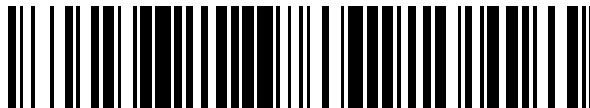


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 386 338**

51 Int. Cl.:
C07K 14/245 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **10159280 .6**
96 Fecha de presentación: **26.03.2004**
97 Número de publicación de la solicitud: **2228383**
97 Fecha de publicación de la solicitud: **15.09.2010**

54 Título: **Colección de mutantes de toxinas y métodos para su uso**

45 Fecha de publicación de la mención BOPI:
17.08.2012

45 Fecha de la publicación del folleto de la patente:
17.08.2012

73 Titular/es:
**MOLECULAR TEMPLATES, INC.
111 W. COOPERATIVE WAY, SUITE 201
GEORGETOWN, TX 78626, US**

72 Inventor/es:
**Gariepy, Jean y
Wei, Xin**

74 Agente/Representante:
de Elizaburu Márquez, Alberto

ES 2 386 338 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

Colección de mutantes de toxinas y métodos para su uso.

Antecedentes de la invención

5 Esta solicitud se refiere a colecciones de mutantes de toxinas y a métodos para usar las mismas en el desarrollo de agentes terapéuticos dirigidos contra tipos celulares específicos.

10 Las toxinas vegetales y bacterianas tienen una organización estructural con dos o más dominios o subunidades polipeptídicas responsables de distintas funciones, denominadas A y B. Las toxinas pueden denominarse toxinas AB_x, en donde x representa el número de subunidades B idénticas u homólogas en la toxina. Esta familia de toxinas relacionadas con el armazón, incluye ejemplos tales como las toxinas Shiga y de tipo Shiga, las enterotoxinas termo-
15 sensibles de *E. coli*, la toxina del cólera, la toxina diftérica, la toxina pertussis, la exotoxina A de *Pseudomonas aeruginosa* (Olsnes, S. y Sandvik, K. (1988) en Immunotoxins pp. 39-73, Kluwer Academic, Boston; Sandvik, K., Dubini-
na, E., Garred, O. y col. (1992) Biochem. Soc. Trans. 20:724) así como toxinas vegetales tales como ricina y abrina. En algunos casos las toxinas son heterómeras, porque las cadenas B son en realidad entidades separadas que se conectan con la cadena tóxica A mediante un enlace no covalente. En otros casos, la toxina es monómera, puesto que la cadena B forma parte de la misma proteína cuando la toxina se produce en la naturaleza.

20 Basándose en su capacidad para bloquear la síntesis de proteínas, proteínas tales como las toxinas Shiga y de tipo Shiga, así como la ricina, abrina, gelonina, crotina, proteína antivírica de fitolaca americana, saporina, momordina, modicina, sarcina, toxina diftérica y exotoxina A, se han denominado proteínas inactivadoras de ribosomas (RIP, del inglés "ribosome-inactivating proteins"). La potencia de las RIPs es extremadamente alta, habiéndose demostrado que una molécula de la cadena A de la toxina diftérica (Yamaizumi y col. (1978) Cell 15:245-250) o de la cadena A de la ricina (Eiklid y col. (1980) Exp. Cell Res. 126:321-326) es suficiente para destruir una célula eucariótica.

25 El documento de Publicación de Patente Internacional nº WO99/40185 describe colecciones de toxinas mutantes en las que se introducen las mutaciones en el dominio de unión para alterar el tipo de células al que se suministran las especies tóxicas. Las nuevas proteínas se obtienen mutando una subunidad de unión de la proteína citotóxica heterómera de tipo silvestre, para crear una colección de clones de microorganismos que producen proteínas mutantes que, a continuación, se someten a un escrutinio para estudiar su capacidad para unirse y destruir específicamente un tipo de célula diana.

30 El documento de Patente de los EE.UU. nº 5.552.144 describe una variante de la toxina II de tipo Shigella en la que se introduce una mutación en la cadena A en la posición 167, para cambiar el aminoácido en esta posición por uno con diferente carga. Esto dio como resultado una toxina con menor actividad enzimática asociada con la toxicidad.

35 El documento de Patente de los EE.UU. nº 6.593.132 describe proteínas tóxicas recombinantes que son tóxicas específicamente para células enfermas, pero que no dependen para su especificidad de la acción de un componente específico de la unión a la célula. Las proteínas recombinantes de la patente '132 tienen una cadena A de una toxina de tipo ricina enlazada a una cadena B mediante una secuencia enlazadora sintética que puede ser escindida específicamente por una proteasa localizada en las células o tejidos afectados por una enfermedad específica, para liberar la cadena tóxica A, inhibiendo o destruyendo así selectivamente las células o los tejidos enfermos.

40 El documento de Patente de los EE.UU. nº 6.649.742 describe proteínas inactivadoras de ribosomas (RIPs) de Tipo I y análogos de RIPs que tienen una cisteína disponible para formar un puente disulfuro con moléculas diana. Las RIPs y los análogos de RIPs se usan como componentes de agentes terapéuticos citotóxicos para eliminar selectivamente cualquier tipo celular al que se dirija el componente RIP, mediante la capacidad de unión específica del segundo componente del agente.

Sumario de la invención

45 La presente proporciona colecciones combinatorias de proteínas que comprenden una pluralidad de especies proteicas, en las que cada especie proteica comprende una cadena A de una proteína tóxica heterómera en la que se ha introducido un inserto. Según la invención, el inserto es un polipéptido con una secuencia de aminoácidos variable que tiene una longitud de 2 o más residuos de aminoácido, por ejemplo, de 3 a 200 residuos de aminoácido; y el inserto se introduce en el bucle sensible a proteasa de la secuencia de la cadena A. El resultado de la introducción del inserto crea un dominio de unión artificial dentro de la cadena A, de tal forma que la cadena A desarrolla una especificidad tóxica que es independiente y diferente de la especificidad normal asociada con los dominios de unión de las cadenas B. El escrutinio de la colección permite la selección e identificación de toxinas mutantes que son específicas de tipos celulares diferentes, incluyendo los tipos de células cancerosas. En una realización de la invención, la colección combinatoria comprende especies proteicas que se forman introduciendo el inserto en la cadena A de una toxina I de tipo Shiga, por ejemplo, en la región entre los aminoácidos 242 y 261, tal y como se define con referencia a SEQ ID NO: 1.

55 La invención proporciona también una colección combinatoria de expresión que comprende una pluralidad de espe-

cies de sistemas de expresión, expresando cada especie una especie proteica que comprende una cadena A de una proteína tóxica heterómera en la que se ha introducido el inserto tal y como se ha descrito anteriormente. La expresión de proteínas a partir de la colección combinatoria de expresión da como resultado la formación de una colección combinatoria de proteínas.

5 En un aspecto adicional, la invención proporciona una composición para tratar melanomas y métodos para usar tal composición. La composición comprende una especie proteica que comprende una cadena A de una proteína tóxica heterómera, en la que un polipéptido que tiene una longitud de 2 o más residuos de aminoácido, por ejemplo de 3 a 200 residuos de aminoácido, se introduce en el bucle sensible a proteasa de la secuencia de la cadena A. El inserto se selecciona de tal forma que la especie proteica tenga actividad tóxica frente a células de melanoma. La especie
10 proteica se usa en el tratamiento de melanomas, administrándola a un paciente diagnosticado con melanoma en una cantidad suficiente para producir la reducción del número de células de melanoma vivas.

En un aspecto adicional, la invención proporciona una composición para tratar otros tipos de cáncer. La composición comprende una especie proteica que comprende una cadena A de una proteína tóxica heterómera en la que un polipéptido que tiene una longitud de 2 o más residuos de aminoácido, por ejemplo de 3 a 200 residuos de aminoácido,
15 se introduce en el bucle sensible a proteasa de la secuencia de la cadena A. El inserto se selecciona de tal forma que la especie proteica tenga actividad tóxica frente a las células cancerosas. En una realización específica, el inserto se selecciona para que se una a receptores MUC-1. La especie proteica se usa en el tratamiento del melanoma, administrándola a un paciente diagnosticado con melanoma en una cantidad suficiente para que se produzca una reducción del número de células de melanoma vivas.

20 En un aspecto adicional, la invención proporciona un método para identificar ligandos que se unen a dianas/receptores específicos, tales como marcadores tumorales que se sabe que existen en las células cancerosas. En este método, la toxina de la colección combinatoria actúa como informadora, y una colección combinatoria de proteínas según la invención, se somete a escrutinio en busca de células conocidas por poseer la diana/el receptor. Las proteínas que presentan toxicidad frente a las células, se evalúan para determinar la secuencia de la región insertada. Los péptidos con esta secuencia pueden usarse seguidamente, en combinación con una toxina u otras moléculas,
25 para dirigir compuestos a células que poseen la diana/el receptor.

En aún otro aspecto adicional de la invención, se proporciona un método para identificar sustancias tóxicas específicas de un marcador celular conocido. En esta realización de la invención, la toxina no necesita actuar como informadora. Por tanto, las células que tienen el marcador o una diana/un receptor aislado, cuando está disponible, se exponen a la colección combinatoria de proteínas. En realizaciones preferidas, las células o la diana/el receptor aislado, se inmovilizan sobre un soporte sólido, tal como pocillos de plástico. Las proteínas capturadas de la colección se someten seguidamente a un nuevo escrutinio para confirmar su toxicidad y su especificidad hacia las células que expresan la diana/el receptor y su idoneidad para usarse como agente terapéutico.
30

Breve descripción de los dibujos

35 La Fig. 1 es un diagrama esquemático que representa los dominios A1 y A2 de SLT-1. La cadena A está compuesta por 293 aminoácidos. La cadena se escinde con furina para producir un fragmento catalítico A1 y una cola C-terminal A2 asociada de manera no covalente al pentámero B. Un bucle sensible a proteasa (área rallada) queda definido por los dos únicos residuos de cisteína en la cadena A (Cys 242 y 261). Tyr77, Glu167, Arg170 y Trp203 representan residuos decisivos para la actividad catalítica del dominio A1 (flechas).

40 La Fig. 2A es una representación esquemática de la cadena A de SLT-1 (1-293) con el epítipo de MUC1 asociado a cáncer de mama, PDTRPAP (secuencia testigo reconocida por el AcMo Onc M27) insertado entre los residuos 245 y 246 y un marcador de 6 histidinas en su extremo N-terminal.

La Fig. 2B es una representación de nuestra estructura artificial de la colección de la cadena A de SLT-1-tripéptidos, en donde las tres posiciones clave del epítipo de MUC1 reconocidas por el AcMo Onc M27 se combinaron aleatoriamente (región XXX). La colección de tripéptidos se insertó en una región en forma de bucle de la cadena A presente en la naturaleza, creada por la presencia de un puente disulfuro entre Cys 242 y Cys 261.
45

La Fig. 3 muestra un conjunto de datos de ELISA representativos, obtenidos a partir de un escrutinio de 96 variantes distintas con una única cadena A procedentes de nuestra colección de cadena A de SLT-1-tripéptidos con el AcMo Onc M27. La variante de toxina nº 41 (Tablas 1 y 2) presentaba una fuerte señal en el ensayo ELISA y tenía el epítipo esperado.
50

La Fig. 4 muestra un diagrama esquemático de un segmento al azar de 7 aminoácidos insertado entre los residuos 245 y 246 de la cadena A.

La Fig. 5 muestra los resultados de los ensayos con siete variantes de toxina que se identificaron como destructores repetitivos de la línea celular de melanoma humano 518A2. El eje de abscisas representa el logaritmo de la concentración de toxina usada para tratar las células y el eje de ordenadas representa el porcentaje observado de células que son viables después de 48 horas. Los triángulos negros representan el efecto de la toxina de tipo silvestre sobre las células 518A2, mientras que las variantes de la cadena A más eficaces se denominaron SAMn°3 (cuadrados sin
55

relleno) y SAMn⁵ (símbolos X).

Descripción detallada de la invención

La presente invención se refiere a una colección combinatoria de proteínas que comprende una pluralidad de especies proteicas, comprendiendo cada especie proteica una cadena A de una proteína tóxica heterómera en la que se ha introducido un inserto. El inserto es un polipéptido con una secuencia de aminoácidos variable que tiene una longitud de 2 o más residuos de aminoácido, por ejemplo de 3 a 200 residuos de aminoácido; y se introduce en el bucle sensible a proteasa de la secuencia de la cadena A. La colección proporciona un conjunto de especies proteicas que pueden someterse a escrutinio para seleccionar proteínas individuales que sean tóxicas frente a tipos celulares específicos, tales como tipos específicos de células cancerosas. Las especies proteicas individuales así seleccionadas se usan adecuadamente en el tratamiento del cáncer.

Tal y como se utiliza en la memoria descriptiva y en las reivindicaciones de esta solicitud, la expresión "colección combinatoria" se refiere a una mezcla de especies, cada una de las cuales tiene una porción común y una porción variable. En el caso de una "colección combinatoria de proteínas", cada una de las especies es una proteína o un péptido, y las porciones comunes y las porciones variables son cada una de ellas secuencias de aminoácidos. En el caso de una colección combinatoria de expresión, las especies son microorganismos, vectores de expresión o polinucleótidos que, cuando se expresan, producen proteínas o péptidos que tienen porciones comunes y porciones variables. En este caso, las porciones comunes y las porciones variables son cada una de ellas secuencias de nucleótidos. Puesto que el objetivo de la colección combinatoria es proporcionar múltiples variantes con el fin de ser sometidas a escrutinio, la colección combinatoria contiene preferiblemente al menos 100 especies distintas de unidades proteicas o de expresión, más preferiblemente al menos 1000 especies distintas.

Tal y como se usa en la memoria descriptiva y en las reivindicaciones de esta solicitud, la expresión "proteína tóxica heterómera" se refiere a la clase de toxinas proteicas con la característica de organización común de ser heterómeras por naturaleza, con dos o más dominios o subunidades polipeptídicas responsables de distintas funciones (Meritt, E.A. y Hol, W.G.J. (1995) *Curr. Opin. Struct. Biol.* 5:165-171). En tales proteínas, las dos o más subunidades o dominios podrían denominarse A y B, y las toxinas AB_x en donde x representa el número de subunidades B idénticas u homólogas en la toxina. Esta familia de toxinas relacionadas con el armazón incluye ejemplos tales como las toxinas Shiga y de tipo Shiga, las enterotoxinas termosensibles de *E. coli*, la toxina del cólera, la toxina diftérica, la toxina pertussis y la exotoxina A de *Pseudomonas aeruginosa*, así como toxinas vegetales tales como ricina y abrina. Basándose en su capacidad para bloquear la síntesis de proteínas, proteínas tales como las toxinas Shiga y de tipo Shiga, así como la ricina, abrina, gelonina, crotina, proteína antivirica de fitolaca americana, saporina, momordina, modicina, sarcina, toxina diftérica y exotoxina A se han denominado proteínas inactivadoras de ribosomas (RIP). En estas proteínas tóxicas heterómeras existentes en la naturaleza, la cadena A es la porción tóxica, mientras que las cadenas B forman un resto de fijación que se une a un receptor en una célula susceptible a la toxina, suministrándose de esta forma la cadena A a la célula.

Un ejemplo específico de la cadena A de una proteína tóxica heterómera es la cadena A de SLT-1 que tiene la secuencia proporcionada como SEQ ID NO: 1. La cadena A de SLT-1 comprende 293 aminoácidos, abarcando el dominio enzimático (tóxico) desde los residuos 1 al 239. Un bucle sensible a proteasa que abarca los residuos 242 a 261 queda normalmente expuesto, y es un sitio adecuado para insertar una secuencia peptídica.

SLT-1 es una proteína inactivadora de ribosomas de tipo II, producida por una cepa patógena de *Escherichia coli* (0157:H7) (24). SLT-1 es un complejo AB₅ de aproximadamente 70 kD (O'Brien, A. D. y Holmes, R. K. (1987) "Shiga and Shiga-like toxins". *Microbiol Rev* 51, 206-220). La subunidad A catalítica de 32 kD de cadena sencilla está asociada de manera no covalente con un pentámero de cinco subunidades B idénticas de 7,7 kD. El pentámero de subunidades B reconoce el glicolípido globotriaosilceramida (también conocido como CD77 o Gb3) en la superficie de las células diana (Lingwood, C. A. (1993) "Verotoxins and their glycolipid receptors". *Adv Lipid Res* 25, 189-211; Jacewicz, y col. (1986) "Pathogenesis of shigella diarrhea: XI. Isolation of a shigella toxin-binding glycolipid from rabbit jejunum and HeLa cells and its identification as globotriaosylceramide". *J Exp Med* 163, 1391-1404). Un bucle sensible a proteasa situado entre Cys242 y Cys261 en el extremo C-terminal de la cadena A se escinde con furina durante el encaminamiento celular (Fig. 1). La cadena A permanece asociada con su pentámero de subunidades B gracias a un enlace disulfuro intracatenario entre Cys242 y Cys261 a medida que viaja hacia el lumen del retículo endoplasmático (Sandvig y col. (1989) "Endocytosis from coated pits of Shiga toxin: a glycolipid-binding protein from *Shigella dysenteriae* 1". *J Cell Biol* 108, 1331-1343; Garred, y col. (1995) "Role of processing and intracellular transport for optimal toxicity of Shiga toxin and toxin mutants". *Exp Cell Res*: 218, 39-49). El puente disulfuro es finalmente reducido en el lumen del retículo endoplasmático y la cadena A1 (primeros 251 aa) se libera y se retrotransloca subsiguientemente al citosol, en donde inactiva los ribosomas (O'Brien y col. (1992) "Shiga toxin: biochemistry, genetics, mode of action, and role in pathogenesis". *Curr Top Microbiol Immunol* 180, 65-94). Más específicamente, la cadena A de SLT-1 es una N-glicosidasa que escinde catalíticamente un nucleótido de adenina específico (4324) procedente del ARNr 28 S (Brigotti y col. (1997) "The RNA-N-glycosidase activity of Shiga-like toxin I: kinetic parameters of the native and activated toxin". *Toxicon* 35, 1431-1437). Este suceso lleva a la inhibición de la síntesis proteica, impidiendo la unión de los aminoacil-ARNt al ribosoma y deteniendo la elongación de la proteína. Estudios de mutagénesis, así como análisis estructurales realizados sobre las cadenas A de ST y ricina, han definido los residuos clave conservados, implicados en la actividad catalítica (Deresiewicz y col. (1992) "Mutations affecting the activity of the

Shiga-like toxin I A-chain". *Biochemistry* 31, 3272-3280; Ready y col. (1991) "Site-directed mutagenesis of ricin A chain and implications for the mechanism of action". *Proteins* 10, 270-278). Los residuos decisivos para la actividad catalítica de SLT-1 son la tirosina 77, el ácido glutámico 167, la arginina 170 y el triptófano 203 (Hovde y col. (1988) "Evidence that glutamic acid 167 is an active-site residue of Shiga-like toxin I". *Proc Natl Acad Sci USA*. 85, 2568-2572; Yamasaki y col. (1991) "Importance of arginine at position 170 of the A subunit of Vero toxin I produced by enterohemorrhagic *Escherichia coli* for toxin activity". *Microb Pathog* 11, 1-9). Además, la unión de la toxina a la superficie celular es decisiva para su introducción dentro de la célula y, por tanto, para su actividad tóxica. Debido a esto, la cadena A en solitario no es considerablemente tóxica.

Además de la cadena A de SLT-1, pueden usarse también otras toxinas para formar colecciones y composiciones según la invención, y pueden usarse en los métodos de la invención. Específicamente, pueden usarse las toxinas Shiga y otras de tipo Shiga, así como ricina, abrina, gelonina, crotina, proteína antivírica de fitolaca americana, saporina, momordina, modicina, sarcina, toxina diftérica y exotoxina A, y otras proteínas inactivadoras de ribosomas (RIP) relacionadas funcionalmente.

Con el fin de preparar la colección combinatoria de la invención, se insertan en este bucle sensible a proteasa secuencias cortas de aminoácidos de al menos 2 aminoácidos, por ejemplo de 3 a 200 residuos de aminoácidos de longitud. El número de aminoácidos en el inserto define el número de posibles variantes al azar que puede haber en la colección. Por ejemplo, cuando el número de aminoácidos en el inserto es 3, el número máximo de variantes es 20^3 u 8000 variantes. Insertos más largos proporcionan un mayor número correspondiente de posibles variantes.

Como alternativa al uso de insertos con secuencias puramente al azar, pueden diseñarse insertos basados en un molde conocido. Por ejemplo, tal y como se describe más abajo en el contexto de Muc-1, para identificar un inserto que asegura la optimización de las propiedades tóxicas de la estructura artificial proteica, pueden usarse variaciones de una secuencia que se sabe que proporciona propiedades de unión a un receptor para un tipo celular en particular. Esta misma optimización puede realizarse en una secuencia individual aislada mediante el escrutinio de una colección combinatoria mayor. Se apreciará, sin embargo, que el inserto en los ensayos de la prueba de concepto del Ejemplo 1, se usa un inserto que es la diana/el receptor, mientras que en el caso actual el inserto se basaría en la secuencia de un ligando conocido, que se va a mejorar para conseguir una eficacia y especificidad máximas.

Este acercamiento, en el que un tipo de receptor específico y conocido es la diana, ilustra un aspecto adicional de la invención, a saber, un método para identificar ligandos peptídicos que se unen a dianas/receptores específicos, tales como marcadores tumorales que se sabe que existen sobre las células cancerosas. En este método, una colección combinatoria de proteínas según la invención, se somete a escrutinio frente a células que se sabe que poseen la diana/el receptor. La toxina actúa como informadora, de tal forma que las proteínas que muestran ser tóxicas para las células, se valoran para determinar la secuencia de la región insertada. Los péptidos de esta secuencia pueden usarse seguidamente, en combinación con una toxina u otras moléculas, para dirigir compuestos a las células que poseen la diana/el receptor. Otros péptidos en reposo de la secuencia de la región insertada pueden ser apropiados para confirmar que son un ligando para la diana/el receptor específicos, en contraposición con otros receptores en el tipo celular. Esto puede realizarse usando ensayos de unión con receptor aislado, cuando esté disponible.

La invención proporciona también un método para identificar sustancias tóxicas específicas para un marcador celular conocido, y particularmente marcadores que están disponibles de forma aislada. En esta realización de la invención, la toxina no necesita actuar como informadora. Así, las células que tienen el marcador, o una diana/un receptor aislados, cuando están disponibles, se exponen a la colección combinatoria de proteínas. En realizaciones preferidas, las células o la diana/el receptor aislados se inmovilizan sobre un soporte sólido, tal como en pocillos de plástico. Las proteínas capturadas desde la colección se someten a continuación nuevamente a un escrutinio frente a las células, para confirmar su toxicidad y especificidad hacia las células que expresan la diana/el receptor y su idoneidad para usarse como agente terapéutico. Este método puede usarse para identificar toxinas con insertos que se unen, específicos para cualquier marcador tumoral o receptor celular, incluyendo sin limitaciones marcadores tumorales tales como las mucinas, como MUC-1 y sus glicofomas, Her-2, Her2-Neu, marcadores de cinasa de tirosina, EGFR, GD2 y GD3.

Por tanto, según este aspecto específico de la invención, se proporciona un método para aislar una toxina específica para una diana/un receptor conocidos, que comprende las etapas de:

(a) exponer la diana/el receptor a una colección combinatoria de proteínas que comprende una pluralidad de especies proteicas, comprendiendo cada especie proteica una cadena A de una proteína tóxica en la que se ha introducido un inserto, en donde,

el inserto es un polipéptido con una secuencia variable de aminoácidos que tiene una longitud de al menos 2 residuos de aminoácido; y

el inserto se introduce en el bucle sensible a proteasa de la secuencia de la cadena A; y

(b) aislar al menos una especie proteica desde la colección combinatoria de proteínas capturada mediante la unión a la diana/el receptor. Tal como se emplea en la memoria descriptiva y las reivindicaciones de la presente, el término "aislar" se refiere a cualquier mecanismo para obtener una composición que contiene la proteína separada del en-

torno en el que se expresa, en forma adecuada para análisis adicionales. Esto incluiría la liberación de la diana/el receptor después de la captura (por ejemplo, por exposición a un agente de unión competitivo) o el aislamiento a partir de un cultivo de un clon que expresa la proteína que se va a capturar. El método puede comprender además la etapa de someter a escrutinio la proteína aislada frente a células que expresan la diana/el receptor, para confirmar su toxicidad para las células que expresan la diana/el receptor. Los procedimientos adecuados para este escrutinio se describen en el Ejemplo 3. Como se ha señalado anteriormente, en este método, la diana/el receptor puede ser una diana/un receptor purificado y puede inmovilizarse sobre un soporte sólido. La diana/el receptor puede estar también en la superficie de células, que se pueden inmovilizar. Cuando la diana/el receptor está en la superficie de células, la toxina puede actuar como informadora, y la muerte de las células es indicativa de la unión al receptor.

Nuestra experiencia en la construcción de colecciones de SLT-1 nos ha señalado varias cuestiones prácticas que se consideran apropiadas en la selección de un molde proteico para crear colecciones combinatorias. Un factor importante es elegir una proteína de cadena sencilla, preferiblemente una proteína bacteriana de menos de 300 aminoácidos (si las colecciones van a expresarse en procariontes). El uso de una toxina más pequeña aumenta su potencial para penetrar en tumores sólidos y disminuye su inmunogenicidad. En segundo lugar, el molde proteico debería plegarse espontáneamente en solución en su forma activa, de manera que hay una necesidad mínima de chaperonas del hospedante. Deberían evitarse, por ejemplo, estructuras que contienen múltiples residuos de cisteína, implicados normalmente en los puentes disulfuro. Además, una proteína de cadena sencilla, en contraposición con un complejo de múltiples subunidades, se puede exportar más fácilmente desde las bacterias. En tercer lugar, el molde proteico debería poseer una actividad enzimática, que se pueda medir fácilmente para confirmar el pegamiento apropiado de las variantes peptídicas que contienen mutaciones sencillas o múltiples dirigidas al sitio. En cuarto lugar, debería incorporarse un acercamiento simple de escrutinio al diseño de colecciones combinatorias. Tales búsquedas deberían ser susceptibles de permitir el uso de métodos de escrutinio de alto rendimiento. La cadena A catalítica de SLT-1 (residuos 1 a 293) cumple estos criterios porque es una cadena sencilla que no tiene ninguna función conocida de unión a receptores, y que tiene una estructura y un sitio catalítico bien definidos.

Un aspecto adicional de la invención es una colección combinatoria de expresión que comprende una pluralidad de especies de sistemas de expresión. Cada especie en la colección de expresión expresa una especie proteica que comprende una cadena A de una proteína tóxica heterómera en la que se ha introducido un inserto. El inserto es un polipéptido con una secuencia de aminoácidos variable que tiene una longitud de 2 o más residuos de aminoácido, por ejemplo de 3 a 200 residuos de aminoácido, y se introduce en el bucle sensible a proteasa de la secuencia de la cadena A. Los sistemas de expresión adecuados incluyen plásmidos y vectores víricos.

La invención se describirá ahora adicionalmente con referencia a los siguientes ejemplos no limitantes.

Ejemplo 1

Prueba de concepto: diseño y evaluación de una colección prototipo de cadena A-tripéptidos

Originalmente creamos una colección simple de tripéptidos insertados en el bucle sensible a proteasa del extremo C-terminal de la cadena A de STL-1 (Figs. 2A y B). Esta región en forma de bucle de la cadena A está restringida de forma natural debido a la presencia de un único puente disulfuro que forma un puente desde Cys242 a Cys261. Puede calcularse, por tanto, que la diversidad máxima de esta colección será de 20^3 ó 8000 permutaciones de una secuencia tripeptídica. Como prueba de concepto de que las colecciones de cadena A pueden someterse a escrutinio fácilmente para detectar una nueva actividad de unión a un receptor, seleccionamos más de 3000 colonias a partir de esta colección de cadena A-tripéptidos y purificamos la toxina mutante producida por cada clon. En este estudio, observamos enseguida que el nivel de expresión del mutante de la cadena A aumentaba espectacularmente cuando se expresaba en presencia de la subunidad B de SLT-1 de tipo silvestre. Por tanto, las formas mutantes de la cadena A se expresaron y purificaron inicialmente como variantes de la toxina AB5. Puesto que todas las subunidades A llevan una secuencia marcadora de poliHis para su purificación, resulta relativamente fácil eliminar la subunidad B con agentes desnaturizantes (por ejemplo, urea) y recuperar la cadena A en columnas o perlas de afinidad con metales. Las transferencias Western realizadas con clones bacterianos seleccionados al azar, indicaron que >70% de estas colonias producían cantidades significativas de estos mutantes de la cadena A.

Estas variantes de toxina se revistieron seguidamente en pocillos individuales en placas de 96 pocillos y se sometieron a escrutinio mediante ELISA para estudiar su capacidad para unirse al anticuerpo monoclonal Onc M27 (Linsley y col. (1988) "Monoclonal antibodies reactive with mucin glycoproteins found in sera from breast cancer patients". *Cancer Res* 48, 2138-2148), dirigido contra el epítipo tripeptídico bien caracterizado de cáncer de mama Thr-Arg-Pro de la repetición en tándem de MUC1 humano (Gendler y col. (1988) "A highly immunogenic region of a human polymorphic epithelial mucin expressed by carcinomas is made up of tandem repeats". *J Biol Chem* 263, 12820-12823; Girling y col. (1989) "A core protein epitope of the polymorphic epithelial mucin detected by the monoclonal antibody SM-3 is selectively exposed in a range of primary carcinomas". *Int J Cancer* 43, 1072-1076). Tal y como se muestra en las Tablas 1 y 2, la mayoría de los mutantes de la cadena A no codificaban el inserto tripeptídico que era complementario al epítipo diana de Onc M27, ni eran reconocidos por el anticuerpo. Sin embargo, en dos ocasiones, una variante de toxina que llevaba el epítipo exacto (Tabla 1, Fig. 3) presentó una fuerte señal en el ensayo de ELISA, comparable a la observada con nuestra cadena A testigo que llevaba el epítipo tripeptídico de MUC1, y tenía la secuencia del epítipo esperada en su región tripeptídica al azar. En la Fig. 3 se presenta un conjunto de datos

de ELISA típico para 96 mutantes de la cadena A, que subraya el hecho de que la mayoría de los mutantes de la cadena A no reconocían el AcMo Onc M27, excepto una variante de la cadena A (mutante nº 41 en las Tablas 1 y 2). Estos resultados mostraron claramente que las colecciones de cadenas A pueden construirse fácilmente y someterse a escrutinio para encontrar variantes de la toxina capaces de dirigirse específicamente a un receptor dado, en este caso un sitio de combinación con antígeno.

Ejemplo 2

Preparación de una Colección Combinatoria de A de SLT-1-heptapéptidos

La diversidad de las colecciones representa un parámetro crucial en el escrutinio de colecciones combinatorias para buscar ligandos capaces de unirse específicamente y con alta afinidad a una diana en particular. La colección de cadenas A de SLT-1-tripéptidos descrita en el Ejemplo 1 tiene una diversidad máxima de 20^3 o 8000 cadenas A mutadas posibles. Este repertorio de muestras es pequeño pero resultó útil en la elaboración de mapas de un epítipo tripéptido para establecer nuestra prueba de concepto. Una colección de siete residuos (20^7 o $1,3 \times 10^9$ posibles mutantes) representa un nivel de diversidad mínimo más típico, utilizado normalmente en el diseño de colecciones de expresión de fagos, así como de péptidos sintéticos. Por tanto, como punto de partida, construimos una colección de cadenas A de SLT-1 con una secuencia al azar de 7 aminoácidos de longitud, insertada en su extremo C-terminal. Esta colección se instaló en el bucle sensible a proteasa de la cadena A, una región en forma de bucle restringida naturalmente por un puente disulfuro. Esta colección proporcionó suficiente diversidad para asegurar que pueden identificarse variantes de toxina de cadena A que se dirigen hacia receptores interiorizados nuevos o conocidos en células cancerosas. Todos los elementos de esta colección (así como todas las otras colecciones propuestas) contienen una secuencia marcadora His N-terminal para purificar rápidamente los mutantes de la cadena A. La colección (Fig. 4) se generó usando una estrategia de PCR con megacebadores (Sarkar G. y Sommers S., (1990)) "The megaprimer method of site-directed mutagenesis". *Biotechniques* 8, 404-407). La estrategia del megacebador se emplea ampliamente para introducir mutaciones en una secuencia de ADN diana, a través de dos rondas de PCR que usan dos cebadores flanqueantes y un cebador interno mutagénico. La descripción del diseño de la colección para la colección de cadenas A de SLT-1-heptapéptidos, servirá como ejemplo para otras colecciones futuras de cadenas A sencillas. La colección de cadenas A de SLT-1-heptapéptidos (Fig. 5) lleva una inserción al azar de 7 aminoácidos entre los aminoácidos 245 y 246 de la cadena A, sitio idéntico al usado para construir nuestra colección de cadenas A de SLT-1-tripéptidos (Tablas 1, 2, Fig. 3). Brevemente, dos cebadores flanqueantes A (GTT ACT GTG ACA GCT GAA GCT TTA CGT TTT CG (SEQ ID NO: 2) y B (GAG AAG AAG AGA CTG CAG ATT CCA TCT GTT G (SEQ ID NO: 3) que portaban los sitios de restricción HindIII y PstI, respectivamente, se reasociaron en los extremos 5' y 3' del operón de SLT-1. Se sintetizó una colección de oligonucleótidos F que contenía los siete aminoácidos al azar (NNS) así como una larga secuencia complementaria para reasociarse con el molde. En la síntesis del oligonucleótido al azar, la representación relativa de cada aminoácido se mejoró restringiendo la tercera posición de cada codón a G o T (Noren, K. A. y Noren, C. J. (2001) "Construction of high-complexity combinatorial phage display peptide libraries". *Methods* 23, 169-178). Este tipo de restricción reduce la complejidad global de la secuencia de ADN así como las discrepancias en la codificación entre los residuos (Reidhaar-Olson y col. (1991) "Random mutagenesis of protein sequences using oligonucleotide cassettes". *Methods Enzymol* 208, 564-586). Esta estrategia minimiza también la aparición de codones de detención (TAA y TGA), mientras que el codón de detención (TAG) se elimina usando una cepa bacteriana supE, que especifica la inserción de un residuo Gln cuando se traducen codones TAG. La primera reacción de PCR se realizó usando los cebadores A y F y el producto resultante se purificó mediante electroforesis en gel de poliacrilamida, en condiciones desnaturizantes. Este producto sirvió a continuación como megacebador con el cebador B para una segunda reacción de PCR, para amplificar el ADN al azar. El ADN final de la colección (producto de la PCR) se digirió seguidamente con HindIII y PstI y se clonó en la cadena principal de un vector de expresión pECHE9a (MTI, Toronto). El vector pECHE resultante se utilizó subsiguientemente para transformar la cepa de *E. coli* JM101 y se seleccionaron las colonias bacterianas aisladas, se lisaron y el material sobrenadante se analizó para detectar la expresión de toxinas de cadena A sencilla o se dispusieron en una capa sobre células cancerosas y se sometieron a escrutinio usando SRB para someter a ensayo la citotoxicidad celular.

Ejemplo 3

Evaluación de la colección combinatoria formada por A de SLT-1-heptapéptidos frente a líneas de células cancerosas usando un ensayo de citotoxicidad

Sometimos a escrutinio nuestra colección formada por A de SLT-1-heptapéptidos usando la función citotóxica de la cadena A como señal informadora. La citotoxicidad es una propiedad a medir que proporciona más información que la unión a un receptor, ya que implica que la toxina es interiorizada, procesada y suministrada cerca de los ribosomas, un suceso que ocurre claramente en múltiples etapas. El ensayo de citotoxicidad se realizó esencialmente tal y como se ha descrito previamente (Bray y col. (2001) "Probing the surface of eukaryotic cells using combinatorial toxin libraries". *Current Biology* 11, 697-701). Brevemente, la estrategia para someter a escrutinio todas nuestras colecciones de cadena A se basó en los siguientes principios. Se cultivaron líneas de células cancerosas establecidas, tales como SK-BR-3 (cáncer de mama humano), CAMA-1 (cáncer de mama humano), 518A2 (melanoma humano), PC3 (cáncer de próstata humano) y B16 (melanoma de murido) en placas de 96 pocillos y se usaron como dianas en las primeras etapas del escrutinio. Estas líneas celulares se seleccionaron inicialmente para nuestras búsquedas de colecciones de holotoxinas (Bray, véase más arriba) basándose en su adherencia (plástico), sus propiedades de

tinción según la viabilidad de las células (SRB) en un ejemplo de escrutinio de alto rendimiento, así como en su falta de receptor y sensibilidad hacia SLT-1 natural (para asegurar un nivel reducido de falsos positivos). Se seleccionaron colonias bacterianas aisladas a partir de cada colección y se cultivaron en placas de 96 pocillos profundos. Se recogieron las células, se lisaron y sus lisados se aclararon. Puesto que todas las variantes expresadas de la cadena A de SLT-1 tienen una secuencia marcadora de 6 histidinas en su extremo N-terminal, se purificó cada una de ellas a partir de su lisado, usando perlas de afinidad con níquel (en formato de 96 pocillos) y se dispusieron en capa sobre las células diana. Las placas que contenían las células diana tratadas con las variantes de la cadena A se incubaron seguidamente a 37°C durante 48 horas, y a continuación se procedió a la fijación y tinción con Sulfurodamina B (SRB). El ensayo con SRB es un ensayo de indicadores colorimétricos que cuantifica las células viables tiñendo su contenido proteico celular (Skehan y col., (1990) "New colorimetric cytotoxicity assay for anticancer-drug screening". J Natl Cancer Inst 82, 1107-1112). El ensayo con SRB ha sido adoptado por NCI/NIH para su escrutinio de alto rendimiento de candidatos a fármacos sobre líneas de células cancerosas. Los ensayos de viabilidad se repitieron con cualquier extracto bacteriano que conducía a la muerte celular. Se realizó una primera ronda de escrutinio sobre más de 5000 clones bacterianos (equivalentes a 5000 toxinas de cadena A distintas) y se identificaron 7 variantes de toxina como destructores repetitivos de la línea celular de melanoma humano 518A2 (Fig. 5). El eje de abscisas representa el logaritmo de la concentración de toxina usada para tratar las células y el eje de ordenadas representa el porcentaje observado de células que son viables después de 48 horas. Los triángulos rellenos representan el efecto de la toxina de tipo silvestre sobre las células 518A2, mientras que las dos variantes más eficaces de la cadena A se denominaron SAMn°3 (cuadrados sin relleno) y SAMn°5 (símbolos X).

Siete variantes prometedoras de la cadena A se sometieron nuevamente a escrutinio frente a un panel de líneas celulares (Vero [mono, riñón normal]; PC-3 [humano, cáncer de próstata]; HepG2 [humano, hepatoma]; SiHa [humano, cáncer de cuello uterino]; PanC [humano, cáncer pancreático]; SKBR-3 [humano, cáncer de mama]; 518-A2 (humano, melanoma); U87 [humano, glioma]; B16-F10 [ratón, melanoma]; HS-216 [humano, fibroblasto normal]; CAMA-1 [humano, cáncer de mama]; OVCAR-3 [humano, cáncer de ovario]). Entre estas siete, se observó que cuatro tenían actividad frente a una línea de células cancerosas, dos hacia melanoma humano 518-A2, una hacia SiHa (células de cáncer de cuello uterino humano) y otra hacia U87-A (células de cáncer cerebral humano; glioma).

Los genes que codificaban las dos toxinas de la cadena A (SAM3 y SAM5) que dieron como resultado la toxicidad de las líneas celulares de melanoma humano, se secuenciaron para determinar las secuencias de aminoácidos insertadas entre los residuos 245 y 246 de la cadena A de tipo silvestre. Las secuencias, incluyendo el marcador His, se incluyen en la lista de SEQ ID NOs: 4 y 5, respectivamente.

TABLA 1: Secuencias de ADN de clones seleccionados al azar a partir de la colección de cadenas A de SLT-1-tripéptidos. Las bases mutadas están en negrita. El mutante n°41 se identificó en nuestro escrutinio mediante ELISA por ligarse fuertemente al AcMo Onc M27 (Fig. 3)

Secuencia de nucleótidos de la cadena A de SLT-1 Variante	Diversidad (cambios de nucleótidos en la región mutada)
epítipo MUC1	CCA GAC ACG CGA CCA GCT CCA
Mutante n° 1	CCA GAC GGG ATC GGG GCT CCA
Mutante n° 2	CCA GAC CTG GAG ATG GCT CCA
Mutante n° 3	CCA GAC CCC CGT GGG GCT CCA
Mutante n° 4	CCA GAC GAC GAC TTG GCT CCA
Mutante n° 5	CCA GAC GTC CGG TGG GCT CCA
Mutante n° 6	CCA GAC CAG CGC TGG GCT CCA
Mutante n° 7	CCA GAC CTC AGG ATG GCT CCA
Mutante n° 8	CCA GAC TCC CAG GAG GCT CCA
Mutante n° 9	CCA GAC TCC GAC CCC GCT CCA
Mutante n° 41	CCA GAC ACG CGC CCC GCT CCA

TABLA 2: Alineamiento de las secuencias de aminoácidos de clones seleccionados al azar a partir de la colección de cadenas A de SLT-1-tripéptidos y la señal en el ensayo ELISA de las variantes de la cadena A de SLT-1 purificadas detectadas con un AcMo (Onc M27) obtenido contra el epítipo de MUC1 Thr-Arg-Pro. La región tripeptídica mutada está en negrita. El mutante n°41 se identificó en nuestro escrutinio mediante ELISA por presentar una fuerte unión al AcMo Onc M27

Secuencia de aminoácidos deducida de la cadena A de SLT-1 Variante	Lecturas de ELISA (405 nm)
epítipo MUC1	CHHHPD TRP APASRVARMASDEFPSMC
Mutante n° 1	CHHHPD GIG APASRVARMASDEFPSMC
Mutante n° 2	CHHHPD LQM APASRVARMASDEFPSMC
Mutante n° 3	CHHHPD PRG APASRVARMASDEFPSMC

ES 2 386 338 T3

Mutante nº 4	CHHHPD DDL APASRVARMASDEFPSMC	0,06
Mutante nº 5	CHHHPD VRW APASRVARMASDEFPSMC	0,07
Mutante nº 6	CHHHPD QRL APASRVARMASDEFPSMC	0,06
Mutante nº 7	CHHHPD LRM APASRVARMASDEFPSMC	0,11
Mutante nº 8	CHHHPD SQE APASRVARMASDEFPSMC	0,13
Mutante nº 9	CHHHPD SDP APASRVARMASDEFPSMC	0,07
Mutante nº 41	CHHHPD TRP APASRVARMASDEFPSMC	1,25

LISTADO DE SECUENCIAS

<110> Molecular Templates, Inc.

<120> COLECCIÓN DE MUTANTES DE TOXINAS Y MÉTODOS PARA SU USO

<130> M2798 EP/1 S3

5 <150> EP04785852.7
<151> 26-03-2004

<160> 7

<170> PatentIn versión 3.2

<210> 1

10 <211> 299

<212> PRT

<213> Escherichia coli

<220>

<221> misc_feature

15 <223> cadena A de SLT-1 de tipo silvestre

<400> 1

Ile Glu Gly Arg Ala Ser Lys Glu Phe Thr Leu Asp Phe Ser Thr Ala
1 5 10 15

Lys Thr Tyr Val Asp Ser Leu Asn Val Ile Arg Ser Ala Ile Gly Thr
20 25 30

Pro Leu Gln Thr Ile Ser Ser Gly Gly Thr Ser Leu Leu Met Ile Asp
35 40 45

Ser Gly Ser Gly Asp Asn Leu Phe Ala Val Asp Val Arg Gly Ile Asp
50 55 60

Pro Glu Glu Gly Arg Phe Asn Asn Leu Arg Leu Ile Val Glu Arg Asn
65 70 75 80

Asn Leu Tyr Val Thr Gly Phe Val Asn Arg Thr Asn Asn Val Phe Tyr
85 90 95

Arg Phe Ala Asp Phe Ser His Val Thr Phe Pro Gly Thr Thr Ala Val
100 105 110

Thr Leu Ser Gly Asp Ser Ser Tyr Thr Thr Leu Gln Arg Val Ala Gly
115 120 125

Ile Ser Arg Thr Gly Met Gln Ile Asn Arg His Ser Leu Thr Thr Ser
130 135 140

ES 2 386 338 T3

Tyr Leu Asp Leu Met Ser His Ser Gly Thr Ser Leu Thr Gln Ser Val
145 150 155 160

Ala Arg Ala Met Leu Arg Phe Val Thr Val Thr Ala Glu Ala Leu Arg
165 170 175

Phe Arg Gln Ile Gln Arg Gly Phe Arg Thr Thr Leu Asp Asp Leu Ser
180 185 190

Gly Arg Ser Tyr Val Met Thr Ala Glu Asp Val Asp Leu Thr Leu Asn
195 200 205

Trp Gly Arg Leu Ser Ser Val Leu Pro Asp Tyr His Gly Gln Asp Ser
210 215 220

Val Arg Val Gly Arg Ile Ser Phe Gly Ser Ile Asn Ala Ile Leu Gly
225 230 235 240

Ser Val Ala Leu Ile Leu Asn Cys His His His Ala Ser Arg Val Ala
245 250 255

Arg Met Ala Ser Asp Glu Phe Pro Ser Met Cys Pro Ala Asp Gly Arg
260 265 270

Val Arg Gly Ile Thr His Asn Lys Ile Leu Trp Asp Ser Ser Thr Leu
275 280 285

Gly Ala Ile Leu Met Arg Arg Thr Ile Ser Ser
290 295

5 <210> 2
<211> 32
<212> ADN
<213> Artificial

<220>
<221> Fuente
<223> /nota = "Descripción de la Secuencia Artificial: Cebador"

10 <400> 2

gttactgtga cagctgaagc tttacgtttt cg 32

15 <210> 3
<211> 31
<212> ADN
<213> Artificial

<220>
<221> Fuente
<223> /nota = "Descripción de la Secuencia Artificial: Cebador"

<400> 3

20 gagaagaaga gactgcagat tccatctggt g 31

<210> 4
<211> 302
<212> PRT

ES 2 386 338 T3

<213> Artificial

<220>

<221> Fuente

<223> /nota = "Descripción de la Secuencia Artificial: secuencia de la proteína libn^o3 de la cadena A de SLT-1 (SAM3)"

5

<400> 4

Lys Gly Met Arg Ser His His His His His His His His Ile Glu Gly
1 5 10 15

Arg Ala Ser Lys Glu Phe Thr Leu Asp Phe Ser Thr Ala Lys Thr Tyr
20 25 30

Val Asp Ser Leu Asn Val Ile Arg Ser Ala Ile Gly Thr Pro Leu Gln
35 40 45

Thr Ile Ser Ser Gly Gly Thr Ser Leu Leu Met Ile Asp Ser Gly Ser
50 55 60

Gly Asp Asn Leu Phe Ala Val Asp Val Arg Gly Ile Asp Pro Glu Glu
65 70 75 80

Gly Arg Phe Asn Asn Leu Arg Leu Ile Val Glu Arg Asn Asn Leu Tyr
85 90 95

Val Thr Gly Phe Val Asn Arg Thr Asn Asn Val Phe Tyr Arg Phe Ala
100 105 110

Asp Phe Ser His Val Thr Phe Pro Gly Thr Thr Ala Val Thr Leu Ser
115 120 125

Gly Asp Ser Ser Tyr Thr Thr Leu Gln Arg Val Ala Gly Ile Ser Arg
130 135 140

Thr Gly Met Gln Ile Asn Arg His Ser Leu Thr Thr Ser Tyr Leu Asp
145 150 155 160

Leu Met Ser His Ser Gly Thr Ser Leu Thr Gln Ser Val Ala Arg Ala
165 170 175

Met Leu Arg Phe Val Thr Val Thr Ala Glu Ala Leu Arg Phe Arg Gln
180 185 190

ES 2 386 338 T3

Ile Gln Arg Gly Phe Arg Thr Thr Leu Asp Asp Leu Ser Gly Arg Ser
 195 200 205

Tyr Val Met Thr Ala Glu Asp Val Asp Leu Thr Leu Asn Trp Gly Arg
 210 215 220

Leu Ser Ser Val Leu Pro Asp Tyr His Gly Gln Asp Ser Val Arg Val
 225 230 235 240

Gly Arg Ile Ser Phe Gly Ser Ile Asn Ala Ile Leu Gly Ser Val Ala
 245 250 255

Leu Ile Leu Asn Cys His His His Ile Tyr Ser Asn Lys Leu Met Ala
 260 265 270

Ser Arg Val Ala Arg Met Ala Ser Asp Glu Phe Pro Ser Met Cys Pro
 275 280 285

Ala Asp Gly Arg Val Arg Gly Ile Thr His Asn Lys Ile Leu
 290 295 300

<210> 5

<211> 319

5 <212> PRT

<213> Artificial

<220>

<221> Fuente

10 <223> /nota = "Descripción de la Secuencia Artificial: secuencia de la proteína libn^o5 de la cadena A de SLT-1 (SAM5)"

<400> 5

Lys Gly Met Arg Ser His His His His His His His His Ile Glu Gly
 1 5 10 15

Arg Ala Ser Lys Glu Phe Thr Leu Asp Phe Ser Thr Ala Lys Thr Tyr
 20 25 30

Val Asp Ser Leu Asn Val Ile Arg Ser Ala Ile Gly Thr Pro Leu Gln
 35 40 45

Thr Ile Ser Ser Gly Gly Thr Ser Leu Leu Met Ile Asp Ser Gly Ser
 50 55 60

Gly Asp Asn Leu Phe Ala Val Asp Val Arg Gly Ile Asp Pro Glu Glu
 65 70 75 80

Gly Arg Phe Asn Asn Leu Arg Leu Ile Val Glu Arg Asn Asn Leu Tyr
 85 90 95

ES 2 386 338 T3

Val Thr Gly Phe Val Asn Arg Thr Asn Asn Val Phe Tyr Arg Phe Ala
 100 105 110

Asp Phe Ser His Val Thr Phe Pro Gly Thr Thr Ala Val Thr Leu Ser
 115 120 125

Gly Asp Ser Ser Tyr Thr Thr Leu Gln Arg Val Ala Gly Ile Ser Arg
 130 135 140

Thr Gly Met Gln Ile Asn Arg His Ser Leu Thr Thr Ser Tyr Leu Asp
 145 150 155 160

Leu Met Ser His Ser Gly Thr Ser Leu Thr Gln Ser Val Ala Arg Ala
 165 170 175

Met Leu Arg Phe Val Thr Val Thr Ala Glu Ala Leu Arg Phe Arg Gln
 180 185 190

Ile Gln Arg Gly Phe Arg Thr Thr Leu Asp Asp Leu Ser Gly Arg Ser
 195 200 205

Tyr Val Met Thr Ala Glu Asp Val Asp Leu Thr Leu Asn Trp Gly Arg
 210 215 220

Leu Ser Ser Val Leu Pro Asp Tyr His Gly Gln Asp Ser Val Arg Val
 225 230 235 240

Gly Arg Ile Ser Phe Gly Ser Ile Asn Ala Ile Leu Gly Ser Val Ala
 245 250 255

Leu Ile Leu Asn Cys His His His Ala Ala Phe Ala Asp Leu Ile Ala
 260 265 270

Ser Arg Val Ala Arg Met Ala Ser Asp Glu Phe Pro Ser Met Cys Pro
 275 280 285

Ala Asp Gly Arg Val Arg Gly Ile Thr His Asn Lys Ile Leu Trp Asp
 290 295 300

Ser Ser Thr Leu Gly Ala Ile Leu Met Arg Arg Thr Ile Ser Ser
 305 310 315

<210> 6

<211> 7

<212> PRT

5 <213> Artificial

<220>

<221> Fuente

<223> /nota = "Descripción de la Secuencia Artificial: primer inserto activo de melanoma"

<400> 6

Ile Tyr Ser Asn Lys Leu Met
 1 5

10

ES 2 386 338 T3

<210> 7
<211> 7
<212> PRT
<213> Artificial

5 <220>
<221> Fuente
<223> /nota = "Descripción de la Secuencia Artificial: segundo inserto activo de melanoma"
<400> 7

Ala Ala Phe Ala Asp Leu Ile
1 5

REIVINDICACIONES

- 5 1. Una colección combinatoria de proteínas que comprende una pluralidad de especies proteicas, comprendiendo cada especie de proteína una cadena A de una proteína tóxica ABx que comprende un bucle sensible a proteasa en el cual se ha introducido un inserto, en donde el inserto es un polipéptido de secuencia de aminoácidos variable que tiene una longitud de 2-7 residuos de aminoácidos, y en donde cuando la pluralidad de especies proteicas se ponen en contacto con al menos una diana/receptor candidato, una especie proteica que comprende un inserto se une a una diana/un receptor específico, identificando de este modo el inserto como un ligando de la diana/el receptor específicos.
- 10 2. Una colección combinatoria de proteínas que comprende una pluralidad de especies proteicas, comprendiendo cada especie de proteína una cadena A de una proteína tóxica ABx que comprende un bucle sensible a proteasa en el cual se ha introducido un inserto, en donde el inserto es un polipéptido de secuencia de aminoácidos variable que tiene una longitud de 2-7 residuos de aminoácidos, y en donde cuando la pluralidad de las especies proteicas se ponen en contacto con al menos una diana/un receptor candidato, localizado en la superficie de una célula una especie proteica que comprende un inserto, se une a una diana/un receptor específico y es tóxico para la célula, identificando de este modo un inserto que comprende un ligando de la diana/el receptor.
- 15 3. La colección combinatoria de proteínas de acuerdo con la reivindicación 1, en donde la diana/el receptor se localiza en la superficie de una célula.
4. La colección combinatoria de proteínas de acuerdo con una cualquiera de las reivindicaciones 1-3, en donde la colección comprende al menos 100 especies proteicas.
- 20 5. La colección combinatoria de proteínas de acuerdo con una cualquiera de las reivindicaciones 1-4, en donde las especies proteicas se forman introduciendo el inserto en una cadena A de una toxina I de tipo Shiga, que comprende la secuencia de aminoácidos de SEQ ID NO: 1.
- 25 6. La colección combinatoria de proteínas de acuerdo con la reivindicación 5, en donde la especie proteica se forma introduciendo el inserto entre los aminoácidos 242 y 261, tal y como se define haciendo referencia a SEQ ID NO: 1.
7. La colección combinatoria de proteínas de acuerdo con la reivindicación 6, en donde la especie proteica se forma introduciendo el inserto entre los aminoácidos 245 y 246, tal y como se define haciendo referencia a SEQ ID NO: 1.
- 30 8. La colección combinatoria de proteínas de acuerdo con la reivindicación 5, en donde la especie proteica se forma introduciendo el inserto antes o después de los aminoácidos 1-239 de la cadena A de la toxina I de tipo Shiga, tal y como se define haciendo referencia a SEQ ID NO: 1.
9. La colección combinatoria de proteínas de acuerdo con una cualquiera de las reivindicaciones 1-8, en donde el inserto tiene una longitud de 7 aminoácidos.
- 35 10. Una colección combinatoria de proteínas que comprende una pluralidad de especies de sistemas de expresión, expresando cada especie una especie proteica de acuerdo con una cualquiera de las reivindicaciones 1-9.
11. Una proteína mutante que comprende una cadena A de una proteína tóxica ABx en la que se ha introducido un inserto, en donde
 - (a) el inserto es un polipéptido con una secuencia de aminoácidos variable que tiene una longitud de 2-7 residuos de aminoácidos; y
 - 40 (b) el inserto se introduce en el bucle sensible a proteasa de la secuencia de la cadena A.
12. La proteína mutante de acuerdo con la reivindicación 11, en donde la cadena A de la proteína tóxica ABx es una cadena A de la toxina I de tipo Shiga.
13. La proteína mutante de acuerdo con la reivindicación 12, en donde el inserto se introduce entre los aminoácidos 242 y 261, tal y como se define haciendo referencia a SEQ ID NO: 1.
- 45 14. La proteína mutante de acuerdo con la reivindicación 13, en donde el inserto se introduce entre los aminoácidos 245 y 246, tal y como se define haciendo referencia a SEQ ID NO: 1.
15. La proteína mutante de acuerdo con la reivindicación 14, en donde el inserto comprende la secuencia IYSNKLM (SEQ ID NO: 6).
- 50 16. La proteína mutante de acuerdo con la reivindicación 14, en donde el inserto comprende la secuencia AA-FADLI (SEQ ID NO: 7).

17. La proteína mutante de acuerdo con la reivindicación 12, en donde el inserto se introduce antes o después de los aminoácidos 1-239 de la cadena A de la toxina I de tipo Shiga, tal y como se define haciendo referencia a SEQ ID NO: 1.
- 5 18. La proteína mutante de acuerdo con cualquiera de las reivindicaciones 11-17, en donde el inserto tiene una longitud de 7 aminoácidos.
19. Un método para identificar un ligando que se une a una diana/un receptor específico, que comprende las etapas de:
- (a) exponer las células que se sabe que tienen la diana/el receptor a miembros de una colección combinatoria de proteínas de acuerdo con cualquiera de las reivindicaciones 1-9; y
- 10 (b) seleccionar miembros de la colección de proteínas que se observan que son tóxicos para las células.
20. El método de acuerdo con la reivindicación 19, que comprende adicionalmente la etapa de evaluar los miembros seleccionados de la colección proteica para determinar la secuencia de la región insertada, identificándose de ese modo un péptido de la secuencia de la región insertada como posible ligando de una diana/un receptor sobre la célula.
- 15 21. El método de acuerdo con una cualquiera de las reivindicaciones 19-20, que comprende adicionalmente la etapa de someter a ensayo adicionalmente los miembros seleccionados de la colección proteica para confirmar que son ligandos de la diana/el receptor específico.
22. Un método para aislar una toxina específica de una diana/un receptor que comprende las etapas de:
- 20 (a) exponer la diana/el receptor a una colección combinatoria de proteínas de acuerdo con una cualquiera de las reivindicaciones 1-9; y
- (b) aislar al menos una especie proteica de la colección de proteínas capturada por la unión a la diana/el receptor.
23. El método de acuerdo con la reivindicación 22, que comprende adicionalmente la etapa de escrutar la proteína aislada frente a células que expresan la diana/el receptor para confirmar su toxicidad para células que expresan la diana/el receptor.
- 25 24. El método de acuerdo con la reivindicación 22, en el que la diana/el receptor es una diana/el receptor purificado y se inmoviliza sobre un soporte sólido.
25. El método de acuerdo con la reivindicación 23, en donde la diana/el receptor está en la superficie de las células.
26. El método de acuerdo con la reivindicación 25, en donde las células se inmovilizan sobre un soporte sólido.
- 30 27. El método de acuerdo con una cualquiera de las reivindicaciones 25-26, en donde la toxina sirve como informadora y la muerte de las células es indicativa de la unión al receptor.
28. Un método para identificar una toxina para una célula que expresa una diana/un receptor específico, que comprende las etapas de:
- 35 (a) exponer las células que se sabe que tienen la diana/el receptor a miembros de una colección combinatoria de proteínas de acuerdo con una cualquiera de las reivindicaciones 1-9; y
- (b) seleccionar miembros de la colección de proteínas que se observan que son tóxicos para las células,
- en donde los miembros de la colección proteica que se observan que son tóxicos para las células, son toxinas para las células que expresan la diana/el receptor específico.
- 40 29. El método de acuerdo con la reivindicación 28, que comprende adicionalmente la etapa de identificar un ligando que se une a una diana/un receptor específico, que comprende las etapas de:
- (a) exponer las células que se sabe que tienen la diana/el receptor a miembros de una colección combinatoria de proteínas de acuerdo con una cualquiera de las reivindicaciones 1-9; y
- (b) seleccionar los miembros de la colección de proteínas que se observan que son tóxicos para las células,
- 45 en donde los miembros de la colección proteica que se observan que son tóxicos para las células, son ligandos que se unen a la diana/el receptor específico.
30. El método de acuerdo con una cualquiera de las reivindicaciones 28-29, que comprende adicionalmente una etapa o etapas de mutación de un inserto para incrementar adicionalmente la toxicidad de la especie proteica en

comparación con la especie proteica que comprende el inserto original.

31. Un método para incrementar la toxicidad de una especie proteica de acuerdo con la reivindicación 11, que comprende las etapas de:

(a) mutar un inserto conocido que posee toxicidad para las células que expresan una diana/un receptor específico; y

5 (b) comparar la toxicidad de la especie proteica que comprende el inserto mutado con la especie proteica que comprende el inserto original.

32. El método de acuerdo con una cualquiera de las reivindicaciones 28-31, que comprende adicionalmente una etapa de evaluación de los miembros seleccionados de la colección proteica para determinar la secuencia de la región insertada.

10 33. El método de acuerdo con una cualquiera de las reivindicaciones 28-32, que comprende adicionalmente una etapa de someter a ensayo adicionalmente los miembros seleccionados de la colección proteica para confirmar que son tóxicos para una célula que expresa una diana/un receptor específico.

34. El método de acuerdo con una cualquiera de las reivindicaciones 19-21, 23 o 25-33, en donde las células son células cancerosas.

15 35. El método de acuerdo con la reivindicación 34, en donde las células cancerosas se seleccionan entre el grupo consistente en células de melanoma, células de cáncer de cuello uterino o células de cáncer cerebral.

36. El método de acuerdo con la reivindicación 35, en donde las células cancerosas son células de melanoma.

37. El método de acuerdo con la reivindicación 35, en donde las células cancerosas son células de cáncer de cuello uterino.

20 38. El método de acuerdo con la reivindicación 35, en donde las células cancerosas son células de cáncer cerebral.

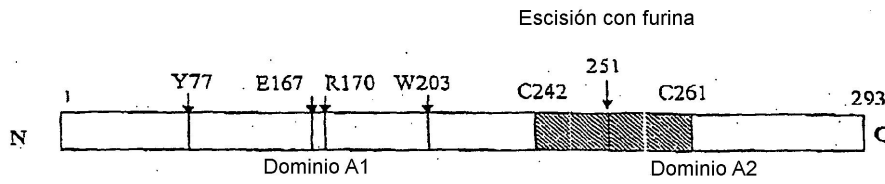


Fig. 1

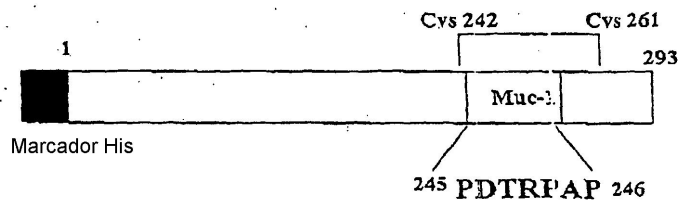


Fig. 2A

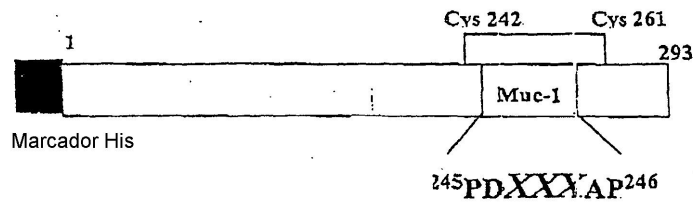


Fig. 2B

Conjunto de datos representativos de ELISA
procedentes del escrutinio de 96 variantes
de toxinas distintas

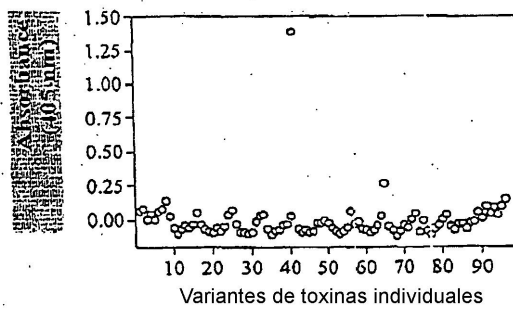


Fig. 3

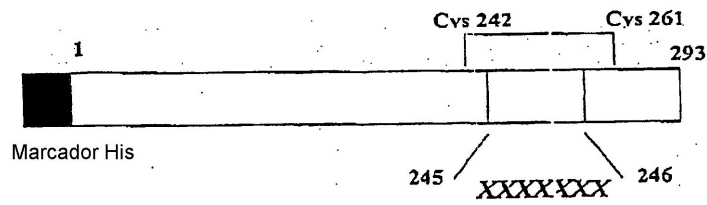


Fig. 4

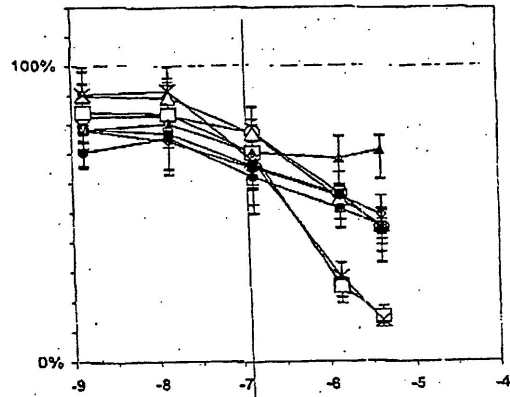


Fig 5