



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 362 424**

51 Int. Cl.:

C12Q 1/70 (2006.01)

C12Q 1/68 (2006.01)

G01N 33/53 (2006.01)

C07H 21/04 (2006.01)

C07K 7/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **04700617 .6**

96 Fecha de presentación : **07.01.2004**

97 Número de publicación de la solicitud: **1587907**

97 Fecha de publicación de la solicitud: **26.10.2005**

54

Título: **Biblioteca del dominio Kunitz.**

30

Prioridad: **07.01.2003 US 438491 P**

45

Fecha de publicación de la mención BOPI:
05.07.2011

45

Fecha de la publicación del folleto de la patente:
05.07.2011

73

Titular/es: **DYAX CORPORATION**
300 Technology Square, 8th Floor
Cambridge, Massachusetts 02139, US

72

Inventor/es: **Ladner, Robert, Charles y**
Nixon, Andrew

74

Agente: **Carvajal y Urquijo, Isabel**

ES 2 362 424 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

Biblioteca del dominio Kunitz

5 Antecedentes

La metodología de expresión de fagos puede ser utilizada para identificar ligandos de proteína que interactúan con un objetivo particular. Esta técnica utiliza partículas de bacteriófago como vehículos para el enlazamiento de ligandos de proteína candidatos con los ácidos nucleicos que los codifican. El ácido nucleico codificador se empaqueta dentro del bacteriófago, y generalmente la proteína codificada sobre la superficie del fago. La metodología de expresión de fagos está descrita, por ejemplo, en Ladner et al., Patente Estadounidense No. 5.223.409; Smith (1985) Science 228: 1315 - 1317; WO 92/18619; WO 91/17271; WO 92/20791; WO 92/15679; WO 93/01288; WO 92/01047; WO 92/09690; WO 90/02809; de Haard et al. (1999) J. Biol. Chem 274: 18218 - 30; Hoogenboom et al. (1998) Immunotechnology 4: 1 - 20; y en Hoogenboom et al. (2000) Immunol Today 2: 371 - 8. Otros métodos de evaluación de proteína, por ejemplo, arreglos de proteína, pueden ser utilizados también para identificar ligandos útiles.

Se pueden utilizar una variedad de armazones de proteína como molde para identificar ligandos útiles. Los armazones de proteína útiles pueden incluir posiciones de aminoácidos que contribuyen a una estructura estable del armazón de proteína y a otras posiciones que pueden variarse para producir un sitio de enlazamiento que interactúa con un objetivo.

Además, US 6 423 498 B1 se relaciona con una biblioteca variada de péptidos con el dominio Kunitz y con los usos del mismo.

También, Markland W et al. (Biochemistry 18 Jun 1996, vol. 35, no. 24) divulga la "optimización repetitiva de inhibidores de proteasas de alta afinidad utilizando expresión de fagos. 1. Plasmina".

Además, Clackson T y Wells J (Trends in Biotechnology, vol. 12, 1994, página 173) describe la "Selección *in vitro* de bibliotecas de péptido y de proteína".

Resumen

Se divulgan bibliotecas, vectores, partículas de fago, células huésped, y métodos para presentar un dominio Kunitz. Una biblioteca de ejemplo es una biblioteca de fagos que incluye partículas de fago que contienen suficientes genes de fago para producir partículas de fago y una proteína de expresión que tiene un dominio Kunitz modificado. Un dominio Kunitz puede variar en al menos dos bucles de interacción con respecto a otros miembros de la biblioteca. En una modalidad, se pueden presentar dominios Kunitz modificados sobre un fago filamentoso en una valencia baja. En una modalidad, la valencia baja es proporcionada por un ácido nucleico de fago que incluye una secuencia que codifica la proteína de expresión fusionada a un dominio funcional de una proteína de recubrimiento menor y una secuencia que produce una proteína homóloga que también incluye el dominio funcional.

En un aspecto la invención se caracteriza por una biblioteca que incluye una pluralidad de partículas filamentosas de fago como se define en las reivindicaciones. Cada partícula de fago de la pluralidad incluye (i) una proteína de expresión que comprende un dominio Kunitz, (ii) un ácido nucleico que incluye (a) opcionalmente, genes de fago suficientes para producir una partícula infecciosa de fago y (b) una secuencia que codifica la proteína de expresión. El dominio Kunitz incluye al menos una posición variada de aminoácido en cada uno de los dos bucles de interacción. Las posiciones en los bucles respectivos se modifican entre las partículas de la pluralidad. El número promedio de copias del dominio Kunitz por partículas de fago de la pluralidad es menor a 2,0.

El dominio Kunitz puede ser al menos 70, 75, 80, 85, 90, 95, 97, 98, 99, ó 100% idéntico a un dominio Kunitz humano (por ejemplo, LACI-K1) en posiciones de aminoácidos que no se modifican. En una modalidad, se modifican entre una y quince, por ejemplo, cinco y doce posiciones de aminoácidos entre las partículas de la pluralidad. Por ejemplo, se modifican al menos las posiciones de aminoácidos correspondientes a las posiciones de los

aminoácidos 16, 17, 18, 19, 34, y 39 de LACI-K1 humano entre las partículas de la pluralidad. En otro ejemplo, se modifican al menos las posiciones de los aminoácidos correspondientes a las posiciones de los aminoácidos 11, 13, 15, 16, 17, 18, 19, 34, y 39 de LACI-K1 humano entre las partículas de la pluralidad. En aún otro ejemplo, se modifican al menos seis, siete u ocho posiciones de aminoácidos correspondientes a las posiciones de los aminoácidos 11, 13, 15, 16, 17, 18, 19, 34, 39, 40, y 46 de LACI-K1 humano entre las partículas de la pluralidad. Únicamente se pueden variar Las posiciones de los aminoácidos correspondientes a las posiciones de los aminoácidos 11, 13, 15, 16, 17, 18, 19, 34, 39, y 40 de LACI-K1 humano entre las partículas de la pluralidad, o únicamente 16, 17, 18, 19, 34, y 39, o únicamente 11, 13, 15, 16, 17, 18, 19, 34, y 39, u otras combinaciones.

10 Cada dominio Kunitz en la pluralidad puede incluir al menos 75, 80, 85, 90, 95, 97, 98, 99, ó 100% de los dominios Kunitz en la pluralidad puede incluir residuos Kunitz conservados o residuos Kunitz altamente conservados en algunas (por ejemplo, 50, 60, 80, ó 90%) o en todas las posiciones modificadas. Cada dominio Kunitz en la pluralidad puede incluir al menos 75, 80, 85, 90, 95, 97, 98, 99, ó 100% de los dominios Kunitz en la pluralidad puede incluir residuos Kunitz conservados o residuos Kunitz altamente conservados en algunas (por ejemplo, 50, 60, 80, ó 90%) o en todas las posiciones que no varían. En una modalidad, los aminoácidos en las posiciones 32 y 46 no varían.

20 El grado de variaciones puede diferir en diferentes posiciones modificadas. Por ejemplo, las posiciones de los aminoácidos correspondientes a la posición del aminoácido 40 pueden ser modificadas entre G y A. Si no se modifica la posición 40, entonces puede ser restringida a G o A, por ejemplo. Se pueden modificar otras posiciones cambiadas, en formas diversas, entre todos los aminoácidos, todos los aminoácidos que no son cisteína, todos los aminoácidos excepto C y P, todos los aminoácidos excepto C, P y G, hidrófobos, alifáticos, hidrofílicos, y cargados.

25 En una modalidad, la proteína de expresión incluye un dominio funcional de una proteína menor de recubrimiento fusionada al dominio Kunitz, y los genes del fago incluyen un gen que codifica la proteína menor de recubrimiento que está (i') no fusionada a una secuencia de aminoácidos no viral de al menos cinco aminoácidos o (ii') no fusionada a una secuencia modificada de aminoácidos. Por ejemplo, los genes del fago pueden incluir una copia de tipo silvestre del gen que codifica la proteína menor endógena de recubrimiento del bacteriófago. Generalmente, los genes del fago pueden ser copias de tipo silvestre o variantes funcionales de los mismos.

30 Además de la pluralidad de partículas de fago caracterizadas anteriormente, pueden estar presentes otras partículas de fago, incluyendo por ejemplo, partículas inactivas que carecen de un componente de ácido nucleico.

35 En otro aspecto, se describe una biblioteca que incluye una pluralidad de partículas de fago. Cada partícula de fago de la pluralidad incluye (i) una proteína de expresión que contiene un dominio Kunitz y al menos una porción de la proteína de recubrimiento de fago del gen III, (II) una proteína funcional de recubrimiento de fago del gen III, y (iii) un ácido nucleico que incluye (a) opcionalmente, genes de fago suficientes para producir una partícula infecciosa de fago y (b) una secuencia que codifica la proteína de expresión. El dominio Kunitz puede incluir una secuencia que sea al menos 50, 60, 70, 80, 85, 87, 90, 92, 94, 95, 96, 97, 98, 100% idéntica a

40 MHSFCAFKADX₁₁GX₁₃CX₁₅X₁₆X₁₇X₁₈X₁₉RFFFNIFTRQCEEFX₃₄YGGCX₃₉X₄₀ QNRFESLEECKKMCTRDGA, en posiciones a parte de X (SEQ ID NO: 10). Por ejemplo,

X₁₁ es uno de: A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

X₁₃ es uno de: A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

X₁₅ es uno de: A, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, Y;

45 X₁₆ es uno de: A, G, E, D, H, T;

X₁₇ es uno de: A, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, Y;

X₁₈ es uno de: A, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, Y;

X₁₉ es uno de: A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

X₃₄ es uno de: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

50 X₃₉ es uno de: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y; y

X₄₀ es uno de: G, A.

Al menos dos de X₁₁, X₁₃, X₁₅, X₁₆, X₁₇, X₁₈, X₁₉, X₃₄, X₃₉, y X₄₀ pueden variar entre partículas de la pluralidad. Al menos dos de las posiciones modificadas están típicamente en el primero y en el segundo bucles de interacción del

dominio Kunitz de manera que la variación puede estar presente en ambos bucles. Si una posición X no varía, puede, por ejemplo, ser fijada a uno de los aminoácidos enlistados anteriormente para una posición respectiva. Si una posición X varía, puede variar entre las posiciones de los aminoácidos enlistados anteriormente, o entre otras posibles combinaciones.

5 En una modalidad, la proteína de expresión incluye al muñón de la proteína de recubrimiento del gen III. Los genes del fago incluyen un gen que codifica una proteína de recubrimiento del gen III de tipo silvestre.

10 En una modalidad, la biblioteca tiene una diversidad teórica de diversidad teórica de al menos 10^7 , 10^9 , 10^{10} , ó 10^{11} dominios Kunitz diferentes y/o menor a 10^{15} , 10^{14} , 10^{13} , 10^{12} , 10^{11} , ó 10^{10} dominios Kunitz diferentes. En una modalidad, la diversidad teórica está entre 10^5 - 10^{11} o entre 10^3 y 10^{15} dominios Kunitz diferentes.

15 En una modalidad, se modifican al menos las posiciones de aminoácidos 15, 16, 17, 18, 34, y 39. En otra modalidad, se modifican al menos las posiciones de los aminoácidos 11, 13, 15, 16, 17, 18, 19, 34, 39, y 40. En otra modalidad, se modifican únicamente esas posiciones, por ejemplo, únicamente 15, 16, 17, 18, 34, y 39, ó 11, 13, 15, 16, 17, 18, 19, 34, 39, y 40, ó 13, 15, 16, 17; 18, 19, 34, 39, y 40.

20 Se puede utilizar una biblioteca descrita aquí en un método para suministrar un ácido nucleico que codifique un dominio Kunitz que interactúe con un objetivo. El método incluye, por ejemplo, el suministro de la biblioteca; poner el contacto las partículas de fago de la biblioteca con un objetivo; y opcionalmente, recuperar el ácido nucleico de las partículas que interactúan con el objetivo. Por ejemplo, se inmoviliza el objetivo, y la etapa de recuperación incluye la separación de las partículas que interactúan con el objetivo de las partículas que no interactúan con el objetivo. Los ejemplos de objetivos incluyen proteasas.

25 En otro aspecto, la invención se caracteriza por un ácido nucleico que incluye un marco de lectura abierto y un promotor operativamente enlazado con el marco de lectura abierto. El marco de lectura abierto codifica una proteína de expresión que incluye: (i) un primer elemento que contiene un dominio Kunitz; y (ii) un segundo elemento que contiene una porción de una proteína de recubrimiento de fago, en donde la porción es suficiente para asociar físicamente la proteína de expresión con una partícula de fago.

30 El ácido nucleico puede ser un vector que contiene suficiente información genética para producir partículas infecciosas de fago en ausencia de un fago auxiliar. Por ejemplo, el ácido nucleico puede incluir un grupo de genes de fago suficiente para producir partículas infecciosas de fago. Por ejemplo, el ácido nucleico es un vector fago.

35 En una modalidad, el promotor es un promotor regulable, por ejemplo, un promotor inducible, por ejemplo, el promotor lac.

40 En una modalidad, la proteína de recubrimiento es una proteína menor de recubrimiento, por ejemplo, la proteína del gen III. La porción de la proteína de recubrimiento puede incluir una porción que se une físicamente a una partícula de fago, tal como un dominio ancla. En una modalidad preferida, la porción de la proteína de recubrimiento es el dominio ancla del gen III, o "muñón". En otra modalidad, la proteína de recubrimiento es una proteína de recubrimiento principal, por ejemplo, la proteína del gen VIII. En una modalidad preferida, el segundo elemento contiene la proteína de recubrimiento madura del gen VIII, de longitud completa. En una modalidad, la porción de la proteína de recubrimiento de (ii) se deriva de una entre: la proteína del gen IV, la proteína del gen VII, la proteína del gen IX.

45 En una modalidad, el dominio Kunitz tiene un residuo Kunitz conservado o un residuo Kunitz altamente conservado en al menos 70, 75, 80, 85, 90 ó 95% de las posiciones de los aminoácidos (o en todas las posiciones).

50 En una modalidad, el dominio Kunitz tiene al menos 30% de identidad con un dominio Kunitz de una proteína de ocurrencia natural, por ejemplo, con una proteína que incluye un dominio Kunitz mencionado aquí. El dominio Kunitz puede tener al menos 40%, 50%, 60%, 70 %, 80%, 90%, 95%, o 98% de identidad con un dominio Kunitz de una proteína de ocurrencia natural, por ejemplo, de un mamífero, por ejemplo, de un primate, por ejemplo, una proteína humana, por ejemplo, LAC-I.

Por ejemplo, al menos dos cisteínas están presentes en el dominio Kunitz, y se puede formar un disulfuro entre las cisteínas. Por ejemplo, si están presentes cuatro cisteínas, se forman dos disulfuros. Típicamente, están presentes seis cisteínas y se forman tres disulfuros entre las cisteínas.

5 El dominio Kunitz comprende la siguiente secuencia: X₁-X₂-X₃-X₄-C₅-X₆-X₇-X₈-X₉-X_{9a}-X₁₀-X₁₁-X₁₂-X₁₃-C₁₄-X₁₅-X₁₆-X₁₇-X₁₈-X₁₉-X₂₀-X₂₁-X₂₂-X₂₃-X₂₄-X₂₅-X₂₆-X₂₇-X₂₈-X₂₉-X_{29a}-X_{29b}-X_{29c}-C₃₀-X₃₁-X₃₂-X₃₃-X₃₄-X₃₅-X₃₆-X₃₇-C₃₈-X₃₉-X₄₀-X₄₁-X₄₂-X_{42a}-X_{42b}-X₄₃-X₄₄-X₄₅-X₄₆-X₄₇-X₄₈-X₄₉-X₅₀-C₅₁-X₅₂-X₅₃-X₅₄-C₅₅-X₅₆-X₅₇-X₅₈ (SEQ ID NO: 2). Al menos cuatro o seis de C₅, C₁₄, C₃₀, C₃₈, C₅₁, y C₅₅, son independientemente cisteína. Por ejemplo, todas las seis son cisteína. Si cuatro son cisteína, el resto de C₅, C₁₄, C₃₀, C₃₈, C₅₁, y C₅₅ pueden ser un aminoácido diferente de cisteína, o estar ausente.

10 Cada uno de X₁-X₄ es cualquier aminoácido, o estar ausente. Cada uno de X₆-X₁₃ es cualquier aminoácido pero preferiblemente no Cys. X_{9a} es cualquier aminoácido pero preferiblemente no Cys, o estar ausente. X₁₅-X_{29b} es cualquier aminoácido pero preferiblemente no Cys. Cada uno de X_{29a}, X_{29b}, X_{29c} es cualquier aminoácido, o estar ausente. Cada uno de X₃₁-X₃₇ es cualquier aminoácido pero preferiblemente no Cys. Cada uno de X₃₉-X₅₀ es cualquier aminoácido pero preferiblemente no Cys. Cada uno de X₅₂-X₅₄ es cualquier aminoácido pero preferiblemente no Cys. Cada uno de X₅₆-X₅₈ es cualquier aminoácido, o estar ausente.

15 En algunas modalidades, el número de residuos aminoácidos entre las cisteínas de un dominio Kunitz aumenta o disminuye en aproximadamente menos de 5 aminoácidos, por ejemplo, en 5, 4, 3, 2, ó 1 aminoácido. Por ejemplo, se pueden insertar o remover residuos en o entre X₆-X₁₃, X₁₅-X_{29c}, X₃₁-X₃₇, X₃₉-X₅₀, y X₅₂-X₅₄.

20 En una modalidad, en la cual ya sea C₁₄ o C₃₀ no es cisteína, entonces ninguna es cisteína. En una modalidad, en donde C₁₄ y X_{9a} están ausentes, X₁₂ puede ser G. En una modalidad, X₃₇ es G. En una modalidad, X₃₃ es F o Y. En una modalidad, X₄₅ es F o Y. Por ejemplo, al menos cuatro cisteínas están presentes y se pueden formar puentes Cys-Cys entre dos de C₅ y C₅₅, C₁₄ y C₃₈, y C₃₀ y C₅₁. Típicamente, están presentes seis cisteínas; y se pueden formar puentes Cys-Cys entre C₅ y C₅₅, C₁₄ y C₃₈, y C₃₀ y C₅₁.

25 El grupo de genes fago puede incluir un gen que codifique la proteína de recubrimiento del fago de (ii), que es además la copia que está enlazada con el dominio Kunitz (codificado en el segundo elemento). Por ejemplo, el grupo de genes puede incluir una copia de la proteína de recubrimiento de longitud completa. En una modalidad, el ácido nucleico que codifica la porción de la proteína de recubrimiento es (ii) contiene nucleótidos que han sido alterados para prevenir la recombinación con las secuencias relacionadas, por ejemplo, una copia del gen.

30 En una modalidad, cada gen del fago del grupo está operativamente enlazado con el promotor endógeno a cada gen.

35 En una modalidad, C₅, C₁₄, C₃₀, C₃₈, C₅₁, y C₅₅ están presentes, y X_{9a}, X_{29a}, X_{29b}, X_{29c}, X_{42a}, X_{42b}, están ausentes.

40 En otra modalidad, X₁₂ es G, X₃₃ es F, y X₃₇ es G

En una modalidad, el dominio Kunitz puede incluir una o más de las siguientes propiedades: X₂₁ es uno de F, Y, y W; X₂₂ es uno de F y Y. X₂₃ es uno de F y Y; X₃₅ es uno de Y y W; X₃₆ es uno de G y S. X₄₀ es uno de G y A. X₄₃ es uno de G y N. X₄₅ es uno de F y Y.

45 En una modalidad, el dominio Kunitz puede incluir una o más de las siguientes propiedades: X₁ es M; X₂ es H; X₃ es S; X₄ es F; X₆ es A; X₇ es F; X₈ es K; X₉ es A; X₁₀ es D; X₂₀ es R; X₂₁, X₂₂, X₂₃ son cada uno F; X₂₄ es N; X₂₅ es I; X₂₆ es F; X₂₇ es T; X₂₈ es R; X₂₉ es Q; X₃₁ es E; X₃₅ es Y; X₃₆ es G; X₄₁ es N; X₄₂ es Q; X₄₃ es N; X₄₄ es R; X₄₅ es F; X₄₇ es S; X₄₈ es L; X₄₉ y X₅₀ son cada uno E; X₅₂ y X₅₃ son cada uno K; X₅₄ es M; X₅₆ es T; X₅₇ es R; y X₅₈ es D.

50 En una modalidad, X₁₅, X₁₇, X₁₈, X₄₀, X₄₆ son cada uno cualquier aminoácido excepto prolina; y X₁₆ es uno de A, G, E, D, H, T.

En una modalidad, X₃₂ es E; X₃₄ es I; X₃₉ es E; X₄₀ es G; y X₄₆ es E.

En una modalidad, el ácido nucleico incluye además un marcador seleccionable (por ejemplo, un gen de resistencia a los antibióticos, el gen amp).

5 En otro aspecto, se describe una biblioteca que comprende una pluralidad de ácidos nucleicos, en donde cada ácido nucleico de la pluralidad puede incluir una o más de las características descritas aquí. En una modalidad, la secuencia del dominio Kunitz varía entre los ácidos nucleicos de la pluralidad.

10 La pluralidad puede contener ácidos nucleicos que codifican al menos 10^7 , 10^9 , 10^{10} , ó 10^{11} dominios Kunitzs diferentes y/o menos de 10^{15} , 10^{14} , 10^{13} , 10^{12} , 10^{11} , ó 10^{10} dominios Kunitzs diferentes. En una modalidad, la pluralidad contiene ácidos nucleicos que codifican al menos 10^5 - 10^{11} dominios Kunitzs diferentes. La pluralidad se puede caracterizar por una diversidad teórica de al menos 10^7 , 10^9 , 10^{10} , ó 10^{11} dominios Kunitzs diferentes y/o menos de 10^{15} , 10^{14} , 10^{13} , 10^{12} , 10^{11} , ó 10^{10} dominios Kunitzs diferentes. En una modalidad, la diversidad teórica está entre 10^5 - 10^{11} dominios Kunitzs diferentes.

15 En una modalidad, la invención se caracteriza por una biblioteca que contiene una pluralidad de ácidos nucleicos, en donde C₅, C₁₄, C₃₀, C₃₈, C₅₁, y C₅₅ están presentes, y X_{9a}, X_{29a}, X_{29b}, X_{29c}, X_{42a}, X_{42b}, están ausentes en cada ácido nucleico, y en donde la secuencia del dominio Kunitz varía al menos en dos de las posiciones correspondientes a X₃₂, X₃₄, X₃₉, X₄₀ y X₄₆ de (SEQ ID NO: 2) entre los miembros de la pluralidad.

20 En una modalidad, la secuencia del dominio Kunitz es invariante es una o más de las posiciones correspondientes a X₁₁, X₁₂, X₁₅, X₁₆, X₁₇, X₁₈, y X₁₉ (de la SEQ ID NO: 2). En otra modalidad, la secuencia del dominio Kunitz varía en una o más de las posiciones correspondientes a X₁₁, X₁₂, X₁₅, X₁₆, X₁₇, X₁₈, y X₁₉ de la SEQ ID NO: 2 entre los miembros de la pluralidad. En una modalidad, la secuencia del dominio Kunitz es invariante en una o más de las posiciones correspondientes a X₃₂, X₃₄, X₃₉, X₄₀ y X₄₆ de la SEQ ID NO: 2.

25 En una modalidad, la proteína de expresión incluye una pluralidad de dominios Kunitz, por ejemplo, una variedad de dominios Kunitz modificados, o un dominio Kunitz modificado y al menos otra secuencia modificada.

30 En otro aspecto, la invención se caracteriza por una célula huésped que comprende un ácido nucleico, en donde el ácido nucleico contiene una o más de las características descritas aquí. La célula huésped puede ser una célula bacteriana, por ejemplo, una célula de *E. coli*.

35 En otro aspecto, la invención se caracteriza por una biblioteca que comprende una pluralidad de células huésped, en donde cada célula huésped de la pluralidad contiene un ácido nucleico que tiene una o más de las características descritas aquí. En una modalidad, la secuencia del dominio Kunitz varía entre los ácidos nucleicos de la pluralidad.

40 En un aspecto, se divulga una partícula de fago que contiene un ácido nucleico, en donde el ácido nucleico puede incluir una o más de las características descritas aquí. En una modalidad, la partícula contiene el dominio Kunitz físicamente unido a la superficie.

45 En otro aspecto, la invención se caracteriza por una biblioteca que contiene una pluralidad de partículas de fago, en donde cada partícula de fago de la pluralidad comprende un ácido nucleico que puede incluir una o más de las características descritas aquí. En una modalidad, la secuencia del dominio Kunitz es diferente entre los ácidos nucleicos de la pluralidad.

50 En una modalidad, la pluralidad de las partículas de fago contiene al menos 10^3 partículas que incluyen cada una un ácido nucleico que codifica una proteína de expresión diferente. Preferiblemente, la pluralidad de partículas de fago comprende al menos 10^6 partículas que incluyen cada una un ácido nucleico que codifica una proteína de expresión diferente. Más preferiblemente, la pluralidad de partículas de fago comprende al menos 10^9 partículas que incluyen cada una un ácido nucleico que codifica una proteína de expresión diferente. Esta medición real de la diversidad puede ser menor a la diversidad teórica de la biblioteca.

En una modalidad, al menos 20%, 40%, 60%, ó 80% de las partículas de fago de la pluralidad de partículas de fago incluye la proteína de expresión de la respectiva partícula de fago unida físicamente a la partícula de fago. En una

5 modalidad, el número promedio de copias de la proteína de expresión es menor a 3, por ejemplo, menor a 2.5, 2.0, 1.7, 1.5, ó 1.2. Si, sin embargo, las partículas que no incluyen una proteína de expresión pero incluyen un componente apropiado de ácido nucleico están incluidas en la pluralidad, el número promedio de copias de la proteína de expresión puede ser menor a 2, 1.5, 1.2, 1.1, 1.0, ó 0.9. Por ejemplo, el número promedio de copias está entre 2.4 y 0.5, ó 1.8 y 0.5, ó 1.4 y 0.5.

10 En algunos casos, puede ser útil evaluar una pluralidad de partículas de fago que excluya partículas que no incluyan una proteína de expresión. Por lo tanto, el número promedio de copias podría no ser menor a uno. En tales casos, el número promedio de copias de la proteína de expresión es menor a 3, por ejemplo, menor a 2.5, 2.0, 1.7, 1.5, o 1.2, por ejemplo, entre 2.5 y 1.0 ó 1.5 y 1.0.

15 Una biblioteca de partículas de fago puede ser preparada en una composición líquida, por ejemplo, una composición acuosa. La composición puede suministrar un ambiente oxidante, por ejemplo, favoreciendo la formación de disulfuro dentro de los dominios Kunitzs. La composición puede ser no viscosa o tener una viscosidad suficientemente baja para permitir la manipulación del líquido (por ejemplo, pipeteando 10, 5, o menos microlitros). Cuando se seleccionan los dominios Kunitzs en forma aleatoria, la biblioteca puede incluir al menos 30, 40, 50, 70, 75, 80, 85, 90, ó 95% dominios que están plegados. Un método para determinar la fracción de dominios plegados es expresar los dominios Kunitzs en una célula con una etiqueta de epítipo, y llevar a cabo transferencias tipo Western sobre la fracción soluble de lisados crudos de las células. Niveles detectables de la etiqueta del epítipo (con el peso molecular apropiado) en las fracciones solubles indica que se produjo un dominio Kunitz plegado. Ver, por ejemplo, Davidson et al. (1994) Proc. Natl. Acad. Sci. USA. 91(6): 2146 - 50.

25 En otro aspecto, la invención se caracteriza por un método que incluye: (i) proporcionar la biblioteca que incluye una o más de las características descritas aquí, (ii) seleccionar un grupo de partículas de fago que se enlacen con un objetivo utilizando la proteína de expresión. El método puede ser utilizado para seleccionar un fago que codifique una proteína que enlace un objetivo a partir de una pluralidad de partículas de fago.

30 En diferentes modalidades, al menos 10%, 20%, ó 40%, más preferiblemente, al menos 60%, más preferiblemente, al menos 80% de las partículas de fago presentan la proteína de expresión sobre la superficie.

La selección puede incluir: (a) la formación de una mezcla que contiene las partículas de fago, un objetivo, y un soporte, y (b) la separación del fago que no se enlace con el objetivo a partir de los complejos objetivo inmovilizados por el fago.

35 En una modalidad, el objetivo es una proteasa, por ejemplo, una proteasa activa o una proteasa inactiva. Las proteasas inactivas incluyen proteasas con alteraciones de aminoácidos (por ejemplo, sustituciones, inserciones, o supresiones) que reduzcan parcial o completamente la actividad de la proteasa.

40 En un aspecto, se describe un método que incluye: (a) proporcionar una pluralidad de ácidos nucleicos, en donde la pluralidad incluye una o más de las características descritas aquí; (b) introducir al menos algunos ácidos nucleicos de la pluralidad dentro de las células huésped; y (c) el ensamblaje de las partículas de fago que empacan los ácidos nucleicos introducidos bajo condiciones, en donde al menos algunas partículas incorporan la proteína de expresión codificada por el ácido nucleico introducido respectivo. El método puede ser utilizado para proporcionar una biblioteca de fagos.

45 En modalidades en las cuales el ácido nucleico contiene un promotor regulable, el método puede incluir además la propagación de la biblioteca bajo condiciones en las cuales se reprime el promotor regulable.

50 En una modalidad, no se introducen los fagos auxiliares en las células huésped.

En otro aspecto, la invención se caracteriza por un método que incluye: (i) proporcionar una biblioteca de fagos, en donde la biblioteca puede incluir una o más de las características descritas aquí; (ii) seleccionar una partícula de fago que presente una proteína de expresión que se enlace con el objetivo; y (iii) recuperar el ácido nucleico de la partícula de fago seleccionada, identificando así una proteína de expresión que enlace un objetivo. En una

- modalidad, el método incluye además expresar un polipéptido de enlace que incluya al dominio Kunitz de la proteína de expresión identificada. El método puede incluir además purificar el polipéptido de enlace, formulando el polipéptido de enlace como una composición farmacéutica, y administrar el polipéptido de enlace (por ejemplo, tal como una composición farmacéutica) a un individuo, por ejemplo, un mamífero, por ejemplo, un humano. El método puede ser utilizado para la identificación de una proteína de expresión que enlace un objetivo a partir de una pluralidad de proteínas de expresión. La información sobre una proteína identificada puede ser transmitida, por ejemplo, en forma digital, o recibida. El receptor puede producir una proteína con base en la información.
- 5
- 10 En otro aspecto, la invención se caracteriza por un ácido nucleico que incluye: (a) un grupo de genes fago suficientes para producir una partícula infecciosa de fago, (b) un marco de lectura abierto y (c) un promotor operativamente enlazado al marco de lectura abierto. El marco de lectura abierta codifica una proteína de expresión que incluye: (i) un primer elemento que contiene un dominio Kunitz; y (ii) un segundo elemento que incluye uno o más aminoácidos que puede asociar físicamente la proteína de expresión con una partícula de fago. El segundo
- 15 elemento puede ser, por ejemplo, una cisteína que puede formar un disulfuro con una cisteína sobre la partícula de fago, una secuencia de aminoácido que interactúa en forma no covalente con la partícula de fago (por ejemplo, para interacción fos-jun), o todo o parte de una proteína de recubrimiento de fago. El ácido nucleico puede ser utilizado como se describe aquí.
- 20 También se divulga una proteína que contiene un dominio Kunitz, por ejemplo, una proteína que incluye un dominio Kunitz identificado por medio de un método descrito aquí. Por ejemplo, la proteína puede ser menor a 200, 100, ó 70 aminoácidos de longitud. Puede incluir un solo dominio Kunitz o múltiples dominios Kunitz. El dominio Kunitz de la proteína puede incluir:
- 25 MHSFCAFKADX11GX₁₃CX₁₆X₁₆X₁₇X₁₈X₁₉RFFFNIFTRQCEEFX₃₄YGGCX₃₉X₄₀NQNR ESLEECKMCTRDGA, en posiciones diferentes a X (SEQ ID NO: 10), pero difieren de la secuencia de LACI-K1 en al menos un residuo aminoácido, por ejemplo, al menos cinco, seis, siete, u ocho de las posiciones 11, 13, 15, 16, 17, 18, 19, 34, y 39. Por ejemplo,
- 30 X₁₁ es uno de: A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
X₁₃ es uno de: A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
X₁₅ es uno de: A, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, Y;
X₁₆ es uno de: A, G, E, D, H, T;
X₁₇ es uno de: A, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, Y;
X₁₈ es uno de: A, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, Y;
- 35 X₁₉ es uno de: A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
X₃₄ es uno de A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
X₃₉ es uno de: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y; y
X₄₀ es uno de: G, A.
- 40 El término "polipéptido" se refiere a un polímero de tres o más aminoácidos enlazados por medio de un enlace peptídico. El polipéptido puede incluir uno o más aminoácidos no naturales. Típicamente, el polipéptido incluye únicamente aminoácidos naturales. El término "péptido" se refiere a un polipéptido que está entre tres y treinta y dos aminoácidos de longitud.
- 45 Una "proteína" puede incluir una o más cadenas de polipéptido. Por lo tanto, el término "proteína" abarca polipéptidos y péptidos. Una proteína o polipéptido puede incluir también una o más modificaciones, por ejemplo, una glicosilación, amidación, fosforilación, y así sucesivamente.
- 50 El término "proteína de expresión" se refiere a una proteína, diferente a una proteína de recubrimiento viral no modificada, físicamente asociada con el ácido nucleico que la codifica. En el caso de la tecnología de expresión de fagos, la proteína de expresión está físicamente asociada con la partícula de fago que empaqueta al ácido nucleico que la codifica, y también incluye una secuencia de aminoácidos de al menos tres aminoácidos que es heteróloga al bacteriófago. La región heteróloga, más típicamente, incluye un dominio de andamio, por ejemplo, un dominio Kunitz. La asociación física puede ser mediada, por ejemplo, por medio de uno o más enlaces covalentes, por

ejemplo, uno o más enlaces peptídicos o un enlace disulfuro. En una modalidad, una proteína de expresión incluye al menos un dominio funcional de una proteína de recubrimiento de fago, de tal manera que la proteína de expresión se incorpora dentro de la partícula del fago.

- 5 El término "particular viral" abarca virus que incluyen una información genética suficiente para producir partículas de descendencia en una célula huésped y partículas virales que pueden entrar en las células, pero no pueden producir descendencia. El término "bacteriopartícula de fago" abarca bacteriófagos que incluyen una información genética suficiente para producir partículas de descendencia en una célula huésped y partículas de fago que pueden entrar en las células, pero no pueden producir descendencia. Por lo tanto, el término "bacteriopartícula de fago" incluye tanto partículas que empaquetan un fagémido como partículas que empaquetan un genoma de fago completo.

- 10 Una "biblioteca del dominio Kunitz humano" se refiere a una biblioteca que incluye diferentes dominios Kunitz o ácidos nucleicos que codifican tales dominios, en donde las posiciones invariables de los aminoácidos en la biblioteca son al menos 85% idénticas a un dominio Kunitz humano particular. Las bibliotecas del dominio Kunitz humano pueden ser al menos 85, 87, 90, 92, 93, 94, 95, 96, 97, 98, ó 100% idénticas a un dominio Kunitz humano particular en las posiciones invariables de los aminoácidos.

- 20 Una "biblioteca del dominio LACI-K1" se refiere a una biblioteca que incluye diferentes dominios Kunitz o ácidos nucleicos que codifican tales dominios, en donde las posiciones invariables de los aminoácidos en la biblioteca son al menos 85% idénticas a LACI-K1 humano. Las bibliotecas del dominio LACI-K1 pueden ser al menos 85, 87, 90, 92, 93, 94, 95, 96, 97, 98, ó 100% idénticas a LACI-K1 en las posiciones invariables de los aminoácidos. Por ejemplo, una biblioteca 100% del dominio LACI-K1 se refiere a una biblioteca en la cual las posiciones invariables de los aminoácidos en la biblioteca es 100% idéntica al LACI-K1 humano. Se puede utilizar la misma nomenclatura para referirse a bibliotecas que tienen una relación correspondiente con otros dominios Kunitz, por ejemplo, otros dominios Kunitz humanos, por ejemplo, un dominio descrito aquí.

- 25 Un "sistema de expresión" es una configuración de secuencias de ácido nucleico que incluye un marco de lectura abierto y un promotor de tal manera que el marco de lectura abierto está operativamente enlazado al promotor y el marco de lectura abierto puede ser expresado como un transcrito que puede ser traducido.

- 30 Los cálculos de homología o de identidad de secuencia entre secuencias (los términos se usan en forma intercambiable aquí) se llevan a cabo de la siguiente manera. El "porcentaje de identidad" entre las dos secuencias es una función del número de posiciones idénticas compartidas por las secuencias, teniendo en cuenta el número de huecos, y la longitud de cada hueco, que debe ser introducido para una alineación óptima de las dos secuencias. La comparación de las secuencias y la determinación del porcentaje de identidad entre dos secuencias se pueden lograr utilizando un algoritmo matemático. El porcentaje de identidad entre dos secuencias de aminoácidos o de nucleótidos se puede determinar utilizando el algoritmo de Needleman y Wunsch ((1970) J. Mol. Biol. 48: 444 - 453) algoritmo que ha sido incorporado en el programa GAP en el paquete de software GCG, utilizando una matriz Blossum 62 y un peso de hueco de 12, una penalización por extensión de hueco de 4, y una penalización por hueco que desplaza el marco de 5.

- 40 Generalmente, para determinar el porcentaje de identidad de dos secuencias de aminoácidos, o de dos secuencias de ácido nucleico, se alinean las secuencias para propósitos de una comparación óptima (por ejemplo, se pueden introducir huecos en una o en ambas de una primera y una segunda secuencia de ácidos nucleicos o de aminoácidos para alineación óptima y se pueden ignorar secuencias no homólogas para propósitos de comparación). En una modalidad preferida, la longitud de una secuencia de referencia alineada para propósitos de comparación es al menos 30%, preferiblemente al menos 40%, más preferiblemente al menos 50%, 60%, e incluso más preferiblemente al menos 70%, 80%, 90%, 100% de la longitud de la secuencia de referencia (por ejemplo, al menos 51, 55, 57 ó 58 aminoácidos). Se comparan entonces los residuos aminoácidos o los nucleótidos en las correspondientes posiciones de los aminoácidos o en las posiciones de los nucleótidos. Cuando una posición en la primera secuencia está ocupada por el mismo residuo aminoácido o nucleótido que la correspondiente posición en segunda secuencia, entonces las moléculas son idénticas en esa posición (como se lo utiliza aquí, "identidad" de ácido nucleico o aminoácido es equivalente a "homología" de ácido nucleico o aminoácido).

Se utiliza el programa GAP para identificar la alineación óptima entre una secuencia de consulta y una secuencia del dominio Kunitz de referencia (por ejemplo, la secuencia LACI-K1) para identificar las "correspondientes" posiciones de los aminoácidos.

5 Los "bucles de interacción" de un dominio Kunitz se refieren al primer bucle de interacción que incluye las posiciones de los aminoácidos correspondientes a las posiciones de los aminoácidos 11 a 19 de LACI-K1, y el segundo bucle de interacción que incluye las posiciones de los aminoácidos correspondientes a las posiciones de los aminoácidos 32 a 40 de LACI-K1.

10 Los detalles de una o más modalidades de la invención están expuestos en los dibujos acompañantes y la descripción que viene a continuación. Otras características, propósitos, y ventajas de la invención serán evidentes a partir de la descripción y de los dibujos, y de las reivindicaciones.

Descripción de los dibujos

15 La FIG. 1 es un diagrama de un ejemplo de un vector fago, DY3P82, que puede ser utilizado para una biblioteca del dominio Kunitz derivada de LACI-K1.

20 La FIG. 2 es un diagrama del diseño de una primera biblioteca de ejemplo. La secuencia de aminoácidos está enlistada como la SEQ ID NO: 4. Una interpretación aproximada de la secuencia de ácido nucleico está enlistada como la SEQ ID NO: 5. (Sin embargo, el uso de codones modificados de trinucleótidos puede no estar incluida en el listado de secuencia de ácido nucleico).

25 La FIG. 3 es un diagrama del diseño de una segunda biblioteca de ejemplo. La secuencia de aminoácidos está enlistada como la SEQ ID NO: 6. Una interpretación aproximada de la secuencia de ácido nucleico está enlistada como la SEQ ID NO: 7. (Nuevamente, el uso de codones modificados de trinucleótidos puede no estar incluida en el listado de secuencia de ácido nucleico).

30 Para las FIG. 2 y 3, los números de la posición de los aminoácidos están enlistados encima de cada aminoácido. Los nucleótidos correspondientes, los sitios de la enzima de restricción, y los sitios de la variación están enlistados debajo de cada aminoácido. Las posiciones variables también contienen un segundo número, indicando el número posible de aminoácidos permitidos en esa posición.

35 Las FIG. 4A, B, C y D ilustran la secuencia de ADN (SEQ ID NO: 8) del casete de expresión del ejemplo del vector fago DY3P82_LACIK1, y la secuencia de aminoácidos (SEQ ID NO: 9) de la proteína de expresión modificada. (El uso de codones modificados de trinucleótidos puede no estar incluido en el listado de secuencia del ácido nucleico). Las bases 7244 hasta 7415 contienen al promotor PlacZ y un sitio de enlazamiento del ribosoma. Las bases 7416 - 7469 codifican la secuencia señal de 18 aminoácidos de M13 iii. La peptidasa señal escinde después de Ala18. Las bases 7470 - 7475 codifican los aminoácidos Ala - Glu, marcados aquí "a" y "b". Estos permiten una escisión eficiente por medio de la peptidasa señal L. Las bases 7476 - 7649 codifican al dominio LACI K1, mostrado aquí con la secuencia de tipo silvestre y numerado de 1 a 58. Se muestran las posiciones modificadas (11, 13, 15, 16, 17, 18, 19, 34, 39, y 40). Los sitios de restricción mostrados son únicos dentro de DY3P82_LACIK1.

Descripción detallada

45 En un aspecto, la invención proporciona una biblioteca de partículas de fago en la cual se modifican entre sí los dominios Kunitz presentados de los miembros respectivos como se define en las reivindicaciones. Las bibliotecas de expresión de fagos pueden ser utilizadas para identificar un dominio Kunitz que puede enlazarse con un objetivo particular, por ejemplo, una proteasa objetivo. El dominio Kunitz puede inhibir también la actividad de la proteasa objetivo.

50 Las bibliotecas de expresión de fagos son colecciones de partículas de fagos que (i) incluyen una proteína de expresión modificada sobre la superficie de la partícula y (ii) contienen el ácido nucleico que codifica la proteína de

expresión. La asociación física entre la proteína de expresión y su ácido nucleico que la codifica es conveniente para el aislamiento rápido de las proteínas que se enlazan con el objetivo.

5 Las partículas de fago pueden presentar la proteína de expresión con una valencia deseada. Por ejemplo, las partículas pueden presentar la proteína de expresión con una valencia limitada, por ejemplo, en un número de copia que es menor al número de copia máximo posible. Por ejemplo, si la proteína de expresión está unida a la partícula de fago en al menos una porción de la proteína del gen III, el número de copia máximo posible es aproximadamente cinco. Por lo tanto, la valencia del dominio Kunitz expresado sobre la partícula del fago puede estar limitada, por ejemplo, a menos de cinco, cuatro, tres, o dos copias por partícula.

10 En una biblioteca, la valencia media o promedio puede ser, por ejemplo, menor o igual a 4, 3, 2, 1.5, 1.4, 1.25, 1.1 ó 1.0, o entre 0.25 - 2, 0.3 y 1.8, 0.3 y 1.5, ó 0.5 y 1.5. La valencia reducida favorece la selección de enlazadores altamente específicos. Se puede utilizar una expresión monovalente (por ejemplo, una valencia promedio menor a 1.4).

15 Un método para lograr una valencia limitada es utilizar una proteína de expresión que incluya al menos una porción funcional de una proteína de recubrimiento menor. Otra copia de la proteína de recubrimiento menor (o una porción funcional) de la misma - una proteína homóloga - es expresada también mientras están siendo producidas partículas de fago de tal manera que, en promedio, se incorporan menos proteínas de expresión en una partícula dada que el número de copias máximo posible de la proteína de recubrimiento menor. Por ejemplo, si se utiliza la proteína de recubrimiento del gen III del fago filamentoso, el número de copias máximo posible típico es aproximadamente de cinco. La expresión de proteínas de expresión como fusiones con al menos el dominio ancla de la proteína de recubrimiento del gen III en forma simultánea con la expresión de la proteína de recubrimiento del gen III de tipo silvestre puede resultar en partículas de fago con menos de cinco, cuatro, tres, o dos proteínas de expresión por partícula.

25 La secuencia de ácido nucleico que codifica la expresión puede estar operativamente enlazada a un promotor que produce un nivel deseado de expresión con relación a otros componentes del fago. Por ejemplo, se puede utilizar un promotor regulable, por ejemplo, para permitir el control de la relación de expresión de la proteína de expresión con respecto a su homóloga.

30 Sin querer vincularse a ninguna teoría en particular, la proteína de expresión y su homóloga pueden competir por la incorporación dentro de la partícula del fago. Una consecuencia del control de la proporción de la expresión de la proteína de expresión y de su homóloga es el control de la valencia de la proteína de expresión.

35 Las secuencias que codifican los correspondientes dominios funcionales de la proteína de expresión y de la proteína homóloga pueden ser diferentes. Por ejemplo, si los correspondientes dominios funcionales tienen idénticas secuencias de aminoácidos, se puede modificar la escogencia del codón para producir diferentes secuencias de codificación. Por ejemplo, una de las dos puede incluir codones sintéticos seleccionados para prevenir la recombinación entre la secuencia de ácido nucleico que codifica la proteína de expresión y la secuencia de ácido nucleico que codifica la proteína homóloga, que puede utilizar codones naturales. El escenario puede ser invertido también, por ejemplo, el ácido nucleico que codifica la proteína de expresión puede utilizar codones sintéticos para codificar la proteína de recubrimiento o un fragmento de la misma. Por ejemplo, las dos secuencias pueden diferir entre 5 y 95, 5 a 60, ó 20 y 50% de los codones.

45 Ejemplos de dominios Kunitz

50 Como se lo utiliza aquí, un "dominio Kunitz" es un dominio del polipéptido que incluye al menos 51 aminoácidos y que contiene al menos dos, y más típicamente tres, enlaces disulfuro. El dominio se pliega de tal manera que, si está presente, la primera y la sexta cisteínas, la segunda y la cuarta, y la tercera y la quinta cisteínas forman enlaces disulfuro. Por ejemplo, en un ejemplo de un dominio Kunitz que tiene 58 aminoácidos, se forman disulfuros entre las cisteínas en la posición 5 y 55, 14 y 38, y 30 y 51.

En implementaciones en las cuales están presentes dos disulfuros, se pueden formar enlaces disulfuro entre un subgrupo correspondiente de cisteínas. El espaciamento entre las respectivas cisteínas puede estar entre 7, 5, 4, 3 ó 2 aminoácidos del siguiente espaciamento: 5 a 55, 14 a 38, y 30 a 51.

5 En la SEQ ID NO: 2, los enlaces disulfuro enlazan al menos dos de: 5 a 55, 14 a 38, y 30 a 51, de la siguiente manera: X₁-X₂-X₃-X₄-C₅-X₆-X₇-X₈-X₉-X_{9a}-X₁₀-X₁₁-X₁₂-X₁₃-C₁₄-X₁₅-X₁₆-X₁₇-X₁₈-X₁₉-X₂₀-X₂₁-X₂₂-X₂₃-X₂₄-X₂₅-X₂₆-X₂₇-X₂₈-X₂₉-X_{29a}-X_{29b}-X_{29c}-C₃₀-X₃₁-X₃₂-X₃₃-X₃₄-X₃₅-X₃₆-X₃₇-C₃₈-X₃₉-X₄₀-X₄₁-X₄₂-X_{42a}-X_{42b}-X₄₃-X₄₄-X₄₅-X₄₆-X₄₇-X₄₈-X₄₉-X₅₀-C₅₁-X₅₂-X₅₃-X₅₄-C₅₅-X₅₆-X₅₇-X₅₈ (SEQ ID NO: 2).

10 El número de disulfuros puede reducirse en uno, pero, generalmente, ninguna de las cisteínas estándar se dejará desapareada. Por lo tanto, si se cambia una cisteína, entonces se añade una cisteína de compensación en una posición adecuada o la cisteína emparejada es también reemplazada por uno que no sea cisteína. Por ejemplo, el inhibidor de proteasa de la glándula secundaria de la *Drosophila funebris* macho no tiene cisteína en la posición 5, pero tiene una cisteína en la posición -1 justo antes de la posición 1; presumiblemente esta forma un disulfuro con Cys₅₅.

15 En algunas modalidades, se pueden cambiar C₁₄ y C₃₀ con aminoácidos diferentes a la cisteína, pero preferiblemente si se cambia uno de estos residuos, se cambian ambos. Si C₁₄ está presente y X_{9a} está ausente, entonces X₁₂ puede ser Gly. En algunas modalidades, X₃₇ es Gly. En algunas modalidades, X₃₃ es Phe o Tyr, y X₄₅ es Phe o Tyr. En algunas modalidades, se reemplazan Cys₁₄ y Cys₃₈, y la exigencia de Gly₁₂, (Gly o Ser)₃₇, y Gly₃₆ se cae.

25 Se pueden ubicar desde cero hasta muchos residuos en cualquier extremo de un dominio Kunitz. Estos residuos pueden constituir, por ejemplo, uno o más dominios (por ejemplo, otros dominios de expresión Kunitz y no Kunitz) u otras secuencias de aminoácidos.

30 Los dominios Kunitz naturales son generalmente dominios muy estables. Los dominios Kunitz pueden enlazarse en los sitios activos de sus respectivos objetivos de proteasa de tal manera que un enlace peptídico (el "enlace escindible") del dominio Kunitz es: 1) no escindido, 2) escindido muy lentamente, o 3) escindido hasta que no hay efecto debido a que la estructura del inhibidor impide la liberación o la separación de los segmentos escindidos. Los enlaces disulfuro generalmente actúan para mantener la proteína unida incluso si los enlaces peptídicos expuestos son escindidos.

35 A partir del residuo sobre el lado amino del enlace escindible, y alejándose del enlace, los residuos son convencionalmente llamados P1, P2, P3, etc., los residuos que van después del enlace escindible son llamados P1', P2', P3', etc. Generalmente se acepta que cada serina proteasa tiene sitios (que contienen varios residuos) S1, S2, etc., que reciben los grupos laterales y los átomos de la cadena principal de P1, P2, etc., del sustrato o del inhibidor y sitios S1', S2', etc., que reciben los grupos laterales y los átomos de la cadena principal de P1', P2', etc., del sustrato o del inhibidor. Esto es, las interacciones entre los sitios S y los grupos laterales P y los átomos de la cadena principal que le dan especificidad a la proteasa con respecto a sustratos y la especificidad de los inhibidores con respecto a proteasas. Debido a que el fragmento que tiene el nuevo terminal amino deja primero la proteasa, muchos diseños de inhibidores de proteasa de molécula pequeña se han concentrado en compuestos que enlazan sitios S1, S2, S3, etc.

45 Típicamente, X₁₅ o la posición del aminoácido correspondiente a la posición 15 de LACI-K1 es equivalente a P1 (descrito anteriormente), X₁₆ o la posición del aminoácido correspondiente a la posición 16 de LACI-K1 es equivalente a P1', X₁₇ o la posición del aminoácido correspondiente a la posición 17 de LACI-K1 es equivalente a P2', X₁₈ o la posición del aminoácido correspondiente a la posición 18 de LACI-K1 es equivalente a P3', y X₁₉ o la posición del aminoácido correspondiente a la posición 19 de LACI-K1 es equivalente a P4'. Como se discute más adelante, se pueden modificar uno o más de S1, S2, S3, P1', P2', P3', y P4'.

50 Los ejemplos de dominios Kunitz humanos incluyen los tres dominios Kunitz de LACI (Wun et al., (1988) J. Biol. Chem. 263(13): 6001 - 6004; Girard et al., (1989) Nature, 338: 518 - 20; Novotny et al, (1989) J. Biol. Chem., 264(31): 18832 - 18837); los dos dominios Kunitz del inhibidor Inter- α -Tripsina, APPI (inhibidor del precursor de

proteína β amiloide de Alzheimer) (Kido et al., (1988) J. Biol. Chem., 263(34): 18104 - 18107), un dominio Kunitz de colágeno, y los tres dominios Kunitz de TFPI-2 (Sprecher et al., (1994) Proc. Nat. Acad. USA, 91: 3353 - 3357).

LACI es una fosfoglicoproteína de suero humano que contiene tres dominios Kunitz:

5

```

MIYTMKKVHA LWASVCLLLN LAPAPLNADS EEDEEHTIIT DTELPPLKLM
HSFCAFKADD GPCKAIMKRF FFNIFTRQCE EFIYGGCEGN QNRFESLEEC
KKMCTRDNAN RIIKTTLQQE KPDFCFLEED PGICRGYITR YFYNNQTKQC
ERFKYGGCLG NMNMFETLEE CKNICEDGPN GFQVDNYGTQ LNAVNNSLTP
QSTKVPSLFE FHGPSWCLTP ADRGLCRANE NRFYNSVIG KCRPFKYSGC
GGNENNFTSK QECLRACKKG FIQRISKGGL IKTKRKRKKQ RVKIAYEEIF
VKNM (SEQ ID NO:3)

```

10 La secuencia de la señal está localizada en los aminoácidos 1 - 28. Los tres dominios Kunitz dentro de LACI se denominan como LACI-K1 (residuos 50 a 107 de la SEQ ID NO: 3), LACI-K2 (residuos 121 a 178 de la SEQ ID NO: 3), y LACI-K3 (213 a 270 de la SEQ ID NO: 3). La secuencia de ADNc de LACI está reportada en Wun et al. (J. Biol. Chem., 1988, 263(13): 6001 - 6004). Girard et al. (Nature, 1989, 338: 518 - 20) reporta estudios de mutaciones en los cuales los residuos P1 de cada uno de los tres dominios Kunitz fueron alterados. LACI-K1 inhibe al Factor VIIa cuando el Factor VIIa forma complejo con el factor del tejido. LACI-K2 inhibe al Factor Xa.

15 Aquí, los residuos de los ejemplos de los dominios Kunitz están numerados con referencia al LACI-K1 (dominio Kunitz 1 de LACI):

```

MHSFCAFKAD DGPKAIMKR FFFNIFTRQC EEFIYGGCEG
NQNRFESLEE CKKMCTRD (SEQ ID NO:1)

```

20 El primer residuo de cisteína del dominio Kunitz LACI-K1 es el residuo 5 y la última cisteína es el residuo 55. Las posiciones de los aminoácidos en un dominio Kunitz también pueden ser referenciadas con respecto a su correspondencia con los aminoácidos en LACI-K1, utilizando la alineación óptima suministrada por el programa GAP (ver más arriba).

25 Los dominios Kunitz descritos aquí pueden ser al menos 30, 40, 50, 60, 70, 80, ó 90% idénticos a LACI-K1. Otros dominios Kunitz descritos aquí son homólogos (por ejemplo, al menos 30, 40, 50, 60, 70, 80, ó 90% idénticos) a otros dominios Kunitz de ocurrencia natural (por ejemplo, un dominio Kunitz descrito aquí), particularmente a otros dominios Kunitz humanos.

30 Se conocen las estructuras moleculares tridimensionales de muchos dominios Kunitz con alta resolución. Ver, por ejemplo, Eigenbrot et al. (1990) Protein Engineering 3: 591 - 598 y Hynes et al. (1990) Biochemistry 29: 10018 - 10022. Un ejemplo de un modelo estructural por rayos X del dominio Kunitz BPTI está depositado en el Brookhaven Protein Data Bank como "6PTI".

35 Se conocen más de setenta secuencias del dominio Kunitz. Las proteínas que contienen ejemplos de dominios Kunitz incluyen las siguientes, con los Números de Acceso del SWISS-PROT entre paréntesis: A4_HUMAN (P05067), A4_MACFA (P53601), A4_MACMU (P29216), A4_MOUSE (P12023), A4_RAT (P08592), A4_SAISC (Q95241), AMBP_PLEPL (P36992), APP2_HUMAN (Q06481), APP2_RAT (P15943), AXP1_ANTAF (P81547), AXP2_ANTAF (P81548), BPT1_BOVIN (P00974), BPT2_BOVIN (P04815), CA17_HUMAN (Q02388), CA36_CHICK (P15989), CA36_HUMAN (P12111), CRPT_BOOMI (P81162), ELAC_MACEU (O62845), ELAC_TRIVU (Q29143), EPPI_HUMAN (O95925), EPPI_MOUSE (Q9DA01), HTIB_MANSE (P26227), IBP_CARCR (P00993), IBPC_BOVIN

(P00976), IBPI_TACTR (P16044), IBPS_BOVIN (P00975), ICS3_BOMMO (P07481), IMAP_DROFU (P11424), IP52_ANESU (P10280), ISC1_BOMMO (P10831), ISC2_BOMMO (P10832), ISH1_STOHE (P31713), ISH2_STOHE (P81129), ISIK_HELPO (P00994), ISP2_GALME (P81906), IVB1_BUNFA (P25660), IVB1_BUNMU (P00987), IVB1_VIPAA (P00991), IVB2_BUNMU (P00989), IVB2_DABRU (P00990), IVB2_HEMHA (P00985), IVB2_NAJNI (P00986), IVB3_VIPAA (P00992), IVBB_DENPO (P00983), IVBC_NAJNA (P19859), IVBC_OPHHA (P82966), IVBE_DENPO (P00984), IVBI_DENAN (P00980), IVBI_DENPO (P00979), IVBK_DENAN (P00982), IVBK_DENPO (P00981), IVBT_ERIMA (P24541), IVBT_NAJNA (P20229), MCPI_MELCP (P82968), SBPI_SARBU (P26228), SPT3_HUMAN (P49223), TKD1_BOVIN (Q28201), TKD1_SHEEP (Q29428), TXCA_DENAN (P81658), UPTI_PIG (Q29100), AMBP_BOVIN (P00978), AMBP_HUMAN (P02760), AMBP_MERUN (Q62577), AMBP_MESAU (Q60559), AMBP_MOUSE (Q07456), AMBP_PIG (P04366), AMBP_RAT (Q64240), IATR_HORSE (P04365), IATR_SHEEP (P13371), SPT1_HUMAN (O43278), SPT1_MOUSE (Q9R097), SPT2_HUMAN (O43291), SPT2_MOUSE (Q9WU03), TFP2_HUMAN (P48307), TFP2_MOUSE (O35536), TFPI_HUMAN (P10646), TFPI_MACMU (Q28864), TFPI_MOUSE (O54819), TFPI_RABIT (P19761), TFPI_RAT (Q02445), y YN81_CAEEL (Q03610).

Un "residuo conservado de Kunitz" en una posición de un aminoácido particular es un aminoácido que está presente, en esa posición, en al menos 5 secuencias del listado anterior. Un "residuo conservado de Kunitz" en una posición de un aminoácido particular es un aminoácido que está presente, en esa posición, en al menos 30% de las secuencias del listado anterior. Más de un residuo conservado o altamente conservado de Kunitz puede estar disponible en una posición particular. Las posiciones se basan en la alineación CLUSTALW óptima del listado anterior, y se referencian de acuerdo con la numeración de los aminoácidos de LACI-K1.

Se pueden utilizar una variedad de métodos para identificar un dominio Kunitz de una base de datos de secuencias. Por ejemplo, se pueden buscar una secuencia conocida de aminoácidos de un dominio Kunitz, una secuencia de consenso, o un motivo (por ejemplo, el Motivo ProSite) contra las bases de datos de secuencias del GENBANK® (National Center for Biotechnology Information, National Institutes of Health, Bethesda MD), por ejemplo, utilizando BLAST; contra la base de datos Pfam de los HMM (Hidden Markov Models) (por ejemplo, utilizando parámetros predeterminados para la búsqueda en Pfam; contra la base de datos SMART™; o contra la base de datos ProDom. Ver, por ejemplo, Sonhammer et al. (1997) *Proteins* 28(3): 405 - 420; Gribskov et al. (1990) *Meth. Enzymol.* 183: 146 - 159; Gribskov et al. (1987) *Proc. Natl. Acad. Sci. USA* 84: 4355 - 4358; Krogh et al. (1994) *J. Mol. Biol.* 235: 1501 - 1531; Stultz et al. (1993) *Protein Sci.* 2: 305 - 314; Schultz et al. (1998), *Proc. Natl. Acad. Sci. USA* 95: 5857 Schultz et al. (2000) *Nucl. Acids Res* 28: 231; y Corpet et al. (1999), *Nucl. Acids Res.* 27: 263 - 267). Prosite enlista el dominio Kunitz como un motivo e identifica proteínas que incluyen un dominio Kunitz. Ver, por ejemplo, Falquet et al. *Nucleic Acids Res.* 30: 235 - 238(2002).

Modificación de Dominios Kunitz

Las bibliotecas de expresión incluyen una variación en una o más posiciones en el dominio Kunitz expresado. Por ejemplo, se pueden variar entre una y 20 ó 5 y 12 posiciones. Las posiciones modificadas se pueden localizar en uno de los dos bucles de interacción del dominio Kunitz. El primer "bucle de interacción" incluye P1, P1', P2', P3', y P4' y otras posiciones de aminoácidos correspondientes a los aminoácidos 11 a 19 de LACI-K1. El segundo "bucle de interacción" incluye las posiciones de los aminoácidos correspondientes a los aminoácidos 32 a 40 de LACI-K1.

La biblioteca puede incluir una variación en uno o en ambos bucles de interacción.

El tamaño teórico de la biblioteca, por ejemplo, el número de proteínas de expresión únicas que pueden ser codificadas por una biblioteca puede ser grande (por ejemplo, entre 103 a 1019, 103 a 1015, 105 a 1014, ó 107 a 1012 proteínas de expresión diferentes, y/o por ejemplo, al menos 105, 106, 108, 109, ó 1010. El tamaño teórico se refiere al número total de secuencias únicas de aminoácidos que podrían ser codificadas por la biblioteca en su forma completamente representada, independientemente de una implementación real. La diversidad teórica es generalmente el producto del número de variaciones en cada posición. Por ejemplo, la diversidad teórica de modificar únicamente dos posiciones entre todos los veinte aminoácidos es de 20 x 20, ó 400. Una biblioteca con gran diversidad es muy útil a pasar que la biblioteca real utilizada pueda muestrear únicamente un pequeño subgrupo de diversidad teórica.

5 Sin embargo, las bibliotecas con diversidad limitada, por ejemplo, menor a 10^{15} , 10^{14} , 10^{13} , 10^{12} , 10^{11} , 10^{10} , ó 10^9 proteínas diferentes son también útiles. Si se diseñan apropiadamente, tales bibliotecas pueden incluir también una fracción mayor de proteínas plegadas. Las bibliotecas con diversidad limitada también facilitan una evaluación rigurosa de un espacio de secuencia particular.

10 Diversidad sintética. Las bibliotecas pueden incluir regiones de diversas secuencias de ácido nucleico que se originan a partir de secuencias sintetizadas artificialmente. Las secuencias sintéticas de aminoácidos incluyen variantes de secuencias de ocurrencia natural, por ejemplo, variantes que son al menos 30, 50, 70, 80, 90, 95 ó 98% idénticas a una secuencia de ocurrencia natural.

15 Típicamente, se sintetiza la biblioteca a partir de una o más poblaciones de oligonucleótidos degenerados que incluyen una distribución de nucleótidos en una pluralidad de posiciones seleccionadas. La inclusión de un nucleótido dado es aleatoria con respecto a la distribución. Un ejemplo de una fuente degenerada de diversidad sintética es un oligonucleótido que incluye NNN en donde N es cualquiera de los cuatro nucleótidos en igual proporción. El oligonucleótido degenerado también incluye posiciones invariables que codifican posiciones de aminoácidos invariables del molde del dominio Kunitz.

20 La diversidad sintética también puede ser más limitada, por ejemplo, para limitar el número de codones en una secuencia de ácido nucleico en un trinucleótido dado con una distribución que sea más pequeña que NNN. Por ejemplo, tal distribución puede ser limitada utilizando menos de cuatro nucleótidos (por ejemplo, tres o dos) en algunas posiciones del codón. Además, la tecnología de adición de trinucleótidos puede ser utilizada para limitar más la distribución.

25 La así llamada "tecnología de adición de trinucleótidos" está descrita, por ejemplo, en Wells et al. (1985) Gene 34: 315 - 323, Patente Estadounidense No. 4.760.025 y 5.869.644. Los oligonucleótidos se sintetizan sobre un soporte en fase sólida, un codón (es decir, un trinucleótido) a la vez. El soporte incluye muchos grupos funcionales para síntesis de tal manera que se sintetizan en paralelo muchos oligonucleótidos. Se expone primero el soporte a una solución que contiene una mezcla del grupo de codones para la primera posición.

30 La unidad está protegida así que no se añaden unidades adicionales. La solución que contiene la primera mezcla se elimina por lavado y se desprotege el soporte sólido de tal manera que se puede añadir una segunda mezcla que contiene un grupo de codones para una segunda posición, a la unidad adjuntada primero.

35 Se repite el proceso para ensamblar secuencialmente múltiples codones. La tecnología de adición de trinucleótidos permite la síntesis de un ácido nucleico que en una posición dada puede codificar una cantidad de aminoácidos. La frecuencia de estos aminoácidos puede ser regulada por medio de la proporción de codones en la mezcla. Además, la escogencia de aminoácidos en la posición dada no está restringida a cuadrantes de la tabla de codones como es el caso si se añaden las mezclas de nucleótidos individuales durante la síntesis.

40 En algunas modalidades, se pueden limitar las variaciones de las secuencias de aminoácidos en diversas regiones, por ejemplo, a subgrupos de aminoácidos que tienen cadenas laterales similares. Los ejemplos de variaciones limitadas en las secuencias de aminoácidos incluyen, por ejemplo, todos los aminoácidos excepto cisteína; aminoácidos con cadenas laterales básicas (por ejemplo, lisina, arginina, histidina); aminoácidos con cadenas laterales ácidas (por ejemplo, ácido aspártico, ácido glutámico); aminoácidos con cadenas laterales polares no cargadas (por ejemplo, glicina, asparagina, glutamina, serina, treonina, tirosina); aminoácidos con cadenas laterales no polares (por ejemplo, alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptófano), aminoácidos con cadenas laterales aromáticas (por ejemplo, tirosina, fenilalanina, triptófano, histidina). Las variaciones en una posición particular pueden ser limitadas también a los residuos aminoácidos conservados o altamente conservados del dominio Kunitz.

50 Diversidad natural. Las bibliotecas pueden incluir regiones de diversas secuencias de ácidos nucleicos que originan (o se sintetizadas con base en) a partir de diferentes secuencias de ocurrencia natural, por ejemplo, diferentes dominios Kunitz de ocurrencia natural. Para algunas bibliotecas están incluidas tanto diversidad natural como sintética.

Mutagénesis. En una modalidad, se utiliza la tecnología de la biblioteca de expresión en una forma reiterada. Se utiliza una primera biblioteca de expresión para identificar uno o más ligandos para un objetivo. Estos ligandos identificados son luego modificados utilizando un método de mutagénesis para formar una segunda biblioteca de expresión. Se seleccionan luego ligandos de mayor afinidad a partir de la segunda biblioteca, por ejemplo, por medio del uso de mayor rigurosidad o de condiciones más competitivas más enlazamiento y de lavado.

En algunas implementaciones, la mutagénesis de un dominio Kunitz está dirigida a regiones conocidas o probablemente que están en la interfaz de enlazamiento, por ejemplo, una o más posiciones, descritas aquí, por ejemplo, una o más posiciones en los bucles de interacción. Además, la mutagénesis puede ser dirigida a regiones marco cercanas o adyacentes a los bucles de interacción. En el caso de dominios Kunitz, la mutagénesis puede estar limitada también a uno o dos de los bucles de interacción, a una o dos de las posiciones de los aminoácidos allí. La mutagénesis enfocada puede facilitar mejoras precisas paso a paso.

Algunos ejemplos de técnicas de mutagénesis incluyen: PCR propenso a errores (Leung et al. (1989) *Technique* 1: 11 - 15), recombinación, arrastre de ADN utilizando escisión aleatoria (Stemmer (1994) *Nature* 389 - 391; llamado "arrastre de ácido nucleico"), RACHITT™ (Coco et al. (2001) *Nature Biotech.* 19:354), mutagénesis dirigida al sitio (Zooler et al. (1987) *Nucl Acids Res* 10: 6487 - 6504), mutagénesis de casete (Reidhaar-Olson (1991) *Methods Enzymol.* 208: 564 - 586) y la incorporación de oligonucleótidos degenerados (Griffiths et al. (1994) *EMBO J* 13: 3245). La mutagénesis puede ser utilizada también para preparar una biblioteca inicial de dominios Kunitz variados.

En un ejemplo de selección reiterada, se usan los métodos descritos aquí para identificar primero un ligando de proteína a partir de una biblioteca de expresión que enlaza un compuesto objetivo con al menos una especificidad de enlazamiento mínima por un objetivo o una actividad mínima, por ejemplo, una constante de equilibrio de disociación para el enlazamiento de menos de 100 nM, 10 nM, ó 1 nM. La secuencia de ácido nucleico que codifica al ligando inicialmente identificado de la proteína es utilizada como un ácido nucleico molde para la introducción de variaciones, por ejemplo, para identificar un segundo ligando de proteína que tiene propiedades mejoradas (por ejemplo, afinidad de enlazamiento, cinéticas, o estabilidad) con relación al ligando de la proteína inicial.

Expresión de fagos

La expresión de fagos utiliza bacteriófagos para expresar polipéptidos modificados. La proteína de expresión puede ser enlazada a una proteína de recubrimiento del bacteriófago con enlaces covalentes, no covalentes y no peptídicos. Ver, por ejemplo, la Patente Estadounidense No. 5.223.409, Cramer et al. (1993) *Gene* 137: 69 y WO 01/05950. El enlace puede ser el resultado de la traducción de un ácido nucleico que codifica al componente modificado fusionado a la proteína de recubrimiento. El enlace puede incluir un enlazador peptídico flexible, un sitio para la proteasa, o un aminoácido incorporado como resultado de la supresión de un codón de detención.

La expresión de fagos está descrita, por ejemplo, en Ladner et al., Patente Estadounidense No. 5.223.409; Smith (1985) *Science* 228: 1315 - 1317; WO 92/18619; WO 91/17271; WO 92/20791; WO 92/15679; WO 93/01288; WO 92/01047; WO 92/09690; WO 90/02809; de Haard et al. (1999) *J. Biol. Chem* 274: 18218 - 30; Hoogenboom et al. (1998) *Immunotechnology* 4: 1 - 20; Hoogenboom et al. (2000) *Immunol Today* 2: 371 - 8; Fuchs et al. (1991) *Bio/Technology* 9: 1370 - 1372; Hay et al. (1992) *Hum Antibod Hybridomas* 3: 81 - 85; Huse et al. (1989) *Science* 246: 1275 - 1281; Griffiths et al. (1993) *EMBO J* 12: 725 - 734; Hawkins et al. (1992) *J Mol Biol* 226: 889 - 896; Clackson et al. (1991) *Nature* 352: 624 - 628; Gram et al. (1992) *PNAS* 89: 3576 - 3580; Garrard et al. (1991) *Bio/Technology* 9: 1373 - 1377; Rebar et al. (1996) *Methods Enzymol.* 267: 129 - 49; Hoogenboom et al. (1991) *Nucl Acid Res* 19: 4133 - 4137; y Barbas et al. (1991) *PNAS* 88: 7978 - 7982.

Los sistemas de expresión de fagos han sido desarrollados para el fagos filamentos Ff (fago fl, fd, y M13) así como para otro bacteriófago (por ejemplo, el bacteriófago T7 y los fagos lambdoides; ver, por ejemplo, Santini (1998) *J. Mol. Biol.* 282: 125 - 135; Rosenberg et al. (1996) *Innovations* 6: 1 - 6; Houshmet et al. (1999) *Anal Biochem* 268: 363 - 370).

Se han descrito ácidos nucleicos adecuados para expresión de fagos, por ejemplo, vectores fago. Ver, por ejemplo, Armstrong et al. (1996) *Academic Press*, Kay et al., Ed. páginas 35 - 53; Corey et al. (1993) *Gene* 128 (1): 129 - 34; Cwirla et al. (1990) *Proc Natl Acad Sci USA* 87 (16): 6378 - 82; Fowlkes et al. (1992) *Biotechniques* 13 (3): 422 - 8;

Hoogenboom et al. (1991) *Nucleic Acids Res* 19 (15): 4133 - 7; McCafferty et al. (1990) *Nature* 348 (6301): 552 - 4; McConnell et al. (1994) *Gene* 151 (1 - 2): 115 - 8; Scott y Smith (1990) *Science* 249 (4967): 386 - 90.

5 Fagémidos. Una configuración alternativa de expresión de fagos utiliza un vector fagémido. En un sistema fagémido, se suministra el ácido nucleico que codifica la proteína fago sobre un plásmido, típicamente de una longitud menor a 6000 nucleótidos. El plásmido incluye un origen de replicación del fago de tal manera que el plásmido se incorpora en partículas de bacteriófago cuando las células bacterianas que soportan al plásmido son infectadas con fago auxiliar, por ejemplo M13K01. Los fagémidos, sin embargo, carecen de un conjunto suficiente de genes de fago con el propósito de producir partículas de fago estables. Estos genes de fago pueden ser suministrados por un fago 10 auxiliar. Típicamente, el fago auxiliar proporciona una copia intacta del gen III y de otros genes de fago requeridos para la replicación y ensamblaje del fago. Debido a que el fago auxiliar tiene un origen defectuoso, el genoma del fago auxiliar no es incorporado en forma eficiente dentro de las partículas del fago con relación al plásmido que tiene un origen de tipo silvestre. Ver, por ejemplo, la Patente Estadounidense No. 5.821.047. El genoma del fagémido contiene un gen marcador seleccionable, por ejemplo *Amp^R* o *Kan^R* para la selección de células que son infectadas 15 por un miembro de la biblioteca.

20 Vectores fago. Otra configuración de expresión de fagos utiliza vectores que incluyen un grupo de gens de fago suficientes para producir una partícula infecciosa de fago cuando se expresa, una señal de empaquetamiento de fago, y una secuencia de replicación autónoma. Por ejemplo, el vector puede ser un genoma de fago que ha sido modificado para incluir una secuencia que codifica la proteína de expresión. Los vectores de expresión de fago pueden incluir además un sitio dentro del cual una secuencia foránea de ácido nucleico puede ser insertada, tal como un sitio de clonación múltiple que contiene sitios de digestión para la enzima de restricción. Las secuencias foráneas de ácido nucleico, por ejemplo, que codifican proteínas de expresión en vectores fago, pueden estar enlazadas con un sitio de enlazamiento ribosomal, una secuencia señal (por ejemplo, una secuencia señal M13), y 25 una secuencia terminadora de la transcripción.

30 Los vectores pueden ser construidos por medio de técnicas estándar de clonación para contener una secuencia que codifica un polipéptido que incluye un dominio Kunitz y una porción de una proteína de recubrimiento de fago, y que está operativamente enlazada a un promotor regulable. En algunas modalidades, un vector de expresión del fago incluye dos secuencias de ácido nucleico que codifican la misma región de una proteína de recubrimiento de fago. Por ejemplo, el vector incluye una secuencia que codifica tal región en una posición operativamente enlazada a la secuencia que codifica la proteína de expresión, y otra secuencia que codifica tal región en el contexto del gen funcional del fago (por ejemplo, un gen del fago de tipo silvestre) que codifica la proteína de recubrimiento.

35 Una ventaja de los vectores fago es que ellos no requieren del uso del fago auxiliar, simplificando así la preparación de la biblioteca, reduciendo posibles sesgos de la biblioteca, y produciendo bibliotecas de fagos libres de partículas que empaquetan ácido nucleico del fago auxiliar.

40 Los vectores de expresión de fagos pueden incluir también un marcador seleccionable tal como marcadores de resistencia a fármacos, por ejemplo, un gen de resistencia a la ampicilina. Sin embargo, a diferencia de los fagémidos, también es posible y algunas veces conveniente utilizar vectores fago que no incluyen tal marcador seleccionable.

Proteínas de recubrimiento

45 Los sistemas de expresión de fagos utilizan típicamente fago filamentoso Ff. En implementaciones que utilizan fago filamentoso, por ejemplo, la proteína de expresión está físicamente unida a un dominio ancla de una proteína de recubrimiento de fago. La co-expresión de la proteína de expresión con otro polipéptido que tenga el mismo dominio ancla, por ejemplo, una copia endógena de la proteína de recubrimiento, resultará en la competición por la expresión sobre la superficie de la partícula. 50

Las proteínas de recubrimiento de fago que pueden ser utilizadas para la expresión de proteína incluyen (i) proteínas de recubrimiento menores de fago filamentoso, tales como la proteína en el gen, y (ii) proteínas de recubrimiento mayores de fago filamentoso tales como la proteína del gen VIII. También se pueden utilizar fusiones con otras

proteínas de recubrimiento del fago tales como la proteína del gen VI, la proteína del gen VII, o la proteína del gen IX (ver, por ejemplo, WO 00/71694).

5 También se pueden utilizar porciones (por ejemplo, dominios o fragmentos) de estas proteínas. Las porciones útiles incluyen dominios que son incorporados en forma estable dentro de la partícula de fago, por ejemplo, de tal manera que la proteína de fusión permanezca en la partícula durante todo un procedimiento de selección.

10 En una modalidad, se usa el dominio ancla o dominio "muñón" de la proteína del gen III (ver, por ejemplo, la Patente Estadounidense No. 5.658.727 para una descripción de un ejemplo de un dominio muñón de la proteína del gen III). Como se utiliza aquí, un "dominio ancla" se refiere a un dominio que está incorporado en un empaquetamiento genético (por ejemplo, un fago). Un dominio ancla típico de fago está incorporado dentro del recubrimiento o cápside del fago.

15 En otra modalidad, se utiliza la proteína del gen VIII. Ver, por ejemplo, la Patente Estadounidense No. 5.223.409. La proteína del gen VIII de longitud completa puede estar enlazada con la proteína de expresión.

20 Los sistemas de expresión del fago filamentoso utilizan típicamente fusiones de proteína para unir físicamente la secuencia heteróloga de aminoácido con una proteína de recubrimiento de fago o dominio ancla. Por ejemplo, el fago puede incluir un gen que codifica una secuencia señal, la secuencia heteróloga de aminoácidos, y el dominio ancla, por ejemplo un dominio ancla de proteína del gen III.

25 También es posible utilizar otros formatos de expresión para seleccionar bibliotecas de dominios Kunitz, por ejemplo bibliotecas cuya variación se diseña como se describe aquí. Los ejemplos de otros formatos de expresión incluyen la expresión con base en la célula (por ejemplo, expresión de levadura) y fusiones de proteína - ácido nucleico. Ver, por ejemplo, la Patente Estadounidense No. 6.207.466 y WO 03/029456. También se pueden utilizar arreglos de proteína. Ver, por ejemplo, la Patente Estadounidense No. 6.207.466 y WO 03/029456. También se pueden utilizar arreglos de proteína. Ver, por ejemplo, WO 01/40803, WO 99/51773, y US2002-0192673-A1.

Promotores para expresión de proteína de expresión

30 Se pueden utilizar promotores regulables para controlar la valencia de la proteína de expresión. Se puede utilizar la expresión regulada para producir un fago que tenga una valencia baja de la proteína de expresión. Se conocen muchas secuencias promotoras regulables (por ejemplo, inducibles y/o reprimibles). Tales secuencias incluyen promotores regulables cuya actividad puede ser alterada o regulada por medio de la intervención del usuario, por ejemplo, por medio de la manipulación de un parámetro ambiental.

35 Por ejemplo, se puede añadir un compuesto químico exógeno para regular la transcripción de algunos promotores. Los promotores regulables pueden contener sitios de enlazamiento para uno o más activadores transcripcionales o proteínas represoras. Los promotores sintéticos que incluyen sitios de enlazamiento del factor de transcripción pueden ser construidos y pueden ser utilizados también como promotores regulables.

40 Los ejemplos de promotores regulables incluyen promotores sensibles a un parámetro ambiental, por ejemplo, cambios térmicos, hormonas, metales, metabolitos, antibióticos, o agentes químicos. Los promotores regulables apropiados para uso en *E. coli* incluyen promotores que contienen sitios de enlazamiento para el factor de transcripción de las secuencias operadoras *lac*, *tac*, *trp*, *trc*, y *tet*, u operones, el promotor de la fosfatasa alcalina (45 *pho*), un promotor de arabinosa tal como un promotor *araBAD*, el promotor de ramnosa, los promotores en sí mismos, o fragmentos funcionales de los mismos (ver, Elvin et al., 1990, Gene 37 : 123 - 126; Tabor y Richardson, 1998, Proc. Natl. Acad. Sci. U. S. A. 1074 - 1078; Chang et al., 1986, Gene 44 : 121 - 125; Lutz y Bujard, March 1997, Nucl. Acids. Res. 25: 1203 - 1210; D. V. Goeddel et al., Proc. Nat. Acad. Sci. U.S.A., 76: 106 - 110, 1979; J. D. Windass et al. Nucl. Acids. Res., 10: 6639 - 57, 1982; R. Crowl et al., Gene, 38: 31 - 38, 1985; Brosius, 1984, Gene 50 27 : 161 - 172 ; Amanna y Brosius, 1985, Gene 40 : 183 - 190; Guzman et al., 1992, J. Bacteriol., 174: 7716 - 7728; Haldimann et al., 1998, J. Bacteriol., 180: 1277 - 1286). El promotor *tac* es un ejemplo de un promotor sintético.

El promotor *lac*, por ejemplo, puede ser inducido por medio de lactosa o por moléculas estructuralmente relacionadas tales como isopropil-beta-D-tiogalatóside (IPTG) y es reprimido por medio de glucosa. Algunos

promotores inducibles son inducidos por medio de un proceso de desrepresión, por ejemplo, inactivación de una molécula represora.

5 Una secuencia promotora regulable puede ser también regulada en forma indirecta. Los ejemplos de promotores que pueden ser modificados por medio de ingeniería genética para regulación indirecta incluyen: al fago lambda P_R , - P_L , el fago T7, SP6, y los promotores T5. Por ejemplo, la secuencia reguladora es reprimida o activada por un factor cuya expresión está regulada, por ejemplo, por un parámetro ambiental. Un ejemplo de tal promotor es un promotor T7. La expresión de la T7 ARN polimerasa puede ser regulada por medio de un promotor ambientalmente sensible tal como el promotor *lac*. Por ejemplo, la célula puede incluir un ácido nucleico heterólogo que incluye una secuencia
10 que codifica a la T7 ARN polimerasa y una secuencia reguladora (por ejemplo, el promotor *lac*) que es regulada por un parámetro ambiental. La actividad de la T7 ARN polimerasa también puede ser regulada por la presencia de un inhibidor natural de la ARN polimerasa, tal como T7 lisozima.

15 En otra configuración, el P_L lambda puede ser modificado por medio de ingeniería genética para ser regulado por medio de un parámetro ambiental. Por ejemplo, la célula puede incluir una secuencia de ácido nucleico que codifica una variante sensible a la temperatura del represor lambda. El aumento de las células no tolerantes a la temperatura libera al promotor P_L de la represión.

20 Las propiedades reguladoras de un promotor o secuencia reguladora transcripcional pueden ser fácilmente analizadas enlazando operativamente al promotor o la secuencia a una secuencia que codifica una proteína reportera (o cualquier proteína detectable). Se introduce esta construcción en una célula bacteriana y se evalúa la abundancia de la proteína reportera bajo una variedad de condiciones ambientales. Un promotor útil o secuencia es una que sea selectivamente activada o reprimida en ciertas condiciones.

25 En algunas modalidades, se usan promotores no regulables. Por ejemplo, se puede seleccionar un promotor que produzca una cantidad apropiada de transcripción bajo las condiciones relevantes. Un ejemplo de un promotor no regulable es el promotor del gen III.

30 En una modalidad, se disponen los promotores como se describe en 10/723.981.

Producción y selección de fagos

35 Se pueden utilizar bibliotecas para expresión de fagos para identificar dominios Kunitz que interactúen con un objetivo, por ejemplo, un compuesto objetivo. En una modalidad, el método incluye la amplificación de un miembro de una biblioteca de fagos recuperado en una selección para enlazadores de un compuesto objetivo.

Un ejemplo de un método de selección y amplificación de fagos incluye lo siguiente:

- 40 a. Poner en contacto una pluralidad de diversos fagos de expresión con un compuesto objetivo;
- b. Separar el fago que se enlaza con el compuesto objetivo del fago no enlazado;
- c. Recuperar el fago que se enlazó al compuesto objetivo;
- d. Infectar células huésped con el fago que enlazó al compuesto objetivo;
- e. Producir réplicas de fagos a partir de las células infectadas;
- f. opcionalmente, repetir a. hasta d. una o más veces, por ejemplo, una a seis veces;
- 45 g. Recuperar el fago enlazado o el ácido nucleico dentro del fago, por ejemplo, para caracterización individual.

El método puede ser adaptado para uso ya sea con un fago que contenga genomas de fago o con un fago que contenga fagémidos.

50 Para producir el fago (por ejemplo, en la etapa e., o antes de la etapa a.) se mantienen las células huésped bajo condiciones que proporcionen un nivel seleccionado de actividad transcripcional del promotor regulable, por ejemplo, inducible durante la producción del fago. En un ejemplo en el cual el promotor inducible es un promotor *lac*, se puede incluir un inductor *lac* (por ejemplo, IPTG), o un agente que inhiba la actividad de un promotor *lac* (por

ejemplo, glucosa) en el medio de crecimiento. En una modalidad, se utilizan altas concentraciones de glucosa (por ejemplo, > 1%). En otra modalidad, se utilizan bajas concentraciones de glucosa (por ejemplo, < 0,1%).

5 La regulación de la expresión de una proteína de expresión, por ejemplo, que contenga un dominio Kunitz, puede proporcionar un medio de regulación de la valencia del dominio sobre la superficie de la partícula del fago. Para implementaciones en las cuales el dominio Kunitz se expresa como una fusión con una porción de una proteína de recubrimiento menor, por ejemplo, el dominio ancla de la proteína del gen III, está bajo el control del promotor *lac*, y se co-expresa con otra copia de la proteína de recubrimiento, por ejemplo, una proteína del gen VIII de longitud completa, la valencia de la proteína de expresión puede ser variada de la siguiente manera: En presencia de glucosa, el promotor *lac* será reprimido y la proteína de expresión será expresada con una valencia baja, por ejemplo, 0 - 1 copias por partícula de fago; en presencia de IPTG, el promotor *lac* será inducido, y la proteína de expresión será expresada con una valencia superior, por ejemplo, al menos 1.0, 1.5, 1.8, ó 2.0 copias por partícula de fago, por ejemplo, en promedio.

15 En algunas implementaciones, la proteína de expresión está enlazada a la proteína de recubrimiento mayor (VIII). El fago produce grandes cantidades de proteína del gen VIII (VIII). La secreción parcial de la proteína de expresión enlazada a la VIII madura puede ser menor que la producción de VIII de tipo silvestre. Por lo tanto, también se puede lograr una valencia reducida de la proteína de expresión por medio del enlazamiento de la proteína de expresión con VIII, por ejemplo, VIII madura.

20 Las condiciones para la producción del fago pueden incluir un cambio en la temperatura. La disminución de la temperatura de incubación durante un intervalo de tiempo especificado durante la producción del fago puede facilitar el plegamiento de la secuencia del aminoácido de expresión. Un ejemplo de un procedimiento para el cultivo de células huésped durante la producción del fago incluye un período de incubación de 20 minutos a 37°C seguido por un período de incubación de 25 minutos a 30°C.

25 Después de cualquier ciclo dado de selección, se puede analizar un fago individual por medio del aislamiento de las colonias de células infectadas bajo condiciones de infección de baja multiplicidad. Cada colonia bacteriana se cultiva bajo condiciones que resultan en la producción del fago, por ejemplo, en pozos de microtitulación. Se cosechan los fagos de cada cultivo y se usan en un ensayo de ELISA. El compuesto objetivo se enlaza a un pozo de una placa de microtitulación y se pone en contacto con el fago. Las placas se lavan y se detecta la cantidad de fagos enlazados utilizando, por ejemplo, un anticuerpo para el fago.

30 La selección del fago que enlaza una molécula objetivo incluye poner en contacto al fago con la molécula objetivo. La molécula objetivo puede estar enlazada a un soporte sólido, ya sea directa o indirectamente. Las partículas de fago que se enlazan con el objetivo son luego inmovilizadas y separadas de los miembros que no se enlazan con el objetivo. Las condiciones de la etapa de separación pueden variar en rigurosidad. Por ejemplo, se pueden variar el pH y la fuerza iónica de las condiciones fisiológicas. Se pueden llevar a cabo múltiples ciclos de enlazamiento y separación.

35 Se pueden utilizar métodos covalente y no covalentes para unir moléculas objetivo a un soporte sólido o insoluble. Tales soportes pueden incluir una matriz, cuentas, resina, superficie plana, o inmunotubo. En un ejemplo de un método no covalente de unión, se unen moléculas objetivo a un miembro de un par de enlazamiento. El otro miembro del par de enlazamiento se une a un soporte. Estreptavidina y biotina son un ejemplo de un par de enlazamiento que interactúa con alta afinidad. Otros pares de enlazamiento no covalentes incluyen glutatona-S-transferasa y glutatona (ver, por ejemplo, la Patente Estadounidense No. 5.654.176), hexahistidina y Ni²⁺ (ver, por ejemplo, la Patente Alemana No. DE 19507 166), y un anticuerpo y un epítipo de péptido (ver, por ejemplo, Kolodziej y Young (1991) *Methods Enz.* 194: 508 - 519 para los métodos generales de suministrar una etiqueta de epítipo).

40 Los métodos covalentes de unión de compuestos objetivo incluyen métodos de entrelazamiento químico. Los reactivos pueden crear enlaces covalentes entre grupos funcionales sobre la molécula objetivo y el soporte. Los ejemplos de grupos funcionales que pueden reaccionar químicamente son grupos amino, tiol y carboxilo. N-

etilmaleimida, iodoacetamida, N-hidrosuccinimida, y glutaraldehído son ejemplos de reactivos que reaccionan con grupos funcionales.

5 Los miembros de la biblioteca de expresión pueden ser seleccionados o capturados con una variedad de métodos. El fago puede ser capturado por medio de adherencia a un recipiente, tal como una placa de microtitulación, que está recubierta con una molécula objetivo. Alternativamente, el fago puede hacer contacto con moléculas objetivo que son inmovilizadas dentro de una cámara de flujo, tal como una columna de cromatografía. Las partículas de fago pueden ser capturadas también por medio de partículas magnéticamente sensibles tales como cuentas paramagnéticas. Las cuentas pueden estar recubiertas con un reactivo que puede enlazar al compuesto objetivo (por ejemplo, un anticuerpo), o un reactivo que puede enlazar indirectamente un compuesto objetivo (por ejemplo, cuentas recubiertas con estreptavidina que se enlazan con compuestos objetivo biotinilados).

10 La identificación de miembros útiles de expresión puede ser automatizada. Ver, por ejemplo, US-2003-0129659-A1. Los dispositivos adecuados para automatización incluyen sistemas de transporte de placas de múltiples pozos, procesadores de partículas de cuentas magnéticas, unidades para manejo de líquidos, unidades para recolección de colonias, y otros dispositivos robóticos. Estos dispositivos pueden ser contruidos sobre las especificaciones de los clientes o adquiridos a proveedores comerciales, tales como Autogen (Framingham MA), Beckman Coulter (EUA), Biorobotics (Woburn MA), Genetix (New Milton, Hampshire, RU), Hamilton (Reno NV), Hudson (Springfield NJ), LabSystems (Helsinki, Finlandia), Packard Bioscience (Meriden CT), y Tecan (Mannedorf, Suiza).

20 En algunos casos, los métodos descritos aquí incluyen un proceso automatizado para el manejo de partículas magnéticas. Se inmoviliza el compuesto objetivo sobre las partículas magnéticas. El sistema KingFisher™, un procesador de partículas magnéticas de Thermo LabSystems (Helsinki, Finlandia), por ejemplo, puede ser utilizado para seleccionar miembros de la biblioteca de expresión contra el objetivo. Se pone en contacto la biblioteca de expresión con las partículas magnéticas en un tubo. Se mezclan las cuentas y la biblioteca. Luego, un rodillo magnético, cubierto por una envoltura desechable, recupera las partículas magnéticas y las transfiere a otro tubo que incluye una solución de lavado. Se mezclan las partículas con la solución de lavado. De esta forma, se puede utilizar el procesador de partículas magnéticas para transferir en forma serial las partículas magnéticas a múltiples tubos para lavar a los miembros de la biblioteca enlazados débilmente o en forma no específica de las partículas.

25 Después del lavado, se pueden transferir las partículas a un recipiente que incluya un medio que soporte la amplificación de miembros de la biblioteca de expresión. En el caso de la expresión de fagos, el recipiente puede incluir también células huésped.

35 En algunos casos, por ejemplo, para expresión de fagos, el procesador puede separar también células huésped infectadas de las partículas previamente utilizadas. El procesador puede añadir también un nuevo suministro de partículas magnéticas para una ronda adicional de selección.

40 El uso de la automatización para llevar a cabo la selección puede incrementar la reproducibilidad del proceso de selección así como el rendimiento del procesamiento.

45 Un ejemplo de una partícula de respuesta magnética es el DYNABEAD® que se adquiere con DYNAL BIOTECH (Oslo, Noruega). DYNABEADS® proporciona una superficie esférica de tamaño uniforme, por ejemplo, de 2 µm, 4.5 µm, y 5.0 µm de diámetro. Las cuentas incluyen Fe₂O₃ gamma y Fe₃O₄ como material magnético. Las partículas son súper paramagnéticas ya que tienen propiedades magnéticas en un campo magnético, pero carecen de magnetismo residual fuera del campo. Las partículas se encuentran disponibles con una variedad de superficies, por ejemplo, hidrofílicas con una superficie carboxilatada e hidrófoba con una superficie activada con tosilo. Las partículas también pueden ser bloqueadas con un agente de bloqueo, tal como BSA o caseína para reducir el enlazamiento no específico y el acoplamiento de compuestos diferentes al objetivo con la partícula.

50 El objetivo se une a la partícula paramagnética directa o indirectamente. Se pueden adquirir una variedad de moléculas objetivo en una forma enlazada con partículas paramagnéticas. En un ejemplo, se acopla químicamente un objetivo a una partícula que incluye un grupo reactivo, por ejemplo, un entrelazador (por ejemplo, el éster N-hidroxi-succinimidilo) o un tiol.

5 En otro ejemplo, se enlaza el objetivo a la partícula utilizando un miembro de un par de enlazamiento específico. Por ejemplo, el objetivo puede ser acoplado a biotina. El objetivo es luego enlazado a partículas paramagnéticas que están recubiertas con estreptavidina (por ejemplo, DYNAPARTICLES® M-270 y M-280 recubiertas con Estreptavidina que pueden adquirirse con DYNAL BIOTECH, Oslo, Noruega). En una modalidad, se pone en contacto el objetivo con la muestra antes de la unión del objetivo con las partículas paramagnéticas.

10 En algunas implementaciones, se utiliza también automatización para analizar miembros de la biblioteca de expresión identificados en el proceso de selección. A partir de la muestra final, se pueden obtener clones individuales de cada miembro de expresión. El dominio Kunitz de cada miembro puede ser analizado individualmente, por ejemplo, para evaluar una propiedad funcional. Por ejemplo, el dominio puede ser evaluado para determinar si afecta la actividad enzimática de una proteasa objetivo *in vitro* o *in vivo*. Los métodos para evaluar la actividad de la proteasa y su cinética son bien conocidos. Por ejemplo, se puede evaluar la digestión de sustratos marcados.

15 Los ejemplos de propiedades funcionales incluyen: un parámetro cinético (por ejemplo, para enlazamiento con el compuesto objetivo), un parámetro de equilibrio (por ejemplo, avidéz, afinidad, y así sucesivamente, por ejemplo, para enlazamiento con el compuesto objetivo), una propiedad estructural o bioquímica (por ejemplo, estabilidad térmica, estado de oligomerización, solubilidad y así sucesivamente), y una propiedad fisiológica (por ejemplo, excreción renal, toxicidad, especificidad del tejido objetivo, y así sucesivamente) y así sucesivamente. Los métodos
20 para analizar los parámetros de enlazamiento incluyen ELISA, ensayos de enlazamiento homogéneo, y SPR. Por ejemplo, los ELISA sobre una proteína expresada, por ejemplo, que contiene un dominio Kunitz modificado, pueden ser llevados a cabo directamente, por ejemplo, en el contexto del fago o de otro vehículo de expresión, o la proteína expresada removida del contexto del fago u otro vehículo de expresión.
25 Cada miembro puede también ser secuenciado, por ejemplo, para determinar la secuencia de aminoácidos del dominio Kunitz que es expresado.

Compuestos objetivo

30 En un aspecto, el método se relaciona con la selección de un fago que enlace una molécula objetivo. Cualquier compuesto puede servir como molécula objetivo. La molécula objetivo puede ser una molécula pequeña (por ejemplo, una molécula orgánica o inorgánica pequeña), un polipéptido, un ácido nucleico, un polisacárido, y así sucesivamente. A manera de ejemplo, se describen para objetivos una cantidad de ejemplos y configuraciones. Desde luego, se pueden utilizar los compuestos objetivo diferentes a, o que tienen propiedades diferentes a aquellas enlistadas más adelante.

35 Se pueden utilizar bibliotecas del dominio Kunitz para seleccionar polipéptidos que sean capaces de inhibir proteasas. Por ejemplo, el método puede ser utilizado aquí para identificar dominios Kunitz, particularmente efectivamente dominios Kunitz humanos, que se enlazan y/o que inhiben plasmina, tripsina, quimotripsina, elastasa, y otras proteasas.

40 Se pueden utilizar formas activas e inactivas de la proteasa. Por ejemplo, las formas activas incluyen formas que han sido activadas a partir de zimógeno, por ejemplo, por medio de la remoción de un prodominio. Las proteasas secretadas son típicamente procesadas también para remover una secuencia señal.

45 Las formas inactivas incluyen proteasas químicamente modificadas y proteasas genéticamente modificadas. Incluso otras formas inactivas incluyen proteasas en forma de zimógeno y proteasas que son enlazadas por un inhibidor u otra molécula de inactivación. Las proteasas genéticamente modificadas pueden incluir alteraciones genéticas (por ejemplo, una sustitución, inserción, o supresión) que disminuyen la actividad al menos 20% (por ejemplo, al menos 30, 40, 50, 60, 70, 80, 90, 95 ó 99%). Una alteración puede estar en o cerca del sitio activo, por ejemplo, en un
50 residuo del sitio activo, por ejemplo, un miembro de una triada catalítica.

Por ejemplo, se pueden utilizar proteasas genéticamente modificadas para proporcionar formas inactivas en las cuales el sitio activo no está ocluido. Esta molécula puede ser utilizada, por ejemplo, en una criba o selección inicial, para encontrar dominios Kunitz que se enlacen con el sitio activo, incluso si tales dominios son susceptibles de

escisión por parte de la proteasa objetivo. Se pueden utilizar formas inactivas en las cuales se ocluye el sitio activo (por ejemplo, por medio del enlazamiento de un inhibidor) para descartar dominios Kunitz que interactúan con la proteasa objetivo, pero no en el sitio activo.

5 Las moléculas objetivo de proteína pueden tener una conformación física específica, por ejemplo, una forma plegada o no plegada, o una forma activa o inactiva. En una modalidad, la proteína tiene más de una conformación específica. Por ejemplo, los priones pueden adoptar más de una conformación. Ya sea la conformación nativa o la enferma pueden ser un objetivo deseable, por ejemplo, para aislar agentes que estabilicen la conformación nativa o que identifiquen o elijan como objetivo la conformación enferma.

10 En algunas modalidades, se asocia la proteína objetivo con una enfermedad, por ejemplo, enfermedades y trastornos neoplásicos, cardiovasculares, neurológicos, inflamatorios y pulmonares.

Composiciones farmacéuticas

15 Se suministra una composición que incluya una proteína que contenga un dominio Kunitz que se enlace con un objetivo, por ejemplo, una célula objetivo o molécula (por ejemplo, una proteína objetivo, por ejemplo, una proteasa). La composición puede ser una composición farmacéuticamente aceptable. Por ejemplo, la proteína que contiene al dominio Kunitz puede ser formulada junto con un portador farmacéuticamente aceptable. Como se lo utiliza aquí,
20 "composiciones farmacéuticas" abarca composiciones diagnósticas marcadas (por ejemplo, para formaciones de imágenes *in vivo*) así como composiciones terapéuticas.

25 Como se lo utiliza aquí, "portador farmacéuticamente aceptable" incluye todos y cada uno de los solventes, medios de dispersión, recubrimientos, agentes antibacterianos y antihongos, agentes isotónicos y que retrasan la absorción, y similares que sean fisiológicamente compatibles. Preferiblemente, el portador es adecuado para administración intravenosa, intramuscular, subcutánea, parenteral, espinal o epidérmica (por ejemplo, por medio de inyección o de infusión). Dependiendo de la vía de administración, la proteína que contiene al dominio Kunitz puede estar recubierta en un material para proteger al compuesto de la acción de ácidos y de otras condiciones naturales que puedan inactivarlo.

30 Una "sal farmacéuticamente aceptable" se refiere a una sal que retiene la actividad biológica deseada del compuesto del cual se deriva y no imparte ningún efecto toxicológico indeseado (ver, por ejemplo, Berge, S. M., et al. (1977) J. Pharm. Sci. 66: 1 - 19). Los ejemplos de tales sales incluyen sales de adición ácida y sales de adición básica. Las sales de adición básica incluyen a aquellas derivadas de ácidos inorgánicos no tóxicos, tales como ácido clorhídrico,
35 nítrico, fosfórico, sulfúrico, bromhídrico, yodhídrico, fosforoso y similares, así como las derivadas de ácidos orgánicos no tóxicos tales como ácidos mono y dicarboxílicos, alifáticos, ácidos alcanóicos fenil sustituidos, ácidos hidroxil alcanóicos, ácidos aromáticos, ácidos sulfónicos alifático y aromático y similares. Las sales de adición básica incluyen a aquellas derivadas de metales alcalinotérreos, tales como sodio, potasio, magnesio, calcio y similares, así como de aminas orgánicas no tóxicas, tales como N,N'-dibenciletildiamina, N-metilglucamina, cloroprocaína,
40 colina, dietanolamina, etilendiamina, procaína y similares.

Las composiciones que incluyen una proteína que contiene un dominio Kunitz pueden estar en una variedad de formas. Estas incluyen, por ejemplo, formas de dosificación líquidas, semisólidas y sólidas, tales como soluciones líquidas (por ejemplo, soluciones inyectables y de infusión), dispersiones o suspensiones, tabletas, píldoras, polvos,
45 liposomas y supositorios. Las composiciones típicas están en la forma de soluciones inyectables o para infusión, tales como composiciones similares a aquellas utilizadas para administración de anticuerpos a humanos. Una forma común de administración es en forma parenteral (por ejemplo, intravenosa, subcutánea, intraperitoneal, intramuscular). En una modalidad, se administra la proteína que contiene al dominio Kunitz por medio de una infusión o inyección intravenosa. En otra modalidad, se administra la proteína que contiene el dominio Kunitz por
50 medio de una inyección subcutánea o intramuscular.

La composición puede ser formulada como una solución, microemulsión, dispersión, liposoma, u otra estructura ordenada adecuada para una concentración alta de fármaco. Las soluciones inyectables estériles pueden ser preparadas por medio de la incorporación de la proteína que contiene el dominio Kunitz en la cantidad requerida en

un solvente apropiado con una o una combinación de ingredientes enumerados más arriba, según se requiera, seguido por esterilización por medio de filtración. Generalmente, se preparan las dispersiones por medio de la incorporación del compuesto activo en un vehículo estéril que contiene un medio de dispersión básico y los otros ingredientes requeridos a partir de aquellos enumerados anteriormente. La composición puede ser preparada por medio de un método que incluya secado o deshidratación (por ejemplo, secado al vacío y liofilización), filtración estéril, dispersión de partículas, y adición de tensoactivo. Se pueden utilizar polímeros biodegradables, biocompatibles, tales como acetato de vinil etileno, polianhídridos, ácido poliglicólico, colágeno, poliortoésteres, y ácido poliláctico. Muchos métodos para la preparación de tales formulaciones están patentados o son generalmente conocidos. Ver, por ejemplo, Sustained and Controlled Release Drug Delivery Systems, J. R. Robinson, ed., Marcel Dekker, Inc., New York, 1978.

La proteína que contiene al dominio Kunitz puede ser administrada por medio de una variedad de métodos, por ejemplo, inyección o infusión intravenosa. Por ejemplo, para aplicaciones terapéuticas, se puede administrar la proteína que contiene al dominio Kunitz por medio de infusión intravenosa a una velocidad menor a 30, 20, 10, 5 ó 1 mg/min hasta alcanzar una dosis aproximadamente de 1 a 100 mg/m² ó de 7 a 25 mg/m². La ruta y/o el modo de administración variarán dependiendo de los resultados deseados.

Las composiciones farmacéuticas pueden ser administradas por medio de un dispositivo médico, por ejemplo, dispositivos para inyección hipodérmica sin aguja, implantes, bombas implantables (por ejemplo, bombas de micro infusión), inhaladores, y supositorios. Ver, por ejemplo, las Patentes Estadounidenses Nos. 5.399.163, 5.383.851, 5.312.335, 5.064.413, 4.941.880, 4.790.824, 4.596.556, 4.487.603, 4.486.194, 4.447.233, 4.447.224, 4.439.196, y 4.475.196. Se puede utilizar, por ejemplo una bomba implantable de micro infusión para dispensar una composición a una velocidad controlada.

Los regímenes de dosificación se ajustan para proporcionar la respuesta óptima deseada (por ejemplo, una respuesta terapéutica). Por ejemplo, se puede administrar un solo bolo, se pueden administrar varias dosis divididas en un lapso de tiempo o se puede reducir o incrementar proporcionalmente la dosis según las exigencias de la situación terapéutica. Se pueden formular composiciones parenterales en forma de dosis unitarias para una administración fácil y uniforme de la dosis. Una forma unitaria de dosificación como se la utiliza aquí, se refiere a unidades físicamente discretas adecuadas como dosis unitarias para los individuos que están siendo tratados; cada unidad contiene una cantidad predeterminada de compuesto activo calculada para producir el efecto terapéutico deseado junto con el portador farmacéutico requerido.

Un ejemplo, de un rango no limitante para una cantidad terapéutica o profilácticamente efectiva de una proteína que contiene al dominio Kunitz es de 0,1 - 20 mg/kg, 0,02 - 2 mg/kg, 0,1 - 5 mg/kg, o 1 - 10 mg/kg. La proteína que contiene al dominio Kunitz puede ser administrada por medio de infusión intravenosa a una velocidad de menos de 20, 10, 5, 1, ó 0,3 mg/min para alcanzar una dosis aproximadamente de 1 a 100 mg/m², aproximadamente 5 a 30 mg/m², o aproximadamente 0,5 a 7 mg/m². Los valores de las dosis pueden variar con el tipo y la severidad de la condición que va a ser aliviada. Se pueden ajustar los regímenes de dosis específicos a lo largo del tiempo.

Las composiciones farmacéuticas pueden incluir una cantidad terapéutica o profilácticamente efectiva de una proteína que contiene el dominio Kunitz. Tales cantidades se refieren a una cantidad efectiva, en las dosis y durante los períodos de tiempo necesarios, para lograr el resultado terapéutico o profiláctico deseado. Una "dosis terapéuticamente efectiva" modula preferiblemente un parámetro medible o un síntoma de un trastorno relevante al menos aproximadamente en un 20%, 40%, 60%, u 80% con relación a individuos no tratados.

La habilidad de una proteína que contiene al dominio Kunitz para modular un trastorno puede ser evaluada en un sistema de un modelo animal. Por ejemplo, la habilidad de una proteína que contiene al dominio Kunitz para inhibir al menos un síntoma del cáncer puede ser evaluada en un modelo animal que tenga tumores humanos xenoinjertados. Se pueden utilizar también ensayos *in vitro*.

En una modalidad, la proteína que contiene al dominio Kunitz se conjuga con un agente, por ejemplo, un fármaco citotóxico, o un radioisótopo.

También están contemplados kits que incluyen una proteína que contiene al dominio Kunitz e instrucciones de uso, por ejemplo, el tratamiento, uso profiláctico o diagnóstico. En una modalidad, las instrucciones para las aplicaciones diagnósticas incluyen el uso de una proteína que contiene al dominio Kunitz para detectar un objetivo, *in vitro*, por ejemplo, en una muestra, por ejemplo, una biopsia o células de un paciente que tiene un cáncer o un trastorno neoplásico, o *in vivo*. En otra modalidad, las instrucciones para aplicaciones terapéuticas incluyen dosis sugeridas y/o modos de administración en un paciente con un cáncer o un trastorno neoplásico. El kit puede contener además un reactivo adicional, tal como un agente terapéutico o de diagnóstico, por ejemplo, un agente terapéutico o de diagnóstico como el descrito aquí, y/o una o más proteínas adicionales que incluyen un dominio Kunitz para enlazamiento del objetivo, formulado según sea conveniente, en una o más preparaciones farmacéuticas separadas.

Tratamientos

Las proteínas que contienen al dominio Kunitz identificadas por medio del método descrito aquí y/o detallado aquí tienen utilidad terapéutica o profiláctica. Por ejemplo, estas proteínas pueden ser administradas a células en cultivo, por ejemplo, *in vitro* o *ex vivo*, o en un individuo, por ejemplo, *in vivo*, para tratar, prevenir, y/o diagnosticar una variedad de trastornos, tales como cánceres y una enfermedad cardiovascular.

Como se lo utiliza aquí, el término “tratar” o “tratamiento” se define como la aplicación o la administración de una proteína que contiene el dominio Kunitz a un individuo o una célula o tejido del individuo para prevenir, mejorar, o curar el trastorno o al menos un síntoma de un trastorno. Como se lo utiliza aquí, el término “individuo” incluye un humano, por ejemplo, un paciente que tenga el trastorno, y animales no humanos.

La proteína que contiene al dominio Kunitz puede ser administrada a un individuo humano, por ejemplo, para propósitos terapéuticos, o a un individuo no humano, por ejemplo, para propósitos veterinarios o como un modelo animal de una enfermedad humana.

El método puede ser utilizado para tratar un cáncer incluidos todos los tipos de crecimientos cancerosos o procesos oncogénicos, tejidos con metástasis o células, tejidos, órganos transformados en forma maligna, independientemente del tipo histopatológico o etapa de invasividad.

Los ejemplos de trastornos cancerosos incluyen, pero no se limitan a, tumores sólidos, tumores de tejido blando, y lesiones metastásicas, y cánceres de origen hematopoyético. Los ejemplos de tumores sólidos incluyen neoplasias malignas, por ejemplo, sarcomas, adenocarcinomas y carcinomas, de los diferentes sistemas de órganos, tales como aquellos que afectan al pulmón, mama, linfodeo, gastrointestinal (por ejemplo, colon), y al tracto genitourinario (por ejemplo, células uroteliales, renales), faringe, próstata, y ovario.

El método puede ser utilizado para tratar un trastorno caracterizado por excesiva actividad de una proteasa que la proteína que contiene al dominio Kunitz inhibe. Por ejemplo, la proteasa puede ser una proteasa que puede modificar un factor de coagulación o un componente de la matriz extracelular.

Usos diagnósticos

Las proteínas que contienen al dominio Kunitz también tienen utilidad diagnóstica *in vitro* e *in vivo*. Pueden ser utilizadas, por ejemplo, en un método diagnóstico para detectar la presencia de un objetivo, *in vitro* (por ejemplo, una muestra biológica, tal como tejido, biopsia, por ejemplo, un tejido canceroso) o *in vivo* (por ejemplo, formación de imágenes *in vivo* en un individuo).

El método incluye: (i) poner en contacto una muestra con la proteína que contiene al dominio Kunitz, y (ii) detectar la formación de un complejo entre la proteína que contiene al dominio Kunitz y la muestra. El método puede incluir también poner en contacto una muestra de referencia (por ejemplo, una muestra de control) con la proteína que contiene al dominio Kunitz, y determinar el grado de formación del complejo entre la proteína que contiene al dominio Kunitz y la muestra con relación a la misma para la muestra de referencia. Un cambio, por ejemplo, un cambio estadísticamente significativo, en la formación del complejo en la muestra o el individuo con relación a la muestra de control o al individuo puede ser indicativo de la presencia de un objetivo en la muestra.

Otro método incluye: (i) la administración de la proteína que contiene al dominio Kunitz a un individuo; y (ii) detectar la formación de un complejo entre la proteína que contiene al complejo Kunitz y el individuo. La detección puede incluir determinar la localización o el tiempo de formación del complejo.

- 5 La proteína que contiene al dominio Kunitz puede ser etiquetada directa o indirectamente con una sustancia detectable para facilitar la detección. Las sustancias detectables adecuadas incluyen diferentes enzimas, grupos prostéticos, materiales fluorescentes, materiales luminiscentes y materiales radioactivos. Los ejemplos de moléculas fluorescentes incluyen colorantes de xanteno, por ejemplo, fluoresceína y rodamina, y naftilaminas. Los ejemplos de etiquetas útiles para formación de imágenes diagnósticas de acuerdo con la presente invención incluyen marcadores radioactivos tales como ^{131}I , ^{111}In , ^{123}I , $^{99\text{m}}\text{Tc}$, ^{32}P , ^{125}I , ^3H , ^{14}C , y ^{188}Rh , etiquetas fluorescentes tales como fluoresceína y rodamina, etiquetas activas de resonancia magnética nuclear, isótopos que emiten positrones detectables por medio de un escáner de tomografía de emisión de positrones ("PET"), compuestos quimioluminiscentes tales como luciferina, y marcadores enzimáticos tales como peroxidasa o fosfatasa. Los ejemplos de tales agentes de contraste incluyen agentes paramagnéticos y ferromagnéticos o agentes súper paramagnéticos. Se pueden utilizar quelatos (por ejemplo, los quelatos EDTA, DTPA y NTA) para la unión (y toxicidad reducida) de algunas sustancias paramagnéticas (por ejemplo, Fe^{+3} , Mn^{+2} , Gd^{+3}). Las proteínas que contienen al dominio Kunitz también pueden ser etiquetadas con un grupo indicador que contenga al átomo de ^{19}F activo para RMN.

20 Ejemplos

Ejemplo 1: Generación de insertos de biblioteca de dominio Kunitz

- 25 Se preparó la biblioteca utilizando los oligonucleótidos usando fosforamiditas de trinucleótido activadas para los codones abigarrados y la adición normal de nucleótidos individuales para las regiones constantes. La secuencia de aminoácidos de esta biblioteca del dominio Kunitz está descrita en la FIG. 4.

Tabla 1: Diseño de un primer ejemplo de biblioteca

Posición	Variabilidad	Comentario
11	19	todo excepto C
13	19	todo excepto C
15	18	todo excepto C y P
16	6	AGEDHT
17	18	todo excepto C y P
18	18	todo excepto C y P
19	19	todo excepto C
34	19	todo excepto C
39	19	todo excepto C
40	2	AG
Total	1.73E+11	

- 30 **Ejemplo 2:** Generación de un primer ejemplo de una biblioteca de dominio Kunitz

- Después del montaje de la PCR del inserto de biblioteca, se llevó a cabo una digestión de restricción utilizando enzimas *Nco*I y *Xba*I. Se purificó el inserto de biblioteca y luego se lo ligó dentro del vector de expresión del fago monovalente DY3P82 que había sido digerido y purificado en forma similar. Se transformó el ADN ligado en células electrocompetentes DH5- α resultando en un total de $2,8 \times 10^9$ transformantes.

- El vector de expresión, DY3P82, es resistente a la ampicilina, contiene una copia de longitud completa del gen *iii* y también una copia truncada del gen *iii* como ancla para el dominio Kunitz expresado (Fig. 1). La expresión de la fusión de expresión/gen *III* es controlada por el promotor/operador *lac*. El uso del promotor *lac* permite el control del nivel de expresión y en consecuencia del nivel de expresión por medio de la adición de IPTG (inducción de la expresión) o glucosa (represión de la expresión).

El fago DY3P82 es un derivado de M13mp18. DY3P82 fue construido cambiando el sitio KasI de M13mp18 dentro de un sitio BamHI. El segmento de este sitio BamHI hasta el sitio Bsu36I fue reemplazado con el ADN mostrado. Este contiene un gen *bla* (obtenido a partir de pGemZ3f y modificado por remoción de los sitios de restricción ApaLI y BssSI) y el casete de expresión. Este ejemplo de casete de expresión incluye: a) al promotor PlacZ (entre XhoI y PflMI), b) un sitio de enlazamiento del ribosoma secuencia arriba de la secuencia señal M13, c) una secuencia señal M13 modificada que contiene los sitios de restricción NcoI y EagI, d) partes del dominio LACI-K1 que incluye los sitios NsiI, MluI, AgeI, y XbaI, e) un enlazador que incluye un sitio NheI, f) el tercer dominio del gen III para M13 (el ADN codifica la secuencia de aminoácidos del dominio 3 de M13, pero muchos de los codones son escogidos para que sean diferentes de aquellos del gen III endógeno), g) dos codones de detención, h) un sitio AvrII, i) el terminador *trp*, y j) el sitio NsiI (de los cuales hay dos en el vector).

Ejemplo 3: Expresión por medio de una biblioteca del dominio Kunitz

Se prepararon bibliotecas que contienen al fago ya sea en presencia de glucosa (sin expresión) o en presencia de IPTG (expresión) y se llevó a cabo un ELISA utilizando una preparación de anticuerpo policlonal anti-DX-88. Esta biblioteca también había mejorado las propiedades de manipulación, por ejemplo, una viscosidad mejorada.

DX-88 es un dominio Kunitz derivado de LACI-K1. Los anticuerpos policlonales anti-DX-88 reaccionan en forma cruzada con aislados de bibliotecas de LACI-K1 tanto en transferencias tipo Western como los ELISA.

Se recubrieron los pozos de una placa de microtitulación con un anticuerpo anti-gen VIII para facilitar la captura del fago. Se bloquearon luego los pozos con un reactivo apropiado (tal como BSA) para evitar el enlazamiento no específico y finalmente se lavó con PBS/Tween-20 al 0,05% (PBST). Se aplicaron luego los fagos de la biblioteca a la placa de microtitulación y se permitió que se enlazaran durante 1 hora a 37°C. Se retiraron por lavado los fagos no enlazados, utilizando PBST y un anticuerpo secundario apropiado conjugado con HRP añadido. Después de la incubación con el anticuerpo secundario y el lavado, se desarrolló el ELISA utilizando TMB.

Se incubaron los fagos que expresan DX-88 en forma polivalente como control positivo. Estos fagos no son regulables con IPTG o glucosa. El cambio de las condiciones de inducción (más IPTG) por condiciones de represión (más glucosa) no debe tener efecto sobre el fago DX-88. Los clones R1F4, R2D1, R2F6, R2F8 son anticuerpos aislados clonados dentro del vector fago DY3P82 y se incluyen como controles negativos.

Ejemplo 4: Identificación de inhibidores de la serina proteasa con una biblioteca de dominio Kunitz

La biblioteca Kunitz monovalente ha sido utilizada para identificar enlazadores (y presumiblemente inhibidores) con una serina proteasa recombinante (rSerProteasa-1). Se llevaron a cabo tres rondas de selección. El enlazamiento del fago con el objetivo rSerProteasa-1 biotinilada fue en solución durante dos horas seguido por la captura del complejo fago - objetivo sobre cuentas magnéticas recubiertas con estreptavidina. El análisis ELISA de aislados de fago de la tercera ronda fue llevado a cabo utilizando placas recubiertas con rSerProteasa-1 y un anticuerpo anti-gen VIII para detectar al fago. Se aislaron una cantidad de fagos que se enlazan específicamente con rSerProteasa-1.

Ejemplo 5: Un segundo ejemplo de una biblioteca del dominio Kunitz

Se construyó también una biblioteca de un dominio Kunitz similar a la biblioteca del dominio Kunitz en el Ejemplo 4. La secuencia de aminoácidos de esta biblioteca es mostrada en la FIG. 5. Esta biblioteca contiene sitios adicionales de variación en las posiciones 32, 34, 39, y 40. El tamaño teórico de esta biblioteca es de $2,64 \times 10^{14}$ secuencias únicas de aminoácidos.

En una modalidad, la biblioteca utiliza una proteína de recubrimiento del gen *iii* como ancla. Los dominios Kunitz se expresan a razón aproximadamente de cinco copias por fago. Aunque la proteólisis en el periplasma puede reducir la valencia, cada fago tiene dos o más copias que pueden conducir a efectos de avidéz no deseados.

Tabla 2: Diseño de un segundo ejemplo de biblioteca

Posición	Variabilidad	Comentario
11	19	todo excepto C
13	19	todo excepto C
15	18	todo excepto C y P
16	6	AGEDHT
17	18	todo excepto C y P
18	18	todo excepto C y P
19	19	todo excepto C
32	19	todo excepto C
34	19	todo excepto C
39	19	todo excepto C
40	18	todo excepto C y P
46	18	todo excepto C y P
Total	5.33E+14	

Ejemplo 6

5 Los siguientes son ejemplos de secuencias de vectores

DY3P82

AATGCTACTACTATTAGTAGAATTGATGCCACCTTTTCAGCTCGCGCCCCAAATGAAA
 ATATAGCTAAACAGGTTATTGACCATTTGCGAAATGTATCTAATGGTCAAACATAATC
 TACTCGTTTCGCAGAATTGGGAATCAACTGTTATATGGAATGAAACTTCCAGACACCGT
 ACTTTAGTTGCATATTTAAAACATGTTGAGCTACAGCATTATATTCAGCAATTAAGCT
 CTAAGCCATCCGCAAAAATGACCTCTTATCAAAAGGAGCAATTAAGGTACTCTCTAA
 TCCTGACCTGTTGGAGTTTGCTTCCGGTCTGGTTCGCTTTGAAGCTCGAATTTAAACG
 CGATATTTGAAGTCTTTCGGGCTTCCCTTAAATCTTTTTGATGCAATCCGCTTTGCTT
 CTGACTATAATAGTCAGGGTAAAGACCTGATTTTTGATTTATGGTCATTCTCGTTTTTC
 TGAAGTGTAAAGCATTTGAGGGGGATTCAATGAATATTTATGACGATTCCGCAGTA
 TTGGACGCTATCCAGTCTAAACATTTTACTATTACCCCTCTGGCAAACTTCTTTTG
 CAAAAGCCTCTCGCTATTTTGGTTTTTATCGTCGTCCTGGTAAACGAGGGTTATGATAG
 TGTGCTCTTACTATGCCTCGTAATTCCTTTTGGCGTTATGTATCTGCATTAGTTGAA
 TGTGGTATTCCTAAATCTCAACTGATGAATCTTTCTACCTGTAATAATGTTGTTCCGT
 TAGTTCGTTTTATTAACGTAGATTTTTCTTCCCAACGTCCTGACTGGTATAATGAGCC
 AGTTCTTAAAATCGCATAAGGTAATTCACAATGATTAAAGTTGAAATTAACCCTCTC
 AAGCCCAATTTACTACTCGTTCCTGGTGTCTCTCGTCAGGGCAAGCCTTATTCACTGAA
 TGAGCAGCTTTGTTACGTTGATTTGGGTAATGAATATCCGGTCTTGTCAAGATTACT
 CTTGATGAAGGTCAGCCAGCCTATGCGCCTGGTCTGTACACCGTTCATCTGTCCCTCTT
 TCAAAGTTGGTCAGTTCGGTTCCTTATGATTGACCGTCTGCGCCTCGTTCGGCTAA
 GTAACATGGAGCAGGTCGCGGATTTTCGACACAATTTATCAGGCGATGATACAAATCTC
 CGTTGTACTTTGTTTCGCGCTTGGTATAATCGCTGGGGGTCAAAGATGAGTGTTTTAG
 TGTATTCTTTTGCCTCTTTCGTTTTAGGTTGGTGCCTTCGTAGTGGCATTACGTATTT
 TACCCGTTTAAATGGAACTTCCTCATGAAAAGTCTTTAGTCCTCAAAGCCTCTGTAG
 CCGTTGCTACCCTCGTTCGGATGCTGTCTTTCGCTGCTGAGGGTGACGATCCCGCAA

AGCGGCCTTTAACTCCCTGCAAGCCTCAGCGACCGAATATATCGGTTATGCGTGGGCG
 ATGGTTGTTGTCATTGTCGGCGCAACTATCGGTATCAAGCTGTTAAGAAATTCACCT
 CGAAAGCAAGCTGATAAACCGATACAATTAAGGCTCCTTTTGGAGCCTTTTTTTTTTG
 GAGATTTTCAACGTGAAAAAATTATTATTCGCAATTCCTTTAGTTGTTCTTTCTATT
 CTCCTCCGCTGAAACTGTTGAAAGTTGTTTAGCAAAATCCCATACAGAAAATTCATT
 TACTAACGTCTGGAAAGACGACAAAACCTTTAGATCGTTACGCTAACTATGAGGGCTGT
 CTGTGGAATGCTACAGGCGTTGTAGTTTGTACTGGTGACGAAACTCAGTGTACGGTA
 CATGGGTTCTTATTGGGCTTGCTATCCCTGAAAATGAGGGTGGTGGCTCTGAGGGTGG
 CGGTTCTGAGGGTGGCGGTTCTGAGGGTGGCGGTAATAACCTCTGAGTACGGTGAT
 ACACCTATTCCGGGCTATACTTATATCAACCCTCTCGACGGCACTTATCCGCTGGTA
 CTGAGCAAAACCCCGCTAATCCTAATCCTTCTCTTGAGGAGTCTCAGCCTCTTAATAC
 TTTTCATGTTTTCAGAATAATAGGTTCCGAAATAGGCAGGGGGCATTAACTGTTTATACG
 GGCCTGTTACTCAAGGCACTGACCCCGTTAAACTTATTACCAGTACACTCCTGTAT
 CATCAAAGCCATGTATGACGCTTACTGGAACGGTAAATTCAGAGACTGCGCTTTCCA
 TTCTGGCTTTAATGAGGATTTATTTGTTTGTGAATATCAAGGCCAATCGTCTGACCTG
 CCTCAACCTCCTGTCAATGCTGGCGGCGCTCTGGTGGTGGTTCTGGTGGCGGCTCTG
 AGGGTGGTGGCTCTGAGGGTGGCGGTTCTGAGGGTGGCGGCTCTGAGGGAGGCGGTTT
 CGGTGGTGGCTCTGGTCCGGTGATTTTGATTATGAAAAGATGGCAAACGCTAATAAG
 GGGGCTATGACCGAAAATGCCGATGAAAACGCGCTACAGTCTGACGCTAAAGGCAAAC
 TTGATTCTGTGCTACTGATTACGGTGTCTATCGATGGTTTCATTGGTGACGTTTCT
 CGGCCTTGCTAATGGTAATGGTGCTACTGGTGATTTTGCTGGCTCTAATTCCCAAATG
 GCTCAAGTCGGTGACGGTGATAATTCACCTTTAATGAATAATTTCCGTCAATATTTAC
 CTTCCCTCCCTCAATCGGTTGAATGTGCGCCTTTTGTCTTTGGCGCTGGTAAACCATA
 TGAATTTTCTATTGATTGTGACAAAATAAACTTATTCCGTGGTGTCTTTGCGTTTCTT
 TTATATGTTGCCACCTTTATGTATGTATTTTCTACGTTTGCTAACATACTGCGTAATA
 AGGAGTCTTAATCATGCCAGTTCCTTTTGGGTATTCCGTTATTATTGCGTTTCTCCTCGGT
 TTCTTCTGGTAACCTTTGTTCCGGCTATCTGCTTACTTTTCTTAAAAAGGGCTTCGGTA
 AGATAGCTATTGCTATTTTATTGTTTCTTGCTCTTATTATTGGGCTTAACTCAATTCT
 TGTGGGTTATCTCTGATATTAGCGCTCAATTACCCTCTGACTTTGTTTCAGGGTGT
 CAGTTAATTCTCCCGTCTAATGCGCTTCCCTGTTTTTATGTTATTCTCTCTGTAAAGG
 CTGCTATTTTTCATTTTTGACGTTAAACAAAAAATCGTTTCTTATTGATTGGGATAA
 ATAATATGGCTGTTTATTTTGTAACTGGCAAATAGGCTCTGGAAAGACGCTCGTTAG
 CGTTGGTAAGATTCAGGATAAAATGTAGCTGGGTGCAAAATAGCAACTAATCTTGAT
 TTAAGGCTTCAAACCTCCCGCAAGTCGGGAGGTTGCTAAAACGCCTCGCGTTCTTA
 GAATACCGGATAAGCCTTCTATATCTGATTTGCTTGCTATTGGGCGCGGTAATGATTC
 CTACGATGAAAATAAAAACGGCTTGCTTGTTCTCGATGAGTGCGGTACTTGGTTTAAAT
 ACCCGTTCTTGGAATGATAAGGAAAGACAGCCGATTATTGATTGGTTTCTACATGCTC
 GTAAATTAGGATGGGATATTATTTTTCTTGTTTCAAGACTTATCTATTGTTGATAACA
 GGCGCGTTCTGCATTAGCTGAACATGTTGTTTATTGTCGTCGCTGGACAGAATTACT
 TTACCTTTTGTGCGTACTTTATATTCTCTTATTACTGGCTCGAAAATGCCTCTGCCTA
 AATTACATGTTGGCGTTGTTAAATATGGCGATTCTCAATTAAGCCCTACTGTTGAGCG
 TTGGCTTTTACTGGTAAGAATTTGTATAACGCATATGATACTAAACAGGCTTTTTCT
 AGTAATTATGATTCCGGTGTATTCTTATTAAACGCCTATTTTATCACACGGTCGGT
 ATTTCAAACCATTAAATTTAGGTCAGAAGATGAAATTAATAAAATATATTTGAAAAA
 GTTTTCTCGCGTTCTTTGTCTTGCGATTGGATTTGCATCAGCATTACATATAGTTAT
 ATAACCCAACCTAAGCCGGAGGTTAAAAGGTAGTCTCTCAGACCTATGATTTTGATA
 AATTCATATTGACTCTTCTCAGCGTCTAATCTAAGCTATCGCTATGTTTTCAAGGA
 TTCTAAGGGAAAATTAATTAATAGCGACGATTTACAGAAGCAAGGTTATTCATCACA
 TATATTGATTTATGTAAGTTTCCATTAAAAAAGGTAATTCAAATGAAATTGTAAAT
 GTAATTAATTTGTTTTCTTGATGTTTGTTCATCATCTTCTTTTGCTCAGGTAATTG

AAATGAATAAATTCGCTCTGCGGATTTTGTAACTTGGTATTCAAAGCAATCAGGCGA
 ATCCGTTATTGTTTCTCCCGATGTAAAAGGTACTGTTACTGTATATTCATCTGACGTT
 AAACCTGAAAATCTACGCAATTTCTTTATTTCTGTTTTACGTGCAAATAAATTTTGATA
 TGGTAGGTTCTAACCTTCCATAATTCAGAAGTATAATCCAAACAATCAGGATTATAT
 TGATGAATTGCCATCATCTGATAATCAGGAATATGATGATAAATCCGCTCCTTCTGGT
 GGTTTCTTTGTTCCGCAAATGATAATGTTACTCAAACCTTTAAAATTAATAACGTTT
 GGGCAAAGGATTTAATACGAGTTGTGCAATGTTTGTAAAGTCTAATACTTCTAAATC
 CTCAAATGTATTATCTATTGACGGCTCTAATCTATTAGTTGTTAGTGCTCCTAAAGAT
 ATTTTAGATAACCTTCCCTCAATTCCTTTCAACTGTTGATTTGCCAACTGACCAGATAT
 TGATTGAGGGTTTGATATTTGAGGTTGAGCAAGGTGATGCTTTAGATTTTTCATTTGC
 TGCTGGCTCTCAGCGTGGCACTGTTGCAGGCGGTGTTAATACTGACCGCTCACCTCT
 GTTTTATCTTCTGCTGGTGGTTCGTTCCGTTATTTTAAATGGCGATGTTTTAGGGCTAT
 CAGTTCGCGCATTAAAGACTAATAGCCATTCAAAAATATTTGTCTGTGCCACGTATTCT
 TACGCTTTCAGGTCAGAAGGGTCTATCTCTGTTGGCCAGAATGTCCCTTTTACT
 GGTCGTGTGACTGGTGAATCTGCCAATGTAAATAATCCATTTGAGACGATTGAGCGTC
 AAAATGTAGGTATTTCCATGAGCGTTTTTCTGTTGCAATGGCTGGCGGTAATATTGT
 TCTGGATATTACCAGCAAGGCCGATAGTTGAGTCTTCTACTCAGGCAAGTGATGTT
 ATTACTAATCAAAGAAGTATTGCTACAACGGTTAATTTGCGTGATGGACAGACTCTTT
 TACTCGGTGGCCTCACTGATTATAAAAACACTTCTCAGGATTCTGGCGTACCGTTCCT
 GTCTAAAATCCCTTTAATCGGCCTCCTGTTTAGCTCCCGCTCTGATTCTAACGAGGAA
 AGCACGTTATACGTGCTCGTCAAAGCAACCATAGTACGCGCCCTGTAGCGGCGCATTAA
 AGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAG
 CGCCCGCTCCTTTGCTTTCTTCCCTTCTTTCTCGCCACGTTGCGCGGCTTTCCCG
 TCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTC
 GACCCCAAAAACCTTGATTTGGGTGATGGTTGGCCATCGCCCTGATAGACGGTTTTTC
 GCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAAC
 AACACTCAACCCTATCTCGGGCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTG
 GAACCACCATCAAACAGGATTTTTCGCTGCTGGGGCAAACCAGCGTGGACCGCTTGCT
 GCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACTGGTG
 AAAAGAAAAACCACCCTGGATCCAAGCTTGCAGGTGGCACTTTTTCGGGGAAATGTGCG
 CGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGA
 CAATAACCCTGATAAATGCTTCAATAATATGAAAAAGGAAGAGTATGAGTATTCAAC
 ATTTCCGTGTGCGCCCTTATTCCCTTTTTTTCGCGCATTTTGCCTTCTGTTTTTGCTCA
 CCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGCGCACTAGTGGGT
 TACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAAC
 GTTTTCCAATGATGAGCACTTTTAAAGTCTTGCTATGTGGCGCGGTATTATCCCGTAT
 TGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTT
 GAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTAT
 GCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGAT
 CGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACCTCGC
 CTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCA
 CGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCGAACTACTTAC
 TCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCA
 CTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTTATTGCTGATAAATCTGGAGCCGGTG
 AGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTAT
 CGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATC
 GCTGAGATAGGTGCCTCACTGATTAAGCATTTGGTAACTGTCAGACCAAGTTTACTCAT
 ATATACTTTAGATTGATTTAAAACCTCATTTTTTAAATTTAAAAGGATCTAGGTGAAGAT
 CCTTTTTGATAATCTCATGACCAAAATCCCTTAAACGTGAGTTTTCGTTCCTGTACG
 TAAGACCCCAAGCTTGTGCACTGAATGGCGAATGGCGCTTTGCCTGGTTTCCGGCAC

CAGAAGCGGTGCCGGAAAGCTGGCTGGAGTGCATCTTCCTGACGCTCGAGCGCAACG
CAATTAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCC
GGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTCACACAGGAAACAGCTAT
GACCATGATTACGCCAAGCTTTGGAGCCTTTTTTTTGGAGATTTTCAACGTGAAAAA
TTATTATTTCGCAATTCCTTTAGTTGTTCCCTTTCTATTCCATGGCGGCCGAGATGCATT
CATTCATTTTACGCGTCAGTGCAGGAAAACCGGTTTCGAGTCTCTAGAGGAATGTAA
GAAGATGTGCACTCGTGATTCTGCTAGCTCTGCTAGTGGCGACTTCGACTACGAGAAA
ATGGCTAATGCCAACAAAGGCGCCATGACTGAGAACGCTGACGAGAATGCTTTGCAA
GCGATGCCAAGGGTAAGTTAGACAGCGTCGCGACCGACTATGGCGCCGCCATCGACGG
CTTTATCGGCGATGTCAGTGGTTTGGCCAACGGCAACGGAGCCACCGGAGACTTCGCA
GGTTCGAATTCAGATGGCCAGGTTGGAGATGGGGACAACAGTCCGCTTATGAACA
ACTTTAGACAGTACCTCCGTCTCTCCGCAGAGTGTGAGTGCCGTCCATTCGTTTT
CGGTGCCGGCAAGCCTTACGAGTTCAGCATCGACTGCGATAAGATCAATCTTTTCCGC
GGCGTTTTTCGCTTTCTTGCTATACGTCGCTACTTTCATGTACGTTTTTCAGCACTTTCG
CCAATATTTTACGCAACAAAGAAAGCTAGTGATCTCCTAGGAAGCCCGCCTAATGAGC
GGGCTTTTTTTTTCTGGTATGCATCCTGAGGCCGATACTGTCGTCGTCCCTCAA
GGCAGATGCACGGTTACGATGCGCCCATCTACACCAACGTGACCTATCCCATTACGGT
CAATCCGCCGTTTGTTCACGGAGAATCCGACGGGTTGTTACTCGCTCACATTTAAT
GTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTTTTTTGATGGCGTTCCTA
TTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAATGCGAATTTTAACAAAAATTA
ACGTTTACAATTTAAATATTTGCTTATAACAATCTTCCTGTTTTTTGGGGCTTTTCTGAT
TATCAACCGGGGTACATATGATTGACATGCTAGTTTTACGATTACCGTTCATCGATTC
TCTTGTTTGCTCCAGACTCTCAGGCAATGACCTGATAGCCTTTGTAGATCTCTCAAAA
ATAGCTACCCTCTCCGGCATTAAATTTATCAGCTAGAACGGTTGAATATCATATTGATG
GTGATTTGACTGTCTCCGGCCTTTCTCACCTTTTGAATCTTTACCTACACATTACTC
AGGCATTGCATTTAAAATATATGAGGGTCTAAAAATTTTTATCCTTGCGTTGAAATA
AAGGCTTCTCCCGCAAAGTATTACAGGGTCATAATGTTTTTTGGTACAACCGATTTAG
CTTTATGCTCTGAGGCTTTATTGCTTAATTTTGCTAATCTTTGCCTTGCTGTATGA
TTTATTGGATGTT

DY3P82_LACIK1

AATGCTACTACTATTAGTAGAATTGATGCCACCTTTTCAGCTCGCGCCCCAA
ATGAAAATATAGCTAAACAGGTTATTGACCATTTGCGAAATGTATCTAATG
GTCAAACTAAATCTACTCGTTCGCAGAATTGGGAATCAACTGTTATATGGA
ATGAAACTTCCAGACACCGTACTTTAGTTGCATATTTAAAACATGTTGAGCT
ACAGCATTATATTCAGCAATTAAGCTCTAAGCCATCCGCAAAAATGACCTC
TTATCAAAGGAGCAATTAAGGTAAGCTCTCTAATCCTGACCTGTTGGAGTTT
GCTTCCGGTCTGGTTCGCTTTGAAGCTCGAATTAACGCGATATTTGAAGT
CTTTCGGGCTTCTCTTAATCTTTTTGATGCAATCCGCTTTGCTTCTGACTAT
AATAGTCAGGGTAAAGACCTGATTTTTGATTTATGGTCATTCTCGTTTTCTG
AACTGTTTAAAGCATTGAGGGGGATTCAATGAATATTTATGACGATTCCG
CAGTATTGGACGCTATCCAGTCTAAACATTTACTATTACCCCTCTGGCAA
AACTTCTTTTGCAAAGCCTCTCGCTATTTGGTTTTTATCGTCGTCTGGTAA
ACGAGGGTTATGATAGTGTTGCTCTTACTATGCCTCGTAATTCCTTTGGCG
TTATGTATCTGCATTAGTTGAATGTGGTATTCCTAAATCTCAACTGATGAAT
CTTCTACCTGTAATAATGTTGTTCCGTTAGTTCGTTTTATTAACGTAGATTT
TTCTCCCAACGTCCTGACTGGTATAATGAGCCAGTTCTTAAAATCGCATAA
GGTAATTCACAATGATTAAAGTTGAAATTAACCATCTCAAGCCCAATTA

C T A C T C G T T C T G G T G T T T C T C G T C A G G G C A A G C C T T A T T C A C T G A A T G A G C A
 G C T T T G T T A C G T T G A T T T G G G T A A T G A A T A T C C G G T T C T T G T C A A G A T T A C T
 C T T G A T G A A G G T C A G C C A G C C T A T G C G C C T G G T C T G T A C A C C G T T C A T C T G T
 C C T C T T T C A A A G T T G G T C A G T T C G G T T C C C T T A T G A T T G A C C G T C T G C G C C T
 C G T T C C G G C T A A G T A A C A T G G A G C A G G T C G C G G A T T T C G A C A C A A T T T A T C
 A G G C G A T G A T A C A A A T C T C C G T T G T A C T T T G T T T C G C G C T T G G T A T A A T C G C
 T G G G G G T C A A A G A T G A G T G T T T A G T G T A T T C T T T T G C C T C T T T C G T T T T A G
 G T T G G T G C C T T C G T A G T G G C A T T A C G T A T T T T A C C C G T T T A A T G G A A A C T T C
 C T C A T G A A A A A G T C T T T A G T C C T C A A A G C C T C T G T A G C C G T T G C T A C C C T C G
 T T C C G A T G C T G T C T T T C G C T G C T G A G G G T G A C G A T C C C G C A A A A G C G G C C T T
 T A A C T C C C T G C A A G C C T C A G C G A C C G A A T A T A T C G G T T A T G C G T G G G C G A T
 G G T T G T T G T C A T T G T C G G C G C A A C T A T C G G T A T C A A G C T G T T T A A G A A A T T C
 A C C T C G A A A G C A A G C T G A T A A A C C G A T A C A A T T A A A G G C T C C T T T T G G A G C
 C T T T T T T T T T G G A G A T T T T C A A C G T G A A A A A T T A T T A T T C G C A A T T C C T T T A
 G T T G T T C C T T T C T A T T C T C A C T C C G C T G A A A C T G T T G A A A G T T G T T T A G C A A
 A A T C C C A T A C A G A A A A T T C A T T T A C T A A C G T C T G G A A A G A C G A C A A A A C T T
 T A G A T C G T T A C G C T A A C T A T G A G G G C T G T C T G T G G A A T G C T A C A G G C G T T G
 T A G T T T G T A C T G G T G A C G A A A C T C A G T G T T A C G G T A C A T G G G T T C C T A T T G G
 G C T T G C T A T C C C T G A A A A T G A G G G T G G T G G C T C T G A G G G T G G C G G T T C T G A
 G G G T G G C G G T T C T G A G G G T G G C G G T A C T A A A C C T C C T G A G T A C G G T G A T A C
 A C C T A T T C C G G G C T A T A C T T A T A T C A A C C C T C T C G A C G G C A C T T A T C C G C C T
 G G T A C T G A G C A A A C C C C G C T A A T C C T A A T C C T T C T C T T G A G G A G T C T C A G
 C C T C T T A A T A C T T T C A T G T T T C A G A A T A A T A G G T T C C G A A A T A G G C A G G G G
 G C A T T A A C T G T T T A T A C G G G C A C T G T T A C T C A A G G C A C T G A C C C C G T T A A A
 A C T T A T T A C C A G T A C A C T C C T G T A T C A T C A A A A G C C A T G T A T G A C G C T T A C T
 G G A A C G G T A A A T T C A G A G A C T G C G C T T T C C A T T C T G G C T T T A A T G A G G A T T T
 A T T T G T T T G T G A A T A T C A A G G C C A A T C G T C T G A C C T G C C T C A A C C T C C T G T C
 A A T G C T G G C G G C G G C T C T G G T G G T G G T T C T G G T G G C G G C T C T G A G G G T G G T
 G G C T C T G A G G G T G G C G G T T C T G A G G G T G G C G G C T C T G A G G G A G G C G G T T C C
 G G T G G T G G C T C T G G T T C C G G T G A T T T T G A T T A T G A A A A G A T G G C A A A C G C T
 A A T A A G G G G G C T A T G A C C G A A A A T G C C G A T G A A A A C G C G C T A C A G T C T G A
 C G C T A A A G G C A A A C T T G A T T C T G T C G C T A C T G A T T A C G G T G C T G C T A T C G A T
 G G T T C A T T G G T G A C G T T T C C G G C C T T G C T A A T G G T A A T G G T G C T A C T G G T G
 A T T T T G C T G G C T C T A A T T C C C A A A T G G C T C A A G T C G G T G A C G G T G A T A A T T C
 A C C T T T A A T G A A T A A T T T C C G T C A A T A T T T A C C T T C C C T C C C T C A A T C G G T T G
 A A T G T C G C C C T T T T G T C T T T G G C G C T G G T A A A C C A T A T G A A T T T T C T A T T G A
 T T G T G A C A A A A T A A A C T T A T T C C G T G G T G T C T T T G C G T T T C T T T T A T A T G T G
 C C A C C T T T A T G T A T G T A T T T T C T A C G T T T G C T A A C A T A C T G C G T A A T A A G G A
 G T C T T A A T C A T G C C A G T T C T T T T G G G T A T T C C G T T A T T A T T G C G T T T C C T C G G
 T T T C C T T C T G G T A A C T T T G T T C G G C T A T C T G C T T A C T T T T C T T A A A A A G G G C T
 T C G G T A A G A T A G C T A T T G C T A T T T C A T T G T T T C T T G C T C T T A T T A T T G G G C T T
 A A C T C A A T T C T T G T G G G T T A T C T C T C T G A T T A T A G C G C T C A A T T A C C C T C T G
 A C T T T G T T C A G G G T G T T C A G T T A A T T C T C C C G T C T A A T G C G C T T C C C T G T T T
 T A T G T T A T T C T C T C T G T A A A G G C T G C T A T T T T C A T T T T T G A C G T T A A A C A A A
 A A A T C G T T T C T T A T T T G G A T T G G G A T A A A T A A T A T G G C T G T T T A T T T T G T A A
 C T G G C A A A T T A G G C T C T G G A A A G A C G C T C G T T A G C G T T G G T A A G A T T C A G G
 A T A A A A T T G T A G C T G G G T G C A A A A T A G C A A C T A A T C T T G A T T T A A G G C T T C
 A A A A C C T C C C G C A A G T C G G G A G G T T C G C T A A A A C G C C T C G C G T T C T T A G A A
 T A C C G G A T A A G C C T T C T A T A T C T G A T T T G C T T G C T A T T G G G C G C G G T A A T G A

TTCCTACGATGAAAATAAAAACGGCTTGCTTGTTCGATGAGTGCGGTACT
 TGGTTTAATACCCGTTCTTGGAATGATAAGGAAAGACAGCCGATTATTGAT
 TGGTTTCTACATGCTCGTAAATTAGGATGGGATATTATTTTCTTGTTGAGG
 ACTTATCTATTGTTGATAAACAGGCGCGTTCTGCATTAGCTGAACATGTTGT
 TTATTGTCGTCGTCCTGGACAGAATTACTTTACCTTTTGTGCGTACTTTATATT
 CTCTTATTACTGGCTCGAAAATGCCTCTGCCTAAATTACATGTTGGCGTTGT
 TAAATATGGCGATTCTCAATTAAGCCCTACTGTTGAGCGTTGGCTTTATACT
 GGTAAGAATTTGTATAACGCATATGATACTAAACAGGCTTTTTCTAGTAATT
 ATGATTCCGGTGTTTATTCTTATTTAACGCCTATTTATCACACGGTCGGTAT
 TTCAAACCATTAAATTTAGGTCAGAAGATGAAATTAACATAAAATATATTTG
 AAAAAGTTTCTCGCGTTCTTTGTCTTGCGATTGGATTGTCATCAGCATTTA
 CATATAGTTATATAACCCAACCTAAGCCGGAGGTTAAAAGGTAGTCTCTC
 AGACCTATGATTTTGATAAATCACTATTGACTCTTCTCAGCGTCTTAATCT
 AAGCTATCGCTATGTTTTCAAGGATTCTAAGGGAAAATTAATTAATAGCGA
 CGATTTACAGAAGCAAGGTTATTCACTCACATATATTGATTTATGTACTGTT
 TCCATTAATAAAGGTAATTCAAATGAAATTGTTAAATGTAATTAATTTTGT
 TTCTTGATGTTTGTTCATCATCTTCTTTTGTGTCAGGTAATTGAAATGAATAA
 TTCGCCCTCGCGGATTTTGTAACCTGGTATTCAAAGCAATCAGGCGAATCC
 GTTATTGTTTCTCCCGATGTAAGGTAAGTACTGTTACTGTATATTCACTGACG
 TTAAACCTGAAAATCTACGCAATTTCTTTATTTCTGTTTTACGTGCAAATAA
 TTTTGATATGGTAGGTTCTAACCTTCCATAATTCAGAAGTATAATCCAAC
 AATCAGGATTATATTGATGAATTGCCATCATCTGATAATCAGGAATATGAT
 GATAATCCGCTCCTTCTGGTGGTTTCTTTGTTCCGCAAATGATAATGTTA
 CTCAAACTTTTAAATTAATAACGTTCCGGGCAAAGGATTTAATACGAGTTG
 TCGAATTGTTTGTAAGTCTAATACTTCTAAATCCTCAAATGTATTATCTAT
 TGACGGCTCTAATCTATTAGTTGTTAGTGCTCCTAAAGATATTTTAGATAAC
 CTTCCCTCAATTCCTTCAACTGTTGATTTGCCAACTGACCAGATATTGATTG
 AGGGTTTGATATTTGAGGTTGAGCAAGGTGATGCTTTAGATTTTTCATTTGC
 TGCTGGCTCTCAGCGTGGCACTGTTGCAGGCGGTGTTAATACTGACCGCCTC
 ACCTCTGTTTTATCTTCTGCTGGTGGTTCGTTCCGGTATTTTAAATGGCGATGT
 TTTAGGGCTATCAGTTCGCGCATTAAAGACTAATAGCCATTCAAAAATATT
 GTCTGTGCCACGATTCTTACGCTTTCAGGTCAGAAGGGTCTATCTCTGTT
 GGCCAGAATGTCCCTTTTATTACTGGTTCGTGTGACTGGTGAATCTGCCAATG
 TAAATAATCCATTTGAGACGATTGAGCGTCAAATGTAGGTATTTCCATGA
 GCGTTTTTCTGTTGCAATGGCTGGCGGTAATATTGTTCTGGATATTACCAG
 CAAGGCCGATAGTTTGAAGTCTTCTACTCAGGCAAGTGATGTTATTACTAAT
 CAAAGAAGTATTGCTACAACGGTTAATTTGCGTGATGGACAGACTCTTTTA
 CTCGGTGGCCTCACTGATTATAAAAACACTTCTCAGGATTCTGGCGTACCGT
 TCCTGTCTAAAATCCCTTTAATCGGCCTCCTGTTTAGCTCCCGCTCTGATTCT
 AACGAGGAAAGCACGTTATACGTGCTCGTCAAAGCAACCATAGTACGCGCC
 CTGTAGCGGCGCATTAAAGCGCGGGGGTGTGGTGGTTACGCGCAGCGTGAC
 CGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTGCTTTCTTCCCTTCT
 TTCTCGCCACGTTGCGCGGCTTTCCCGTCAAGCTCTAAATCGGGGGCTCCC
 TTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACCTTGAT
 TTGGGTGATGGTTGGCCATCGCCCTGATAGACGGTTTTTTCGCCCTTTGACGT
 TGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACT
 CAACCCTATCTCGGGCTATTCTTTTATTGTTTATAAGGGATTTTGCCGATTTG
 GAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACC
 GCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGC

CCGTCTCACTGGTGAAAAGAAAAACCACCCTGGATCCAAGCTTGCAGGTGG
 CACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATA
 CATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAAT
 AATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTA
 TTCCCTTTTTTTCGGGCATTTTGCCTTCCTGTTTTTGTCTACCCAGAAACGCTG
 GTGAAAGTAAAAGATGCTGAAGATCAGTTGGGCGCACTAGTGGGTACATC
 GAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAA
 CGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTAT
 CCCGTATTGACGCCGGGCAAGAGCAACTCGGTGCGCCGATACACTATTCTC
 AGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATG
 GCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACA
 CTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCG
 CTTTTTTCACAACATGGGGGATCATGTAACCTCGCCTTGATCGTTGGGAACC
 GGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTG
 TAGCAATGGCAACAACGTTGCGCAAACCTATTAAGTGGCGAACTACTTACTC
 TAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAG
 GACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATC
 TGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGA
 TGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAAC
 TATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAA
 GCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTA
 AAACCTTCATTTTAAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATC
 TCATGACCAAATCCCTTAAACGTGAGTTTTCGTTCCACTGTACGTAAGACCC
 CCAAGCTTGTGCGACTGAATGGCGAATGGCGCTTTGCCTGGTTTCCGGCACC
 AGAAGCGGTGCCGAAAGCTGGCTGGAGTGCATCTTCTGACGCTCGAGC
 GCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTAC
 ACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATT
 TCACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTTGGAGCCTTTT
 TTTTGGAGATTTTCAACGTGAAAAAATTATTATTCGCAATTCCTTTAGTTGT
 TCCTTTCTATTCCATGGCGGCCGAGATGCATTCACTCTGCGCTTCAAAGCT
 GATGACGGTCCGTGTAAAGCTATCATGAAACGTTTCTTCTTCAACATTTTCA
 CGCGTCAATGTGAAGAGTTCATTTACGGTGGTTGTGAAGGTAACCAGAACC
 GGTTTCGAATCTCTAGAGGAATGTAAGAAGATGTGCACTCGTGATTCTGCTA
 GCTCTGCTAGTGGCGACTTCGACTACGAGAAAATGGCTAATGCCAACAAAG
 GCGCCATGACTGAGAACGCTGACGAGAATGCTTTGCAAAGCGATGCCAAG
 GGTAAGTTAGACAGCGTTCGCGACCGACTATGGCGCCGCCATCGACGGCTTT
 ATCGGCGATGTCAGTGGTTTGGCCAACGGCAACGGAGCCACCGGAGACTTC
 GCAGGTTTCGAATTCTCAGATGGCCCAGGTTGGAGATGGGGACAACAGTCCG
 CTTATGAACAACCTTAGACAGTACCTTCCGTCTCTCCGCAGAGTGTGAGT
 GCCGTCCATTGTTTTCCGGTGCCGGCAAGCCTTACGAGTTCAGCATCGACTG
 CGATAAGATCAATCTTTCCGCGGCGTTTTCGCTTTCTTGCTATACGTCGCT
 ACTTTCATGTACGTTTTTCAGCACTTTCGCCAATATTTTACGCAACAAAGAAA
 GCTAGTGATCTCCTAGGAAGCCCGCTAATGAGCGGGCTTTTTTTTTCTGGT
 ATGCATCCTGAGGCCGATACTGTGCTCGTCCCCTCAAACCTGGCAGATGCAC
 GGTTACGATGCGCCATCTACACCAACGTGACCTATCCATTACGGTCAAT
 CCGCCGTTTGTTCACGGAATCCGACGGGTTGTTACTCGCTCACATTTA
 ATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTATTTTTGATG
 GCGTTCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAATGCGAA
 TTTTAACAAAATATTAACGTTTACAATTTAAATATTTGCTTATACAATCTTC

CTGTTTTGGGGCTTTTCTGATTATCAACCGGGGTACATATGATTGACATGC
TAGTTTTACGATTACCGTTCATCGATTCTCTTGTTGCTCCAGACTCTCAGGC
AATGACCTGATAGCCTTTGTAGATCTCTCAAAAATAGCTACCCTCTCCGGCA
TTAATTTATCAGCTAGAACGGTTGAATATCATATTGATGGTGATTTGACTGT
CTCCGGCCTTTCTCACCCCTTTTGAATCITTACCTACACATTACTCAGGCATTG
CATTAAAATATATGAGGGTTCATAAAAATTTTTATCCTTGCGTTGAAATAAA
GGCTTCTCCCGCAAAAGTATTACAGGGTCATAATGTTTTTGGTACAACCGAT
TTAGCTTTATGCTCTGAGGCTTATTGCTTAATTTTGCTAATTCTTTGCCTTG
CCTGTATGATTTATTGGATGTT

5

10

15

20

25

30

35

40

REIVINDICACIONES

- 5 1. Una biblioteca que comprende una pluralidad de partículas de fago filamentosos, cada partícula de fago de la pluralidad incluyendo (i) una proteína de expresión que comprende un dominio Kunitz y al menos una porción de una proteína de recubrimiento de fago del gen III que se une físicamente a la partícula del fago y que se fusiona con el dominio Kunitz, (ii) un ácido nucleico que incluye (a) genes de fago suficientes para producir una partícula infecciosa de fago, dichos genes de fago comprendiendo un gen que codifica una proteína de recubrimiento de fago del gen III que está (i') no fusionada a una secuencia de aminoácidos no viral de al menos cinco aminoácidos o (ii') no fusionada a una secuencia modificada de aminoácidos y (b) una secuencia que codifica la proteína de expresión, en donde el dominio Kunitz incluye al menos una posición modificada de aminoácido en cada uno de los dos bucles de interacción, siendo modificadas las posiciones en los bucles respectivos entre las partículas de la pluralidad, y en donde el número promedio de copias del dominio Kunitz por partículas de fago de la pluralidad es menor a 2,0.
- 10 2. La biblioteca de la reivindicación 1 en donde el dominio Kunitz es al menos 85% idéntico a un dominio Kunitz humano en las posiciones de los aminoácidos que no se modifican.
- 15 3. La biblioteca de la reivindicación 1 en donde dicho dominio Kunitz comprende entre 5 y 12 posiciones modificadas de aminoácidos entre las partículas de la pluralidad.
- 20 4. La biblioteca de la reivindicación 3 en donde dichas posiciones modificadas de los aminoácidos incluyen posiciones de los aminoácidos correspondientes a las posiciones de los aminoácidos 16, 17, 18, 19, 34, y 39 de LACI-K1 humana.
- 25 5. La biblioteca de la reivindicación 4 en donde dichas posiciones modificadas de los aminoácidos incluyen posiciones de los aminoácidos 11, 13, 15, 16, 17, 18, 19, 34, y 39 de LACI-K1 humana.
- 30 6. La biblioteca de la reivindicación 3 en donde dichas posiciones modificadas de los aminoácidos consisten de las posiciones de los aminoácidos correspondientes a las posiciones de los aminoácidos 11, 13, 15, 16, 17, 18, 19, 34, 39 y 40 de LACI-K1 humana.
- 35 7. La biblioteca de la reivindicación 1 en donde la posición del aminoácido correspondiente a la posición del aminoácido 40 es G o A, o está modificada entre G y A.
- 40 8. La biblioteca de la reivindicación 1, en donde el dominio Kunitz comprende una secuencia que es al menos 85% idéntica a MHSFCFAFKADX₁₁GX₁₃CX₁₅X₁₆X₁₇X₁₈X₁₉RFFFNIFTRQCEEFX₃₄YGGCX₃₉X₄₀N QNRFESLEECKKMCTRDGA (SEQ ID NO: 10), en posiciones diferentes a X; y el dominio Kunitz difiere de la secuencia de LACI-K1 en al menos un residuo aminoácido; y
- 45 X₁₁ es uno de: A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
X₁₃ es uno de: A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
X₁₅ es uno de: A, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, Y;
X₁₆ es uno de: A, G, E, D, H, T;
X₁₇ es uno de: A, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, Y;
X₁₈ es uno de: A, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, Y;
45 X₁₉ es uno de: A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
X₃₄ es uno de: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
X₃₉ es uno de: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y; y
X₄₀ es uno de: G, A, y
- 50 en donde al menos dos de X₁₁, X₁₃, X₁₅, X₁₆, X₁₇, X₁₈, X₁₉, X₃₄, X₃₉, y X₄₀ varían entre partículas de la pluralidad, estando al menos dos de las posiciones modificadas en el primero y en el segundo bucles de interacción del dominio Kunitz.
9. La biblioteca de la reivindicación 8 en donde la proteína de expresión incluye al muñón de la proteína de recubrimiento del gen III.

10. La biblioteca de la reivindicación 8 en donde los genes del fago incluyen un gen que codifica una proteína de recubrimiento del gen III de tipo silvestre.
- 5 11. La biblioteca de la reivindicación 8 en donde la biblioteca tiene una diversidad teórica entre 10^3 y 10^{12} .
12. La biblioteca de la reivindicación 11 en donde al menos las posiciones de los aminoácidos 15, 16, 17, 18, 34 y 39 son modificadas.
- 10 13. La biblioteca de la reivindicación 11 en donde las posiciones de los aminoácidos 11, 13, 15, 16, 17, 18, 19, 34, 39 y 40 son modificadas.
14. La biblioteca de la reivindicación 8 en donde el número promedio de copias del dominio Kunitz por partículas de fago de la pluralidad es menor a 1.5.
- 15 15. La biblioteca de la reivindicación 1 en donde el dominio Kunitz es al menos 85% idéntico a LACI-K1 humana en las posiciones de los aminoácidos que no son modificadas.
16. La biblioteca de la reivindicación 1 en donde el dominio Kunitz es idéntico a LACI-K1 humana en las posiciones de los aminoácidos que no son modificadas.
- 20 17. La biblioteca de la reivindicación 8 en donde el dominio Kunitz incluye una secuencia que es 100% idéntica a MHSFCAFKADX₁₁GX₁₃CX₁₅X₁₆X₁₇X₁₈X₁₉RFFFNIFTRQCEEFX₃₄YGGCX₃₉X₄₀N QNRFESLEECKKMCTRDGA (SEQ ID NO: 10), en posiciones diferentes a X.
- 25 18. La biblioteca de la reivindicación 8, cada partícula fago de la pluralidad incluye adicionalmente (iii) una proteína de recubrimiento de fago del gen III funcional, particularmente en donde la biblioteca está definida adicionalmente como en cualquiera de las reivindicaciones 9 - 14 ó 17.
- 30 19. Un método para proporcionar un ácido nucleico que codifica un dominio Kunitz que interactúa con un objetivo, comprendiendo el método:
proporcionar la biblioteca de la reivindicación 1;
poner en contacto las partículas de fago de la biblioteca con un objetivo; y
recuperar el ácido nucleico de las partículas que interactúan con el objetivo.
- 35 20. El método de la reivindicación 19 en donde el objetivo está inmovilizado, y la etapa de recuperación comprende la separación de las partículas que interactúan con el objetivo de las partículas que no interactúan con el objetivo.
- 40 21. Un método para proporcionar ácido nucleico que codifica un dominio Kunitz que interactúa con un objetivo, comprendiendo el método:
proporcionar la biblioteca de la reivindicación 8 ó 18;
poner en contacto las partículas de fago de la biblioteca con un objetivo; y
recuperar el ácido nucleico de las partículas que interactúan con el objetivo.
- 45 22. El método de la reivindicación 21 en donde el objetivo es una proteasa.
23. Un vector fago que comprende:
(a) genes de fago suficientes para producir una partícula infecciosa de fago y (b) una secuencia que codifica una proteína de expresión comprendiendo la proteína de expresión un dominio Kunitz y al menos una porción de una
50 proteína de recubrimiento de fago del gen III que se une físicamente a la partícula de fago y que se fusiona al dominio Kunitz, comprendiendo el dominio Kunitz una secuencia que es al menos 85% idéntica a MHSFCAFKADX₁₁GX₁₃CX₁₅X₁₆X₁₇X₁₈X₁₉RFFFNIFTRQCEEFX₃₄YGGCX₃₉X₄₀N QNRFESLEECKKMCTRDGA (SEQ ID NO: 10), en posiciones diferentes a X;
X₁₁ es uno de: A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

- X₁₃ es uno de: A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
X₁₅ es uno de: A, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, Y;
X₁₆ es uno de: A, G, E, D, H, T;
X₁₇ es uno de: A, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, Y;
5 X₁₈ es uno de: A, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, Y;
X₁₉ es uno de: A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
X₃₄ es uno de: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
X₃₉ es uno de: A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y; y
X₄₀ es uno de: G, A, en donde los genes de fago incluyen un gen que codifica a la proteína de recubrimiento de fago
10 del gen III que no está fusionada a una secuencia heteróloga que codifica más de cinco aminoácidos.

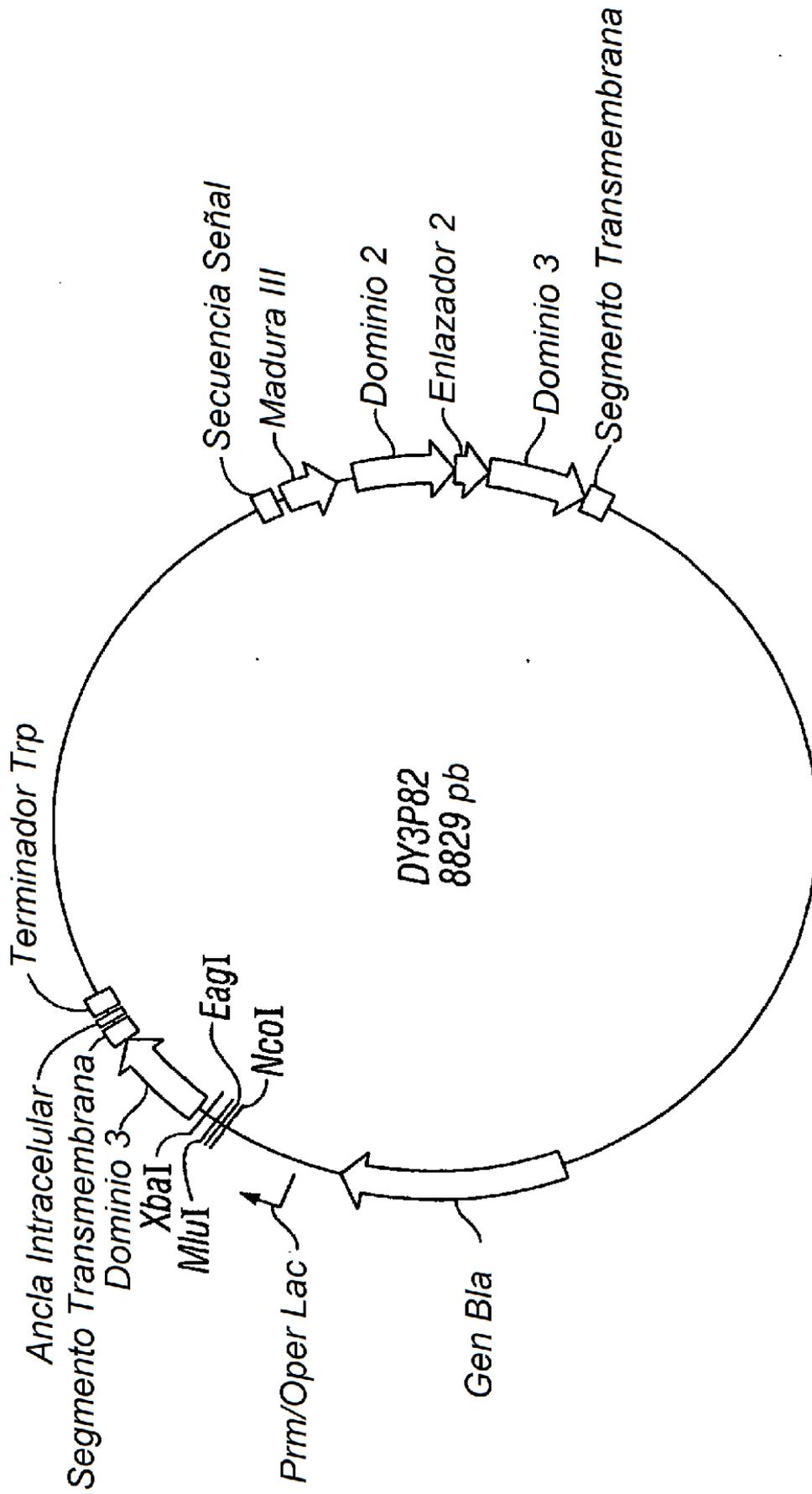


FIG. 1

```

1 2 3 4 5 6 7 8 9
M H S F C A F K A
ATG CAT TCC TTC TGC GCC TTC AAG GCC

10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
D X G X C X X X X X R F F F N
GAC <2> GGT <2> TGT <1> TGT <1> <1> <1> <2> CGT TTC TTC TTC AAC

25 26 27 28 29 30 31 32 33 34 35 36 37 38 39
I F T R Q C E E F X Y G G C X
ATC TTC ACG CGT CAG TGC GAA GAG TTC <2> TAC GGT GGT TGT <2>

40 41 42 43 44 45 46 47 48 49 50 51 52 53 54
X N Q N R F E S L E E C K K M
<4> AAC CAG AAC CGG TTC GAA TCT CTA GAG GAA TGT AAG AAG ATG

55 56 57 58
C T R D SEQ ID NO:3
TGC ACT CGT GAC SEQ ID NO:4

<1> = { A DEFGHIKLMN QRSTVWY } [no C o P]
<2> = { A DEFGHIKLMNPQRSTVWY } [no C]
<3> = { .65 A, .15 G, 0.05(EDHT) }
<4> = { A G}

```

FIG. 2

```

1 2 3 4 5 6 7 8 9
M H S F C A F K A
  ATG CAT TCC TTC TGC GCC TTC AAG GCC

10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
D X G X C X X X X X X R F F F N
  GAC <2> GGT <2> TGT <1> <3> <1> <1> <2> <2> CGT TTC TTC TTC AAC

25 26 27 28 29 30 31 32 33 34 35 36 37 38 39
I F T R Q C E X F X Y G C X
  ATC TTC ACG CGT CAG TGC GAG <2> TTC <2> TAC GGT GGT TGT <2>

40 41 42 43 44 45 46 47 48 49 50 51 52 53 54
X N Q N R F X S L E E C K K M
<1> AAC CAG AAC CGG TTC <1> TCT CTA GAG GAA TGT AAG AAG ATG

55 56 57 58
C T R D      SEQ ID NO:5
TGC ACT CGT GAC      SEQ ID NO:6

<1> = { A DEFGHIKLMN QRSTVWY } [no C o P]
<2> = { A DEFGHIKLMNPQRSTVWY } [no C]
<3> = { .65 A, .15 G, 0.05 (EDHT) }

```

FIG. 3

```

7244          cgcaacgc aattaatgtg agttagctca
7272 ctcattagc acccaggct ttacacttta tgctccggc tcgtatgttg
7322 tgtggaattg tgagcggata acaatttcac acaggaaca gctatgacca
7372 tgattacg   CCaagctt  TGGa gcctttttt  tggagatttt caac
          PflMI.....

1  2  3  4  5  6  7  8  9  10 11 12 13 14 15
M  K  K  L  L  F  A  I  P  L  V  V  P  F  Y
gtg aaa aaa tta tta ttc ttc gca att cct tta gtt gtt cct ttc tat
señal M13, codones 1-15.

16 17 18      a  b  1  2  3  4
S  M  A      A  E  M  H  S  F
tcc ATG Gcg |gcc|gag|ATG|CAT|tca|ttc
          NcoI....EagI.....

5  6  7  8  9  10 11 12 13 14 15 16 17 18 19 20 21
C  A  F  K  A  D  D  G  P  C  K  A  I  M  K  R  F
          <2> <2> <1> <3> <1> <1> <2>
7488 tgc gct ttc aaa gct gat gac GGT CCG tgt aaa gct aTC ATG Aaa cgt ttc

<1> = { A DEFGHIKLMN QRSTVWY } equimolar (no C o P)
<2> = { A DEFGHIKLMNPQRSTVWY } equimolar (no C)
<3> = .65 A, .15 G, 0.05 (EDHT)
<4> = (AG) equimolar

```

FIG. 4A

22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38
 F F N I F T R Q C E E F I Y G G C

<2>

7539 ttc ttc aac att ttc ACG CGT caa tgt gaa gag ttc att tac ggt ggt tgt

MluI...

39 40 41 42 43 44 45 46 47 48 49
 E G N Q N R F E S L E

<2> <4>

7590 gaa GGT AAC Cag aAC CGG Ttc gaa tct CTA GAG

XbaI....

50 51 52 53 54 55
 E C K K M C

7623 gaa tgt aag aag atg tgc

56 57 58
 T R D

S A S S A

7641 act|cgt|gat| tct GCT AGC tct gct

NheI...

FIG. 4B

muñón de III

 Dominio 3 de III

S G D D F D Y E K M A N A N K G A
 7665 agt ggc gac ttc gac tac gag aaa atg gct aat gcc aac aaa ggc gcc

M T E N A D E N A L Q S D A K G
 7713 atG ACT GAG AAC GCT GAC GAG aat gct ttg caa agc gat gcc aag ggt

K L D S V A T D Y G A A I D G F
 7761 aag tta gac agc gTC GCG Acc gac tat GGC GCC gcc ATC GAC ggc ttt
 NruI....

I G D V S G L A N G N G A T G D
 7809 atc ggc gat gtc agt ggt ttG GCC Aac ggc aac gga gcc acc gga gac

F A G S N S Q M A Q V G D G D N
 7857 ttc GCA GGT tcG AAT Tct cag atg gcC CAG gtt gga gaT GGg gac aac
 EcoRI...
 XcmI.....

S P L M N N F R Q Y L P S L P Q
 7905 agt ccg ctt atg aac aac ttt aga cag tac ctt ccg tct ctt ccg cag

S V E C R P F V F G A G K P Y E
 7953 agt gtc gag tgc cgt cca ttc gtt ttc GGT gcc ggc aag cct tac gag

FIG. 4C

```

F S I D C D K I N L F R
8001 ttc aGC Atc gac TGC gat aAg atc aat ctt ttc CGC
      BstAPI.....
      SacII...
      Extremo del Dominio 3

G V F A F L L Y V A T F M Y V F
8037 Ggc gtt ttc gct ttc ttg cta tac gtc gct act ttc atg tac gtt ttc
      inicio del segmento transmembrana

S T F A N I L R N K E S
8085 agc act ttc gcc AAT ATT tta cgc aac aaa gaa agc
      Ancla intracelular.

      .
8121 tag tga tct CCT AGG
      AvrII..

8136 aaG CCC Gcc taa tga GCG GGC ttt ttt ttt ct ggt ATGCAT CCTGAGG
      Tallo ... Bucle ... Tallo ... Multit.....
      | Terminador Trp
      Extremo del casete de expresion
  
```

FIG. 4D

REFERENCIAS CITADAS EN LA DESCRIPCIÓN

Este listado de referencias citado por el solicitante es únicamente para conveniencia del lector. No forma parte del documento europeo de la patente. Aunque se ha tenido gran cuidado en la recopilación, no se pueden excluir los errores o las omisiones y la OEP rechaza toda responsabilidad en este sentido.

Documentos de patente citados en la descripción

- 10 • US 5223409 A, Ladner [0001] [0118] [0119] [0129]
- WO 9218619 A [0001] [0119]
- WO 9117271 A [0001] [0119]
- WO 9220791 A [0001] [0119]
- WO 9215679 A [0001] [0119]
- 15 • WO 9301288 A [0001] [0119]
- WO 9201047 A [0001] [0119]
- WO 9209690 A [0001] [0119]
- WO 9002809 A [0001] [0119]
- US 6423498 B1 [0003]
- US 4760025 A [0111]
- 20 • US 5869644 A [0111]
- WO 0105950 A [0118]
- US 5821047 A [0122]
- WO 0071694 A [0127]
- US 5658727 A [0128]
- 25 • US 6207466 B [0131]
- WO 03029456 A [0131]
- WO 0140803 A [0131]
- WO 9951773 A [0131]
- US 20020192673 A1 [0131]
- US 5654176 A [0150]
- DE 19507166 [0150]
- US 20030129659 A1 [0153]
- US 5399163 A [0176]
- US 5383851 A [0176]
- US 5312335 A [0176]
- US 5064413 A [0176]
- US 4941880 A [0176]
- US 4790824 A [0176]
- US 4596556 A [0176]
- US 4487603 A [0176]
- US 4486194 A [0176]
- US 4447233 A [0176]
- US 4447224 A [0176]
- US 4439196 A [0176]
- US 4475196 A [0176]

Literatura citada en la descripción que no es de patente:

- 30 • Smith. Science, 1985, vol. 228, 1315 - 1317 [0001] [0119]
- de Haard et al. J. Biol. Chem, 1999, vol. 274, 18218 - 30 [0001] [0119]
- Hoogenboom et al. Immunotechnology, 1998, vol. 4, 1 - 20 [0001] [0119]
- Hoogenboom et al. Immunol Today, 2000, vol. 2, 371 - 8 [0001] [0119]
- 35 • Markland W et al. Biochemistry, 18 June 1996, vol. 35 (24) [0004]
- Clackson T ; Wells J. Trends in Biotechnology, 1994, vol. 12, 173 [0005]
- Davidson et al. Proc Natl Acad Sci USA., 1994, vol. 91 (6), 2146 - 50 [0051]
- Needleman ; Wunsch. J. Mol. Biol., 1970, vol. 48, 444 - 453 [0069]
- Wun et al. J. Biol. Chem., 1988, vol. 263 (13), 6001 - 6004 [0094] [0096]
- 40 • Girard et al. Nature, 1989, vol. 338, 518 - 20 [0094] [0096]
- Novotny et al. J. Biol. Chem., 1989, vol. 264 (31), 18832 - 18837 [0094]
- Kido et al. J. Biol. Chem., 1988, vol. 263 (34), 18104 - 18107 [0094]
- Sprecher et al. Proc. Nat. Acad. USA, 1994, vol. 91, 3353 - 3357 [0094]
- Eigenbrot et al. Protein Engineering, 1990, vol. 3, 591 - 598 [0100]
- 45 • Hynes et al. Biochemistry, 1990, vol. 29, 10018 - 10022 [0100]
- Sonhammer et al. Proteins, 1997, vol. 28 (3), 405 - 420 [0103]
- Gribskov et al. Meth. Enzymol., 1990, vol. 183, 146 - 159 [0103]
- Gribskov et al. Proc. Natl. Acad. Sci. USA, 1987, vol. 84, 4355 - 4358 [0103]
- Krogh et al. J. Mol. Biol., 1994, vol. 235, 1501 - 1531 [0103]
- 50 • Stultz et al. Protein Sci., 1993, vol. 2, 305 - 314 [0103]
- Schultz et al. Proc. Natl. Acad. Sci. USA, 1998, vol. 95, 5857 [0103]
- Schultz et al. Nucl. Acids Res, 2000, vol. 28, 231 [0103]
- Corpet et al. Nucl. Acids Res., 1999, vol. 27, 263 - 267 [0103]
- Falquet et al. Nucleic Acids Res., 2002, vol. 30, 235 - 238 [0103]

- Wells et al. *Gene*, 1985, vol. 34, 315 - 323 [0111]
 - Leung et al. *Technique*, 1989, vol. 1, 11 - 15 [0116]
 - Stemmer. *Nature*, 1994, 389 - 391 [0116]
 - Coco et al. *Nature Biotech.*, 2001, vol. 19, 354 [0116]
 - 5 • Zooler et al. *Nucl Acids Res*, 1987, vol. 10, 6487 - 6504 [0116]
 - Reidhaar-Olson. *Methods Enzymol.*, 1991, vol. 208, 564 - 586 [0116]
 - Griffiths et al. *EMBO J*, 1994, vol. 13, 3245 [0116]
 - Cramer et al. *Gene*, 1993, vol. 137, 69 [0118]
 - Fuchs et al. *Bio/Technology*, 1991, vol. 9, 1370 - 1372 [0119]
 - 10 • Hay et al. *Hum Antibod Hybridomas*, 1992, vol. 3, 81 - 85 [0119]
 - Huse et al. *Science*, 1989, vol. 246, 1275 - 1281 [0119]
 - Griffiths et al. *EMBO J*, 1993, vol. 12, 725 - 734 [0119]
 - Hawkins et al. *J Mol Biol*, 1992, vol. 226, 889 - 896 [0119]
 - Clackson et al. *Nature*, 1991, vol. 352, 624 - 628 [0119]
 - 15 • Gram et al. *PNAS*, 1992, vol. 89, 3576 - 3580 [0119]
 - Garrard et al. *Bio/Technology*, 1991, vol. 9, 1373 - 1377 [0119]
 - Rebar et al. *Methods Enzymol.*, vol. 267, 129 - 49 [0119]
 - Hoogenboom et al. *Nuc Acid Res*, 1991, vol. 19, 4133 - 4137 [0119]
 - Barbas et al. *PNAS*, 1991, vol. 88, 7978 - 7982 [0119]
 - 20 • Santini. *J. Mol. Biol.*, 1998, vol. 282, 125 - 135 [0120]
 - Rosenberg et al. *Innovations*, 1996, vol. 6, 1 - 6 [0120]
 - Houshm et al. *Anal Biochem*, 1999, vol. 268, 363 - 370 [0120]
 - Corey et al. *Gene*, 1993, vol. 128 (1), 129 - 34 [0121]
 - Cwirla et al. *Proc Natl Acad Sci USA*, 1990, vol. 87 (16), 6378 - 82 [0121]
 - 25 • Fowlkes et al. *Biotechniques*, 1992, vol. 13 (3), 422 - 8 [0121]
 - Hoogenboom et al. *Nucleic Acids Res*, 1991, vol. 19 (15), 4133 - 7 [0121]
 - McCafferty et al. *Nature*, 1990, vol. 348 (6301), 552 - 4 [0121]
 - McConnell et al. *Gene*, 1994, vol. 151 (1-2), 115 - 8 [0121]
 - Scott ; Smith. *Science*, 1990, vol. 249 (4967), 386 - 90 [0121]
 - 30 • Elvin et al. *Gene*, 1990, vol. 37, 123 - 126 [0134]
 - Tabor ; Richardson. *Proc. Natl. Acad. Sci. U. S. A.*, 1998, 1074 - 1078 [0134]
 - Chang et al. *Gene*, 1986, vol. 44, 121 - 125 [0134]
 - Lutz; Bujard. *Nucl. Acids. Res.*, March 1997, vol. 25, 1203 - 1210 [0134]
 - D. V. Goeddel et al. *Proc. Nat. Acad. Sci. U.S.A.*, 1979, vol. 76, 106 - 110 [0134]
 - 35 • J. D. Windass et al. *Nucl. Acids. Res.*, 1982, vol. 10, 6639 - 57 [0134]
 - R. Cowl et al. *Gene*, 1985, vol. 38, 31 - 38 [0134]
 - Brosius. *Gene*, 1984, vol. 27, 161 - 172 [0134]
 - Amanna ; Brosius. *Gene*, 1985, vol. 40, 183 - 190 [0134]
 - Guzman et al. *J. Bacteriol.*, 1992, vol. 174, 7716 - 7728 [0134]
 - 40 • Haldimann et al. *J. Bacteriol.*, 1998, vol. 180, 1277 - 1286 [0134]
 - Kolodziej ; Young. *Methods Enz.*, 1991, vol. 194, 508 - 519 [0150]
 - Berge, S.M. et al. *J. Pharm. Sci.*, 1977, vol. 66, 1 - 19 [0172]
 - Sustained y Controlled Release Drug Delivery Systems. Marcel Dekker, Inc, 1978 [0174]
- 45