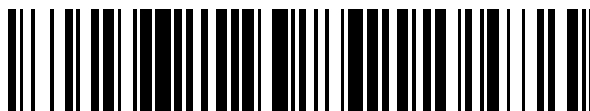


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 379 096**

51 Int. Cl.:
C12N 15/00 (2006.01)
C12P 21/00 (2006.01)
C12N 5/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **06760292 .0**
96 Fecha de presentación: **19.05.2006**
97 Número de publicación de la solicitud: **1885867**
97 Fecha de publicación de la solicitud: **13.02.2008**

54 Título: **Nuevos procedimientos**

30 Prioridad:
20.05.2005 US 683254 P

45 Fecha de publicación de la mención BOPI:
20.04.2012

45 Fecha de la publicación del folleto de la patente:
20.04.2012

73 Titular/es:
GlaxoSmithKline LLC
One Franklin Plaza P.O. Box 7929
Philadelphia, PA 19101, US

72 Inventor/es:
JOHANSON, Kyung Oh y
TRILL, John J.

74 Agente/Representante:
Carpintero López, Mario

ES 2 379 096 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

Nuevos procedimientos

Campo de la invención

5 La presente invención se refiere al uso del dominio B1 de la Proteína G como un marcador de epítipo para la sobreexpresión de proteínas en células de mamíferos.

Antecedentes de la invención

10 El dominio B1 de unión a inmunoglobulina de la proteína G estreptocócica (en el presente documento denominado algunas veces GB1), como se ilustra en Park, y col. *Biochemistry*, 36:14277-14283 (1997), es una proteína pequeña con un núcleo hidrófobo bien desarrollado, que se pliega en una lámina beta de 4 cadenas con una hélice alfa flanqueante. Esta proteína no contiene disulfuros o cisteínas libres, ni prolinas o motivos de unión a metales. La GB1 es muy estable y soluble y es uno de los sistemas de modelos más ampliamente usados en el área del plegamiento y diseño de proteínas.

15 Además, la GB1 contiene en su superficie sitios de unión para el fragmento C-terminal de la cadena pesada de la inmunoglobulina G (IgGFc). Las proteínas que contienen este dominio pueden purificarse por Fc inmovilizado, eluirse de la columna de afinidad mediante ácidos o cortarse proteolíticamente y detectarse mediante reactivos existentes.

20 Hammarstrom, y col. *Protein Sci.*, 11:313-321 (2002) han descrito, comparando niveles de expresión de fusiones de proteínas de expresión solubles con este dominio B1 con fusiones usando MBP, NusA, ZZ, GST, Tioredoxina y 6His y 26 proteínas diferentes, que GB1 proporciona un mayor porcentaje de proteínas solubles expresadas en *E. coli*. Zhou, y col. *Journal of Biomolecular NMR*, 20:11-14 (2001) también han descrito que cuando este dominio se usa como un marcador de epítipo para la fusión con proteínas recombinantes expresadas en *E. coli*, no es necesario eliminarlo para obtener la estructura RMN. Finalmente, Cheng & Patel, *Biochem. Biophys. Res. Commun.*, 30:401-405 (2004) llegaron a la conclusión de que la expresión de proteínas que contenían GB1 en *E. coli* aumentaban los niveles de expresión.

25 En una mayoría de casos, debido a problemas en cuanto al plegamiento y a la solubilidad de las proteínas, muchas proteínas recombinantes no pueden expresarse en organismos procariotas. Por lo tanto, las células de mamífero son el sistema de expresión de elección para proteínas recombinantes que requieren amplias modificaciones post-traduccionales. Sin embargo, incluso con sistemas de expresión de mamíferos, la baja expresión y la baja actividad de las proteínas recombinantes expresadas usando marcadores de epítipo pequeños, tales como 6 Histidinas y marcador Strep-II, han influido negativamente en la purificación a gran escala por cristalografía y HTS. Datos obtenidos a partir de la expresión de la lipasa endotelial indican que sin marcador, con un marcador de 6 Histidinas o con 6 Histidinas/Strep-II, la expresión era muy baja a niveles > 1 mg/litro.

30 Se ha demostrado que el uso de la región IgGFc como un marcador conduce a niveles muy altos de expresión de proteínas. Sin embargo, este marcador puede perjudicar a genes que requieren una orientación particular tal como un dímero inverso. Bloquean el plegamiento correcto y, de esta manera, conducen a una pérdida de actividad.

35 Por tanto, aún existe una necesidad para la expresión o sobreexpresión de polipéptidos seleccionados en células de mamífero. Fue totalmente inesperado descubrir, en la presente invención, un drástico aumento en los niveles de expresión en células de mamífero usando un marcador de GB1 sobre los marcadores usados tradicionalmente y observar una actividad completa en todas las proteínas examinadas.

Sumario de la invención

40 En un aspecto, la presente invención se refiere a un procedimiento para la expresión de un polipéptido (de interés) en una célula huésped recombinante de mamífero que comprende cultivar la célula huésped que comprende un vector de expresión que comprende al menos un polinucleótido GB1 directa o indirectamente unido a un polinucleótido que codifica el polipéptido de interés que produce el polipéptido en una condición de cultivo apropiada y purificar dicho polipéptido.

45 En otro aspecto, la presente invención proporciona polipéptidos fabricados por el procedimiento anterior.

Para la producción del polipéptido, también se describe un vector de expresión para usar en una célula huésped recombinante de mamífero, que comprende al menos un polinucleótido GB1 directa o indirectamente unido a un polinucleótido que codifica el polipéptido.

50 Sin embargo en otro aspecto, para la producción del polipéptido, la presente invención se refiere a una célula huésped recombinante de mamífero que comprende un vector de expresión que comprende al menos un polinucleótido GB1 directa o indirectamente unido a un polinucleótido que codifica el polipéptido.

Descripción detallada

5 Como se usa en el presente documento "polipéptidos GB1" incluyen polipéptidos aislados que comprenden una secuencia de aminoácidos que posee al menos una identidad del 70%, al menos una identidad del 80%, al menos una identidad del 90%, al menos una identidad del 95%, al menos una identidad del 97-99%, con la de la SEC ID N°: 1, 2 ó 3 sobre la longitud completa de la SEC ID N°: 1, 2 ó 3, respectivamente. Dichos polipéptidos incluyen los que comprenden los aminoácidos de la SEC ID N°: 1, 2 ó 3.

DTYKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKTFTVTE (SEC ID N°: 1)

MDTYKLILINGKTLKGETTEAVDAATAEKVFKQYANDNGVDGEWYDDATKTFTVTE (SEC ID N°: 2)

MEILAALPKTDYKLILINGKTLKGETTEAVDAATAEKVFKQYANDNGVDGEWYDDATKTFTVTE (SEC ID N°: 3)

10 Adicionalmente, los polipéptidos GB1 incluyen polipéptidos aislados, en los que la secuencia de aminoácidos poseen al menos una identidad del 70%, al menos una identidad del 80%, al menos una identidad del 90%, al menos una identidad del 95%, al menos una identidad del 97-99%, con la secuencia de aminoácidos de la SEC ID N°: 1, 2 ó 3 sobre la longitud completa de la SEC ID N°: 1, 2 ó 3, respectivamente. Tales polipéptidos incluyen el polipéptido de la SEC ID N°: 1, 2 ó 3.

15 Adicionalmente, los polipéptidos GB1 incluyen polipéptidos aislados codificados por un polinucleótido que comprende la secuencia contenida en la SEC ID N°: 4, 5 ó 6.

20 Como se usa en el presente documento "polipéptidos BG1" incluyen polipéptidos aislados que comprenden una secuencia de nucleótidos que codifica un polipéptido que posee al menos una identidad del 70%, al menos una identidad del 80%, al menos una identidad del 90%, al menos una identidad del 95%, con la secuencia de aminoácidos de la SEC ID N°: 1, 2 ó 3, sobre la longitud completa de la SEC ID N°: 1, 2 ó 3, respectivamente. En este sentido, también se incluyen los polipéptidos que poseen al menos una identidad del 97%, al menos una identidad del 98-99% y al menos una identidad del 99%. Tales polinucleótidos incluyen un polinucleótido que comprende la secuencia de nucleótidos contenida en la SEC ID N°: 4, 5 ó 6, que codifica el polipéptido de la SEC ID N°: 1, 2 ó 3, respectivamente o un polinucleótido que comprende la secuencia de nucleótidos contenida en la SEC ID N°: 4, 5 ó 6 que codifica el polipéptido de la SEC ID N°: 1, 2 ó 3, respectivamente.

25 Adicionalmente, los polinucleótidos de la presente invención incluyen polinucleótidos aislados que comprenden una secuencia de nucleótidos que posee al menos una identidad del 70%, al menos una identidad del 80%, al menos una identidad del 90%, al menos una identidad del 95%, con la SEC ID N°: 4, 5 ó 6 sobre la longitud completa de la SEC ID N°: 4, 5 ó 6, respectivamente. En este sentido, también se incluyen los polinucleótidos que poseen al menos una identidad del 97%, al menos una identidad del 98-99% y al menos una identidad del 99%. Tales polinucleótidos incluyen un polinucleótido que comprende el polinucleótido de la SEC ID N°: 4, 5 ó 6 así como el polinucleótido de la SEC ID N°: 4, 5 ó 6.

30 La invención también proporciona polinucleótidos que son complementarios con todos los polinucleótidos descritos anteriormente:

35 **gacat tacaaattaa tccttaatgg taaaacattg aaaggcgaaa caactactga agctgttgat**
gctgctactg cagaaaaagt ctcaaacaa tacgctaacg acaacgggtg tgacggtgaa
tggacttacg acgatgcgac taagacctt acagttactg aa (SEC ID N°: 4)

40 **atggacact tacaaattaa tccttaatgg taaaacattg aaaggcgaaa caactactga agctgttgat**
gctgctactg cagaaaaagt ctcaaacaa tacgctaacg acaacgggtg tgacggtgaa
tggacttacg acgatgcgac taagacctt acagttactg aa (SEC ID N°: 5)

45 **atggaa atttagctg cattacctaa gactgacact tacaaattaa tccttaatgg taaaacattg**
aaaggcgaaa caactactga agctgttgat gctgctactg cagaaaaagt ctcaaacaa
tacgctaacg acaacgggtg tgacggtgaa tggacttacg acgatgcgac taagacctt
acagttactg aa (SEC ID N°: 6)

Los polinucleótidos que son complementarios con todos los polinucleótidos descritos anteriormente incluyen, pero sin limitación, cualquier polinucleótido que, en condiciones rigurosas, se hibride con las SEC ID N°s: 4, 5 ó 6.

Como se usa en el presente documento, las expresiones "condiciones rigurosas" y una "condición de hibridación rigurosa" significan que la hibridación solo se producirá si entre las secuencias existe una identidad de al menos 70%, al menos 80% y al menos 95%. Un ejemplo de condiciones de hibridación rigurosas es la incubación durante una noche a 42 °C en una solución que comprende: formamida al 50%, 5xSSC (NaCl 150 mM, citrato trisódico 15 mM), fosfato sódico 50 mM (pH 7,6), solución de Denhardt 5x, sulfato de dextrano al 10% y 20 microgramos/ml de ADN de esperma de salmón cortado, desnaturalizado, seguido de lavado de filtros en 0,1x SSC a aproximadamente 65 °C. Las condiciones de hibridación y de lavado se conocen bien y se ilustran en Sambrook y col., Molecular Cloning: A Laboratory Manual, Segunda Edición, Cold Spring Harbor, N.Y., (1989), particularmente en el Capítulo 11 del mismo, cuya descripción se incorpora en el presente documento por referencia en su totalidad.

Identidad", como se conoce en la técnica, es una relación entre dos o más secuencias de polipéptidos o dos o más secuencias de polinucleótidos, determinada comparando las secuencias. En la técnica, "identidad" también significa el grado de relación de secuencia entre secuencias de polipéptidos o de polinucleótidos, según sea el caso, determinado por los emparejamientos entre las cadenas de tales secuencias. La "identidad" y la "similitud" pueden calcularse fácilmente mediante procedimientos conocidos, que incluyen pero sin limitación, los descritos en (Computational Molecular Biology, Lesk, A.M., ed., Oxford University Press, Nueva York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D.W., ed., Academic Press, Nueva York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, Nueva Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; y Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; y Carillo, H., and Lipman, D., SIAM J. Applied Math., 48: 1073 (1988). Para determinar la identidad se diseñan procedimientos preferidos para proporcionar el mayor emparejamiento entre las secuencias ensayadas. Los procedimientos para determinar la identidad y la similitud se codifican en programas informáticos disponibles públicamente. Los procedimientos de programas informáticos preferidos para determinar la identidad y la similitud entre dos secuencias incluyen, pero sin limitación, el paquete de programas CGG (Devereux, J., y col., Nucleic Acids Research 12(1): 387 (1984)), BLASTP, BLASTN y FASTA (Atschul, S.F. y col., J. Molec. Biol. 215: 403-410 (1990). El programa BLAST X se encuentra disponible públicamente en el NCBI (National Center for Biotechnology Information) y otras fuentes (BLAST Manual, Altschul, S., y col., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., y col., J. Mol. Biol. 215: 403-410 (1990). Para determinar la identidad también puede usarse el bien conocido algoritmo de Smith Waterman. Los parámetros para comparar las secuencias de polipéptidos incluyen los siguientes:

- 1) Algoritmo: Needleman y Wunsch, J. Mol Biol. 48: 443-453 (1970)
- Matriz comparativa: BLOSSUM62 de Hentikoff y Hentikoff, Proc. Natl. Acad. Sci. USA. 89:10915-10919 (1992)
- Penalización por hueco: 12
- Penalización por longitud de hueco: 4

Un programa útil con estos parámetros se encuentra disponible públicamente como el programa "gap" (hueco) de Genetics Computer Group, Madison WI. Los parámetros mencionados anteriormente son los parámetros por defecto para comparar péptidos (junto con no penalizaciones por huecos finales).

Los parámetros preferidos para comparar polinucleótidos incluyen los siguientes:

- 1) Algoritmo: Needleman y Wunsch, J. Mol Biol. 48: 443-453 (1970)
- Matriz comparativa: emparejamientos = +10; no emparejamientos = 0
- Penalización por hueco: 50
- Penalización por longitud de hueco: 3
- Disponible como: El programa "gap" de Genetics Computer Group, Madison WI. Estos son los parámetros por defecto para comparar ácidos nucleicos.

A modo de ejemplo, una secuencia de polinucleótidos de la presente invención puede ser idéntica a la secuencia referencia de la SEC ID N°: 4, es decir 100% idéntica, o puede incluir hasta un determinado número entero de modificaciones de nucleótidos en comparación con la secuencia de referencia. Dichas modificaciones se seleccionan del grupo que consiste en al menos una delección, sustitución de nucleótidos, incluyendo transición y transversión, o inserción en el que tales modificaciones pueden producirse en las posiciones 5' o 3' terminales de la secuencia de nucleótidos de referencia o en cualquier parte entre las posiciones terminales, intercaladas individualmente entre los nucleótidos en la secuencia de referencia o en uno o más grupos contiguos dentro de la secuencia de referencia. El número de modificaciones de nucleótidos se determina multiplicando el número total de nucleótidos en la SEC ID N°: 4 por el porcentaje numérico del porcentaje de identidad respectivo (dividido entre 100) y restando ese producto de dicho número total de nucleótidos en la SEC ID N°: 4 o:

$$n_n \leq x_n \cdot (x_n \cdot y)$$

en la que n_n es el número de modificaciones de nucleótidos, x_n es el número total de nucleótidos en la SEC ID N°: 4 e y es, por ejemplo, 0,70 para el 70%, 0,80 para el 80%, 0,85 para el 85%, 0,90 para el 90%, 0,95 para el 95%, etc. y en el que cualquier producto no entero de x_n e y se redondea hacia abajo hacia el número entero más próximo antes de restarlo de x_n . Las modificaciones de una secuencia de polinucleótidos que codifica el polipéptido de la SEC ID N°: 4 puede crear mutaciones sin sentido, de sentido equívoco o con desplazamiento de fase de lectura en

esta secuencia codificante y por lo tanto modificar el polipéptido codificado por el polinucleótido después de tales modificaciones.

De manera similar, una secuencia de polipéptidos de la presente invención puede ser idéntica a la secuencia de referencia de la SEC ID N°: 1, es decir 100% idéntica o puede incluir hasta un determinado número entero de modificaciones de aminoácidos en comparación con la secuencia de referencia de manera que el % de identidad sea menor del 100%. Tales modificaciones se seleccionan del grupo que consiste en al menos una delección, sustitución de aminoácidos, incluyendo sustitución o inserción conservativa y no conservativa, y en el que dichas modificaciones pueden producirse en las posiciones amino- o carboxi- terminales de la secuencia de polipéptido de referencia o en cualquier parte entre las posiciones terminales, intercaladas individualmente entre los aminoácidos en la secuencia de referencia o en uno o más grupos contiguos dentro de la secuencia de referencia. El número de modificaciones de aminoácidos para un % de identidad proporcionado se determinó multiplicando el número total de aminoácidos en la SEC ID N°: 1 por el porcentaje numérico del porcentaje de identidad respectivo (dividido entre 100) y después restando este producto de dicho número total de aminoácidos en la SEC ID N°: 1 o:

$$n_a \leq x_a - (x_a \cdot y)$$

en la que n_a es el número de modificaciones de aminoácidos, x_a es el número total de aminoácidos en la SEC ID N°: 1 e y es, por ejemplo, 0,70 para el 70%, 0,80 para el 80%, 0,85 para el 85%, etc. y en el que cualquier producto no entero de x_a e y se redondea hacia abajo hacia el entero más próximo antes de restarlo de x_a .

La expresión "vector de expresión" se usa para indicar una molécula de ADN, lineal o circular, que comprende un segmento que codifica un polipéptido (de interés) directa o indirectamente fusionado a uno o más polinucleótidos GB1 operativamente unidos a secuencias de control adicionales que permiten su transcripción. Tales segmentos adicionales incluyen secuencias promotoras y terminadoras y también pueden incluir uno o más orígenes de replicación, uno o más marcadores de selección, un potenciador, una señal de poliadenilación, etc. Los vectores de expresión generalmente derivan de ADN plasmídico o viral, o pueden contener elementos de ambos. Algunos ejemplos de vectores de expresión de mamíferos adecuados se encuentran en los documentos EP 307,247; 260,148; y 309,237 307,248.

Esencialmente, en la invención puede usarse cualquier célula de mamífero. La célula puede ser una célula primaria (por ejemplo un hepatocito primario, una célula neuronal primaria, o un mieloblasto primario) o puede ser una célula de una línea celular establecida. No es necesario que la célula sea capaz de someterse a división celular; en la invención puede usarse una célula definitivamente diferenciada. Si se desea, para maximizar la eficacia de infección, el virus puede introducirse en una célula primaria aproximadamente 24 h después de sembrar en placa la célula primaria. Preferentemente, la célula de mamífero es una célula derivada de hígado, tal como una célula HepG2, una célula Hep3B, una célula Huh-7, una célula FTO2B, una célula Hepa1-6 o una célula SK-Hep-1) o una célula de Kupffer; una célula de riñón, tal como una célula de la línea celular de la línea celular de riñón 293, una célula PC12 (por ejemplo una célula PC12 diferenciada inducida por el factor de crecimiento nervioso), una célula COS (por ejemplo una célula COS7), o una célula Vero (una célula de riñón de mono verde Africano); una célula neuronal, tal como una célula neuronal fetal, una célula piramidal cortical, una célula mitral, una célula granular o una célula cerebral (por ejemplo, una célula del córtex cerebral; un astrocito; una célula glial; una célula de Schwann); una célula muscular, tal como un mioblasto o miotubo (por ejemplo una célula C₂ C₁₂); una célula madre embrionaria, una célula de bazo (por ejemplo, un macrófago o linfocito); una célula epitelial, tal como una célula HeLa (una línea epitelial de carcinoma cervical humano); un fibroblasto, tal como una célula NIH3T3; una célula endotelial; una célula WISH; una célula A549 o una célula madre de la médula ósea. Otras células de mamífero preferidas incluyen células CHO/dhfr, células Ramos, Jurkat, HL60 y K-562.

Las células de mamífero pueden transformarse, por ejemplo, de acuerdo con el procedimiento descrito en Saibo Kogaku (Cell Engineering), publicación extra 8, Shin Saibo Kogaku Jikken Protocol (New Cell Engineering Experimental Protocol), 263-267 (1995), publicado por Shujunsha o Virology, 52, 456 (1973). Por tanto, puede obtenerse el transformante transformado con el vector de expresión que contiene el polinucleótido que codifica el polipéptido que se expresará.

Cuando se emplean células de mamífero como el huésped, el transformante se cultiva, por ejemplo, en medio MEM que contiene de aproximadamente del 5 al 20% de suero bovino fetal [Science, 122, 501 (1952)], medio DMEM [Virology, 8, 396 (1959)], medio RPMI 1640 [The Journal of the American Medical Association, 199, 519 (1967)], medio 199 [Proceeding of the Society for the Biological Medicine, 73,1 (1950)], etc. Preferentemente, el pH del medio se ajusta de aproximadamente 6 a aproximadamente 8. El transformante normalmente se cultiva a una temperatura de aproximadamente 30 a 40 °C durante aproximadamente de 15 a 60 horas y, si fuera necesario, el cultivo puede oxigenarse o agitarse.

La presente invención también proporciona un vector de expresión que comprende uno o más polinucleótidos GB1 directa o indirectamente fusionados a un polinucleótido que codifica un polipéptido de interés y elementos relacionados necesarios para la expresión celular de la proteína codificada por un polinucleótido GB1. Por ejemplo, en el vector de expresión, puede incorporarse una secuencia promotora que dirige la transcripción del polinucleótido GB1 en una célula huésped.

Como se usa en el presente documento “directamente fusionado” significa un polinucleótido GB1 que es adyacente a un polinucleótido que codifica un polipéptido de interés. Sin embargo, “indirectamente fusionado” significa un polinucleótido GB1 que no es adyacente al polinucleótido que codifica un polipéptido de interés, sino más bien hay uno o más espaciadores, tal como el sitio de escisión de la proteasa del virus del grabado de tabaco (VGT) y/o entre un polinucleótido GB1 y un polinucleótido que codifica un polipéptido de interés está presente otro marcador (o marcadores) (tal como el marcador de epítipo de 6 Histidinas). Debe ser importante destacar que GB1 puede colocarse antes del extremo 5' o después del extremo 3' del polinucleótido que codifica el polipéptido de interés.

Como se usa en el presente documento “polipéptido de interés” se refiere a un polipéptido que se expresará usando, como marcador, un polipéptido GB1. Los polipéptidos de interés, incluyen, por ejemplo, la proteína Lipasa Endotelial (LE), la proteína de Transferencia de Ésteres de Colesterol (pTEC), etc.

Para su uso en células de mamíferos, las funciones de control sobre los vectores de expresión a menudo se proporcionan por material viral. Por ejemplo, los promotores normalmente usados derivan de los genomas de polioma, Adenovirus2, retrovirus, citomegalovirus y Virus de Simio 40 (SV40). Otros promotores son los procedentes de fuentes heterólogas, por ejemplo, el promotor de beta actina. Los promotores temprano y tardío del virus SV40 son particularmente útiles porque ambos se obtienen fácilmente a partir de virus como un fragmento que también contiene el origen de replicación viral del SV40 [Fiers y col., Nature, 273: 113 (1978)]. También pueden usarse fragmentos más grandes o más pequeños del SV40, siempre que se incluyan los aproximadamente 250 pb de la secuencia que se extiende desde el sitio HindIII hacia el sitio BglI localizado en el origen de replicación viral. El promotor temprano inmediato del citomegalovirus humano se obtiene convenientemente como un fragmento de restricción HindIII. Greenaway, y col., Gene, 18: 355-360 (1982). Además, también es posible, y a menudo deseable, utilizar secuencias promotoras o de control normalmente asociadas con la secuencia génica deseada, siempre que tales secuencias de control sean compatibles con los sistemas de la célula huésped

La transcripción de un ADN que codifica un polipéptido heterólogo deseado (polipéptido de interés) por eucariotas superiores se aumenta insertando en el vector una secuencia potenciadora. El potenciador es un elemento de ADN que actúa en cis, normalmente de aproximadamente 100 a 300 pb, que actúa sobre un promotor para potenciar su actividad de inicio de la transcripción. Se han encontrado potenciadores 5' [Laimins, y col., Proc. Natl. Acad. Sci. USA, 78: 993 (1981)] y 3' [Lusky, y col., Mol. Cell Bio., 3: 1108 (1983)] con orientación y posición relativamente independiente con respecto a la unidad de transcripción, con un intrón [Banerji, y col., Cell. 33: 729 (1983)] así como dentro de la propia secuencia codificante [Osborne, y col., Mol. Cell Bio., 4: 1293 (1984)]. Sin embargo, para la presente invención, el elemento potenciador puede localizarse aguas arriba de la secuencia promotora. Actualmente se conocen muchas secuencias potenciadoras procedentes de genes de mamíferos (globina, elastasa, albúmina, alfa-fetoproteína e insulina). Sin embargo, típicamente, se usará un potenciador de un virus de célula eucariota. Los ejemplos incluyen el potenciador del SV40 en el sitio tardío del origen de replicación (100-270 pb), el potenciador del promotor temprano de citomegalovirus, el potenciador de polioma en el lado tardío del origen de replicación y potenciadores de adenovirus.

Los vectores de expresión usados en células huéspedes de mamíferos también pueden contener sitios de poliadenilación. Son ejemplos de regiones de poliadenilación las derivadas de virus tales como, por ejemplo, el SV40 (tardío o temprano) o el VHB.

Un origen de replicación puede proporcionarse bien por construcción del vector para incluir un origen exógeno, tal como el que puede derivar del SV40 o de otra fuente viral (por ejemplo, Polioma, Adeno, VSV, BPV) o bien puede proporcionarlo la célula huésped. Si el vector está integrado en el cromosoma de la célula huésped, esto último es a menudo suficiente.

Los vectores de expresión pueden contener adecuadamente un gen de selección, denominado también marcador de selección. Un gen de selección codifica una proteína necesaria para la supervivencia o el crecimiento de una célula huésped transformada con el vector. Los ejemplos de marcadores de selección adecuados para células de mamíferos incluyen dihidrofolato reductasa (DHFR), timidina quinasa (TK) o neomicina. Cuando tales marcadores de selección se transfieren con éxito dentro de una célula huésped de mamífero, la célula huésped de mamífero transformada puede sobrevivir si se somete a presión selectiva.

Existen dos categorías diferentes ampliamente usadas de regímenes selectivos. La primera categoría se basa en el metabolismo de una célula y en el uso de una línea celular mutante que carece de la capacidad de crecer independiente de un medio complementado. Dos ejemplos son las células CHO DHFR y las células LTK de ratón. Estas células carecen de la capacidad de crecer sin la adición de nutrientes tales como timidina o hipoxantina. Dado que estas células carecen de determinados genes necesarios para una ruta completa de síntesis de nucleótidos, no pueden sobrevivir a menos que se proporcionen los nucleótidos ausentes en un medio complementado. Una alternativa a la complementación del medio es introducir un gen DHFR o TK intacto en las células que carecen de los genes respectivos, modificando así sus necesidades de crecimiento. Las células individuales no transformadas con el gen DHFR o TK no serán capaces de sobrevivir en un medio no complementado. Por lo tanto, la selección directa de estas células requiere el crecimiento celular en ausencia de nutrientes complementarios.

La segunda categoría es la selección dominante, que se refiere a un esquema de selección que no requiere el uso de una línea celular mutante. Este procedimiento emplea típicamente un fármaco para detener el crecimiento de una célula huésped. Las células que posean un nuevo gen expresarían una proteína transmisora de resistencia a fármacos y sobrevivirían a la selección. Ejemplos de fármacos usados en la selección dominante incluyen neomicina [Southern and Berg, J. Molec. Appl. Genet., 1: 327 (1982)], ácido micofenólico [Mulligan and Berg, Science, 209: 1422 (1980)], o higromicina [Sugden, y col., Mol. Cell. Biol., 5: 410-413 (1985)]. Los tres ejemplos proporcionados anteriormente emplean genes bacterianos bajo control eucariota para transmitir resistencia al fármaco apropiado, es decir, neomicina (G418 o geneticina), xgpt (ácido micofenólico) o higromicina, respectivamente.

Como se usa en el presente documento, la expresión "secuencias de control" se refiere conjuntamente a secuencias promotoras, sitios de unión a ribosoma, señales de poliadenilación, secuencias de terminación de la transcripción, dominios reguladores aguas arriba, potenciadores y similares, que conjuntamente permiten la transcripción y traducción de una secuencia codificante en una célula huésped. No es necesario que todas estas secuencias de control estén siempre presentes en un vector recombinante siempre y cuando la secuencia de ADN de interés sea capaz de transcribirse y traducirse apropiadamente.

Como se usa en el presente documento, la expresión "unido operativamente" se refiere a una disposición de elementos en la que los componentes descritos de esta manera se configuran de manera que realizan su función normal. Por tanto, las secuencias de control unidas operativamente a una secuencia codificante son capaces de efectuar la expresión de la secuencia codificante. No es necesario que las secuencias de control sean contiguas a la secuencia codificante, siempre y cuando funcionen para dirigir la expresión de las mismas. Por tanto, por ejemplo, entre una secuencia promotora y la secuencia codificante, pueden estar presentes secuencias de intervención no traducidas aunque transcritas y la secuencia promotora puede considerarse aún "unida operativamente" a la secuencia codificante.

Además, una secuencia codificante de ácido nucleico está unida operativamente a otra secuencia codificante de ácido nucleico cuando la región codificante de ambas moléculas de ácido nucleico es capaz de expresión en la misma fase de lectura. No es necesario que las secuencias de ácido nucleico sean contiguas, siempre que sean capaces de expresión en la misma fase de lectura. Por tanto, por ejemplo, las secuencias codificantes de intervención y las regiones codificantes de ácido nucleico especificadas pueden aún considerarse "unidas operativamente".

Por tanto, en la presente invención se proporcionan procedimientos para la expresión de un polipéptido en una célula huésped recombinante de mamífero que comprende cultivar al célula huésped en una condición apropiada, en el que dicha célula huésped comprende un vector de expresión que comprende al menos un polinucleótido GB1 unido directa o indirectamente unido al polinucleótido que codifica dicho polipéptido y purificar dicho polipéptido. El polipéptido puede ser la proteína de Transferencia de Ésteres de Colesterol o la Lipasa Endotelial. En algunos casos, el polinucleótido GB1 es la SEC ID N°: 4, 5 ó 6. Dichas células huéspedes pueden ser células de Ovario de Hámster Chino. Determinadas células de Ovario de Hámster Chino pueden comprender el gen E1A de adenovirus. En otro aspecto de la invención, las células huéspedes son células COS o HEK.

En otro aspecto adicional, al menos un polinucleótido GB1 está unido directa o indirectamente al carboxi terminal del polinucleótido que codifica dicho polipéptido. Al menos un polipéptido GB1 puede unirse al carboxi terminal mediante un sitio de escisión selectivo para dicho polinucleótido que codifica dicho polipéptido. El sitio de escisión de selección puede comprender el sitio de escisión de la proteasa del virus del grabado del tabaco. En otro aspecto, se proporcionan procedimientos para escindir dicho sitio de escisión de selección antes de purificar dicho polipéptido.

Se proporcionan vectores de expresión para su uso en una célula huésped recombinante de mamífero que comprenden al menos un polinucleótido GB1 unido, directa o indirectamente, a un polinucleótido que codifica un polipéptido para la producción de dicho polipéptido.

Se proporcionan células huéspedes recombinantes de mamífero que comprenden un vector de expresión que comprende al menos un polinucleótido GB1 unido, directa o indirectamente, a un polinucleótido que codifica un polipéptido para la producción de dicho polipéptido en dicha célula huésped.

En la presente solicitud, salvo que se indique de otra manera, las técnicas utilizadas pueden encontrarse en cualquiera de las diversas referencias bien conocidas tales como Molecular Cloning: A Laboratory Manual (Sambrook, y col. 1989, Cold Spring Harbor Laboratory Press), Gene Expression Technology (Methods in Enzymology, Vol. 185 editado por D. Goeddel, 1991, Academic Press, San Diego, Calif.), "Guide to Protein Purification" in Methods in Enzymology (M. P. Deutscher, ed., (1990) Academic Press, Inc.); PCR Protocols: A Guide to Methods and Applications (Innis, y col. 1990. Academic Press, San Diego, Calif.), Culture of animal Cells: A Manual of Basic Technique, 2nd Ed. (R. I. Freshney, 1987. Liss, Inc. Nueva York, N.Y.) Gene Transfer and Expression Protocols, págs. 109-128, ed. E. J. Murray, The Human Press Inc. Clifton, N.J.), y el Catálogo Ambión (Ambion, Austin, Tex.).

Los siguientes ejemplos pretenden ilustrar realizaciones específicas actualmente conocidas para la realización práctica de la invención, aunque la invención no debe considerarse limitante a ellos.

Ejemplos

Clonación del domino GB1

5 En una realización un polinucleótido GB1 es un polinucleótido de la SEC ID N°: 4. En los siguientes ejemplos se usó el polinucleótido de la SEC ID N°: 4. El polinucleótido de la SEC ID N°: 4 se sintetizó mediante procedimientos convencionales. Para ayudar a futuras clonaciones, para su uso como un marcador C-terminal, se añadió un sitio de restricción Asp718 en el extremo 5' del gen y un codón de terminación con un sitio de restricción Eco RV añadido en el extremo 3' del gen.

Ejemplo 1

Clonación y expresión estable de la Lipasa Endotelial humana en células CHO-E1A

10 Para añadir un sitio de restricción Eco R1 en el extremo 5' del gen de la Lipasa Endotelial (LE) se diseñaron cebadores PCR. El cebador para el extremo 3' del gen se diseñó para eliminar el codón de terminación, añadir un sitio de escisión de la proteasa del VGT e introducir un sitio de restricción Asp 718.

Estos dos fragmentos se ligaron junto con el vector pCDN, que se digirió con EcoR1 y Eco RV, usando técnicas convencionales, para crear pCDN-LE/VGT/GB1.

15 Para crear una línea celular estable, el plásmido pCDN-LE/VGT/GB1, se linealizó usando la enzima de restricción Not 1 y se electroporó en la línea de células CHO-E1A, Acc-317, usando la técnica descrita en Hensley, y col. J. Biol. Chem., 269:23949-23958 (1994). Las células se seleccionaron como una población policlonal en medio MR1-4 1X, que contenía BSA sin nucleósidos. Para proporcionar medio para la purificación, la línea celular policlonal se graduó en 1,5 litros de medio MR1-4 2X en un frasco de 3 litros, 7,5 X 10E+5 células/ml a 34 °C durante 7 días.

20 La Lipasa Endotelial, LE/VGT/GB1 se purificó de la siguiente manera. La LE se capturó del medio sobre resina Fc (flujo rápido Fc humano 4 mg /Sefarosa activada con NHS; GE Healthcare). La resina Fc se lavó cuidadosamente con fosfato de sodio 25 mM, cloruro de sodio 0,15 M y glicerol al 20%, pH=7 y se eluyó con trietilamina 0,1 M pH=11,6. El eluado se neutralice con fosfato de sodio 1M, pH=6. Los rendimientos fueron ~0,9 mg de proteína total/resina FC ml. Los niveles de expresión iniciales superaron los 20 mg/l.

25 Adicionalmente, el plásmido pCDN-LE/VGT/GB1 se linealizó usando la enzima de restricción Not 1 y se electroporó en la línea de células CHO-Lec-E1A, Acc-1169. Las células se seleccionaron como una población clonal en placas de 96 pocillos usando medio MR1-4 1X que contenía BSA, nucleósidos y G418 (Geneticina) 400 ug/ml.

30 Para proporcionar un medio para la purificación, se graduó un solo clon en medio MR1-4 2X con nucleósidos en un frasco de 3 litros, 7,5 X 10E+5 células/ml a 34 °C durante 7 días. Después de la purificación la proteína se sometió a análisis N-term que produjo la secuencia de SPVPGPEGRL (SEC ID N°: 7) y está coincidió con la secuencia prevista. También se realizó MALDI-MS y el P.M. observado, 56302 & 50468 (C-term truncado), coincidió con el P.M. esperado, 56409. El tratamiento con endo H a pH=7 se realizó con éxito y después de la digestión se mantuvo la actividad completa.

35 Para determinar actividad de la enzima, la proteína LE-VGT-GB1 purificada se ensayó para detectar actividad fosfolipasa usando PED 6 como sustrato.

Basándose en los resultados de la purificación de una construcción de pCDN-LE-VGT-His, la línea celular anterior produjo aproximadamente 27 veces más proteína que la construcción LE/His. Ambas construcciones eran activas. Basándose en los resultados de la purificación de la construcción LE-VGT-Fc la construcción Fc produjo aproximadamente 3 veces más proteína; sin embargo, la construcción Fc era inactiva.

40 **Ejemplo 2**

Clonación y expresión estable de la PTEC humana en células CHO-E1A

Para añadir un sitio de restricción Eco R1 en el extremo 5' del gen PTEC (proteína de Transferencia de Ésteres de Colesterol) humana se diseñaron cebadores PCR. El cebador para el extremo 3' del gen se diseñó para añadir un sitio de escisión de la proteasa del VGT e introducir un sitio de restricción Asp 718.

45 Se generaron cebadores PCR para añadir un sitio Asp 718 en el extremo 5' del dominio GB1 y un marcador de epítipo de 6 Histidinas, un codón de terminación seguido de un sitio de restricción Eco RV en el extremo 3' del gen. Estos fragmentos se ligaron junto con el vector pCDN, que se digirió con EcoR1 y Eco RV usando técnicas convencionales para crear el pCDN-PTEC/VGT/GB1-His.

50 Para crear una línea celular estable, el plásmido pCDN-PTEC/VGT/GB1-His, se linealizó usando la enzima de restricción Not 1 y se electroporó en la línea de células CHO-E1A, Acc-317, usando al técnica descrita en Hensley, y col. J. Biol. Chem., 269:23949-23958 (1994). Las células se seleccionaron en placas de 96 pocillos como una población clonal en medio MR1-4 1X que contenía BSA sin nucleósidos.

Para determinar la actividad de los diversos clones celulares que se generaron, el medio acondicionado concentrado se ensayó usando el Kit de Ensayo de Actividad de la Proteína de Transferencia de Ésteres de Colesterol (Roar Biomedical, Inc.) con la PTECr convencional suministrada por Cardiovascular Targets, Inc. Los resultados indican que la proteína es activa.

- 5 Para proporcionar medio para purificación, el clon número 15 se graduó en 2 litros de medio MR1-4 2X con nucleósidos en un frasco de 3 litros, con $7,5 \times 10^5$ células/ml a 34 °C durante 14 días. El medio acondicionado se purificó usando una columna de resina NiNTA con un lavado de imadazol 30 mM seguido de una elución de imadazol 300 mM. Los resultados de purificación indican niveles de expresión de aproximadamente 2,5 mg/l. El análisis N-term indicó una secuencia de SKGTSHEAGIVXRI (SEC ID N°:8), que era idéntica a la secuencia prevista.

10 LISTADO DE SECUENCIAS

<110> Kyung Oh JOHANSON
John J. TRILL

<120> NUEVO PROCEDIMIENTO

<130> PU61420

- 15 <140> desconocido
<141> 19-05-2005

<150> 60/683254
<151> 20-05-2005

<160> 8

- 20 <170> FastSEQ para Windows Versión 4.0

<210> 1
<211> 56
<212> PRT
<213> Streptococcus

- 25 <400> 1

```

Asp Thr Tyr Lys Leu Ile Leu Asn Gly Lys Thr Leu Lys Gly Glu Thr
 1           5           10           15
Thr Thr Glu Ala Val Asp Ala Ala Thr Ala Glu Lys Val Phe Lys Gln
           20           25           30
Tyr Ala Asn Asp Asn Gly Val Asp Gly Glu Trp Thr Tyr Asp Asp Ala
           35           40           45
Thr Lys Thr Phe Thr Val Thr Glu
           50           55
    
```

<210> 2
<211> 57
<212> PRT
30 <213> Streptococcus

<400> 2

```

Met Asp Thr Tyr Lys Leu Ile Leu Asn Gly Lys Thr Leu Lys Gly Glu
 1           5           10           15
Thr Thr Thr Glu Ala Val Asp Ala Ala Thr Ala Glu Lys Val Phe Lys
           20           25           30
Gln Tyr Ala Asn Asp Asn Gly Val Asp Gly Glu Trp Thr Tyr Asp Asp
           35           40           45
Ala Thr Lys Thr Phe Thr Val Thr Glu
           50           55
    
```

- <210> 3
<211> 66
35 <212> PRT
<213> Streptococcus

ES 2 379 096 T3

<400> 3

```

Met Glu Ile Leu Ala Ala Leu Pro Lys Thr Asp Thr Tyr Lys Leu Ile
 1          5          10          15
Leu Asn Gly Lys Thr Leu Lys Gly Glu Thr Thr Thr Glu Ala Val Asp
 20          25          30
Ala Ala Thr Ala Glu Lys Val Phe Lys Gln Tyr Ala Asn Asp Asn Gly
 35          40          45
Val Asp Gly Glu Trp Thr Tyr Asp Asp Ala Thr Lys Thr Phe Thr Val
 50          55          60
Thr Glu
 65

```

<210> 4

<211> 168

5 <212> ADN

<213> Streptococcus

<400> 4

```

gacacttaca aattaatcct taatggtaaa acattgaaag gcgaaacaac tactgaagct 60
gttgatgctg ctactgcaga aaaagtcttc aaacaatagc ctaacgacaa cgggtgttgac 120
ggtgaatgga cttacgacga tgcgactaag acctttacag ttactgaa 168

```

<210> 5

10 <211> 171

<212> ADN

<213> Streptococcus

<400> 5

```

atggacactt acaaattaat ccttaatggt aaaacattga aaggcgaaac aactactgaa 60
gctgttgatg ctgctactgc agaaaaagtc ttcaacaat acgctaacga caacgggtgtt 120
gacggtgaat ggacttacga cgatgcgact aagaccttta cagttactga a 171

```

15 <210> 6

<211> 198

<212> ADN

<213> Streptococcus

<400> 6

```

atggaaattt tagctgcatt acctaagact gacacttaca aattaatcct taatggtaaa 60
acattgaaag gcgaaacaac tactgaagct gttgatgctg ctactgcaga aaaagtcttc 120
aaacaatagc ctaacgacaa cgggtgttgac ggtgaatgga cttacgacga tgcgactaag 180
acctttacag ttactgaa 198

```

20

<210> 7

<211> 11

<212> PRT

<213> humana

25 <400> 7

```

Ser Pro Val Pro Phe Gly Pro Glu Gly Arg Leu
 1          5          10

```

<210> 8

<211> 14

<212> PRT

30 <213> humana

<220>

ES 2 379 096 T3

<221> VARIANTE

<222> 12

<223> Xaa = Cualquier Aminoácido

<400> 8

Ser Lys Gly Thr Ser His Glu Ala Gly Ile Val Xaa Arg Ile
1 5 10

5

REIVINDICACIONES

- 5 1. Un procedimiento para la expresión de un polipéptido en una célula huésped recombinante de mamífero que comprende cultivar la célula huésped en condiciones apropiadas, en el que dicha célula huésped comprende un vector de expresión que comprende al menos un polinucleótido GB1 unido directa o indirectamente a un polinucleótido que codifica dicho polipéptido, y purificar tal polipéptido,
- en el que dicho polinucleótido GB1 codifica una secuencia seleccionada de la SEC ID N°: 1, 2 ó 3 o una secuencia con al menos una identidad del 90% con una secuencia seleccionada de la SEC ID N°: 1, 2 ó 3.
2. El procedimiento de la reivindicación 1, en el que dicho al menos un polinucleótido GB1 es la SEC ID N°: 4, 5 ó 6.
- 10 3. El procedimiento de la reivindicación 1 ó 2, en el que dichas células huéspedes son células de Ovario de Hámster Chino.
4. El procedimiento de la reivindicación 1 ó 2, en el que dichas células huéspedes son células de Ovario de Hámster Chino que contienen el gen E1A de adenovirus.
5. El procedimiento de la reivindicación 1 ó 2, en el que dichas células huéspedes son células COS o HEK.
- 15 6. El procedimiento de una cualquiera de las reivindicaciones anteriores, en el que dicho al menos un polinucleótido GB1 está unido directamente o indirectamente al extremo carboxi terminal del polinucleótido que codifica dicho polipéptido.
7. El procedimiento de una cualquiera de las reivindicaciones 1 a 5, en el que dicho al menos un polinucleótido GB1 está unido mediante un sitio de escisión selectivo a dicho polinucleótido que codifica dicho polipéptido en el extremo carboxi terminal.
- 20 8. El procedimiento de la reivindicación 7, en el que el sitio de escisión de selección comprende el sitio de escisión de la proteasa del virus del grabado del tabaco.
9. El procedimiento de la reivindicación 7 u 8, que comprende adicionalmente escindir dicho sitio de escisión de selección antes de purificar dicho polipéptido.
- 25 10. El procedimiento de una cualquiera de las reivindicaciones anteriores, en el que dicho polipéptido es una proteína de Transferencia de Ésteres de Colesterol o una proteína Lipasa Endotelial.