



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 321 058**

51 Int. Cl.:

C12N 15/62 (2006.01)

C07K 1/22 (2006.01)

C07K 19/00 (2006.01)

C12N 9/16 (2006.01)

G01N 33/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **00966755 .1**

96 Fecha de presentación : **20.09.2000**

97 Número de publicación de la solicitud: **1220933**

97 Fecha de publicación de la solicitud: **10.07.2002**

54

Título: **Purificación de proteínas recombinantes fusionadas con epítopos múltiples.**

30

Prioridad: **08.10.1999 US 415000**

45

Fecha de publicación de la mención BOPI:
02.06.2009

45

Fecha de la publicación del folleto de la patente:
02.06.2009

73

Titular/es: **Sigma-Aldrich Company**
509 West Monroe Street
Highland, Illinois 62249, US

72

Inventor/es: **Brizzard, Bill y**
Hernan, Ron

74

Agente: **Mir Plaja, Mireia**

ES 2 321 058 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

Purificación de proteínas recombinantes fusionadas con epítomos múltiples.

5 **Campo de la invención**

La presente invención se refiere a etiquetas de proteínas y a métodos de purificación de proteínas que usan varias técnicas de ADN recombinante. Más particularmente, la invención trata sobre polipéptidos de identificación nove-
 10 dosos y vectores de ADN que codifican polipéptidos de identificación novedosos que contienen múltiples dominios antigénicos unidos en tándem. Se proporcionan también métodos para usar dichos polipéptidos de identificación para la purificación de péptidos diana y métodos de construcción de vectores de ADN que codifican los polipéptidos de identificación novedosos y vectores de expresión de ADN que codifican los polipéptidos de identificación unidos a un péptido diana.

15 **Antecedentes de la invención**

Las moléculas proteicas tales como enzimas, hormonas, proteínas de almacenamiento, proteínas de unión, pro-
 20 teínas transportadoras y proteínas transductoras de señales se pueden producir y purificar usando varias técnicas de ADN recombinante. Por ejemplo, fragmentos de ADN que codifican una proteína seleccionada, junto con secuencias de ADN apropiadas para un promotor y un sitio de unión al ribosoma se ligan a un plásmido vector. El plásmido se inserta dentro de una célula procariota o eucariota hospedadora. Las células hospedadoras transformadas se identifi-
 25 can, se aíslan y a continuación se cultivan para conseguir la expresión de las moléculas proteicas. Uno de los métodos usados para purificar polipéptidos híbridos es el sistema de poli-arginina en el que un polipéptido híbrido se purifica selectivamente sobre una resina de intercambio catiónico. Véase Sassenfeld, H. M. y Brewer, S. J. *BioTechnology*, 2:76 (1984); patente U.S. n.º 4.532.207. Sassenfeld y Brewer dieron a conocer una extensión del carboxi terminal de cinco residuos de arginina fusionados en una proteína diana. Esta extensión de poliarginina básica permitió la purificación del polipéptido híbrido sobre una resina SP-Sephadex. Un sistema análogo de expresión y purificación de proteínas utiliza una cola o etiqueta de polihistidina en el extremo o bien amino o bien carboxi del polipéptido híbrido. La proteína de fusión se purifica por cromatografía sobre una resina de afinidad metálica con Ni²⁺. Véase Porath,
 30 J., *Protein Expression and Purification*, 3:7995 (1992). Adicionalmente, actualmente se utilizan varios protocolos de purificación por afinidad para facilitar el aislamiento de proteínas de fusión. La cromatografía de afinidad se basa en la capacidad de las proteínas a unirse de forma específica y no covalente con un ligando. Usada por sí sola, puede aislar proteínas de mezclas muy complejas con no solamente un mayor grado de purificación que el posible mediante la cromatografía secuencial de intercambio iónico y en columna de gel, sino también sin pérdidas significativas de actividad. Típicamente, como socio de fusión se selecciona un ligando capaz de unirse con una alta especificidad a una matriz
 35 de afinidad. Por ejemplo, la p-aminofenil-β-D-tiogalactosidil-succinildiaminohexil-Sefarosa se une selectivamente a la β-galactosidasa permitiendo la purificación de proteínas de fusión a β-gal. Véase Germino *et al.*, *Proc. Natl. Acad. Sci. USA* 80:6848 (1983). Otros sistemas de expresión que permiten la purificación por afinidad de proteínas de fusión incluyen proteínas de fusión realizadas con glutatión-S-transferasa, que se recuperan selectivamente en glutatión-agarosa. Véase Smith, D. B. y Johnson., K. S. *Gene* 67:31 (1988). La IgG-Sefarosa se puede usar para purificar por afinidad proteínas de fusión que contienen la proteína estafilocócica A. Véase Uhlen, M. *et al.* *Gene* 23:369 (1983). El dominio de la proteína de unión a maltosa del gen malE de *E. coli* se ha usado como socio de fusión y permite la purificación por afinidad de la proteína de fusión sobre resinas de amilasa.

Otro método usado para detectar y aislar proteínas es mediante el uso de un epítomo etiqueta. El etiquetado con
 45 epítomos utiliza anticuerpos contra péptidos huéspedes para estudiar la localización de la proteína en el nivel celular y niveles subcelulares. Véase Kolodziej, P. A. y Young, RA., *Methods Enzymol.*, 194:508-519 (1991). Usando la tecnología del ADN recombinante, una secuencia de nucleótidos que codifica el epítomo se inserta en la región codificante del gel clonado, y el gen híbrido se introduce en una célula mediante un método tal como transformación. Cuando se expresa el gen híbrido, el resultado es una proteína quimérica que contiene el epítomo como péptido huésped. Si el epítomo queda expuesto en la superficie de la proteína, el mismo queda disponible para su reconocimiento por el anti-
 50 cuerpo específico del epítomo, permitiendo que el investigador observe la proteína dentro de la célula usando técnicas de inmunofluorescencia u otras técnicas de inmunolocalización. Además, las proteínas de fusión marcadas con dichos epítomos etiquetas se usan frecuentemente para purificar proteínas utilizando técnicas de purificación por afinidad.

De este modo, el etiquetado con epítomos se ha convertido en una herramienta poderosa para la detección y puri-
 55 ficación de proteínas expresadas. Véase Kolodziej, P. A. y Young, RA., *Methods Enzymol.*, 194:508-519 (1991). Se han usado muchos tipos de etiquetas, siendo c-myc y las etiquetas FLAG[®] dos de los epítomos más populares usados. Véase Evan *et al.*, *Mol Cell Biol.* 5:3610-3616 (1985). En general, estos epítomos se fusionan con el extremo amino o carboxi de la proteína expresada consiguiendo que los mismos resulten más accesibles al anticuerpo para su detección y que haya menos probabilidad de provocar perturbaciones estructurales o funcionales importantes.

Forsburg, S.L., *et al.*, *Gene*, 191:191-195 (1997) describen una serie de vectores para ser usados en la levadura de
 65 fisión *Schizosaccharomyces pombe* que permiten la fusión de una proteína de interés a un epítomo HA triple o dominio GST. El epítomo HA se puede situar en el extremo N ó el extremo C bajo tres versiones diferentes del promotor *nmt1*, para permitir niveles variables de expresión génica.

Nilsson, J., *et al.* *Protein Expression and Purification*, 11:1-16 (1997) proporcionan una visión general de etiquetas

de afinidad usadas comúnmente que se centran en el uso de fusiones de las etiquetas de afinidad para la detección, purificación, e inmovilización de proteínas recombinantes y que ofrecen algunos ejemplos recientes de aplicaciones específicas para etiquetas de afinidad.

5 Las proteínas de fusión que tienen el octapéptido FLAG[®] Asp-Tyr-Lys-Asp-Asp-Asp-Lys (ID SEC n.º: 1) en el extremo amino se pueden purificar por afinidad sobre una resina de inunoafinidad que contenga un anticuerpo específico para el octapéptido. Véase Hopp, T. P., *et al.* Biotechnology, 6:1204 (1988); Prickett, K. S., *et al.*, BioTechniques, 7:580 (1989); y la patente U.S. n.º 4.851.341. El epítipo etiqueta FLAG[®] se ha usado eficazmente para detectar y purificar proteínas en sistemas de mamíferos y bacterianos. La secuencia FLAG[®] original es reconocida por
10 dos anticuerpos, M1, M2, y una secuencia FLAG[®] con un iniciador metionina fijado es reconocida por un tercer anticuerpo, M5. Los últimos cinco aminoácidos de la secuencia FLAG[®] son un sitio de reconocimiento para la proteasa enteroquinasa, permitiendo de este modo la eliminación del epítipo FLAG[®]. El epítipo FLAG[®] se ha usado en varios sistemas de expresión para la detección y purificación de proteínas heterólogas, por ejemplo, en *E. coli* (Brizzard *et al.*, BioTechniques, 16:730-735 (1994)), *Saccharomyces cerevisiae* (Lee *et al.*, Nature, 372:739-746 (1994); Prickett
15 *et al.*, BioTechniques, 7:580-589 (1989)), *Drosophila* (Xu *et al.*, Development, 117:1223-1237 (1993)), Baculovirus (Dent *et al.*, Mol, Cell Biol, 15:4125-4135 (1995); Ritchie *et al.*, Biochem Journal, 338:305-10 (1999)), y sistemas mamíferos (Overholt *et al.*, Clin. Cancer Res., 3:185-191 (1997); Schulte am Esch *et al.*, Biochemistry, 38:2248-2258 (1999)). No obstante, en muchos sistemas de expresión mamíferos, los niveles de expresión de proteínas son bajos, y una detección eficaz de proteínas foráneas expresadas usando métodos establecidos puede resultar difícil.
20

Por lo tanto, existe una necesidad de un epítipo etiqueta y un sistema de expresión que utilice dichos epítopos etiqueta los cuales permitirían el aumento de la sensibilidad y la detección de proteínas recombinantes.

Resumen de la invención

25 La presente invención afronta uno o más de los problemas anteriores proporcionando métodos y vehículos que se pueden usar para producir altos rendimientos de proteínas recombinantes. Por consiguiente, de entre los varios objetivos de la presente invención se pueden indicar la provisión de un polipéptido de identificación novedoso, una molécula híbrida compuesta por un péptido diana fusionado en el polipéptido de identificación novedoso y vectores
30 de ADN recombinante que codifican los mismos. Se proporcionan también métodos para la purificación del péptido diana en los que se puede utilizar un único ligando o múltiples ligandos, preferentemente anticuerpos, para aislar y purificar sustancialmente todas las moléculas proteicas expresadas por células hospedadoras transformadas, ya sean o no antigénicas. Otro objetivo de la presente invención es proporcionar procesos que se pueden usar para purificar altamente cualquier molécula proteica producida por métodos de ADN recombinante, incluyendo aquellas que no son
35 susceptibles de experimentar procedimientos de cromatografía de afinidad.

Por lo tanto, resumiendo, la presente invención se refiere a un polipéptido de identificación que comprende múltiples copias de un dominio antigénico unidas entre sí en tándem. El polipéptido de identificación puede contener una
40 secuencia de unión que contenga un sitio escindible situado adyacente al péptido diana en donde el sitio escindible no está situado en o interpuesto entre los dominios antigénicos individuales. Cada dominio antigénico es capaz de provocar una respuesta antigénica y se puede unir mediante un ligando, preferentemente, un anticuerpo. Además, cada dominio antigénico está compuesto por una combinación de por lo menos dos, preferentemente tres o más, aminoácidos diferentes.

45 Se proporcionan también proteínas de fusión de la presente invención que comprenden el polipéptido de identificación novedoso fusionado a un péptido diana. El polipéptido de identificación contiene una secuencia de unión que está caracterizada por ser escindible en un residuo aminoácido específico adyacente al péptido diana mediante el uso de un agente proteolítico específico de la secuencia. Dicho sitio escindible está situado adyacente o bien al extremo carboxi o bien al extremo amino del péptido diana, preferentemente situado inmediatamente adyacente al extremo amino del
50 péptido diana. Idealmente, la secuencia de aminoácidos del sitio escindible es única, minimizando de este modo la posibilidad de que el agente proteolítico escinda el péptido diana. En una realización preferida, el sitio escindible comprende aminoácidos específicos para enteroquinasa, trombina o un Factor Xa.

Según este constructo particular de la proteína de fusión, el péptido diana se puede aislar mediante técnicas de
55 cromatografía de afinidad. De este modo, es un objetivo de la invención proporcionar métodos para la purificación del péptido diana. Esto se logra construyendo una columna de afinidad con ligandos inmovilizados específicos para los dominios antigénicos del polipéptido de identificación uniéndose de este modo la proteína de fusión. Se apreciará que, gracias a la presente invención, se pueden usar un único anticuerpo o múltiples anticuerpos para unirse a los dominios antigénicos individuales que comprenden los múltiples dominios antigénicos del polipéptido de identificación.
60 A continuación, la proteína de fusión unida se puede liberar de la columna y el polipéptido de identificación se puede escindir con un agente proteolítico apropiado, liberando de este modo un péptido diana purificado. En una realización preferida, el agente proteolítico usado para escindir el péptido diana del polipéptido de identificación se selecciona de entre el grupo consistente en enteroquinasa, trombina y Factor Xa.

65 Otro objetivo de la presente invención es proporcionar un vector de clonación recombinante que contenga ADN que codifique el polipéptido de identificación. El vector que codifica el polipéptido de identificación incluye también secuencias de ADN que codifican un sitio de clonación múltiple compuesto por sitios de enzimas de restricción múltiples que se pueden situar entre los dominios antigénicos o en cualquiera de los lados de los dominios antigénicos

lo cual permitirá que los expertos en la materia inserten un número cualquiera de secuencias de ADN que codifiquen cualquier proteína deseada. Esta secuencia de ADN se puede insertar dentro de un vector de clonación tal como un plásmido, mediante el uso de endonucleasas y ligasas de restricción apropiadas. El plásmido recombinante se utiliza para transformar células hospedadoras procariotas o eucariotas compatibles para la replicación del plásmido y la expresión de la molécula proteica/dominio de afinidad híbrido. Idealmente, el plásmido tiene un gen marcador fenotípico para la identificación y el aislamiento de células hospedadoras transformadas. En una realización preferida, secuencias de ADN que codifican un péptido señal secretado se unirá o bien al vector de ADN o bien al plásmido permitiendo de este modo que las células hospedadoras transformadas sean fácilmente identificadas y separadas con respecto a células que no experimentan transformación.

Otros objetivos y características en parte resultarán evidentes y en parte serán señalados en lo sucesivo en el presente documento.

Breve descripción de los dibujos

La Fig. 1A es una representación de una secuencia de ADN y proteica del sitio de clonación múltiple 3XFLAG-CMV-7 y secuencias de FLAG.

La Fig. 1B es un mapa del plásmido del p3XFLAG-CMV-7 que muestra el promotor CMV, un sitio de terminación de transcripción y poliadenilación de la hormona del crecimiento humano, el origen de replicación del SV40, el origen de replicación Col E1, y el gen de la β -lactamasa.

La Fig. 2A es un mapa del vector del p3xFLAG-CMV-7-BAP que muestra la inserción de la región codificante del *phoA* en P3xflag-CMV-7.

La Fig. 2B es un mapa del vector de p3XFLAG-ATS-BAP que muestra la inserción de la región codificante del *phoA* en pFLAG-ATS-BAP.

La Fig. 3 es un Western Blot de 3XFLAG-BAP (A) y N-FLAG-BAP (B) purificados usando el anticuerpo anti-FLAG M2. Calle(s) (1) 0,5 ng; (2) 1,0 ng; (3) 2,0 ng; (4) 5,0 ng; y (5) 10 ng. Las cantidades mostradas son las cantidades cargadas en el gel antes de la transferencia.

Abreviaturas y definiciones

Para facilitar la comprensión de la invención, a continuación se definen una serie de términos:

Las bases de nucleótidos se abrevian en el presente documento de la manera siguiente: A representa adenina; C: representa citosina; G representa guanina; T representa timina; U representa uracilo.

Los residuos de aminoácidos se abrevian en el presente documento según sus letras individuales: A representa alanina; R representa arginina; N representa asparagina; D representa ácido aspártico; C representa cisteína; Q representa glutamina; E representa ácido glutámico; G representa glicina; H representa histidina; I representa isoleucina; L representa leucina; K representa lisina; M representa metionina; F representa fenilalanina; P representa prolina; S representa serina; T representa treonina; W representa triptófano; Y representa tirosina; V representa valina.

La expresión “molécula de ADN recombinante” tal como se usa en el presente documento se refiere a una molécula de ADN que está compuesta por segmentos de ADN unidos entre sí por medio de tecnología de ADN recombinante.

La expresión “vector de expresión” tal como se usa en el presente documento hace referencia a secuencias de ácido nucleico que contienen una secuencia de codificación deseada y secuencias de ácido nucleico apropiadas necesarias para la expresión de la secuencia de codificación unida operativamente en un organismo hospedador particular. Las secuencias de ácido nucleico necesarias para la expresión en procariotas incluyen un promotor, un sitio de unión al ribosoma, un codón de iniciación, un codón de parada, opcionalmente una secuencia operadora y posiblemente otras secuencias. Las células eucariotas utilizan promotores, una secuencia Kozak y frecuentemente intensificadores y señales de poliadenilación. Las células procariotas utilizan también un sitio de unión al Ribosoma Shine-Dalgarno. La presente invención incluye vectores o plásmidos que se pueden usar como vehículos para transformar cualquier célula hospedadora viable con el vector de expresión de ADN recombinante.

El término “FLAG” tal como se usa en el presente documento es la marca comercial registrada que se refiere al epítipo etiqueta FLAG[®] ampliamente usado que consiste en la secuencia peptídica sintética de DYKDDDDK (ID SEC n.º: 1) según se describe en las patentes U.S. n.º 4.703.004, 4.782.137, 4.851.341 y 5.011.912.

El término “hidrófilo” cuando se usa en referencia a aminoácidos se refiere a aquellos aminoácidos que tienen cadenas laterales polares y/o cargadas. Los aminoácidos hidrófilos incluyen lisina, arginina, histidina, aspartato (es decir, ácido aspártico), glutamato (es decir, ácido glutámico), serina, treonina, cisteína, tirosina, asparagina y glutamina.

Al término “hidrófobo” cuando se usa en referencia a aminoácidos se refiere a aquellos aminoácidos que tienen cadenas laterales no polares. Los aminoácidos hidrófobos incluyen valina, leucina, isoleucina, cisteína y metionina.

Tres aminoácidos hidrófobos tienen cadenas laterales aromáticas. Por consiguiente, el término “aromático” cuando se usa en referencia a aminoácidos se refiere a los tres aminoácidos hidrófobos aromáticos fenilalanina, tirosina y triptófano.

5 La expresión “sitio escindible” se refiere a una secuencia definida de aminoácidos que permite la escisión de una proteína o péptido que contiene esta secuencia mediante un agente proteolítico selectivo.

10 La expresión “proteína de fusión” tal como se usa en el presente documento se refiere a un polipéptido híbrido que comprende dominios proteicos de por lo menos dos proteínas diferentes. El péptido diana puede estar situado en la parte terminal amino de la proteína de fusión o en la proteína terminal carboxi formando de este modo una “proteína de fusión terminal amino” o una “proteína de fusión terminal carboxi”, respectivamente.

15 La expresión “péptido diana” tal como se usa en el presente documento se refiere al péptido cuya expresión se desea dentro del polipéptido híbrido. En los polipéptidos híbridos de la invención, el péptido diana puede comprender la parte terminal o bien amino o bien carboxi del polipéptido híbrido.

20 El término “endoproteasa” o “endopeptidasa” tal como se usa en el presente documento se refiere a una proteasa capaz de hidrolizar enlaces peptídicos interiores de un polipéptido, por lugares que no sean los enlaces terminales (es decir, los enlaces peptídicos del aminoácido terminal).

25 Las expresiones “molécula de ácido nucleico que codifica”, “secuencia de ADN que codifica” y “ADN que codifica” se refieren al orden o secuencia de desoxirribonucleótidos a lo largo de una cadena de ácido desoxirribonucleico. El orden de estos desoxirribonucleótidos determina el orden de aminoácidos a lo largo de la cadena polipeptídica (proteica). De este modo, la secuencia de ADN codifica la secuencia de aminoácidos.

30 Para todas las secuencias de nucleótidos y aminoácidos dadas a conocer en el presente documento, se entiende que en las secuencias se pueden sustituir nucleótidos y aminoácidos equivalentes sin afectar a la función de las secuencias. Dichas sustituciones quedan incluidas dentro de las habilidades de un profesional con conocimientos habituales en la materia.

35 Los procedimientos dados a conocer en el presente documento que conllevan la manipulación molecular de ácidos nucleicos son conocidos para aquellos expertos en la materia. Véanse en general Fredrick M. Ausubel *et al.* (1995), “Short Protocols in Molecular Biology,” John Wiley and Sons, y Joseph Sambrook *et al.* (1989), “Molecular Cloning, A Laboratory Manual,” segunda ed., Cold Spring Harbor Laboratory Press.

Descripción detallada

40 Según la presente invención, se proporcionan polipéptidos de identificación con una sensibilidad aumentada para la detección y purificación de péptidos diana que se producen usando una tecnología de ADN recombinante. Se proporcionan además moléculas de polipéptidos híbridos compuestas por un polipéptido de identificación y un péptido diana que se producen mediante tecnología de ADN recombinante y se purifican usando cromatografía de afinidad haciendo uso de uno o más ligandos. Por consiguiente, se proporcionan también vectores de expresión de ADN que incluyen segmentos de ADN que codifica el polipéptido de identificación y el péptido diana deseado.

45 Según la presente invención, un péptido diana puede estar compuesto por cualquier sustancia proteica que se pueda expresar en células hospedadoras transformadas. La antigenicidad aumentada del polipéptido de identificación es el resultado de la presencia de múltiples copias de un dominio antigénico en tándem. El polipéptido de identificación contiene también una secuencia de unión escindible que une el polipéptido de identificación al péptido diana produciendo de este modo un polipéptido híbrido. Los vectores de clonación de ADN se pueden replicar y el polipéptido híbrido compuesto por el polipéptido de identificación y un péptido diana se expresa en células procariontas o eucariontas transformadas con el vector. Las células transformadas se aíslan y a continuación se expanden en cultivo u otros medios conocidos en la materia.

50 Las moléculas de polipéptidos híbridos de la presente invención se pueden purificar mediante cromatografía de afinidad. La molécula de polipéptido híbrido que comprende el polipéptido de identificación y el péptido diana se puede purificar usando una resina de afinidad que se une a los dominios antigénicos del polipéptido de identificación. En general, los ligandos específicos para los dominios antigénicos del polipéptido etiqueta de identificación se inmovilizan en una columna de perlas u otro tipo de matriz. Un extracto de las células hospedadoras obtenido a partir del cultivo se aplica a la columna y a continuación los polipéptidos que se unen a la columna son eluidos. Seguidamente el polipéptido de identificación se escinde de la molécula del péptido diana con un agente proteolítico apropiado liberando de este modo el péptido diana en un estado altamente purificado.

Polipéptido de identificación

65 El polipéptido de identificación de la presente invención es una secuencia de residuos de aminoácidos que flanquean al extremo amino o carboxi de un péptido diana. En general, el polipéptido de identificación incluye múltiples copias de un dominio antigénico, en las que cada dominio antigénico es capaz de unir un anticuerpo, una secuencia de unión escindible para unir los dominios antigénicos al péptido diana, y opcionalmente uno o más espaciadores.

ES 2 321 058 T3

Para aumentar la sensibilidad de detección, el polipéptido de identificación contiene preferentemente múltiples copias de un dominio antigénico, es decir, por lo menos dos copias de un dominio antigénico, preferentemente por lo menos tres copias de un dominio antigénico y, en algunas realizaciones, cuatro o más copias de un dominio antigénico. La capacidad de la secuencia de múltiples dominios antigénicos de unirse a un anticuerpo inmovilizado, por ejemplo, en una columna u otra matriz, posibilita el aislamiento y la purificación de péptidos diana.

El polipéptido de identificación incluye una secuencia de unión escindible para unir la secuencia de dominios antigénicos al péptido diana. En general, los residuos de aminoácidos que comprenden la secuencia de unión pueden comprender cualquier secuencia de aminoácidos que serviría para conectar la secuencia de dominios antigénicos al péptido diana. Además, la secuencia de unión contiene un sitio de escisión que comprende una secuencia única de aminoácidos escindible mediante el uso de un agente proteolítico específico de la secuencia. Una vez que el polipéptido híbrido compuesto por el polipéptido de identificación y el péptido diana se ha purificado a partir del extracto del cultivo, el polipéptido de identificación preferentemente se escinde del péptido diana mediante digestión con un agente proteolítico específico para los aminoácidos del sitio de escisión. Alternativamente, el polipéptido de identificación se puede extraer del péptido diana mediante escisión química usando métodos conocidos en la materia.

En general, el sitio escindible puede estar situado en el extremo amino o carboxi del péptido diana. Preferentemente, el sitio escindible es inmediatamente adyacente al péptido diana para permitir la separación del péptido diana con respecto al polipéptido de identificación. Preferentemente, este sitio escindible no aparece en o interpuesto entre los dominios antigénicos o, en el caso de que estén presentes, los dominios espaciadores del polipéptido de identificación. En una realización preferida, el sitio escindible está situado en el extremo amino del péptido diana. Si el sitio escindible está situado en el extremo amino del péptido diana y si existen aminoácidos extraños restantes en el péptido diana después de la escisión con el agente proteolítico, se puede utilizar una endopeptidasa tal como tripsina, clostropaina o furina para eliminar estos aminoácidos restantes, dando como resultado de este modo un péptido diana altamente purificado.

La digestión con un agente proteolítico se puede producir mientras el polipéptido híbrido está todavía unido a la resina de afinidad o alternativamente, el polipéptido híbrido se puede eluir de la resina de afinidad y a continuación digerir con el agente proteolítico para purificar adicionalmente el péptido diana. La eficacia del agente proteolítico o la escisión química del péptido diana recombinante se determina por medio de la secuencia de aminoácidos de la secuencia de unión interpuesta entre la secuencia de dominios antigénicos y el péptido diana.

Idealmente, la secuencia de aminoácidos del sitio de escisión es única, minimizando de este modo la posibilidad de que el agente proteolítico escinda al péptido diana. En una realización preferida, el sitio escindible comprende aminoácidos para un sitio de escisión de enteroquinasa, trombina o Factor Xa.

La enteroquinasa reconoce varias secuencias: Asp-Lys; Asp-Asp-Lys; Asp-Asp-Asp-Lys (ID SEC n.º : 2); y Asp-Asp-Asp-Asp-Lys (ID SEC n.º: 3). Véase Matsushima *et al.*, J. Biochem 125:947-51 (1999). La única aparición natural conocida de Asp-Asp-Asp-Asp-Lys (ID SEC n.º: 3) es en la proteína tripsinógeno que es un sustrato natural para la enteroquinasa bovina y algunas proteínas de levadura. Como tal, interponiendo una secuencia de unión que contenga la secuencia de aminoácidos Asp-Asp-Asp-Asp-Lys (ID SEC n.º: 3) como sitio escindible entre la secuencia de dominios antigénicos y el extremo amino del péptido diana, el péptido diana se puede liberar del polipéptido de identificación mediante el uso de enteroquinasa bovina con una probabilidad muy pequeña de que esta enzima escinda cualquier parte del propio péptido diana.

La trombina se escinde en el lado del terminal carboxi de arginina en la siguiente secuencia: Leu-Val-Pro-Arg-Gly-X (ID SEC n.º: 4), en la que X es un aminoácido no ácido. Véase Chang Eur. J. Biochem., 151:217 (1985). La proteasa Factor Xa (es decir, la forma activada del Factor X) se escinde después de la Arg en las siguientes secuencias: Ile-Glu-Gly-Arg-X (ID SEC n.º: 5), Ile-Asp-Gly-Arg-X (ID SEC n.º: 6), y Ala-Glu-Gly-Arg-X (ID SEC n.º: 7), en las que X es cualquier aminoácido excepto prolina o arginina. Se demostró que una proteína de fusión que comprendía los 31 residuos de terminales amino de la proteína cII, un sitio de escisión de Factor Xa y β -globina humana era escindida por el Factor Xa y generaba β -globina auténtica. Véase Nagai, K. y Thogersen, H. C., Nature, 308: 810-812 (1984). Una de las limitaciones de los sistemas de fusión basados en el Factor Xa es el hecho de que se ha informado de que el Factor Xa se escinde en residuos de arginina que no están presentes dentro de la secuencia de reconocimiento del Factor Xa. Véase Nagai *et al.*, Prot, Expr. and Purif., 2:372 (1991).

En la secuencia de unión también se pueden utilizar otras secuencias de aminoácidos únicas, menos preferidas, para otros sitios escindibles. Por ejemplo, la secuencia de unión puede estar compuesta en parte por un par de aminoácidos básicos, es decir, Lys, Arg o His. Esta secuencia es escindida por calicreínas, una enzima glandular. Además, la parte de unión puede estar compuesta parcialmente por Arg-Gly, ya que se sabe que la enzima trombina se escindirá después de la Arg si a este residuo le sigue Gly. Además, no se requiere que los dominios antigénicos y el sitio escindible sean exclusivos entre sí. Los anticuerpos se pueden unir a aminoácidos que se encuentren en el sitio escindible tal como es el caso correspondiente al anticuerpo monoclonal FLAG[®] M2 que reconoce parte del sitio escindible Asp-Asp-Asp-Asp-Lys (ID SEC n.º: 3) para la enteroquinasa.

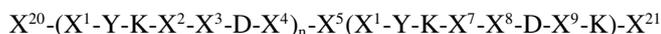
Aunque en general se prefiere que cada dominio antigénico sea inmediatamente adyacente a otro dominio antigénico (es decir, sin secuencias interpuestas), los dominios antigénicos se pueden separar entre sí por medio de un dominio espaciador. También se puede insertar un dominio espaciador entre las múltiples copias del dominio antigénico y la

secuencia de unión. La inserción de un dominio espaciador preferentemente no da como resultado la inserción de una segunda copia del sitio escindible entre los dominios antigénicos del polipéptido de identificación. Se prefiere que el número de residuos de aminoácidos en cada dominio espaciador sea mínimo, preferentemente que consista en no más de diez residuos de aminoácidos, más preferentemente, no más que aproximadamente seis residuos de aminoácidos, y todavía más preferentemente dos o incluso un residuo(s) de aminoácidos en cuanto a longitud.

Si se utiliza un dominio espaciador, el mismo puede estar diseñado para comunicar una o más propiedades deseadas al polipéptido de identificación. En una de las realizaciones, el(los) aminoácido(s) del dominio espaciador se seleccionan de entre aminoácidos hidrófilos para aumentar el carácter hidrófilo del polipéptido de identificación. Alternativamente, el(los) aminoácido(s) del dominio espaciador se puede(n) seleccionar para comunicar un plegamiento deseado al polipéptido de identificación, aumentando de este modo la accesibilidad al anticuerpo; por ejemplo, el dominio espaciador puede comprender residuos de glicina lo cual da como resultado una conformación de plegamiento proteico que permite una accesibilidad mejorada al anticuerpo. Véase Dan *et al.*, J. Bio. Chem. 271:30717-30724 (1996); Borjigin, J. y Nathans, J., J. Biol. Chem. 269:14715-147622 (1994).

Es bien sabido en la técnica que ciertos residuos de aminoácidos tales como histidina presentan una afinidad para unir o quelar iones metálicos inmovilizados. Por consiguiente, el diseño de un polipéptido de identificación con una secuencia quelante de metal compuesta por residuos de histidina múltiples o alternados en el dominio espaciador o que flanqueen cada lado de la secuencia de dominios antigénicos permitiría que el polipéptido híbrido se uniese a un ión metálico inmovilizado en una resina u otra matriz. En una realización preferida, una secuencia quelante de metal que flanquee a las múltiples copias del dominio antigénico o en un dominio espaciador puede comprender por lo menos un residuo de histidina, por lo menos un residuo de glicina o una combinación de residuos de histidina alternados o múltiples de la fórmula: $-(\text{His-X})_m-$, en la que m es de 1 a 6 y X se selecciona del grupo consistente en Gly, His, Tyr, Trp, Val, Leu, Ser, Lys, Phe, Met, Ala, Glu, Ile, Thr, Asp, Asn, Gln, Arg, Cys, y Pro, la cual se puede usar en técnicas de purificación por afinidad usando una resina metal de unión a Ni^{2+} . Véanse, por ejemplo, las patentes U. S. n.º 4.569.794, 5.310.663, 5.284.933 y 5.594.115. Preferentemente, los aminoácidos del dominio espaciador no incluyen una segunda copia del sitio escindible según se describe en el presente documento. Una vez que el polipéptido híbrido se ha unido a la resina metal, el polipéptido híbrido se puede liberar por protonación de su ligando de unión a iones metálicos, asociado. La disociación se logra reduciendo el pH del medio tampón circundante, un método común conocido en la técnica para eluir proteínas unidas.

En una realización de la presente invención, el polipéptido de identificación comprende múltiples copias de un dominio antigénico en correspondencia en general con la secuencia peptídica FLAG[®] unida a una secuencia de unión que contiene un único sitio de escisión de enteroquinasa. Dicho polipéptido de identificación en general se corresponde con la secuencia:



en la que:

D, Y y K son sus aminoácidos representativos;

X^{20} y X^{21} son de forma independiente un hidrógeno o un enlace;

cada X^1 y X^4 es de forma independiente un enlace o por lo menos un residuo de aminoácido, preferentemente, si no es un enlace, por lo menos un residuo de aminoácido seleccionado de grupo consistente en residuos de aminoácidos aromáticos y residuos de aminoácidos hidrófilos, más preferentemente por lo menos un residuo de aminoácido hidrófilo, y todavía más preferentemente por lo menos un residuo de aspartato;

cada X^2 , X^3 , X^7 y X^8 es de forma independiente un residuo de aminoácido, preferentemente un residuo de aminoácido seleccionado del grupo consistente en residuos de aminoácidos aromáticos y residuos de aminoácidos hidrófilos, más preferentemente un residuo de aminoácido hidrófilo, y todavía más preferentemente un residuo de aspartato;

X^5 es un enlace o un dominio espaciador que comprende por lo menos un aminoácido, preferentemente, si no es un enlace, un residuo de histidina, un residuo de glicina o una combinación de residuos de histidina múltiples o alternados, comprendiendo dicha combinación His-Gly-His, ó $-(\text{His-X})_m-$, en la que m es de 1 a 6 y X se selecciona del grupo consistente en Gly, His, Tyr, Trp, Val, Leu, Ser, Lys, Phe, Ala, Glu, Ile, Thr, Asp, Asn, Gln, Arg, Cys, y Pro;

X^9 es un enlace ó D; y

n es por lo menos 2.

En esta realización, la secuencia de aminoácidos $\text{X}^{20}-(\text{X}^1-\text{Y}-\text{K}-\text{X}^2-\text{X}^3-\text{D}-\text{X}^4)_n$ representa las múltiples copias del dominio antigénico $-\text{X}^1-\text{Y}-\text{K}-\text{X}^2-\text{X}^3-\text{D}-$ unidas en tándem que se unen a una secuencia de unión $(\text{X}^1-\text{Y}-\text{K}-\text{X}^7-\text{X}^8-\text{D}-\text{X}^9-\text{K})$. Los dominios antigénicos pueden ser inmediatamente adyacentes entre sí cuando X^4 es un enlace, opcionalmente,

ES 2 321 058 T3

X^4 puede ser un dominio espaciador interpuesto entre las múltiples copias de dominios antigénicos. La secuencia de unión contiene un único sitio escindible de enteroquinasa que se representa mediante la secuencia $-X^7-X^8-D-X^9-K$, en la X^7 y X^8 pueden ser un residuo de aminoácido o un enlace y X^9 es un enlace o un residuo de aspartato. En una realización preferida, cada X^7 , X^8 y X^9 es de forma independiente un residuo de aspartato, dando como resultado de este modo el sitio escindible de enteroquinasa DDDDK (ID SEC n.º: 3) que está situado preferentemente inmediatamente adyacente al extremo amino del péptido diana. Las múltiples copias de dominios antigénicos pueden ser inmediatamente adyacentes a la secuencia de unión cuando X^5 es un enlace, opcionalmente, X^5 puede ser un dominio espaciador interpuesto entre la secuencia de unión y los dominios antigénicos. Cuando cada X^4 y X^5 es de forma independiente un dominio espaciador, se prefiere que el(los) residuo(s) de aminoácidos de cada X^4 y X^5 comuniquen una o más propiedades deseadas al polipéptido de identificación; por ejemplo, los aminoácidos del dominio espaciador se pueden seleccionar para comunicar un plegamiento deseado al polipéptido de identificación, incrementando de este modo la accesibilidad al anticuerpo. En otra realización preferida, los aminoácidos del dominio espaciador X^4 y X^5 se pueden seleccionar para comunicar una característica de afinidad deseada tal como una combinación de residuos de histidina múltiples o alternados capaces de quelarse a un ión metálico inmovilizado en una resina u otra matriz. Además, estas propiedades deseadas pueden estar diseñadas en otras áreas del polipéptido de identificación; por ejemplo, los aminoácidos representados por X^2 y X^3 se pueden seleccionar para comunicar un plegamiento peptídico deseado o una característica de afinidad deseada para su uso en la purificación por afinidad.

Cuando el polipéptido de identificación está situado en el extremo amino del péptido diana, es deseable diseñar la secuencia de aminoácidos del polipéptido de identificación de tal manera que esté presente un iniciador metionina. Por consiguiente, en una realización preferida de la presente invención, el polipéptido de identificación comprende múltiples copias de un dominio antigénico, una secuencia de unión que contiene un único sitio de escisión de enteroquinasa y en general se corresponde con la secuencia:



en la que:

D, Y, y K son sus aminoácidos representativos;

X^{20} y X^{21} son de forma independiente un hidrógeno o un enlace;

X^{10} es un enlace o un aminoácido, preferentemente, si no es un enlace, un residuo de metionina;

cada X^2 , X^3 , X^7 y X^8 es de forma independiente un residuo de aminoácido, preferentemente un residuo de aminoácido seleccionado del grupo consistente en residuos de aminoácidos aromáticos y residuos de aminoácidos hidrófilos, más preferentemente un residuo de aminoácido hidrófilo, y todavía más preferentemente un residuo de aspartato;

X^5 es un enlace o un dominio espaciador que comprende por lo menos un aminoácido, preferentemente, si no es un enlace, un residuo de histidina, un residuo de glicina o una combinación de residuos de histidina múltiples o alternados, comprendiendo dicha combinación His-Gly-His, ó $-(His-X)_m-$, en la que m es de 1 a 6 y X se selecciona del grupo consistente en Gly, His, Tyr, Trp, Val, -Leu, -Ser, -Lys, -Phe, -Met, Ala, Glu, Ile, Thr, Asp, Asn, Gln, Arg, Cys, y Pro;

X^9 es un enlace ó un residuo de aspartato; y

n es por lo menos 2.

En esta realización, la secuencia de aminoácidos $X^{20}-(D-Y-K-X^2-X^3-D)_n-$ representa las múltiples copias del dominio antigénico D-Y-K- X^2 - X^3 -D en tándem que está flanqueado por una secuencia de unión (D-Y-K- X^7 - X^8 -D- X^9 -K) y un iniciador aminoácido X^{10} , preferentemente metionina. El dominio antigénico D-Y-K- X^2 - X^3 -D con un iniciador metionina es reconocido por el anticuerpo M5. En esta realización, un dominio antigénico es inmediatamente adyacente a otro dominio antigénico, es decir, no hay dominios espaciadores interpuestos, y las múltiples copias del dominio antigénico son inmediatamente adyacentes a la secuencia de unión cuando X^5 es un enlace. La secuencia de unión contiene un sitio escindible de enteroquinasa que está representado por la secuencia de aminoácidos X^7 - X^8 -D- X^9 -K, en la que X^7 y X^8 pueden ser un enlace o un residuo de aminoácido, preferentemente un residuo de aspartato, y X^9 es un enlace o un residuo de aspartato. En una realización preferida, cada X^7 , X^8 y X^9 es de forma independiente un residuo de aspartato, dando como resultado de este modo el sitio escindible de enteroquinasa DDDDK (ID SEC n.º: 3) que es preferentemente adyacente al extremo amino del péptido diana. Opcionalmente, las múltiples copias del dominio antigénico están unidas a la secuencia de unión por un dominio espaciador X^5 cuando X^5 es por lo menos un residuo de aminoácido. Cuando X^5 es un dominio espaciador, se prefiere que el(los) residuo(s) de aminoácidos de X^5 comuniquen una o más propiedades deseadas al polipéptido de identificación; por ejemplo, los aminoácidos del dominio espaciador se pueden seleccionar para comunicar un plegamiento deseado al polipéptido de identificación, incrementando de este modo la accesibilidad al anticuerpo. En otra realización preferida, los aminoácidos del dominio espaciador se pueden seleccionar para comunicar una característica de afinidad deseada tal como una combinación de residuos de histidina múltiples o alternados capaces de quelarse a un ión metálico inmovilizado en una resina u

ES 2 321 058 T3

otra matriz. Además, estas propiedades deseadas pueden estar diseñadas en otras áreas del polipéptido de identificación; por ejemplo, los aminoácidos representados por X^2 y X^3 se pueden seleccionar para comunicar un plegamiento peptídico deseado o una característica de afinidad deseada para su uso en la purificación por afinidad.

5 En otra realización de la presente invención, el polipéptido de identificación comprende múltiples copias de una secuencia antigénica, una secuencia de unión que contiene un único sitio escindible de enteroquinasa y en general se corresponde con la secuencia:



en la que:

15 D, Y, y K son sus aminoácidos representativos;

X^{20} y X^{21} son de forma independiente un hidrógeno o un enlace;

cada X^{11} es un enlace o un aminoácido, preferentemente L;

20 cada X^{12} es un aminoácido, preferentemente seleccionado de grupo consistente en residuos de aminoácidos aromáticos y residuos de aminoácidos hidrófilos, más preferentemente un residuo de aminoácido hidrófilo, y todavía más preferentemente un residuo de aspartato;

25 cada X^{13} es un enlace o por lo menos un aminoácido, preferentemente, si no es un enlace, seleccionado del grupo consistente en residuos de aminoácidos aromáticos y residuos de aminoácidos hidrófilos, más preferentemente un residuo de aminoácido hidrófilo, y todavía más preferentemente un residuo de aspartato;

30 X^{14} es un enlace o un dominio espaciador que comprende por lo menos un aminoácido, preferentemente, si no es un enlace, un residuo de histidina, un residuo de glicina o una combinación de residuos de histidina múltiples o alternados, comprendiendo dicha combinación His-Gly-His, ó $-(\text{His-X})_m-$, en la que m es de 1 a 6 y X se selecciona del grupo consistente en Gly, His, Tyr, Trp, Val, Leu, Ser, Lys, Phe, Ala, Glu, Ile, Thr, Asp, Asn, Gln, Arg, Cys, y Pro;

X^{15} es un enlace ó un residuo de aspartato; y

35 n es por lo menos 2.

40 En esta realización, la secuencia de aminoácidos $X^{20}-(D-X^{11}-Y-X^{12}-X^{13})_n$ representa las múltiples copias del dominio antigénico $D-X^{11}-Y-X^{12}-X^{13}$ en tándem que están unidas a una secuencia de unión $(D-X^{11}-Y-X^{12}-X^{13}-D-X^{15}-K)$. Adicionalmente, un dominio antigénico es inmediatamente adyacente a otro dominio antigénico, es decir, no hay dominios espaciadores interpuestos, y las múltiples copias del dominio antigénico son inmediatamente adyacentes a la secuencia de unión cuando X^{14} es un enlace. La secuencia de unión contiene un único sitio escindible de enteroquinasa que está representado por la secuencia $-X^{12}-X^{13}-D-X^{15}-K$ en la que X^{12} y X^{13} pueden ser un enlace o un residuo de aminoácido, preferentemente un residuo de aspartato, y X^{15} es un enlace o un residuo de aspartato. En una realización preferida, cada X^{12} , X^{13} y X^{15} es de forma independiente un residuo de aspartato, dando como resultado de este modo el sitio escindible de enteroquinasa DDDDK (ID SEC n.º: 3) que es preferentemente adyacente al extremo amino del péptido diana. Opcionalmente, las múltiples copias del dominio antigénico están unidas a la secuencia de unión por un espaciador X^{14} cuando X^{14} es por lo menos un residuo de aminoácido. Cuando X^{14} es un dominio espaciador, se prefiere que el(los) residuo(s) de aminoácidos de X^{14} comuniquen una o más propiedades deseadas al polipéptido de identificación; por ejemplo, los aminoácidos del dominio espaciador se pueden seleccionar para comunicar un plegamiento deseado al polipéptido de identificación, incrementando de este modo la accesibilidad al anticuerpo. En otra realización preferida, los aminoácidos del dominio espaciador X^{14} se pueden seleccionar para comunicar una característica de afinidad deseada tal como una combinación de residuos de histidina múltiples o alternados capaces de quelarse a un ión metálico inmovilizado en una resina u otra matriz.

55 Péptido diana

60 Según la presente invención, el péptido diana puede estar compuesto por cualquier sustancia proteica que se pueda expresar en células hospedadoras transformadas. Por consiguiente, la presente invención se puede utilizar ventajosamente para producir sustancialmente cualquier proteína procariota o eucariota, simple o conjugada, que se pueda expresar mediante un vector en una célula hospedadora transformada. Dichas proteínas incluyen enzimas, ya sean oxidoreductasas, transferasas, hidrolasas, liasas, isomerasas o ligasas.

65 La presente invención contempla también la producción de proteínas de almacenamiento, tales como ferritina u ovoalbúmina o proteínas transportadoras como hemoglobina, albúmina sérica o ceruloplasmina. Se incluyen también los tipos de proteínas que funcionan en sistemas contráctiles y móviles, por ejemplo, actina y miosina.

La presente invención contempla también la producción de proteínas que prestan servicio a una función protectora o de defensa, tales como la proteína sanguínea fibrinógeno. Otras proteínas protectoras incluyen las proteínas de unión, tales como anticuerpos o inmunoglobulinas que se unen a antígenos, y por lo tanto los neutralizan.

La proteína producida por la presente invención también puede abarcar varias hormonas tales como la Hormona del Crecimiento Humano, la somatostatina, la prolactina, la estrona, la progesterona, el melanocito, la tirotropina, la calcitonina, la gonadotropina y la insulina. Otras hormonas de este tipo incluyen aquellas que se hayan identificado como involucradas en el sistema inmune, tales como interleuquina 1, interleuquina 2, factor estimulante de colonias, factor activador de macrófagos e interferón.

La presente invención es también aplicable a la producción de proteínas tóxicas, tales como ricina de la semilla de ricino o gossipina de linaza de algodón.

Las proteínas que actúan como elementos estructurales pueden ser producidas por la presente invención; dichas proteínas incluyen las proteínas fibrosas colágeno, elastina y α -queratina. Otras proteínas estructurales incluyen glicoproteínas, proteínas víricas y muco-proteínas.

Además de las proteínas de origen natural antes indicadas, la presente invención se puede utilizar para producir proteínas sintéticas definidas en general como cualquier secuencia de aminoácidos que no se produzca en la naturaleza.

A partir de una variedad de fuentes procariotas o eucariotas, tales como células vegetales o animales o células bacterianas, se pueden obtener genes que codifican los diversos tipos de moléculas de proteínas antes identificadas. Los genes se pueden aislar a partir del material cromosómico de estas células o a partir de plásmidos de células procariotas utilizando técnicas normalizadas, bien conocidas, una variedad de plásmidos de origen natural y sintetizados que tienen genes que codifican muchas moléculas de proteínas diferentes no están disponibles comercialmente a partir de una variedad de fuentes. El ADN deseado también se puede producir a partir de ARNm usando la enzima transcriptasa inversa. Esta enzima permite la síntesis de ADN a partir de un molde de ARN.

Preparación de vectores de expresión de ADN

Según la presente invención, una vez que se ha aislado, sintetizado u obtenido de otra manera un gen que codifica un péptido diana, el mismo se une a un fragmento de ADN sintético que codifica el polipéptido de identificación.

El gen de polipéptido de identificación se puede sintetizar mediante técnicas bien conocidas. Para una composición seleccionada del polipéptido de identificación, se pueden sintetizar oligómeros de ADN que codifican los aminoácidos deseados del polipéptido de identificación usando un sintetizador de ADN automatizado, disponible comercialmente, según una manera bien conocida en la técnica. Las técnicas y aparatos para sintetizar ADN son comunes y conocidos en la materia; de este modo, la descripción y los detalles para realizar esto no se expondrán de forma completa en el presente documento. Esencialmente, este proceso implica la obtención de pares de oligonucleótidos sintéticos y su digestión con las endonucleasas de restricción apropiadas. Esto producirá la secuencia de nucleótidos correcta que codifica el polipéptido etiqueta de identificación. Después de la digestión, se forman varios fragmentos de ADN con extremos cohesivos o "pegajosos". Aunque puede haber muchas maneras según las cuales realizar dicha construcción, la realización preferida implica la generación de una secuencia de múltiples epítomos FLAG[®] o variaciones de la misma en tándem.

El par de oligonucleótidos usado en la construcción del polipéptido de identificación puede ser de origen natural o generado sintéticamente. En general se prefiere que los pares específicos de oligonucleótidos se hayan generado sintéticamente para producir la secuencia de aminoácidos del polipéptido etiqueta de identificación deseado. Las cadenas de cada oligonucleótido se aparean entre sí y se digieren con una endonucleasa de restricción apropiada tal como *EcoR I* y *Hind III*. Después de la digestión y de la creación de las casetes de nucleótidos, las secuencias se pueden verificar a través de secuenciación de ADN.

Tal como se describe posteriormente, los oligómeros de ADN sintéticos que codifican el polipéptido de identificación se pueden ligar a una secuencia de ADN que codifique la proteína deseada, y a continuación los fragmentos de ADN combinados se ligan a un vector de expresión apropiado para formar un vehículo de clonación para la transformación a una célula hospedadora apropiada.

Además del gen del péptido diana y el gen del polipéptido de identificación, si fuera necesario, el fragmento de ADN híbrido puede incluir un sitio de unión al ribosoma para obtener una traducción de proteínas de alto nivel en una célula hospedadora, un codón de iniciación de la traducción (ATG), y un promotor.

En general, los genes que codifican el péptido diana y el polipéptido de identificación se tratan idealmente con una enzima de restricción apropiada o, alternativamente, se manipulan para tener extremos cohesivos que faciliten la ligación mutua y con un plásmido u otro tipo de vector de clonación. El vector de clonación se digiere preferentemente con la misma endonucleasa de restricción usada para acondicionar los genes foráneos con el fin de formar extremos cohesivos complementarios, (es decir, "extremos pegajosos") antes de la ligación con los genes foráneos. Alternativamente, el uso de ciertas enzimas de restricción (por ejemplo, *Pvu II*, *Bal I*) puede dar como resultado la formación de extremos sin secuencias sobresalientes complementarias, a los que se hace referencia comúnmente como

ES 2 321 058 T3

“extremos cuadrados” o “romos”. Los extremos cuadrados del plásmido se pueden unir a los genes foráneos con una ligasa apropiada. Adicionalmente, se pueden usar varias técnicas para manipular los ácidos nucleicos de los extremos romos con el fin de formar extremos cohesivos, por ejemplo, se pueden usar moléculas ligadoras para adicionar bases de nucleótidos o se pueden usar enzimas apropiadas para eliminar bases de nucleótidos de los extremos nivelados. Los métodos y materiales para lograr esto son bien conocidos en la materia.

La PCR es también una herramienta efectiva para clonar genes conocidos (en sitios romos o pegajosos). Los cebadores pueden codificar entre 25 y 40 bases de secuencia conocida y el producto de la PCR resultante se puede clonar en un vector digerido que tenga extremos romos eliminando cualquier posible 3' que sobresalga con ADN polimerasa T4. Otro método de unir secuencias con el uso de la reacción PCR es crear sitios de restricción en el(los) extremo(s) del ADN amplificado. Estos sitios de restricción se adicionan fácilmente a los extremos 5' de los cebadores usados para la amplificación. La digestión de los productos PCR purificados producirá extremos para la ligación a otro ADN que tenga extremos compatibles.

Debe apreciarse que la digestión del plásmido seleccionado con una(s) endonucleasa(s) de restricción puede dar como resultado la formación de dos o más segmentos de ADN lineales. El segmento a usar para formar el vector de clonación, es decir, el segmento que tiene el gen de identidad fenotípico, el replicón y los otros componentes deseados, se puede identificar mediante técnicas bien conocidas, tales como mediante electroforesis en gel.

El vector de clonación resultante se usa para transformar un microorganismo hospedador. Los transformados se aíslan y se analizan en relación con la presencia de los genes foráneos y en relación con la orientación correcta de los genes dentro del vector. A continuación, los transformados se multiplican en cultivo para conseguir la replicación del vector y la expresión de alto nivel del polipéptido híbrido que se está buscando. Adicionalmente, los vectores de clonación se pueden usar para transformar otras cepas del hospedador seleccionado u otros tipos de hospedadores para una producción a gran escala del polipéptido heterólogo híbrido. Old y Primrose, *Principles of Gene Manipulation*, (2ª Ed. 1981) describen varios procedimientos y materiales para preparar vectores recombinantes, transformar células hospedadoras con los vectores, replicar el vector y expresar el polipéptido y proteínas.

Para llevar a cabo la presente invención, se pueden utilizar varios vectores de clonación. Aunque se prefiere el uso de un plásmido, el vector puede ser un bacteriófago o cósmido. Si la clonación tiene lugar en células de mamíferos o vegetales, como vectores se pueden usar virus. Si se utiliza un plásmido, el mismo se puede obtener a partir de una fuente natural o se puede sintetizar artificialmente. El plásmido particular seleccionado debería ser compatible con las células particulares que actúan como hospedador, ya sean una bacteria tal como *Escherichia coli* (*E. coli*), levadura, u otro microorganismo unicelular. El plásmido debería tener el origen de replicación (replicón) correcto para la célula hospedadora particular seleccionada.

Adicionalmente, el tamaño del plásmido debe ser suficiente para alojar los genes híbridos que codifican tanto el péptido diana como el polipéptido de identificación, aunque también con un peso molecular lo más bajo posible. Los plásmidos de bajo peso molecular son más resistentes a daños por cizalladura y se aíslan más fácilmente con respecto a células hospedadoras. Si se obtienen a partir de fuentes naturales, los mismos habitualmente están presentes en forma de múltiples copias, facilitando de este modo su aislamiento. Además, existe menos probabilidad de que un plásmido de bajo peso molecular tenga sitios de sustratos múltiples para endonucleasas de restricción.

Otro requisito para un plásmido vector de clonación es la presencia de sitios de restricción de manera que el adecuado de entre los enzimas de restricción pueda escindir el plásmido para su ligación subsiguiente con los genes foráneos sin provocar inactivación del replicón. Con este fin, resultaría útil que el plásmido tuviera sitios de sustrato único para un gran número de endonucleasas de restricción.

Tal como se ha indicado anteriormente, puede haber dominios espaciadores de aminoácidos interpuestos entre los dominios antigénicos múltiples del polipéptido de identificación. Variando la secuencia de ADN triplete que representa aminoácidos específicos (es decir, codones) en el diseño de estos dominios espaciadores, es posible crear múltiples sitios enzimáticos de restricción para enzimas que reconocen y escinden dichas secuencias diseñadas sin cambiar la secuencia de aminoácidos del polipéptido de identificación codificado. Se prefiere el uso de secuencias que codifican sitios de reconocimiento para enzimas de restricción que tienen un mínimo de 6 bases en el sitio de reconocimiento, reduciendo de este modo la opción de que haya presentes múltiples sitios de restricción tanto en el vector de ADN como en las secuencias de ADN que codifican el péptido diana.

De forma similar, se usa una secuencia de unión para unir las secuencias de ADN que codifican el péptido diana a las secuencias de ADN que codifican los múltiples dominios antigénicos del polipéptido de identificación. Variando la secuencia de ADN triplete que representa aminoácidos específicos en el diseño de la secuencia de unión, es posible crear sitios de restricción para enzimas que reconocen y escinden dichas secuencias diseñadas sin cambiar la secuencia de aminoácidos del polipéptido de identificación codificado. Se prefiere el uso de secuencias que codifican sitios de reconocimiento para enzimas de restricción que tienen un mínimo de seis bases en el sitio de reconocimiento; esto reduce la opción de que haya presentes múltiples sitios escindibles de enzimas de restricción tanto en el vector como en las secuencias que codifican el péptido diana. Por otra parte, el plásmido debería tener una propiedad fenotípica que permita que las células hospedadoras transformadas se identifiquen y separen fácilmente de células que no experimentan transformación. Dichos genes de selección fenotípica pueden incluir genes que proporcionan resistencia a una sustancia inhibidora del crecimiento, tal como un antibiótico. Los plásmidos que incluyen genes resistentes a

varios antibióticos, tales como tetraciclina, estreptomina, sulfamidas, penicilina y ampicilina, no están disponibles de forma generalizada. Cuando se cultivan células hospedadoras en un medio que contiene uno de estos antibióticos, sobrevivirán solamente los transformados que presenten el gen apropiado de resistencia a antibióticos.

5 En lugar de utilizar una resistencia génica a un compuesto de inhibición del crecimiento para identificar células hospedadoras transformadas, los genes de selección fenotípica también pueden incluir aquellos que proporcionan un factor de crecimiento para permitir que las células transformadas se propaguen en un medio que carezca del factor de crecimiento necesario para las células hospedadoras. Por ejemplo, para auxótrofos de levadura, dichos factores de crecimiento incluyen triptófano o leucina.

10 Alternativamente, se prefiere que una secuencia de ADN que codifica un péptido señal se una a las secuencias que codifican el polipéptido de identificación y el péptido diana. El uso de una secuencia señal secretada permitirá también que las células hospedadoras transformadas sean identificadas y separadas fácilmente con respecto a células que no experimentan transformación. Las señales de secreción son relativamente cortas en la mayoría de especies, compuestas generalmente por entre 16 y 40 aminoácidos. Adicionalmente, las secuencias señales de genes bacterianos o eucariotas se conservan altamente en términos de función. Aunque las secuencias de ADN que codifican estos péptidos señal no se conservan altamente, se ha demostrado que muchas de estas secuencias señal son intercambiables. Véase Grey, G. L. *et al.*, Gene 39:247 (1985).

20 *Transformación del plásmido recombinante*

Una vez que se ha construido un vector de ADN adecuado que codifica el polipéptido híbrido deseado, el vector se introduce en la célula hospedadora deseada. Aunque la célula hospedadora puede ser cualquier célula procariota o eucariota adecuada, preferentemente es una bacteria bien definida, tal como *E. coli* o una cepa de levadura. Ambos hospedadores mencionados se transforman fácilmente y tienen capacidad de crecimiento rápido en cultivos de fermentación. En lugar de *E. coli*, se pueden utilizar otros microorganismos unicelulares, por ejemplo, hongos y algas. Adicionalmente, el *E. coli* se puede sustituir por otras formas de bacterias tales como salmonela o neumococo. Sea cual sea el hospedador seleccionado, debería ser uno que no contenga ningún enzima de restricción que escinda al plásmido recombinante y que disponga de las vías bioquímicas necesarias para la expresión fenotípica y otras funciones para una expresión correcta del polipéptido híbrido.

Las moléculas de ADN se transfectan en hospedadores procariotas y eucariotas usando protocolos normalizados conocidos en la técnica. De forma breve, se hace que las células hospedadoras procariotas sean competentes mediante tratamiento con soluciones de cloruro de calcio (las células bacterianas competentes están disponibles comercialmente y se elaboran fácilmente en el laboratorio). Este tratamiento permite la captación de ADN por la célula bacteriana. Otros medios de introducción del ADN en células bacterianas es la electroporación en la que se usa un impulso eléctrico para permitir la captación de ADN por células bacterianas. De forma similar, protocolos normalizados tales como coprecipitación de ADN con fosfato cálcico, transfección mediada con DEAE-dextrano, electroporación, microinyección, lipofección, fusión de protoplastos, infección retroviral, bombardeo de partículas (por ejemplo, biolística) se usan comúnmente para la introducción de moléculas de ADN en hospedadores eucariotas, incluyendo levadura y eucariotas superiores.

En los protocolos de transformación, únicamente se transforma realmente una pequeña parte de las células hospedadoras, debido a la captación limitada del plásmido por las células. De este modo, antes de aislar los transformados, las células hospedadoras usadas en el protocolo de transformación típicamente se multiplican en un medio apropiado. Las células que realmente sean transformados se pueden identificar colocando el cultivo original en placas de agar que contengan un medio de crecimiento adecuado que contenga el identificador fenotípico, tal como un antibiótico. Únicamente sobrevivirán aquellas células que tengan el gen resistente adecuado. Las células de las colonias que sobrevivan se puede lisar y a continuación el plásmido se puede aislar del lisado. El plásmido así aislado se puede caracterizar para determinar si los genes cointegrados están ligados en la orientación correcta, mediante digestión con endonucleasas de restricción y una electroforesis en gel subsiguiente o también con otros métodos normalizados.

Una vez que se han identificado las células transformadas, las mismas se pueden multiplicar mediante técnicas establecidas, tales como por fermentación. Adicionalmente, los plásmidos recombinantes clonados recuperados se pueden usar para transformar otras cepas de bacterias u otros tipos de células hospedadoras para una replicación y expresión a gran escala del polipéptido híbrido.

Purificación del polipéptido híbrido

60 Las moléculas del polipéptido híbrido expresadas por las células hospedadoras transformadas se separan del medio de cultivo, otro material celular, etcétera, preferentemente mediante un proceso de cromatografía de afinidad. Con este fin, se deben generar anticuerpos contra los dominios antigénicos del polipéptido de identificación del polipéptido híbrido para su uso en una matriz de columna. Para producir dichos anticuerpos, en primer lugar el polipéptido de identificación se sintetiza y a continuación se usa para inmunizar un animal adecuado para la producción de un anticuerpo contra el polipéptido de identificación. Dichos métodos para la producción de anticuerpos se dan a conocer en la patente U. S. n.º 4.851.341. El anticuerpo se puede identificar mediante un ensayo de inmunoabsorción ligado a enzimas (Elisa) u otro ensayo apropiado. A continuación, se puede producir un monoclonal mediante técnicas de hibridoma. Los anticuerpos preferidos son los anticuerpos monoclonales FLAG[®] M1, M2 y M5. Después de la

purificación, el anticuerpo o anticuerpos se unen a la matriz de columna y a continuación un extracto de las células hospedadoras transformadas se aplica a la columna para aislar el polipéptido híbrido. El polipéptido híbrido se eluye de la columna, por ejemplo, por competición con respecto al polipéptido de identificación libre.

5 Adicionalmente, si el polipéptido de identificación contiene histidina, glicina o combinaciones de residuos de histidina múltiple o alternada, se puede usar la Cromatografía de Afinidad con Iones Metálicos Inmovilizados (IMAC) como método alternativo para aislar y purificar péptidos diana. Cuando se produce un polipéptido híbrido que contiene el péptido diana y el polipéptido de identificación, y el mismo se hace pasar a través de una columna que contiene iones metálicos inmovilizados, el polipéptido híbrido se quelará los iones metálicos inmovilizados. El polipéptido híbrido debería quelarse a los iones metálicos inmovilizados durante una cantidad de tiempo suficiente como para permitir su separación de otros materiales. Una vez que el polipéptido híbrido está unido a la resina con iones metálicos, el polipéptido híbrido se puede liberar mediante protonación de su ligando asociado de unión a iones metálicos. La disociación se logra reduciendo el pH del medio tampón circundante, un método común conocido en la técnica para eluir proteínas unidas. A continuación, el péptido diana se puede escindir del polipéptido de identificación tal como se describe adicionalmente en el presente documento.

Se pueden usar otros métodos para detectar, monitorizar o aislar péptidos diana. Dichos métodos incluyen la inmunoprecipitación y el Western blotting tal como se describe en "Principles and Practice of Immunoassay", Price and Newman, eds., Stochton Press, 1991. El uso de la inmunoprecipitación como técnica sensible y específica para detectar y cuantificar antígeno diana en mezclas de proteínas es conocido para los expertos en la materia. Véase Molecular Cloning, A Laboratory Manual, 2ª Edición, Maniatis, T. *et al.* eds. (1989) Cold Spring Harbor Press. Resumiendo, se pueden usar anticuerpos, preferentemente anticuerpos monoclonales FLAG[®], M1, M2 ó M5 capaces de unirse a los dominios antigénicos del polipéptido de identificación para detectar las proteínas usando ensayos de inmunoprecipitación. Tal como se ha descrito anteriormente, las células se transforman con el polipéptido de identificación, cultivadas en medios de cultivo, y lisadas para obtener una solución de material proteico etiquetado producido por las células. Esta solución se incuba con una solución de anticuerpos monoclonales, y cualquier complejo entre la proteína marcada con el polipéptido de identificación formada en la célula y los anticuerpo se determina por precipitación. El complejo proteína/anticuerpo a continuación se puede aislar del precipitado. Seguidamente, la presencia de la proteína marcada se confirma mediante métodos analíticos habituales, por ejemplo, electroforesis en gel de poliacrilamida con SDS con fluorografía, en condiciones que disocian el complejo proteína/anticuerpo.

Adicionalmente, el Western blotting es otra técnica de inmunoensayo usada para detectar el péptido diana. En general, pequeñas cantidades de un péptido diana se someten a electroforesis sobre un gel de poliacrilamida y se transfieren (por blotting) a una lámina o membrana polimérica. A continuación, la membrana se incuba con un primer anticuerpo, preferentemente un anticuerpo monoclonal FLAG[®] que se puede unir a los dominios antigénicos del polipéptido de identificación. A continuación, la membrana que contiene el anticuerpo-antígeno se incuba con un segundo anticuerpo marcado específico para el primer anticuerpo. La proteína etiquetada con el polipéptido de identificación se puede detectar y visualizar mediante métodos conocidos tales como autorradiografía.

40 *Separación de proteína madura con respecto a moléculas de proteína/polipéptido de identificación híbridas, purificadas*

A no ser que se retire mientras está todavía unido a la columna o matriz de afinidad, el polipéptido de identificación se puede escindir de la molécula de proteína y la molécula de proteína se puede separar del polipéptido de identificación, dando como resultado de este modo una proteína purificada. Esto se logra suspendiendo en primer lugar las moléculas de proteína/polipéptido de identificación híbrido en tampón. Después de esto, a la suspensión se le adiciona la enzima proteolítica u otro agente proteolítico químico que sea específico para los residuos de aminoácidos que componen la parte de unión del polipéptido de identificación. La enzima se puede acoplar a una matriz de gel para evitar la contaminación de la solución del producto con la enzima. Tal como se ha descrito anteriormente, la enzima proteolítica o agente proteolítico químico escinde al polipéptido híbrido entre los residuos de aminoácidos adyacentes de la parte de unión del polipéptido de identificación y la molécula de proteína. También tal como se ha indicado anteriormente, como ejemplo no limitativo, los aminoácidos de unión pueden estar compuestos por la secuencia: Asp-Asp-Asp-Asp-Lys (ID SEC n.º: 3). Se sabe que esta secuencia particular de aminoácidos se produce solamente de forma natural en la proteína tripsinógeno, el sustrato para la enteroquinasa de mucosa bovina. De este modo, mediante el uso de esta secuencia de aminoácidos particular, es altamente improbable que la escisión por enzimas de las moléculas de proteína polipéptido de identificación híbrido provoque también la escisión de la propia molécula de proteína.

Después de la incubación, la proteína deseada se purifica de la manera siguiente. Si el agente proteolítico es una enzima fijada a una matriz de gel, la suspensión se centrifuga y el pellet (que contiene el conjugado enzima-gel) se descarta. El sobrenadante contiene solamente el producto proteico, el polipéptido de identificación escindido y posiblemente pequeñas cantidades de molécula péptido/proteína sin escindir, además de sales tampón. En el caso de agentes de escisión química, no habría etapa de centrifugación del gel, y la solución contendría un agente químico residual y subproductos de agente químico además del producto proteico, el polipéptido de identificación y pequeñas cantidades de molécula péptido/proteína sin escindir.

La mayoría de las sustancias contaminante antes mencionadas son mucho menores que el producto proteico y se pueden eliminar eficazmente por medios sencillos, tales como filtración en gel o diálisis. Después de dichas etapas, únicamente quedaría la molécula de proteína/polipéptido de identificación sin escindir para contaminar el producto

ES 2 321 058 T3

5 proteico. Para eliminar la molécula polipéptido/proteína del producto proteico, la mezcla se hace pasar sobre una segunda columna de afinidad, teniendo fijada a ella dicha columna el mismo anticuerpo específico para el polipéptido de identificación que el que se usó para la eliminación de la molécula péptido/proteína del medio de producción original. El anticuerpo une la molécula polipéptido/proteína no deseada, y el eluato de la columna contiene solamente la proteína de producto deseado, en este momento libre de todos los contaminantes.

10 Si se usa una enzima soluble para la escisión proteolítica, entonces el producto proteico puede contener pequeñas cantidades de la enzima, que se pueden eliminar haciendo pasar la solución sobre una columna de afinidad que contenga un sustrato inmovilizado para la enzima. De este modo, la enzima se une a la columna y las moléculas proteicas deseadas se permiten pasar a través de esta última.

15 Tal como se ha indicado anteriormente, algunos productos proteicos poseerán la actividad enzimática deseada con el polipéptido de identificación todavía fijado a ellos. Como consecuencia, no es necesario escindir el polipéptido de identificación de la molécula proteica, y por lo tanto no es necesario realizar las etapas antes descritas de escisión y de purificación subsiguiente.

20 Por otra parte, en situaciones en las que el polipéptido de identificación permanece fijado a la molécula proteica, la parte de unión del polipéptido de identificación no es necesaria. En su lugar, el polipéptido de identificación puede estar compuesto meramente por los dominios antigénicos. En esta situación, la construcción y el método de preparación de los vectores de expresión de ADN, antes detallados, se pueden modificar apropiadamente.

Los siguientes ejemplos están destinados a ilustrar, aunque no limitar, la presente invención.

25 Ejemplos

Ejemplo 1

30 Construcción del p3xFLAG-CMV-7

30 Materiales y métodos

Construcción del P3XFLAG-CMV-7

35 El P3XFLAG-CMV-7 se construyó a partir del vector de expresión de mamíferos, pCMV-5. La secuencia FLAG triple se construyó a partir de dos pares de oligonucleótidos complementarios.

El primer par de oligonucleótidos se sintetizó de la manera siguiente:

40 5'GAAGAATTCACCATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGAT3' (ID SEC n.º: 8)

y

45 5'ATCATGATCTTTATAATCACCGTCATGGTCTTTGTAGTCCATGGTGAATTCTTC3' (ID SEC n.º: 9).

El segundo par se sintetizó con la siguiente secuencia:

50 5'GAAGATATCGATTACAAGGATGACGATGACAAGCTTGGG3' (ID SEC n.º: 10)

y

55 5'CCCAAGCTTGTCATCGTCATCCTTGTAATCGATATCTTC3' (ID SEC n.º: 11).

60 El primer par de oligonucleótidos se apareó conjuntamente y se digirió con *EcoRI*. El segundo par de oligonucleótidos se apareó conjuntamente y se digirió con *EcoRV* y *HindIII*. Los dos pares de casetes de nucleótidos digeridos se ligaron en CMV-5, el cual se ha digerido doblemente con *EcoRI* y *HindIII*. La secuencia se verificó mediante secuenciación de ADN.

Construcción del pFLAG-CMV7-BAP

65 Una versión modificada del gen *F. coli pho A* para el que se suprimieron la secuencia líder y los cuatro aminoácidos N-terminales de la enzima madura se subclonó en el vector p3XFLAG-CMV-7. La secuencia modificada se cortó del pFLAG-ATS-BAP mediante doble digestión con *HindIII* y *BglII*. A continuación, el fragmento se clonó en p3XFLAG CMV-7 que se había digerido doblemente con *HindIII* y *BamHI* para generar p3XFLAG-CMV-7-BAP. La secuencia de nucleótidos en el extremo N de la región codificante *phoA* se verificó.

ES 2 321 058 T3

Construcción del FLAGATS-BAP triple

Se sintetizaron dos oligonucleótidos que codificaban la cadena sentido y antisentido para la secuencia FLAG triple, los mismos se fosforilaron en 5' con polinucleótido quinasa del T4, y se aparearon entre sí. El pFLAG-ATS-BAP se digirió con *Nde*I y *Hind*III y el vector se purificó por electroforesis en gel. El casete apareado se ligó al vector pFLAG-ATS-BAP doblemente digerido con ADN ligasa del T4 y la reacción se llevó a cabo durante la noche a 16°C durante 16 horas. La ligación se enriqueció mediante digestión con *Nru* I y a continuación se transformó en *E. coli* DH5 α . Se aislaron clones y los mismos se verificaron mediante secuenciación.

10 Resultados

Los solicitantes han construido un vector para la expresión de proteínas en células hospedadoras de mamíferos usando una versión modificada del sistema de expresión FLAG, que contiene secuencias 3XFLAG en tándem (Figura 1). Este constructo se diseñó para mejorar el límite de detección de proteínas expresadas en células hospedadoras de mamíferos. Los primeros dos péptidos flag son secuencias FLAG modificadas. El epítipo FLAG[®] original es Asp-Tyr-Lys-Asp-Asp-Asp-Lys (ID SEC n.º: 1) mientras que las dos primeras secuencias de reconocimiento flag son es Asp-Tyr-Lys-Asp-His-Asp (ID SEC n.º: 12) con el dominio espaciador o bien Gly ó Ile entre las dos secuencias. Estas secuencias alternativas surgen de los estudios de presentación en fagos (*phage display*) en los que se determinó un motivo de unión diferente. Véase Miceli *et al.*, J. Immunological Methods 167:279-287 (1994). Esto permite la introducción de sitios de unión a anticuerpos FLAG adicionales sin la adición de sitios de reconocimiento/escindibles de enterocinasa extras.

El vector de expresión p3XFLAG-CMV-7 contiene la región promotora del citomegalovirus humano necesaria para la expresión constitutiva de genes clonados en muchas líneas celulares de mamíferos. En el vector se proporciona la secuencia de consenso Kozak junto con un sitio de clonación múltiple, lo cual permite una variedad de estrategias de clonación. El sitio de clonación múltiple es compatible con los otros vectores de expresión de mamíferos CMV existentes. Adicionalmente, el vector de expresión contiene el origen de replicación del SV40 para una expresión transitoria eficaz de alto nivel y un segmento de ADN de la hormona del crecimiento humano que contiene una secuencia de terminación de transcripción y señales de poliadenilación. El p3XFLAG-CMV-7 contiene el gen de la β -lactamasa para la selección del plásmido en *E. coli*.

Ejemplo 2

35 Expresión bacteriana del p3XFLAG-ATS-BAP y purificación de proteínas

Materiales y métodos

E. coli BL21 (DE3) se transformaron con el plásmido de expresión que contenía el constructo BAP FLAG triple realizado según los métodos del Ejemplo 1. Se cultivaron células en *terrific broth* que contenía 100 μ g/ml de ampicilina a 37°C con agitación. El cultivo se desarrolló a un OD₆₀₀=4,0 y a continuación se indujo con IPTG a una concentración final de 1 mM. El cultivo celular se desarrolló durante 3 horas adicionales a 37°C y a continuación se recolectó mediante centrifugación. El pellet celular se resuspendió en Tris-HCl 50 mM pH 8,0 y las células se disgregaron mediante sonicación y los residuos celulares se eliminaron por centrifugación. El sobrenadante se aplicó a gel de afinidad a M2 y se equilibró con Tris-HCl 50 mM pH 8,0, NaCl (TBS) 150 mM. La resina se lavó con 20 volúmenes de lecho de TBS y a continuación el BAP FLAG triple se eluyó con cinco volúmenes de columna de Glicina 0,1 M pH 3,5. La proteína eluída se agrupó y se ajustó a un pH 7,5 con Tris-HCl 1,0 M pH 8,0. El contenido de proteínas se determinó mediante técnicas tanto Bradford como por absorbancia usando ϵ_{280} =0,7 ml/mg.

50 Western blot

3XFLAG-BAP y N-BAP purificados se diluyeron con tampón Laemlli 2X, se hirvieron durante cinco minutos y a continuación se colocaron en hielo. Las muestras se resolvieron en un SDS-PAGE al 15% usando el método de Laemlli (Laemlli, V., Nature, 227:680-685 (1970)) y a continuación se transfirieron a membranas de nitrocelulosa. La membrana se bloqueó con solución salina tamponada con fosfato que contenía un 3% de leche en polvo desnatada durante 1 hora y a continuación se enjuagó tres veces en TBS, 0,05% Tween 20 (TBS-T). La membrana se incubó con anticuerpo M2 a una concentración final de 10 μ g/ml durante 30 minutos en TBS-T y a continuación se enjuagó tres veces en TBS-T. A continuación, la membrana se incubó durante 30 minutos con un conjugado de IgG de cabra anticonejo (molécula completa) y peroxidasa de Rábano (HRP) diluido 1:10.000 en TBS-T, a continuación se enjuagó tres veces en TBS-T. La proteína etiquetada con FLAG se detectó con los conjugados de HRP y se visualizó mediante detección quimioluminiscente usando ECL (Amersham) y película X-Omat MR Kodak según las directrices del fabricante con exposiciones de entre 1 y 30 minutos.

Resultados

Para afrontar la cuestión de si una proteína de fusión FLAG triple produce una respuesta más sensible que el epítipo FLAG[®] tradicional, se construyó una versión FLAG triple de fosfatasa alcalina bacteriana para su expresión en *E. coli*. El vector p3xFLAG-ATS-BAP se transformó en *E. coli* y el 3XFLAG-BAP se expresó y purificó tal como se describe

ES 2 321 058 T3

en la sección de material y métodos. Adicionalmente, también se expresó y purificó un N-FLAG-BAP que contenía el epítipo FLAG® tradicional DYKDDDDK (ID SEC n.º: 1). La comparación de la sensibilidad del FLAG-BAP sencillo con respecto al triple se obtuvo mediante análisis de western blot tal como se ha descrito anteriormente. La Figura 4 muestra el western blot de un flag sencillo y triple purificado, sondeado con anticuerpo antiFLAG M2 y detectado por quimioluminiscencia. Los resultados indican claramente que se produce un aumento de 10 veces en el límite de detección del FLAG-BAP triple en comparación con la proteína de fusión FLAG-BAP sencilla. Los solicitantes pudieron detectar 500 picogramos de fosfatasa alcalina bacteriana con Flag 3X purificada con exposiciones de un valor tan breve como 1 minuto. Con un aumento del tiempo de exposición, se ha logrado una detección de un valor tan bajo como 100 picogramos pero con el nivel de fondo aumentado. Los solicitantes han demostrado también por lo menos una detección incrementada en 10 veces tanto en el ensayo de transferencia puntual (*dot blot*) como de ELISA.

Ejemplo 3

15 *Expresión de fosfatasa alcalina bacteriana 3X en células COS-7*

Materiales y métodos

Transfección de células COS-7 con p3XFLAG-CMV-7-BAP

20 Se cultivaron células COS-7 sobre placas de 35 mm² en Medio de Eagle Modificado (DME) de Dulbecco, que contenía un 10% de suero bovino fetal, L-glutamina 4 mM, 5 µg/ml de gentamicina. Las células se desarrollaron a 37°C en un incubador de CO₂ humidificado con 5% CO₂. La transfección del plásmido p3XFLAG-CMV-7-BAP se logró usando Lipofectamina (Life Technologies Inc., Gaithersburg MD.) según las directrices del fabricante. Se usaron dos microgramos de ADN vector para la transfección. A las 72 horas post-transfección se realizó una inmunotinción.

Inmunotinción

30 A las 72 horas post-inducción, las células se lavaron con Tris-HCl 50 mM pH 7,4, NaCl 150 mM (TBS). Las células se fijaron con una mezcla de metanol-acetona 1:1 (v/v) durante 1 minuto. Las células fijadas se lavaron cuatro veces con TBS y a continuación se incubaron con 10 µg/ml de conjugado de anticuerpo M2-HRP en TBS durante 1 hora. Las células se lavaron con TBS cinco veces y el conjugado de anticuerpo M2-HRP se visualizó con 0,01 mg/ml de *o*-dianisidina, 0,015% peróxido de hidrógeno recién preparados, en TBS. Las células se tiñeron durante aproximadamente durante aproximadamente 15 minutos.

Resultados

40 El p3XFLAG-CMV-7-BAP (Figura 2) se transfectó en células COS-7 tal como se describe en la sección de materiales y métodos. A las 72 horas post-transfección, las células se analizaron por inmunotinción usando un conjugado de anti-FLAG M2 HRP. En la Figura 3 se muestra una microscopía óptica de las células detectadas con anticuerpo M2, y visualizadas con *o*-dianisidina.

Argumentación

45 Los solicitantes han creado un plásmido de expresión de mamíferos que contiene múltiples epítopos FLAG® en tándem, p3X FLAG CMV-7, diseñados para expresión intracelular con un aumento de la sensibilidad de detección. Este vector contiene el promotor del citomegavirus (CMV) y el origen de replicación del SV40 para obtener una expresión eficaz en células COS-7. Por otra parte, se comparó la detección de BAP etiquetada con FLAG triple, expresada y purificada a partir de *E. coli*, con BAP etiquetada con FLAG sencillo.

50 El epítipo etiqueta FLAG® se ha usado eficazmente para detectar y purificar proteínas en sistemas mamíferos y bacterianos. Los solicitantes han demostrado que la presencia de tres epítopos FLAG hace que aumente significativamente el límite de detección de fosfatasa alcalina bacteriana purificada. Por otra parte, el 3X FLAG-BAP no se puede eluir de un gel de afinidad a M2 anti-FLAG por competición con el péptido FLAG® original. No obstante, el 3XFLAG-BAP y el 1XFLAG-BAP se pueden eluir competitivamente del gel de afinidad al M2 anti-FLAG usando un péptido 3X FLAG. El vector p3XFLAG-CMV-7 se diseñó para la expresión y detección de proteínas heterólogas en células de mamíferos y es compatible con vectores pFLAG-CMV existentes, permitiendo de este modo una subclonación sencilla entre vectores que contienen el FLAG sencillo y el FLAG triple. Los resultados de la inmunotinción muestran que la expresión del gen *pho A* en células COS-7 no se ve perturbada significativamente por la adición de la secuencia 3X FLAG.

60 El anticuerpo M2 reacciona con el FLAG alterno en la secuencia 3X FLAG. Por contraposición, el anticuerpo M5 no consigue mostrar el aumento de sensibilidad que muestra claramente el anticuerpo M2. Resultados recientes que hacen uso de la presentación en fagos (*phage display*) han demostrado que los residuos críticos para la unión M2 y la unión M5 son ligeramente diferentes. El anticuerpo M2 prefiere la secuencia Asp-Tyr-Lys-XXX-XXX-Asp-XXX-XXX (ID SEC n.º: 13) mientras que el M5 prefiere Asp-Tyr-XXX-XXX-Asp-Asp-XXX-XXX (ID SEC n.º: 14). La secuencia FLAG triple Asp-Tyr-Lys-Asp-His-Asp (ID SEC n.º: 12) favorece claramente la unión del M2 con respecto a la del M5 ó incluso el anticuerpo M1.

ES 2 321 058 T3

Ejemplo 4

Análisis de la unión del anticuerpo FLAG M2 a múltiples epítomos FLAG

5 *Materiales y métodos*

El análisis termodinámico de la unión del anticuerpo M2 a los epítomos FLAG se midió por calorimetría de titulación isotérmica usando un calorímetro OMEGA (Microcal). Todas las muestras se dializaron frente a PBS que contenía 0,05% azida sódica y se desgasificaron antes de las mediciones. Todas las mediciones se realizaron a 25°C. La concentración del anticuerpo M2 estaba entre 15 y 50 uM dependiendo de qué muestras se usaban. La concentración de los titulantes fue 605 uM para el 1X BAP, 1.110 uM para el péptido 1X FLAG, 400 uM para el BAP 3X, y 580 uM para el péptido 3X FLAG. Se efectuaron inyecciones entre cada 2,5 y 3,0 minutos lo cual resultó suficiente para lograr una referencia con volúmenes de inyección comprendidos entre 4 y 11 uL. Las inyecciones se llevaron a cabo durante un periodo de entre 4 y 10 segundos mientras se agitaba a 400 rpm.

Se realizaron un análisis y ajuste de datos usando el software Origin suministrado por MicroCal. Las entalpías se obtuvieron por integración numérica de los datos y sustracción de los calores de dilución. Se determinaron valores de K_a , siendo n el número de sitios de unión, ajustando los datos a una curva teórica y manteniendo constante solamente la entalpía durante el proceso de ajuste.

20 *Resultados y argumentación*

En el caso del sistema 1X FLAG, K_a resultó lo suficientemente pequeña de modo que el valor $r < 1.000$ en el que $r = K_a M_t(0)$, con $M_t(0)$, es la concentración inicial del anticuerpo M2 en la célula. Para el sistema 3X FLAG, la K_a fue suficientemente grande de tal manera que no se puede determinar una $r > 1.000$ que indica un sistema de unión ajustado y por lo tanto mediciones precisas sobre la K_a .

Los solicitantes han demostrado que la colocación de tres epítomos en tándem produce un aumento en la constante de asociación que es claramente más de un orden de magnitud mayor que el correspondiente un epítomo, tal como se muestra en la Tabla 1.

TABLA 1

	1XBAP	1X Péptido	3XBAP	3X Péptido
Ka	1,69E+07	1,17E+07	2,09E+08	3,66E+08

Los valores K_a para el péptido de un solo epítomo y el BAP de un solo epítomo son similares lo cual hace alusión a mecanismos de unión comparables. Para los sistemas de tres epítomos, los valores de K_a tanto del epítomo péptido como de los epítomos en BAP sugieren también mecanismos comparables. El aumento de nivel de detección observado en el sistema del FLAG triple es debido principalmente a un aumento en la constante de asociación.

45

50

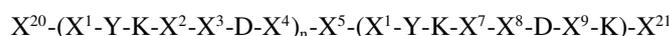
55

60

65

REIVINDICACIONES

1. Polipéptido de identificación para ser usado en la purificación de una molécula de péptido diana, en el que el polipéptido de identificación comprende la secuencia de aminoácidos



en la que:

D, Y y K son sus aminoácidos representativos;

X^{20} es hidrógeno o un enlace;

X^{21} es hidrógeno o un enlace covalente que une el polipéptido de identificación a una molécula de péptido diana;

cada X^1 y X^4 es de forma independiente un enlace o por lo menos un residuo de aminoácido seleccionado del grupo consistente en residuos de aminoácidos aromáticos y lisina, arginina, histidina, aspartato, glutamato, serina, treonina, cisteína, tirosina, asparagina y glutamina;

cada X^2 , X^3 , X^7 y X^8 es de forma independiente un residuo de aminoácido seleccionado del grupo consistente en residuos de aminoácidos aromáticos y lisina, arginina, histidina, aspartato, glutamato, serina, treonina, cisteína, tirosina, asparagina y glutamina;

X^5 es un enlace o un dominio espaciador que comprende por lo menos un residuo de aminoácido;

X^9 es un enlace o un residuo de aspartato; y

n es por lo menos 2, y

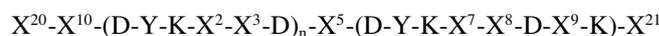
en el que el sitio escindible representado por la secuencia $X^7-X^8-D-X^9-K$ no está situado en el, o interpuesto entre múltiples copias del, dominio antigénico $X^1-Y-K-X^2-X^3-D$.

2. Polipéptido de identificación según la reivindicación 1, en el que:

X^5 es un dominio espaciador que comprende por lo menos un residuo de histidina, por lo menos un residuo de glicina o una combinación de residuos de histidina alternados o múltiples, comprendiendo dicha combinación His-Gly-His, o $-(His-X)_m-$, en la que m es de 1 a 6 y X se selecciona del grupo consistente en Gly, His, Tyr, Trp, Val, Leu, Ser, Lys, Phe, Met, Ala, Glu, Ile, Thr, Asp, Asn, Gln, Arg, Cys, y Pro.

3. Polipéptido de identificación de una cualquiera de las reivindicaciones 1 ó 2, en el que X^1 es Asp y X^4 es un enlace.

4. Polipéptido de identificación de la reivindicación 3 en el que el polipéptido de identificación comprende además X^{10} de tal manera que el polipéptido de identificación comprende la secuencia de aminoácidos



en la que X^{10} es un enlace, cualquier residuo de aminoácido, o un residuo de metionina.

5. Polipéptido de identificación de una cualquiera de las reivindicaciones 1 a 4, en el que cada X^2 y X^3 es de forma independiente una secuencia quelante de metal que comprende por lo menos un residuo de histidina, por lo menos un residuo de glicina o una combinación de residuos de histidina múltiples o alternados, comprendiendo dicha combinación His-Gly-His, o $-(His-X)_m-$, en la que m es de 1 a 6 y X se selecciona del grupo consistente en Gly, His, Tyr, Trp, Val, Leu, Ser, Lys, Phe, Met, Ala, Glu, Ile, Thr, Asp, Asn, Gln, Arg, Cys, y Pro.

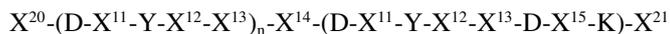
6. Polipéptido de identificación de la reivindicación 5, en el que el polipéptido de identificación comprende la secuencia:



en la que X^{21} es hidrógeno o un enlace covalente que une el polipéptido de identificación a una molécula de péptido diana.

ES 2 321 058 T3

7. Polipéptido de identificación para ser usado en la purificación de una molécula de péptido diana, en el que el polipéptido de identificación comprende la secuencia de aminoácidos



5

en la que:

D, Y y K son sus aminoácidos representativos;

10

X^{20} es hidrógeno o un enlace;

X^{21} es hidrógeno o un enlace covalente que une el polipéptido de identificación a una molécula de péptido diana;

15

cada X^{11} es un enlace, cualquier residuo de aminoácido, o un residuo de lisina;

cada X^{12} es un aminoácido seleccionado de grupo consistente en residuos de aminoácidos aromáticos y lisina, arginina, histidina, aspartato, glutamato, serina, treonina, cisteína, tirosina, asparagina y glutamina;

20

cada X^{13} es un enlace o por lo menos un aminoácido seleccionado del grupo consistente en residuos de aminoácidos aromáticos y lisina, arginina, histidina, aspartato, glutamato, serina, treonina, cisteína, tirosina, asparagina y glutamina;

X^{14} es un enlace o un dominio espaciador que comprende por lo menos un residuo de aminoácido;

25

X^{15} es un enlace o un residuo de aspartato; y

n es por lo menos 2,

30

en el que el sitio escindible representado por la secuencia $X^{12}-X^{13}-D-X^{15}-K$ no está situado en el, o interpuesto entre múltiples copias del, dominio antigénico $D-X^{11}-Y-X^{12}-X^{13}$.

8. Polipéptido de identificación según la reivindicación 7, en el que:

35

X^{14} es un dominio espaciador que comprende por lo menos un residuo de histidina, por lo menos un residuo de glicina o una combinación de residuos de histidina múltiples o alternados, comprendiendo dicha combinación His-Gly-His, o $-(His-X)_m-$, en la que m es de 1 a 6 y X se selecciona del grupo consistente en Gly, His, Tyr, Trp, Val, Leu, Ser, Lys, Phe, Met, Ala, Glu, Ile, Thr, Asp, Asn, Gln, Arg, Cys, y Pro.

40

9. Polipéptido de una cualquiera de las reivindicaciones 1 a 8 en el que X^{21} es hidrógeno.

45

10. Polipéptido de una cualquiera de las reivindicaciones 1 a 9, en el que el polipéptido comprende además un péptido diana unido a la secuencia de unión de tal manera que la secuencia de unión está entre las múltiples copias del dominio antigénico y la molécula del péptido diana.

50

11. Polipéptido de una cualquiera de las reivindicaciones 1 a 8, en el que X^{21} es un enlace covalente que une el polipéptido de identificación a una molécula de péptido diana y el polipéptido comprende además un péptido diana unido a la secuencia de unión de tal manera que la secuencia de unión está entre las múltiples copias del dominio antigénico y la molécula del péptido diana.

12. Segmento de ADN que codifica un polipéptido de identificación según una cualquiera de las reivindicaciones 1 a 11.

55

13. Vector de expresión de ADN que comprende ADN que codifica un polipéptido híbrido que comprende un polipéptido diana y un polipéptido de identificación según una cualquiera de las reivindicaciones 1 a 11.

60

14. Vector de expresión de ADN de la reivindicación 13 en el que el ADN que codifica un polipéptido híbrido comprende además un sitio de clonación múltiple que comprende sitios de reconocimiento de enzimas de restricción múltiples.

15. Método para producir un péptido diana, que comprende:

65

a. se transforman células hospedadoras con el vector de expresión de ADN de una cualquiera de las reivindicaciones 13 a 14;

b. se aísla e identifica dicho péptido diana extrayendo dicho péptido diana de dichas células hospedadoras transformadas; y

ES 2 321 058 T3

c. se purifica el péptido diana mediante el uso de las propiedades de afinidad para ligandos del polipéptido de identificación.

5 16. Método de la reivindicación 15, en el que el péptido diana se aísla e identifica por inmunoprecipitación y Western blotting.

10 17. Método para purificar un polipéptido híbrido que comprende un polipéptido de identificación de una cualquiera de las reivindicaciones 1 a 5 y un péptido diana, en el que se usa una columna que comprende una matriz y un anticuerpo creado contra el dominio antigénico de uno cualquiera de los polipéptidos de identificación de las reivindicaciones 1 a 5.

15

20

25

30

35

40

45

50

55

60

65

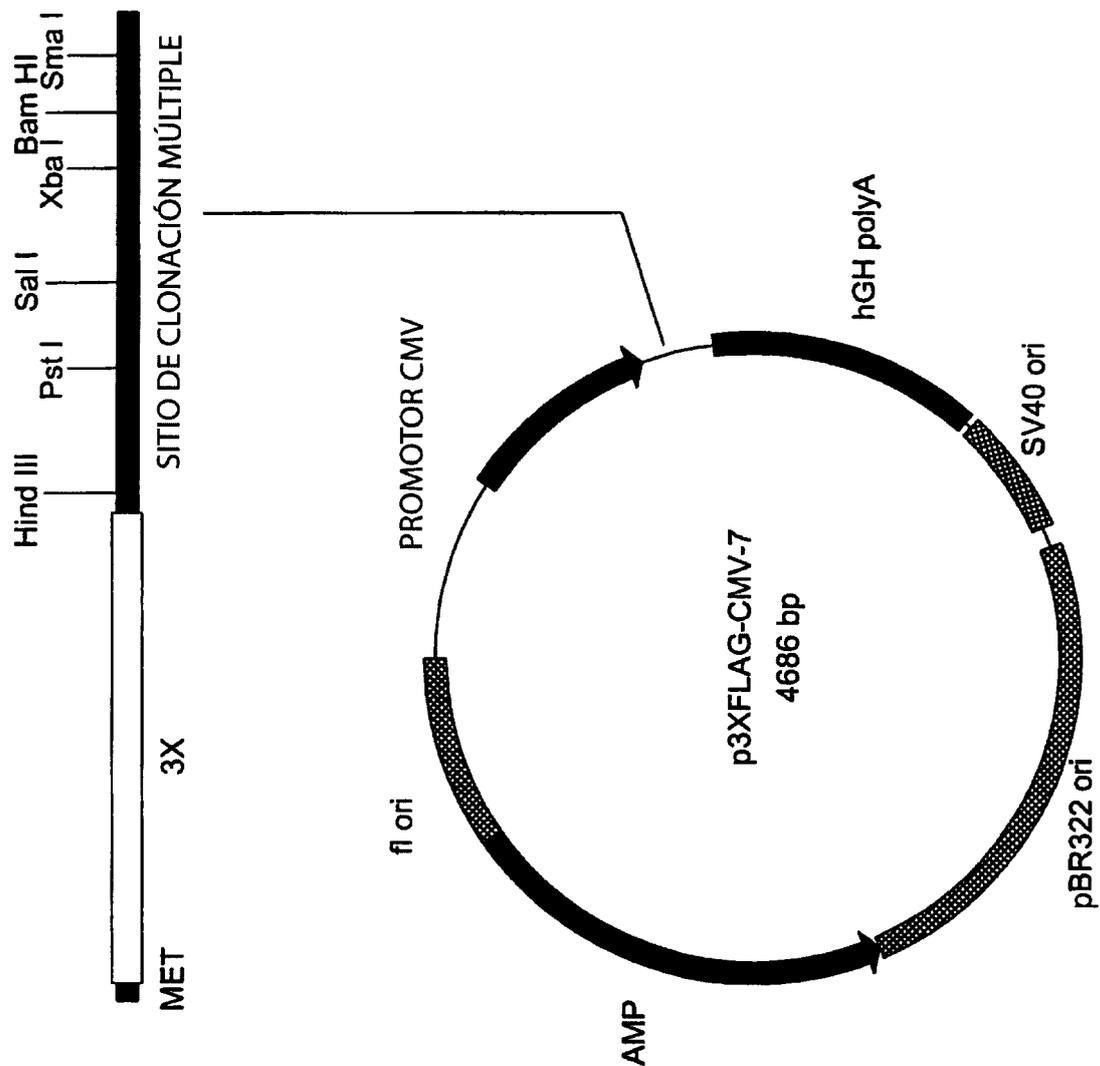


FIG. 1B

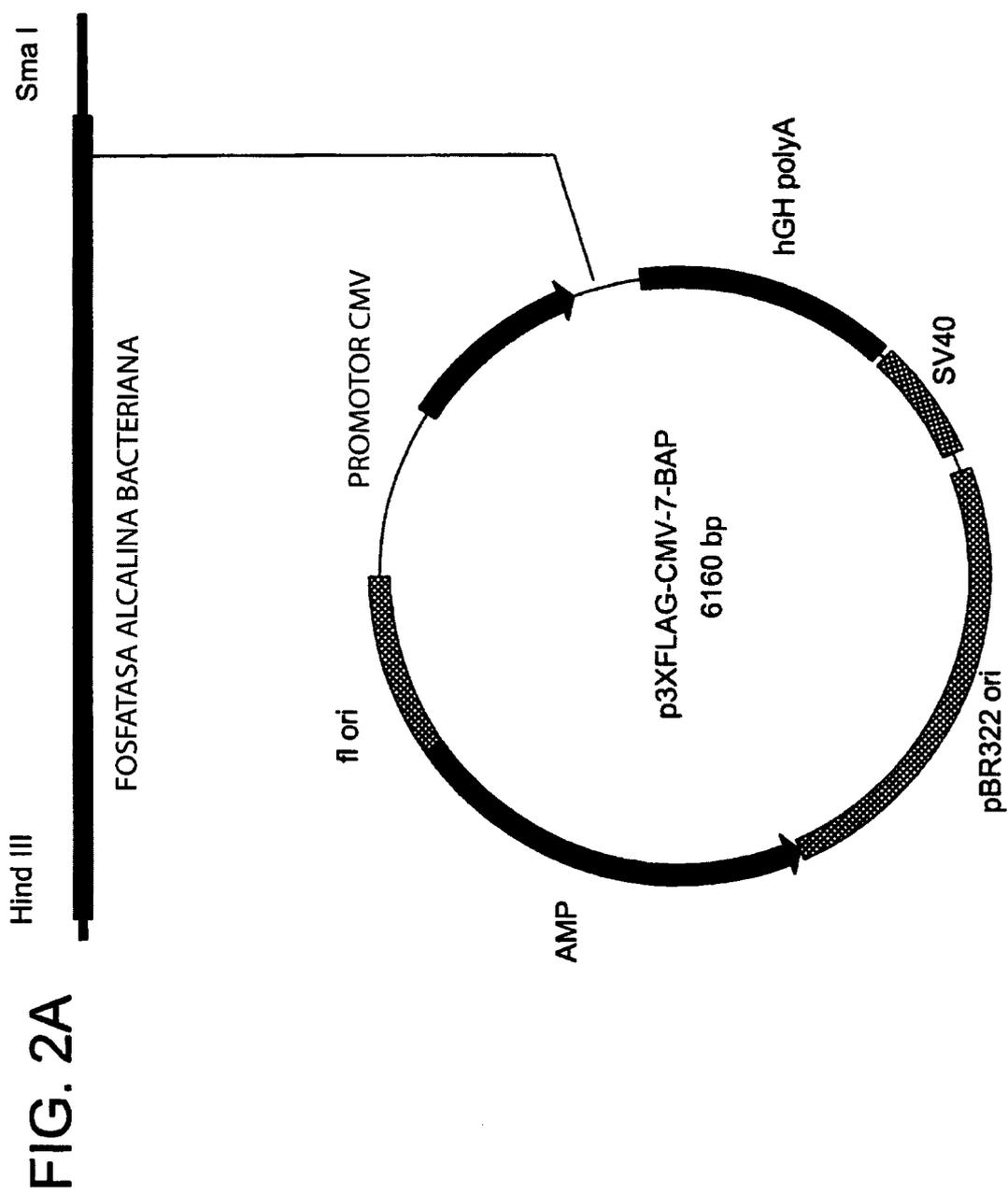


FIG. 2B

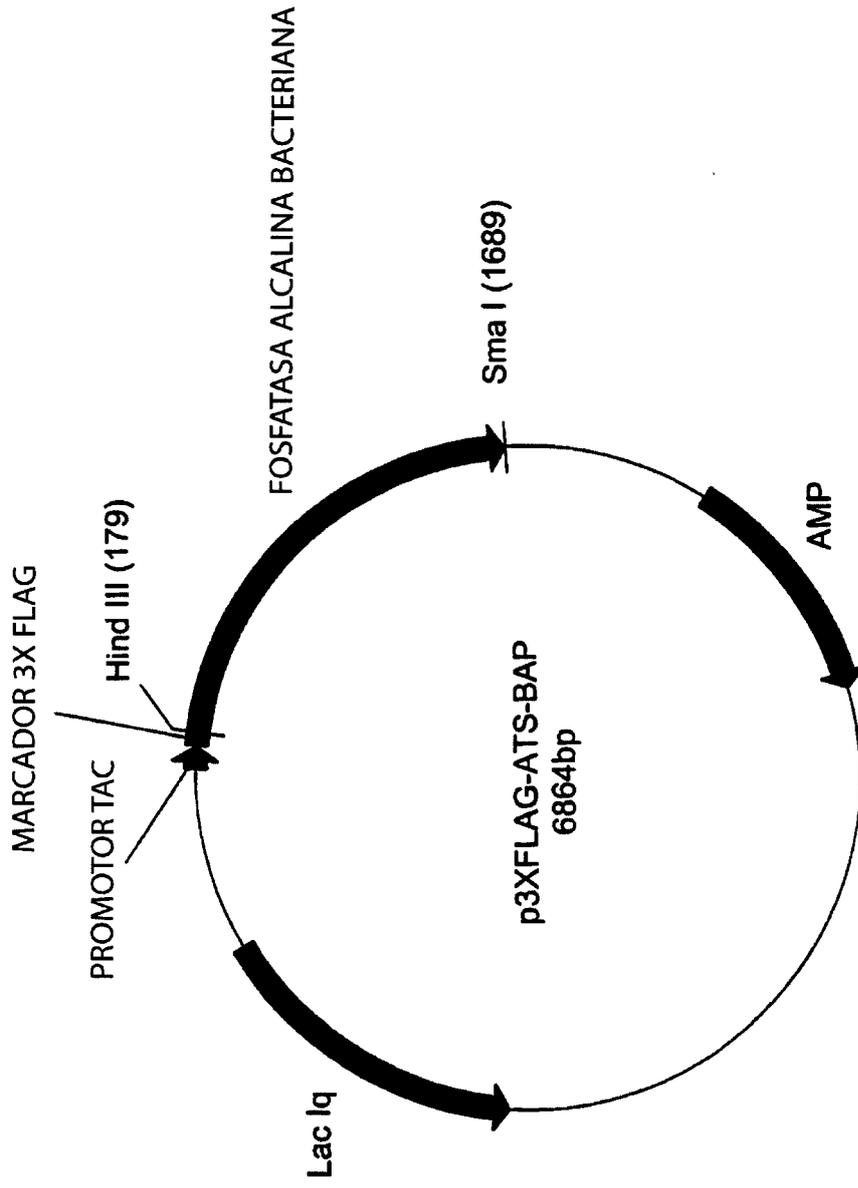
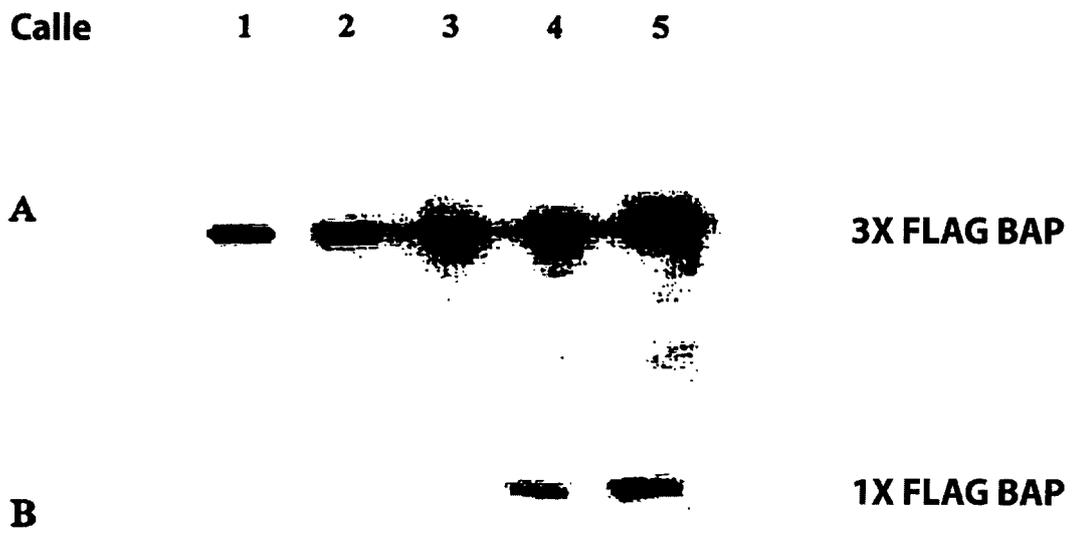


FIG. 3

WESTERN BLOT



ES 2 321 058 T3

LISTA DE SECUENCIAS

<110> BRIZZARD, BILLY L.
HERNAN, RON
5

<120> PURIFICACIÓN DE PROTEÍNAS RECOMBINANTES FUSIONADAS CON MÚLTIPLES EPÍTOPOS

<130> SGM6933
10

<140>
<141>

<150> 09/415,000
<151> 1999-10-08
15

<160> 14
<170> PatentIn Ver. 2.1
20

<210> 1
<211> 8
25
<212> PRT
<213> Secuencia artificial

<220>
30
<223> Descripción de secuencia artificial: secuencia sintetizada

<400> 1
35

Asp Tyr Lys Asp Asp Asp Asp Lys
1 5

<210> 2
40
<211> 4
<212> PRT
45
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: secuencia sintetizada

<400> 2
50

Asp Asp Asp Lys
1
55

<210> 3
60
<211> 5
<212> PRT
<213> Secuencia artificial

<220>
65
<223> Descripción de secuencia artificial: secuencia sintetizada

ES 2 321 058 T3

<222> (7)..(8)

<223> X es cualquier aminoácido

5 <400> 13

Asp Tyr Lys Xaa Xaa Asp Xaa Xaa
1 5

10

<210> 14

<211> 8

15 <212> PRT

<213> Secuencia artificial

<220>

20 <223> Descripción de secuencia artificial: secuencia sintetizada

<220>

<221> DNA_BIND

25 <222> (3)..(4)

<223> X es cualquier aminoácido

<220>

30 <221> DNA BIND

<222> (7)..(8)

<223> X es cualquier aminoácido

35

<400> 14

Asp Tyr Xaa Xaa Asp Asp Xaa Xaa
1 5

40

45

50

55

60

65