUMAN/526-568
CHHFLTQGSCLRAQR	FMGCGWCGMCGRQK	KECIPGSW-QQDHC	Q6DTW4_CANFA/118-160
CRHFLTQWRCLRAQR	FMGCGWCGDRCDRQK	KECIPGSW-QQDHC	RON_MOUSE/528-570
CRHFSTCDRCLRAER	FMGCGWCGNGCTRHH	ECAGPW-VQDSC	Q08757_CHICK/521-563

**Fig. 36. Diseño de bibliotecas de plexina**

Biblioteca del extremo N  
 $SCX_1HX_2X_3X_4CX_5X_6CLAX_7X_8X_9X_{10}X_{11}X_{12}X_{13}CGNCHDKVRSSECLSGINTQOICG$

$X_1$  - SRS (DEGHQR)  
 $X_2$  - TTM (FL)  
 $X_3$  - MNW (HIKLNQRST)  
 $X_4$  - AST (ST)  
 $X_5$  - DSG (AGRSIW)  
 $X_6$  - MDG (KLMQR)  
 $X_7$  - VXS (GILWRSV)  
 $X_8$  - KCT (AS)  
 $X_9$  - 75% de SMA (AEPQ) y 25% de TSG (SW)  
 $X_{10}$  - SVK (ADEGHPQR)  
 $X_{11}$  - KKT (FV)  
 $X_{12}$  - RYS (AIMTV)  
 $X_{13}$  - VRS (DEGHKQRS)

Biblioteca del extremo C  
 $SCRHFQSCQCLSPAFFVCGWCA_1X_2X_3CX_4X_5X_6X_7X_8CX_9X_{10}X_{11}(X_{12})WX_{13}X_{14}X_{15}X_{16}CG$

$X_1$  - VRT (DGHNRS)  
 $X_2$  - RRT (DGNS)  
 $X_3$  - RDG (EGKMRV)  
 $X_4$  - RBT (AGISTV)  
 $X_5$  - 75% de CRM (HQR) y 25% de TGG (W)  
 $X_6$  - 50% de KYE (SLA) y 50% de SAM (DEHQ)  
 $X_7$  - VAW (DEHKQ)  
 $X_8$  - SAM (DEHQ)  
 $X_9$  - VNT (ADHNPT)  
 $X_{10}$  - RVT (ADGNST)  
 $X_{11}$  - VRS (DEGHKQRS)

Fracccionar con  $X_{12}$  - DYT (AFISTV) y sin

$X_{13}$  - MVA (QRT)  
 $X_{14}$  - MAM (HONO)  
 $X_{15}$  - SAM (DEHQ)  
 $X_{16}$  - NRT (ADFHILNPSTVY)

## FIG. 37. Secuencias de cepas aisladas específicas para diana de bibliotecas basadas en plexina

### Biblioteca LMP031

SCXHXXCXXCLXXXXXXXXCGWCHDKCVRSEELSGTWTQQICG

### Biblioteca LMP032

SCRHFQSCSQCLSAPPFVQCGWCXXXCXHXWXXXCG

### Ligantes a DR4

SCHHFISCGRCLRSWHVVDGWCCHKCVRSEELSGTWTQQICG D4.09

SCRHFQSCSQCLSAPPFVQCGWCGDMCARVQQCHDR - WTHHACG D4.01

SCRHFQSCSQCLSAPPFVQCGWCHDKCGHQDECTAS - WRKEACG D4.03

### Ligantes a ErbB2

SCRHFQSCSQCLSAPPFVQCGWCRNMCVQEKQCDDSIWKNQHCG E2.39

SCRHFQSCSQCLSAPPFVQCGWCRDRCSRREDHCPTKTWRNHPCG E2.56

SCRHFQSCSQCLSAPPFVQCGWCNNVCSRHNDCDNN - WQHQNCG E2.57

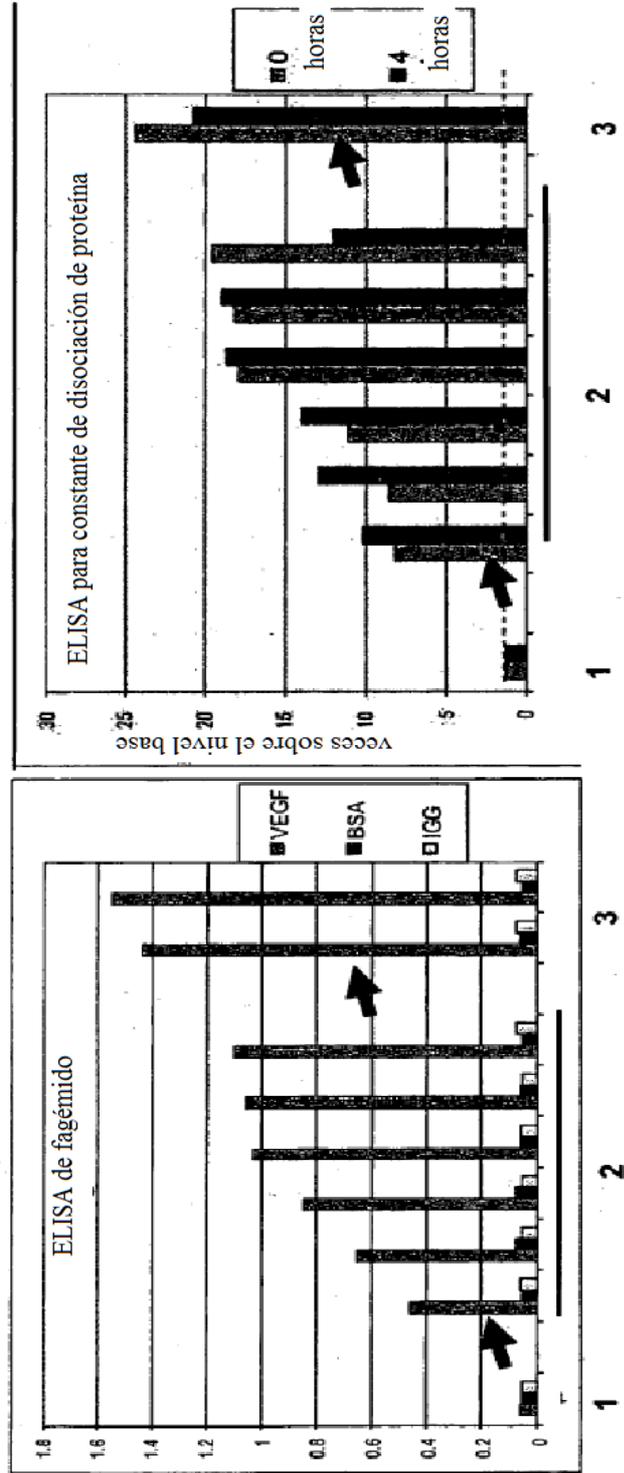
### Ligantes a HGFR

SCRHFQSCSQCLSAPPFVQCGWCNSMCGRAHDCTDH - WQKQHCG CM.33

SCRHFQSCSQCLSAPPFVQCGWCGNMCVRSEECTD - WRHDTCG CM.39

SCRHFQSCSQCLSAPPFVQCGWCNSMCGRAQDCNDR.TWKQHTCG CM.52

**Fig. 38. Unión y expresión de ligantes de 2SS VEGF resultantes de la formación**



**La maduración por afinidad da afinidad mejorada sin pérdida de especificidad**

**Fig. 39. Expresión y secuencias de ligantes de 2SS VEGF resultantes de bibliotecas de formación**

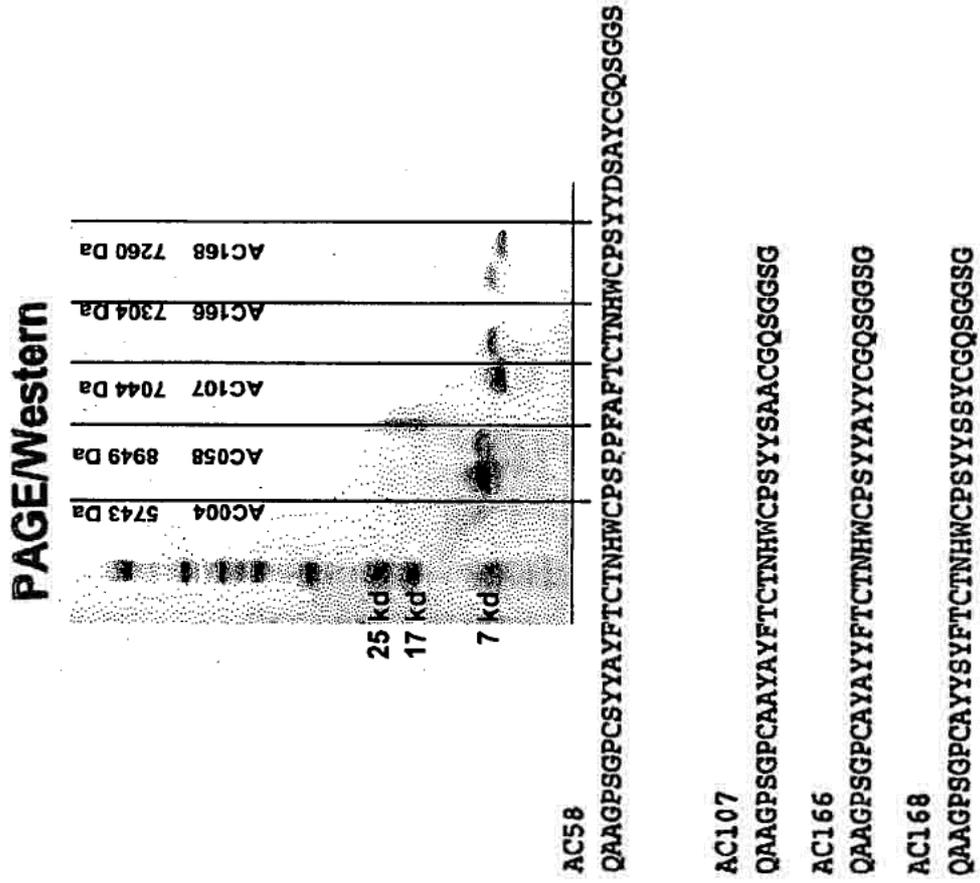
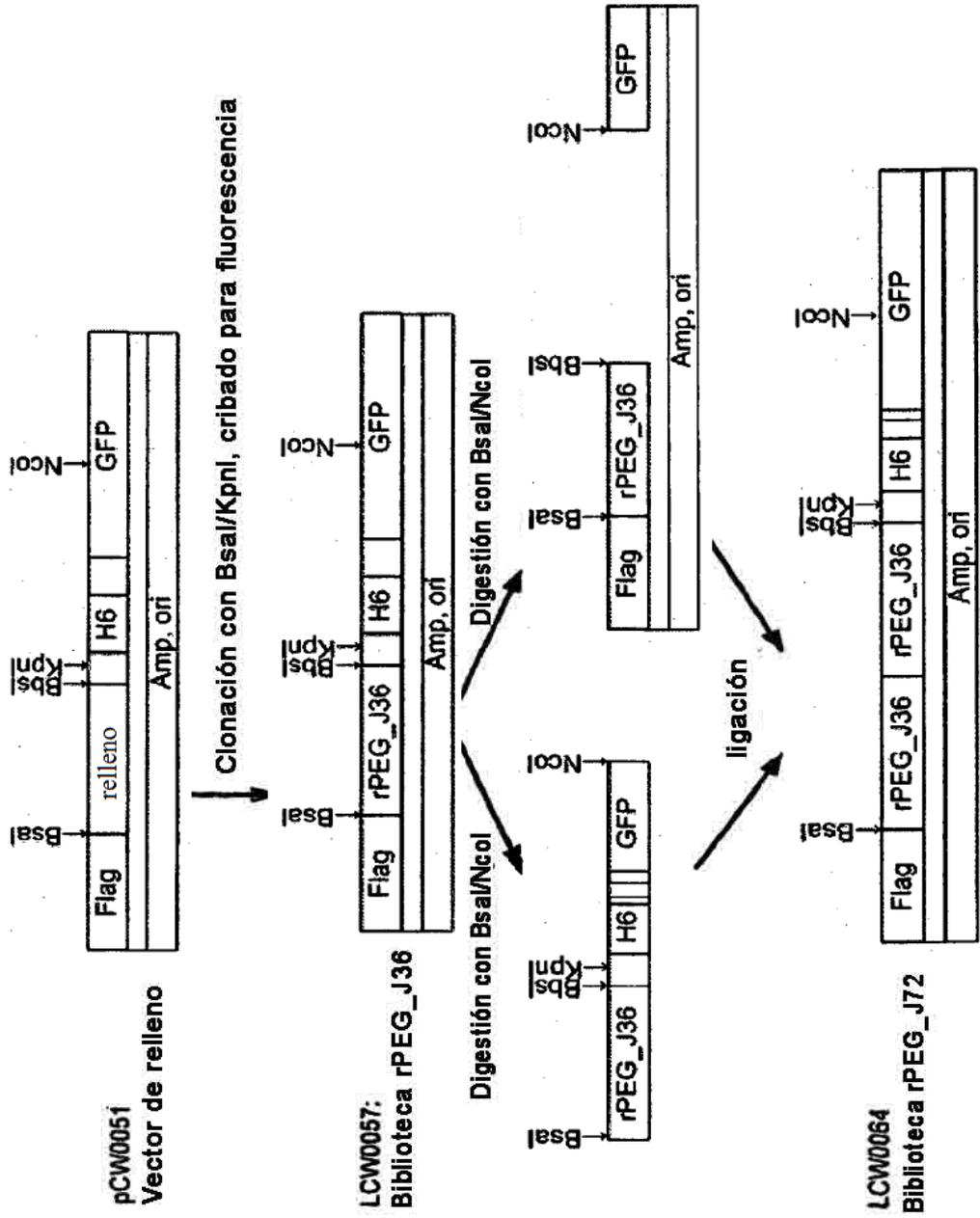
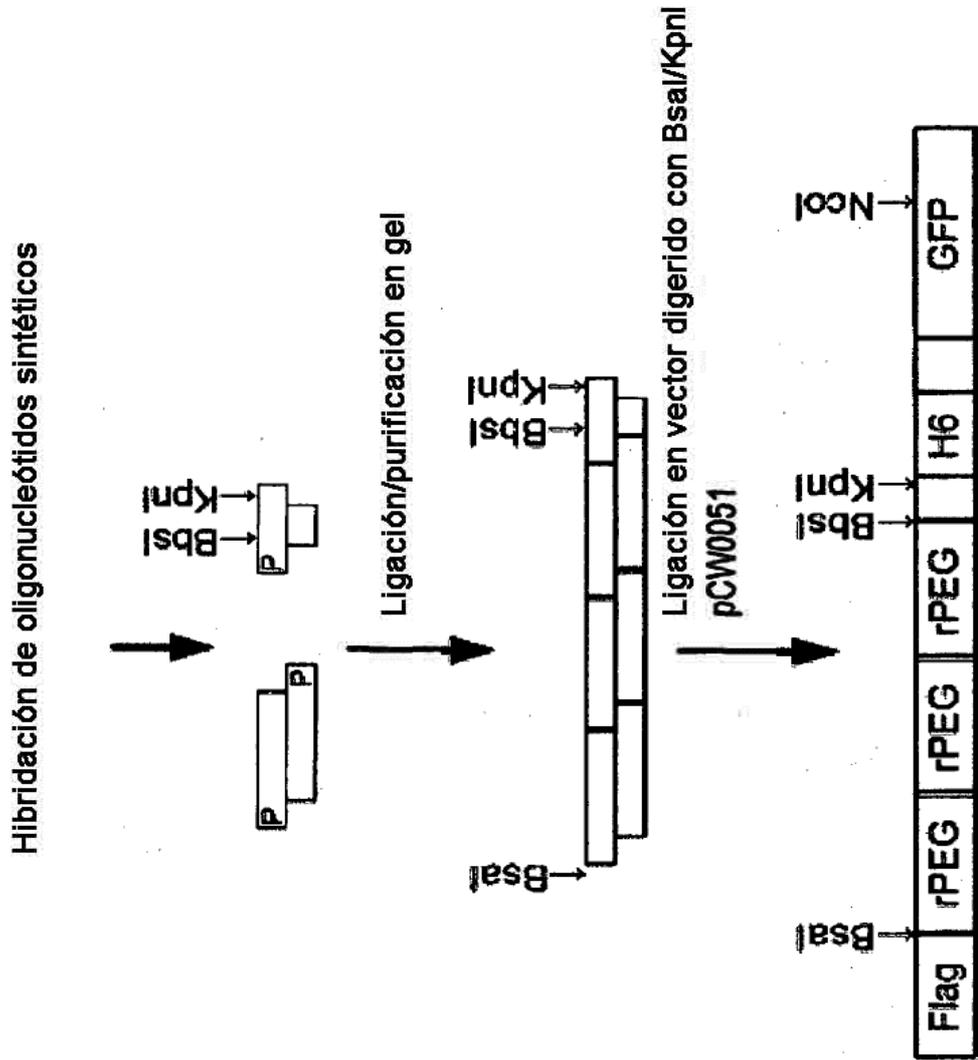


Fig. 40. Construcción de rPEG\_J72



**Fig. 41. Construcción de una biblioteca de codones de rPEG\_J36**

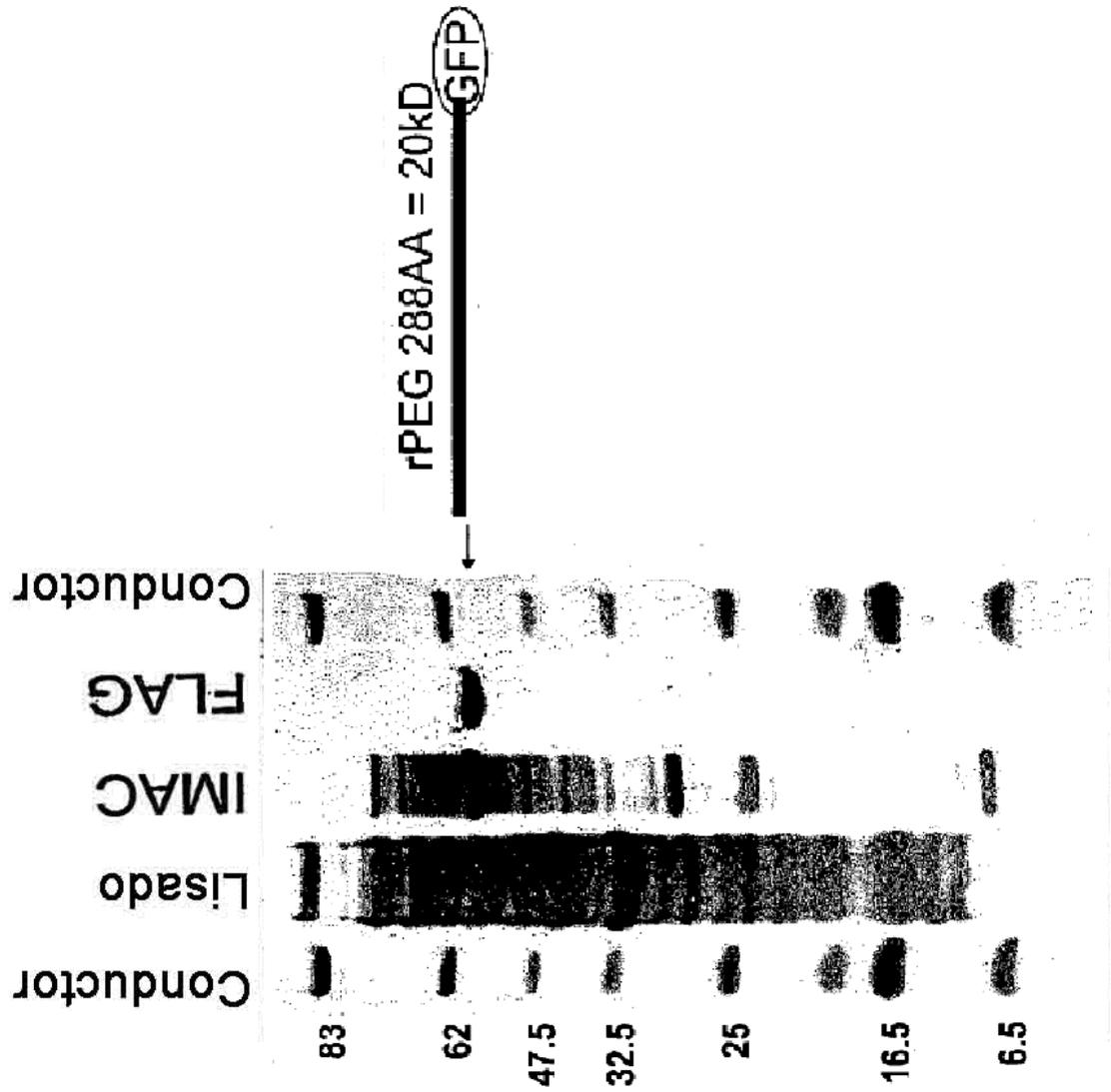


**Fig. 42. Diseño del vector de relleno pCW0051**

Flag	Bsal	BbsI
M D Y K D D D D K G S P G * * P R * * G G S S L E		
ATGGATTATAAGACGATGACGATAAAGGGTCTCCAGGTTAGTAACCTAGGTGATAGGGAGGTT <u>CGTCTTCACTCGAG</u>		
<b>KpnI</b>	<b>Marca 6 x His</b>	
G T H H H H H H E L V P V E K M		
<u>GGTACCCATCACCATCACCATCACGAGCTCGTACCGGTAGAAAAATG</u>		

Las secuencias de reconocimiento de los sitios de restricción están subrayadas. Los nucleótidos protuberantes que se generarán por digestión con Bsal y BbsI se muestran en cursiva. La figura ilustra que la digestión con Bsal y BbsI de pCW0051 genera nucleótidos protuberantes compatibles.

Fig. 43. Purificación de Flag-rPEG\_J288-H6-GFP





**Fig. 45. Expresión de proteínas de fusión entre rPEG\_J288 y módulos efectores humanos**

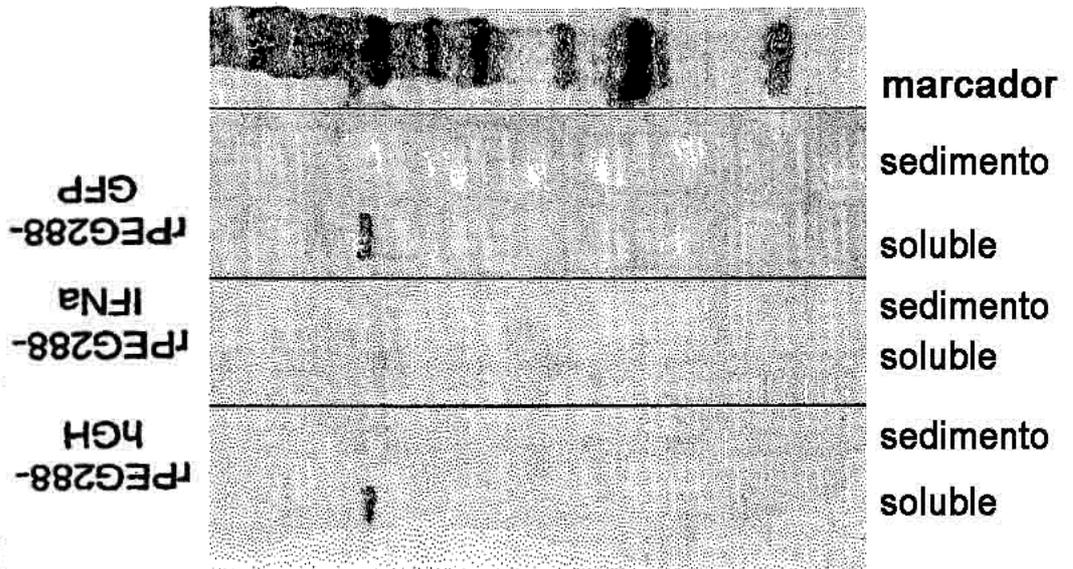


Fig. 46. Optimización de bloqueantes de canales de iones

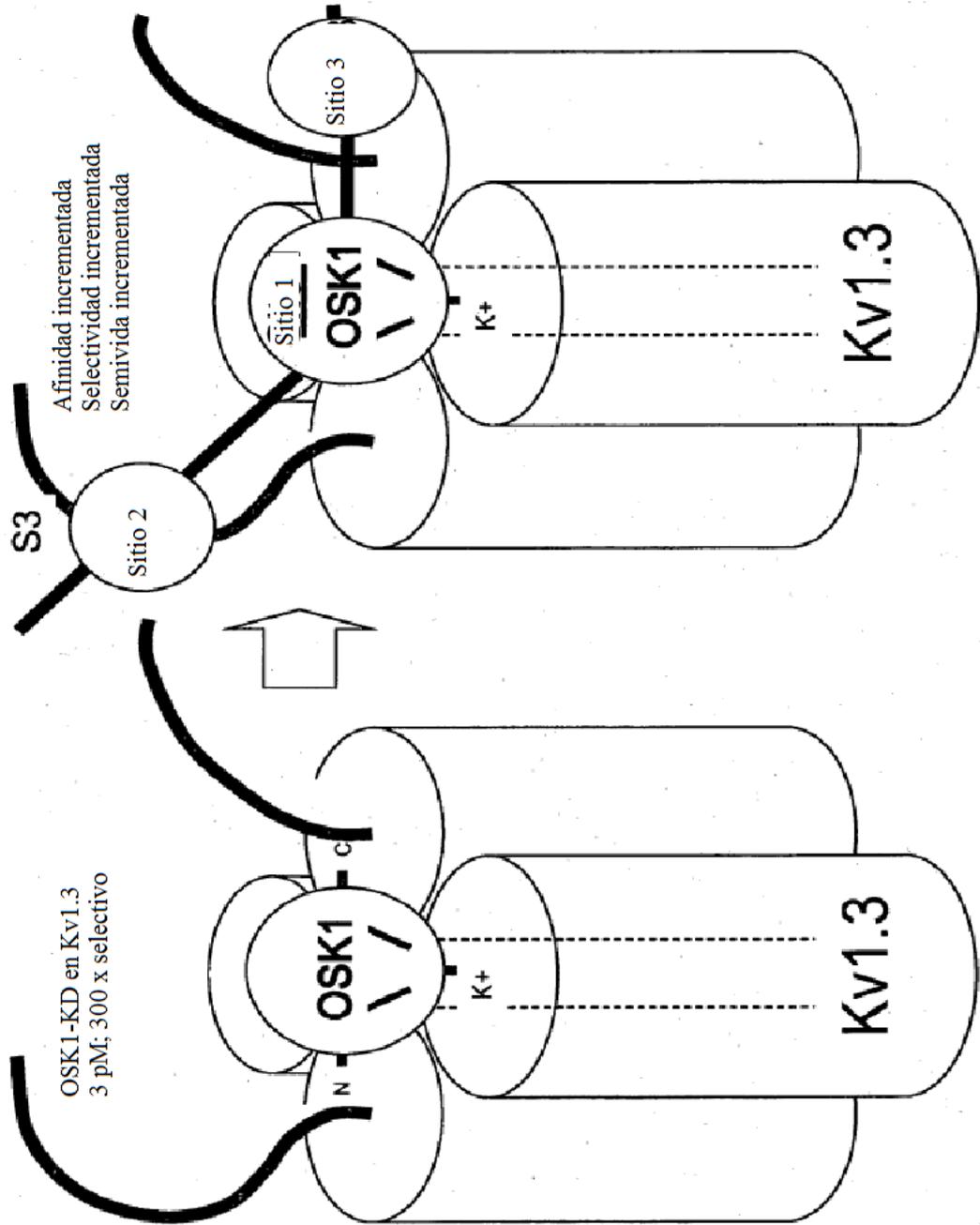


Fig. 47. Formatos de producto

