

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 665 920**

51 Int. Cl.:

C07K 16/00 (2006.01)

C12N 15/09 (2006.01)

C12N 15/67 (2006.01)

C12N 15/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **25.06.2008 PCT/EP2008/005136**

87 Fecha y número de publicación internacional: **08.01.2009 WO09003623**

96 Fecha de presentación y número de la solicitud europea: **25.06.2008 E 08773643 (5)**

97 Fecha y número de publicación de la concesión europea: **14.02.2018 EP 2167540**

54 Título: **Mutante de cadena pesada para mejorar la producción de inmunoglobulina**

30 Prioridad:

29.06.2007 EP 07012774

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

30.04.2018

73 Titular/es:

**F. HOFFMANN-LA ROCHE AG (100.0%)
GRENZACHERSTRASSE 124
4002 BASEL, CH**

72 Inventor/es:

**GOEPFERT, ULRICH;
HANSEN, SILKE;
KNOETGEN, HENDRIK;
KOPETZKI, ERHARD y
PLOETTNER, OLIVER**

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 665 920 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

Mutante de cadena pesada para mejorar la producción de inmunoglobulina

- 5 La presente solicitud describe procedimientos y ácidos nucleicos útiles en la producción de inmunoglobulinas en células de mamífero.

Antecedentes de la invención

- 10 Los sistemas de expresión para la producción de polipéptidos recombinantes se conocen bien y se describen en la literatura del estado de la técnica. Para la producción de polipéptidos usados en aplicaciones farmacéuticas se emplean células de mamífero tales como células CHO, células BHK, células NS0, células Sp2/0, células COS, células HEK, células PER.C6® y similares. El ácido nucleico que codifica el polipéptido se introduce en la célula, por ejemplo, en un plásmido, tal como, por ejemplo, un plásmido de expresión. Los elementos esenciales de un plásmido de expresión son una unidad de propagación de plásmidos procariotas, por ejemplo, para *Escherichia coli*, que comprende un origen de replicación y un marcador de selección, un marcador de selección eucariótico y uno o más casetes de expresión para la expresión del ácido nucleico o ácidos nucleicos de interés, comprendiendo cada uno de ellos un promotor, un gen estructural y un terminador de la transcripción que incluye una señal de poliadenilación. Para la expresión transitoria en células de mamífero se puede incluir un origen de replicación de mamífero, tal como el SV40 Ori u OriP. Como promotor, se puede seleccionar un promotor constitutivo o inducible. Para la transcripción optimizada, se puede incluir una secuencia de Kozak en la región no traducida 5'. Para el procesamiento de ARNm también se puede incluir una señal de poliadenilación.

- 25 Las proteínas, y especialmente las inmunoglobulinas, desempeñan un papel importante en la cartera médica actual. Para la aplicación en humanos, cada sustancia farmacéutica debe cumplir criterios distintos. Para garantizar la seguridad de los agentes biofarmacéuticos en los seres humanos, se deben eliminar especialmente las sustancias que provocarían daños graves.

- 30 El corte y empalme de ARNm se regula por la aparición de un lugar donante de corte y empalme en combinación con un lugar aceptador de corte y empalme, que se localizan en el extremo 5' y en el extremo 3' de un intrón, respectivamente. De acuerdo con Watson *et al.* (Watson *et al.* (Eds), Recombinant DNA: A Short course, Scientific American Books, distribuido por W.H. Freeman and Company, Nueva York, Nueva York, EE. UU. (1983)) son la secuencia de consenso del lugar donante de corte y empalme 5' ag|gtagt (exón|intrón) y del lugar aceptador de corte y empalme 3' (y)_nNcag|g (intrón|exón) (r=base de purina; y=base de pirimidina; n=número entero; N=cualquier base natural).

- 40 En 1980 se publicaron los primeros artículos que tratan sobre el origen de las formas secretadas y fijadas a la membrana de las inmunoglobulinas. La formación de la isoforma secretada (slg) y de la isoforma fijada a la membrana (mlg) resulta del corte y empalme alternativo del pre-ARNm de la cadena pesada. En la isoforma mlg se usan un lugar donante de corte y empalme en el exón que codifica el dominio C-terminal de la forma secretada (es decir, el dominio CH3 o CH4, respectivamente) y un lugar aceptador de corte y empalme localizado a una distancia en dirección 3' del mismo para unir la región constante con los exones en dirección 3' que codifican el dominio transmembrana.

- 45 Un procedimiento para preparar moléculas de ácido nucleico sintético que tiene características transcripcionales inapropiadas o no deseadas cuando se expresa en una célula huésped particular se describe en el documento WO 2002/016944. En el documento WO 2006/042158 se describen moléculas de ácido nucleico modificadas para potenciar la expresión de proteínas recombinantes y/o reducir o eliminar los subproductos de corte y empalme erróneo y/o de lectura íntegra de intrón.

Por lo tanto, existe la necesidad de un procedimiento para la producción recombinante para inmunoglobulinas con subproductos reducidos.

- 55 La BASE DE DATOS EMBL contenía, bajo el n.º de acceso EMBL:DB343679, la secuencia de ARNm del extremo 3' del clon de ADNc de *Homo sapiens* THYMU2022513 y, bajo el n.º de acceso EMBL:AW394005, la secuencia de ARNm de ADNc de *Homo sapiens* QV0-TT0010-091199-053-g03 TT0010.

- 60 Brown, S.L., *et al.* describieron la regulación del desarrollo de ARNm de Ig gamma 2b de membrana y secretora (J. Immunol. 142 (1989) 2072-2080). El efecto de las secuencias del intrón sobre los niveles de expresión de los ADNc de Ig fue descrito por O4: Khamlichi, A.A., *et al.* (Gene 150_(1994) 387-390). Danner, D., *et al.* describieron el papel de un lugar de adición de ácido poliadenílico de escisión de ARN en la producción de ARN mensajero de inmunoglobulina M fijada a la membrana y secretada (Proc. Natl. Acad. Sci. USA 82 (1985) 8658-8662).

65

En el documento WO 2006/122822 se describió la expresión de alto nivel de anticuerpo recombinante en una célula huésped de mamífero. Un procedimiento recombinante para la producción de un anticuerpo monoclonal frente al CD52 para el tratamiento de la leucemia linfocítica crónica se describió en el documento WO 2006/126068. En el documento WO 2006/126069 se describió un procedimiento para la producción de un anticuerpo monoclonal frente al CD20 para el tratamiento del linfoma de linfocitos B.

Sumario de la invención

La invención se define en las reivindicaciones.

En el presente documento se divulga un ácido nucleico que codifica la secuencia de aminoácidos de la parte C-terminal del dominio C_{H3} de una inmunoglobulina de la clase IgA o IgG, o la parte C-terminal del dominio C_{H4} de una inmunoglobulina de la clase IgE o IgM, en el que el dipéptido glicina-lisina comprendido en la secuencia de aminoácidos primaria de la parte C-terminal del dominio C_{H3} o C_{H4} es codificado por el ácido nucleico ggaaaa o el ácido nucleico ggcaaa o el ácido nucleico gggaag o el ácido nucleico ggcaag o el ácido nucleico gggaag.

Un aspecto de la invención es un ácido nucleico que codifica una cadena pesada de inmunoglobulina con organización exón-intrón genómica retenida que comprende un ácido nucleico que codifica la secuencia de aminoácidos de la parte C-terminal del dominio C_{H3} de una inmunoglobulina de la clase IgG, caracterizado por que el dipéptido glicina-lisina en el extremo C de dicha secuencia de aminoácidos de la parte C-terminal del dominio C_{H3} es codificada por el ácido nucleico ggcaaa.

En un modo de realización, el ácido nucleico que codifica el dipéptido glicina-lisina va precedido por el nucleótido g o a.

El segundo aspecto de la presente invención es un plásmido que comprende el ácido nucleico de acuerdo con la invención, y el tercer aspecto de la invención es una célula que comprende el ácido nucleico de acuerdo con la invención.

Un aspecto adicional de la invención es un procedimiento para la producción de una inmunoglobulina en una célula de mamífero que comprende las siguientes etapas:

- a) transfectar una célula de mamífero con un ácido nucleico que codifica una cadena pesada de inmunoglobulina de acuerdo con la invención,
- b) cultivar la célula de mamífero transfectada en condiciones adecuadas para la expresión de la inmunoglobulina,
- c) recuperar la inmunoglobulina del cultivo o de la célula.

En un modo de realización, la célula de mamífero es una célula CHO, una célula BHK, una célula NS0, una célula Sp2/0, una célula COS, una célula HEK o una célula PER.C6[®]. Preferentemente, la célula de mamífero es una célula CHO o una célula BHK o una célula PER.C6[®]. En otro modo de realización, la célula de mamífero se transfecta con dos ácidos nucleicos, en la que el primer ácido nucleico comprende un casete de expresión que codifica una cadena ligera de inmunoglobulina, y en la que el segundo ácido nucleico comprende un casete de expresión que codifica una cadena pesada de inmunoglobulina que comprende un ácido nucleico de acuerdo con la invención.

Descripción detallada de la invención

En el presente documento se divulga un ácido nucleico que codifica la secuencia de aminoácidos de la parte C-terminal del dominio C_{H3} de una inmunoglobulina de la clase IgA o IgG, o la parte C-terminal del dominio C_{H4} de una inmunoglobulina de la clase IgE o IgM, en el que el dipéptido glicina-lisina comprendido en la secuencia de aminoácidos de la parte C-terminal del dominio C_{H3} o C_{H4} es codificado por el ácido nucleico ggaaaa o el ácido nucleico ggcaaa o el ácido nucleico gggaag o el ácido nucleico ggcaag o el ácido nucleico gggaag, y en el que el ácido nucleico que codifica el dipéptido glicina-lisina va precedido opcionalmente por el nucleótido g o por el nucleótido a.

Los procedimientos y técnicas útiles para llevar a cabo la presente invención son conocidos para un experto en la técnica y se describen, por ejemplo, en Ausubel, F.M., ed., Current Protocols in Molecular Biology, Volúmenes I a III (1997) y en Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, segunda edición, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989). Como conoce un experto en la técnica, la producción de numerosos derivados de un ácido nucleico y/o polipéptido hace posible el uso de tecnología de ADN recombinante. Dichos derivados, por ejemplo, se pueden modificar en una posición individual o en varias posiciones mediante sustitución, alteración, intercambio, delección o inserción. La modificación u obtención de

derivados se puede llevar a cabo, por ejemplo, por medio de mutagénesis dirigida al lugar. Dichas modificaciones pueden ser llevadas a cabo fácilmente por un experto en la técnica (véase, por ejemplo, Sambrook, J., *et al.*, Molecular Cloning: A laboratory manual (1999) Cold Spring Harbor Laboratory Press, Nueva York, EE. UU.). El uso de tecnología recombinante hace posible que un experto en la técnica transforme diversas células huéspedes con ácido(s) nucleico(s) heterólogo(s). Aunque el mecanismo de transcripción y traducción, es decir, la expresión, de diferentes células usa los mismos elementos, las células que pertenecen a diferentes especies pueden tener, entre otras cosas, un denominado 'uso de codón' diferente. De este modo, diferente(s) ácido(s) nucleico(s) pueden codificar polipéptidos idénticos (con respecto a la secuencia de aminoácidos). También, debido a la degeneración del código genético, diferentes ácidos nucleicos pueden codificar el mismo polipéptido.

Un "ácido nucleico", como se usa en el presente documento, se refiere a una molécula polimérica que consiste en los nucleótidos individuales (también denominados bases) a, c, g y t (o u en el ARN), por ejemplo, para ADN, ARN o modificaciones de los mismos. Esta molécula de polinucleótido puede ser una molécula de polinucleótido de origen natural o una molécula de polinucleótido sintético o una combinación de una o más moléculas de polinucleótido de origen natural con una o más moléculas de polinucleótido sintético. También se abarcan en esta definición las moléculas de polinucleótido de origen natural en las que se cambian (por ejemplo, mediante mutagénesis), eliminan o añaden uno o más nucleótidos. Un ácido nucleico se puede aislar o integrar en otro ácido nucleico, por ejemplo, en un casete de expresión, un plásmido o el cromosoma de una célula huésped. Un ácido nucleico se caracteriza asimismo por su secuencia de ácido nucleico que consiste en nucleótidos individuales.

Un experto en la técnica conoce bien los procedimientos y técnicas para convertir una secuencia de aminoácidos, por ejemplo, de un polipéptido, en una secuencia de ácido nucleico correspondiente que codifica esa secuencia de aminoácidos. Por lo tanto, un ácido nucleico se caracteriza por su secuencia de ácido nucleico que consiste en nucleótidos individuales y asimismo por la secuencia de aminoácidos de un polipéptido codificado por el mismo.

El término "plásmido" incluye, por ejemplo, plásmidos propagadores y de expresión, así como plásmidos de transfección. El término "vector" se usa como sinónimo de "plásmidos" dentro de la presente solicitud. Típicamente, un "plásmido" también comprenderá un origen de replicación (por ejemplo, el origen de replicación ColE1 u oriP) y un marcador de selección (por ejemplo, un marcador de selección de ampicilina, kanamicina, tetraciclina o cloranfenicol), para la replicación y selección, respectivamente, del plásmido en bacterias.

Un "casete de expresión" se refiere a una construcción que contiene los elementos reguladores necesarios, tales como promotor y lugar de poliadenilación, para la expresión de al menos el ácido nucleico contenido en una célula.

Un "marcador de selección" es un ácido nucleico que permite que las células que transportan el marcador de selección se seleccionen específicamente a favor o en contra, en presencia de un agente de selección correspondiente. Típicamente, un marcador de selección conferirá resistencia a un fármaco o compensará un defecto metabólico o catabólico en la célula huésped. Un marcador de selección puede ser positivo, negativo o bifuncional. Un marcador de selección positivo útil es un gen de resistencia a antibióticos. Este marcador de selección permite que la célula huésped transformada con el mismo se seleccione positivamente en presencia del agente de selección correspondiente, por ejemplo, el antibiótico. Una célula huésped no transformada no puede crecer ni sobrevivir en las condiciones de cultivo selectivo, es decir, en presencia del agente de selección, en el cultivo. Los marcadores de selección positivos permiten la selección de las células que transportan el marcador, mientras que los marcadores de selección negativos permiten que las células que transportan el marcador se eliminen selectivamente. Los marcadores de selección usados con células eucariotas incluyen, por ejemplo, los genes para la aminoglucósido fosfotransferasa (APH), tales como, por ejemplo, los marcadores de selección de higromicina (hyg), neomicina (neo) y G418, dihidrofolato reductasa (DHFR), timidina cinasa (tk), glutamina sintetasa (GS), asparagina sintetasa, triptófano sintetasa (agente de selección indol), histidinol deshidrogenasa (agente de selección histidinol D) y ácidos nucleicos que confieren resistencia a puomicina, bleomicina, fleomicina, cloranfenicol, zeocina y ácido micofenólico. Se describen genes marcadores adicionales, por ejemplo, en los documentos WO 92/08796 y WO 94/28143.

El término "expresión" como se usa en el presente documento, se refiere a procesos de transcripción y/o traducción que se producen dentro de una célula. El nivel de transcripción de una secuencia de ácido nucleico de interés en una célula se puede determinar basándose en la cantidad de ARNm correspondiente que está presente en la célula. Por ejemplo, el ARNm transcrito de una secuencia de interés se puede cuantificar mediante RT-PCR o mediante hibridación Northern (véase Sambrook *et al.*, 1989, *supra*). Los polipéptidos codificados por un ácido nucleico de interés se pueden cuantificar mediante diversos procedimientos, por ejemplo, mediante ELISA, analizando la actividad biológica del polipéptido o empleando ensayos que son independientes de dicha actividad, tales como la inmunoelectrotransferencia o el radioinmunoanálisis, usando inmunoglobulinas que reconocen y se fijan al polipéptido (véase Sambrook *et al.*, 1989, *supra*).

El término "célula" o "célula huésped" se refiere a una célula en la que se puede introducir/transfectar o se introduce/transfecta un ácido nucleico, por ejemplo, que codifica un polipéptido heterólogo. El término "célula" incluye tanto las células procariotas, que se usan para la propagación de plásmidos, como las células eucariotas, que se usan para la expresión de un ácido nucleico. Preferentemente, las células eucariotas son células de mamífero. Preferentemente, la célula de mamífero se selecciona del grupo de células de mamífero que comprende células CHO (por ejemplo, CHO K1, CHO DG44), células BHK, células NS0, células SP2/0, células HEK 293, células HEK 293 EBNA, células PER.C6[®] y células COS. Como se usa en el presente documento, la expresión "célula" incluye la célula en cuestión y su descendencia. Así, las palabras "transformante" y "célula transformada" incluyen la célula del sujeto primario y células en cultivos derivados de la misma sin tener en cuenta el número de transferencias. También se entiende que toda la descendencia puede no ser exactamente idéntica en cuanto al contenido de ADN, debido a mutaciones deliberadas o accidentales. Se incluye la descendencia variante que tiene la misma función o actividad biológica que la cribada en la célula originalmente transformada.

Un "polipéptido" es un polímero que consiste en aminoácidos unidos por enlaces peptídicos, ya sean producidos natural o sintéticamente. Los polipéptidos de menos de aproximadamente 20 residuos de aminoácidos se pueden denominar "péptidos", mientras que las moléculas que consisten en dos o más polipéptidos o que comprenden un polipéptido de más de 100 residuos de aminoácidos se pueden denominar "proteínas". Un polipéptido también puede comprender componentes no aminoacídicos, tales como grupos carbohidrato, iones metálicos o ésteres de ácido carboxílico. Los componentes no aminoacídicos pueden ser añadidos por la célula en la que se expresa el polipéptido y pueden variar con el tipo de célula. Los polipéptidos se definen en el presente documento en términos de la estructura de su esqueleto de aminoácidos o del ácido nucleico que codifica los mismos. En general, las adiciones tales como grupos carbohidrato no se especifican; no obstante, pueden estar presentes.

El término "aminoácido" como se usa dentro de la presente solicitud, denota un grupo de carboxi α -aminoácidos, que directamente o en forma de un precursor pueden ser codificados por un ácido nucleico. Los aminoácidos individuales son codificados por ácidos nucleicos que consisten cada uno en tres nucleótidos, denominados codones o tripletes de bases. Cada aminoácido es codificado por al menos un codón. Por ejemplo, el aminoácido glicina puede ser codificado por cada uno de los cuatro ácidos nucleicos (codones) gga, ggc, ggg y ggt, mientras que el aminoácido lisina solo puede ser codificado por los dos ácidos nucleicos aaa y aag. Este fenómeno se conoce como "degeneración del código genético". El grupo de aminoácidos comprende alanina (código de tres letras: ala, código de una letra: A), arginina (arg, R), asparagina (asn, N), ácido aspártico (asp, D), cisteína (cys, C), glutamina (gln, Q), ácido glutámico (glu, E), glicina (gly, G), histidina (his, H), isoleucina (ile, I), leucina (leu, L), lisina (lys, K), metionina (met, M), fenilalanina (phe, F), prolina (pro, P), serina (ser, S), treonina (thr, T), triptófano (trp, W), tirosina (tyr, Y) y valina (val, V).

Como se usa en el presente documento, el término "inmunoglobulina" denota una proteína que consiste en uno o más polipéptidos sustancialmente codificados por genes de inmunoglobulina. Esta definición incluye variantes tales como formas mutadas, es decir, formas con sustituciones, deleciones e inserciones de uno o más aminoácidos, formas truncadas en el extremo N-terminal, formas fusionadas, formas quiméricas, así como formas humanizadas. Los genes de inmunoglobulina reconocidos incluyen los diferentes genes de la región constante, así como la miríada de genes de la región variable de la inmunoglobulina de, por ejemplo, primates y roedores. Son preferentes las inmunoglobulinas monoclonales. Cada una de las cadenas polipeptídicas pesada y ligera de una inmunoglobulina puede comprender una región constante (en general, la porción carboxiterminal).

El término "inmunoglobulina monoclonal", como se usa en el presente documento, se refiere a una inmunoglobulina obtenida a partir de una población de inmunoglobulinas sustancialmente homogéneas, es decir, las inmunoglobulinas individuales que comprende la población son idénticas excepto por posibles mutaciones de origen natural que pueden estar presentes en cantidades menores. Las inmunoglobulinas monoclonales son altamente específicas, dirigiéndose contra un único lugar antigénico. Además, a diferencia de las preparaciones de inmunoglobulinas policlonales, que incluyen diferentes inmunoglobulinas dirigidas contra diferentes lugares antigénicos (determinantes o epítomos), cada inmunoglobulina monoclonal se dirige contra un único lugar antigénico en el antígeno. Además de por su especificidad, las inmunoglobulinas monoclonales son ventajosas porque se pueden sintetizar sin contaminarse con otras inmunoglobulinas. El modificador "monoclonal" indica el carácter de la inmunoglobulina como obtenida a partir de una población sustancialmente homogénea de inmunoglobulinas y no se debe interpretar como que requiere la producción de la inmunoglobulina mediante cualquier procedimiento particular.

Las formas "humanizadas" de inmunoglobulinas no humanas (por ejemplo, de roedores) son inmunoglobulinas quiméricas que contienen secuencias parciales derivadas de una inmunoglobulina no humana y de una inmunoglobulina humana. En su mayor parte, las inmunoglobulinas humanizadas se derivan de una inmunoglobulina humana (inmunoglobulina receptora), en la que los residuos de una región hipervariable se reemplazan por residuos de una región hipervariable de una especie no humana (inmunoglobulina donante), tal como ratón, rata, conejo o primate no humano, que tiene la especificidad y afinidad deseadas (véase, por

ejemplo, Morrison, S.L., *et al.*, Proc. Natl. Acad. Sci. USA 81 (1984) 6851-6855; documentos US 5.202.238 y US 5.204.244). En algunos casos, los residuos de la región estructural (FR) de la inmunoglobulina humana se reemplazan por los residuos no humanos correspondientes. Además, las inmunoglobulinas humanizadas pueden comprender modificaciones adicionales, por ejemplo, residuos de aminoácidos que no se encuentran en la inmunoglobulina receptora ni en la inmunoglobulina donante. Dichas modificaciones dan como resultado variantes de dicha inmunoglobulina receptora o donante que son homólogas pero no idénticas a la secuencia original correspondiente. Estas modificaciones se hacen para refinar adicionalmente el rendimiento de la inmunoglobulina. En general, la inmunoglobulina humanizada comprenderá sustancialmente todos los dominios variables, al menos uno y típicamente dos, en los que todos o sustancialmente todos los bucles hipervariables se corresponden con los de una inmunoglobulina donante no humana y todas o sustancialmente todas las FR son las de una inmunoglobulina receptora humana. La inmunoglobulina humanizada opcionalmente también comprenderá al menos una porción de una región constante de inmunoglobulina, típicamente la de una inmunoglobulina humana. Los procedimientos para humanizar inmunoglobulinas no humanas se han descrito en la técnica. Preferentemente, una inmunoglobulina humanizada tiene uno o más residuos de aminoácidos introducidos en ella a partir de una fuente que no es humana. Estos residuos de aminoácidos no humanos se denominan a menudo residuos de "importación", que típicamente se toman de un dominio variable de "importación". La humanización se puede realizar esencialmente siguiendo el procedimiento de Winter y colaboradores mediante la sustitución de secuencias de la región hipervariable por las secuencias correspondientes de una inmunoglobulina no humana. En consecuencia, dichas inmunoglobulinas "humanizadas" son inmunoglobulinas quiméricas en las que sustancialmente se ha sustituido menos de un dominio variable humano intacto por la secuencia correspondiente de una especie no humana. En la práctica, las inmunoglobulinas humanizadas son típicamente inmunoglobulinas humanas en las que algunos residuos de la región hipervariable y posiblemente algunos residuos de la región estructural se sustituyen por residuos de lugares análogos en inmunoglobulinas de roedores o de primates no humanos.

Dependiendo de la secuencia de aminoácidos de la región constante de sus cadenas pesadas, las inmunoglobulinas se dividen en las clases: IgA, IgD, IgE, IgG e IgM. Algunas de estas se pueden dividir adicionalmente en subclases (isotipos), por ejemplo, IgG1, IgG2, IgG3 e IgG4, IgA1 e IgA2. De acuerdo con las regiones constantes de cadena pesada, las diferentes clases de cadenas pesadas de inmunoglobulinas se denotan como cadena α , δ , ϵ , γ y μ , respectivamente. La región constante de una cadena pesada de una inmunoglobulina humana de longitud completa de la clase IgA, IgD e IgG está constituida por un dominio constante 1 (denotado en el presente documento como C_{H1}), una región bisagra, un dominio constante 2 (C_{H2}) y un dominio constante 3 (C_{H3}). Las inmunoglobulinas humanas de la clase IgE e IgM comprenden un cuarto dominio constante adicional (C_{H4}). Además, están las inmunoglobulinas de polímeros de la clase IgM que comprenden múltiples inmunoglobulinas, por ejemplo, cinco o seis, unidas covalentemente mediante enlaces disulfuro. La secuencia de aminoácidos del dominio constante C-terminal de estas diferentes clases de inmunoglobulinas humanas se enumera en la tabla 1. El primer aminoácido de SEQ ID NO: 01 a 08 puede estar presente o no, debido a que este aminoácido es codificado por dos exones, el exón que codifica el dominio constante C-terminal y el exón que codifica el dominio precedente.

Tabla 1: Secuencia de aminoácidos C-terminal de las diferentes clases de inmunoglobulina.

Los dipéptidos glicina-lisina aparecen subrayados.

Clase de inmunoglobulina (dominio constante C-terminal)	Secuencia de aminoácidos de la parte C-terminal (en la dirección N-terminal a C-terminal)	SEQ ID NO:
IgA1 (C_{H3}), IgA2 (C_{H3})	GNTFRPEVHL LPPPSEELAL NELVTTLTCLA	01
	RGFSPKDVLLV RWLQGSQELP REKYL	
	TWASR QEPSQGTTTF AVTSILRVAA	
	EDWKKGDTFS CMVGHEALPL AFTQK	
	TIDRL <u>AG</u> KPTHVNVS VVMAEVDGTC Y	

Clase de inmunoglobulina (dominio constante C-terminal)	Secuencia de aminoácidos de la parte C-terminal (en la dirección N-terminal a C-terminal)	SEQ ID NO:
IgD (C _H 3)	AAQAPVKLSLN LLASSDPPEA ASWLLCEVSG FSPPNILLMW LEDQREVNTS GFAPARPPPQ PGSTTFWAWS VLRVPAPPSP QPATYTCVVS HEDSRTLNA SRSLEVS	02
IgE (C _H 4)	GPRAAPEVYA FATPEWPGSR DKRTLACLIQ NFMPELISVQ WLHNEVQLPD ARHSTTQPRK TKGSGFFVFS RLEVTRAWE QKDEFICRAV HEAASPSQTV QRAVSVNPGK	03
IgM (C _H 4)	GVALHRPDVY LLPPAREQLN LRESATITCL VTGFSPADV FVQWMQRGQPL SPEKYVTSAP MPEPQAPGRY FAHSILTVSE EEWNTGETYT CVVAHEALPN RVT ERTVDK STGKPTLYNV SLVMSDTAGT CY	04
IgG1 (C _H 3)	GQPREPQVYT LPPSRDELTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTTTPVLDS DGSFFLYSKL TVDKSRWQQG NVFSCSVMHE ALHNHYTQKS LSLSPGK	05
IgG2 (C _H 3)	GQPREPQVYT LPPSREEMTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTTTPMLDS DGSFFLYSKL TVDKSRWQQG NVFSCSVMHE ALHNHYTQKS LSLSPGK	06

Clase de inmunoglobulina (dominio constante C-terminal)	Secuencia de aminoácidos de la parte C-terminal (en la dirección N-terminal a C-terminal)	SEQ ID NO:
IgG3 (C _H 3)	GQPREPQVYT LPPSREEMTK NQVSLTCLVK GFYPSDIAVE WESSGQPENN YNTTPPMLDS DGSFFLYSKL TVDKSRWQQG NIFSCSV MHE ALHNRFTQKS LSLSPGK	07
IgG4 (C _H 3)	GQPREPQVYT LPPSQEEMTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTTPPVLDSDGSFFLYSRL TVDKSRWQEG NVFSCSV MHE ALHNHYTQKS LSLSLGK	08

El término "parte C-terminal del dominio C_H3 de una inmunoglobulina de la clase IgA o IgG, o la parte C-terminal del dominio C_H4 de una inmunoglobulina de la clase IgE o IgM" denota las secuencias de aminoácidos de una cadena pesada de inmunoglobulina, que se localiza en el extremo C-terminal de una cadena pesada de inmunoglobulina de longitud completa o de origen natural, por lo que el extremo C de dicha parte C-terminal es idéntico al extremo C de la secuencia de aminoácidos primaria de la cadena pesada de inmunoglobulina. El término "secuencia de aminoácidos primaria" denota la secuencia de aminoácidos de una cadena pesada de inmunoglobulina después de la traducción del ARNm correspondiente. Esta secuencia de aminoácidos primaria se puede modificar adicionalmente en la célula de expresión después de la traducción de ARNm, por ejemplo, mediante peptidasas que escinden uno o más aminoácidos C-terminales de la secuencia de aminoácidos primaria. Por lo tanto, la secuencia de aminoácidos primaria y la secuencia de aminoácidos secretada pueden no ser idénticas, sino que pueden diferir en algunos aminoácidos en el extremo C. La parte C-terminal puede comprender al menos los 100 aminoácidos C-terminales de una secuencia de aminoácidos primaria de cadena pesada de inmunoglobulina o, preferentemente, al menos los 50 aminoácidos C-terminales de una secuencia de aminoácidos primaria de cadena pesada de inmunoglobulina o, preferentemente, los al menos 20 aminoácidos C-terminales de una secuencia de aminoácidos primaria de cadena pesada de inmunoglobulina.

Las secuencias de aminoácidos del dominio constante C-terminal de las diferentes inmunoglobulinas humanas son codificadas por las secuencias de ADN correspondientes. En el genoma, estas secuencias de ADN contienen secuencias codificadoras (exónicas) y no codificadoras (intrónicas). Después de la transcripción del ADN al pre-ARNm, el pre-ARNm también contiene estas secuencias intrónicas y exónicas. Antes de la traducción, las secuencias intrónicas no codificadoras se eliminan durante el procesamiento del ARNm mediante corte y empalme de las mismas del transcrito de ARNm primario para generar el ARNm maduro. El corte y empalme del ARNm primario se controla mediante un lugar donante de corte y empalme en combinación con un lugar aceptador de corte y empalme propiamente espaciado. El lugar donante de corte y empalme se localiza en el extremo 5' y el lugar aceptador de corte y empalme se localiza en el extremo 3' de una secuencia intrónica y ambos se eliminan parcialmente durante el proceso de corte y empalme del pre-ARNm.

El término "propiamente espaciado" denota que un lugar donante de corte y empalme y un lugar aceptador de corte y empalme en un ácido nucleico se disponen de tal manera que todos los elementos requeridos para el proceso de corte y empalme están disponibles y se encuentran en una posición apropiada para permitir que tenga lugar el proceso de corte y empalme.

En el presente documento se divulga un ácido nucleico que codifica la secuencia de aminoácidos de la parte C-terminal del dominio C_H3 de una inmunoglobulina de la clase IgA o IgG, o la parte C-terminal del dominio C_H4 de una inmunoglobulina de la clase IgE o IgM, en el que el dipéptido glicina-lisina comprendido en dicha secuencia de aminoácidos de la parte C-terminal del dominio C_H3 o C_H4 es codificado por el ácido nucleico ggaaaa o el ácido nucleico ggcaaa o el ácido nucleico gggaaa.

De manera sorprendente, se ha encontrado que con el ácido nucleico de acuerdo con la invención se puede reducir la formación de subproductos indeseables como consecuencia de sucesos de corte y empalme indeseables.

El término "dipéptido glicina-lisina", como se usa dentro de la presente solicitud, denota el péptido que comprende en dirección N-terminal a C-terminal los dos aminoácidos glicina y lisina unidos por un enlace peptídico. El término "dipéptido glicina-lisina", como se usa dentro de la presente solicitud, denota una fracción dipeptídica de un polipéptido o proteína más grande, que se puede encontrar al comienzo, dentro o al final del polipéptido más grande. El aminoácido lisina puede ser codificado por los ácidos nucleicos aaa y aag.

Las secuencias de ácido nucleico que codifican el dominio constante C-terminal de diferentes clases y subclases de inmunoglobulina humana se enumeran en la tabla 2. Para el dominio C_H3 de las IgG1 y las IgG2 humanas, se conocen dos formas variantes. En un modo de realización, el ácido nucleico codifica una parte del dominio constante C-terminal de una cadena pesada de inmunoglobulina.

Tabla 2: Secuencias de ácido nucleico que codifican la parte C-terminal de la cadena pesada de las diferentes clases de inmunoglobulina humana.

Clase de inmunoglobulina (dominio constante C-terminal)	Secuencia de ácido nucleico que codifica la parte C-terminal de la secuencia de aminoácidos de una cadena pesada de inmunoglobulina	SEQ ID NO:
IgA1 (C _H 3), IgA2 (C _H 3)	ggcttcagcc ccaaggacgt gctggttcgc tggctgcagg ggtcacagga gctgccccgc gagaagtacc tgactggggc atcccggcag gagcccagcc agggcaccac caccttcgct gtgaccagca tactgcgcgt ggcagccgag gactggaaga agggggacac cttctcctgc atggtggggc acgagggcct gccgctggcc ttacacaga agaccatcga ccgcttggcg ggtaaaccac cccatgtcaa tgtgtctgtt gtcattggcg aggtggacgg cacctgctac	09
IgD (C _H 3)	taccaccaa cgtccgtgac tgtcacctgg tacatgggga cacagagcca gcccagaga accttcctg agatacaaag acgggacagc tactacatga caagcagcca gctctccacc cccctccagc agtggcgcca aggcgagtac aaatgcgtgg tccagcacac cgccagcaag agtaagaagg agatcttccg ctggccaggt aggtcgcacc ggagatcacc cagaagggcc ccccaggacc ccagcacct tccactcagg gcctgaccac aaagacagaa gcaagggtg	10
IgE (C _H 4)	tttgcgacgc cggagtggcc ggggagccgg gacaagcgca ccctgcctg cctgatccag aacttcacgc ctgaggacat ctcgtgcag tggctgcaca acgaggtgca gctccggac gcccggcaca gcacgacga gcccgcgaag accaagggt ccggcttctt cgtcttcagc cgctggagg tgaccagggc cgaatgggag cagaaagatg agttcatctg ccgtgcagtc catgaggcag cgagcccctc acagaccgtc cagcgagcgg tgtctgtaaa tcccggtaaa	11

Clase de inmunoglobulina (dominio constante C-terminal)	Secuencia de ácido nucleico que codifica la parte C-terminal de la secuencia de aminoácidos de una cadena pesada de inmunoglobulina	SEQ ID NO:
IgM (C _H 4)	acgggcttct ctcccgcgga cgtcttcgtg cagtggatgc agagggggca gcccttgtcc ccggagaagt atgtgaccag cgcccaatg cctgagcccc agggcccagg ccggtacttc gcccacagca tctgaccgt gtccgaagag gaatggaaca cgggggagac ctacacctgc gtggtggccc atgaggccct gccaacagg gtcaccgaga ggaccgtgga caagtcacc ggtaaacca ccctgtacaa cgtgtccctg gtcattgtccg acacagctgg cacctgtac	12
Variante 1 de IgG1 (C _H 3)	gtgtacacc tgccccatc ccgggatgag ctgaccaaga accaggtcag cctgacctgc ctggtcaaag gcttctatcc cagcgacatc gccgtggagt gggagagcaa tgggcagccg gagaacaact acaagaccac gcctcccgtg ctggactccg acggctcctt cttctctac agcaagtca ccgtggacaa gagcaggtgg cagcagggga acgtcttctc atgtccgtg atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggtaaa	13
Variante 2 de IgG1 (C _H 3)	gggcagcccc gagaaccaca ggtgtacacc ctgccccat cccgggatga gctgaccaag aaccaggta gcctgacctg cctggtcaaa ggcttctatc ccagcgacat cgccgtggag tgggagagca atgggcagcc ggagaacaac tacaagacca cgctcccgt gctggactcc gacggctcct tcttctcta cagcaagctc accgtggaca agagcaggtg gcagcagggg aacgtcttct catgctccgt gatgcatgag gctctgcaca accactacac acagaagagc ctctccctgt ctccgggtaa a	28
Variante 1 de IgG2 (C _H 3)	gtgtacacc tgccccatc ccgggaggag atgaccaaga accaggtcag cctgacctgc ctggtcaaag gcttctatcc cagcgacatc gccgtggagt gggagagcaa tgggcagccg gagaacaact acaagaccac acctccatg ctggactccg acggctcctt cttctctac agcaagtca ccgtggacaa gagcaggtgg cagcagggga acgtcttctc atgtccgtg atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggtaaa	14

Clase de inmunoglobulina (dominio constante C-terminal)	Secuencia de ácido nucleico que codifica la parte C-terminal de la secuencia de aminoácidos de una cadena pesada de inmunoglobulina	SEQ ID NO:
Variante 2 de IgG2 (C _H 3)	gggcagcccc gagaaccaca ggtgtacacc ctgcccccat cccgggagga gatgaccaag aaccaggta cctgacctg cctggtcaaa ggcttctacc ccagcgacat ctccgtggag tgggagagca atgggcagcc ggagaacaac tacaagacca cacctcccat gctggactcc gacggctcct tcttctcta cagcaagctc accgtggaca agagcagggtg gcagcagggg aacgtcttct catgctccgt gatgcatgag gctctgcaca accactacac gcagaagagc ctctccctgt ctccgggtaa a	29
IgG3 (C _H 3)	gtgtacaccc tgcccccatc ccgggaggag atgaccaaga accagggtcag cctgacctgc ctggtcaaag gcttctaccc cagcgacatc gccgtggagt gggagagcag cgggcagccg gagaacaact acaacaccac gcctcccatg ctggactccg acggctcctt cttctctac agcaagctca ccgtggacaa gagcagggtgg cagcagggga acatcttctc atgctccgtg atgcatgagg ctctgcacaa ccgcttcacg cagaagagcc tctccctgtc tccgggtaaa	15
IgG4 (C _H 3)	gtgtacaccc tgcccccatc ccaggaggag atgaccaaga accagggtcag cctgacctgc ctggtcaaag gcttctaccc cagcgacatc gccgtggagt gggagagcaa tgggcagccg gagaacaact acaagaccac gcctcccgtg ctggactccg acggctcctt cttctctac agcaggctaa ccgtggacaa gagcagggtgg caggagggga atgtcttctc atgctccgtg atgcatgagg ctctgcacaa ccactacaca cagaagagcc tctccctgtc tctgggtaaa	16

En el presente documento se divulga un ácido nucleico que codifica una parte del dominio constante C-terminal de una cadena pesada de inmunoglobulina de la clase IgA, IgE, IgM o IgG y se selecciona entre los ácidos nucleicos de SEQ ID NO: 17, 18, 19, 20, 21, 22 o 23 o 30 o 31. En el presente documento se divulga un ácido nucleico que codifica una parte del dominio constante C-terminal de una cadena pesada de inmunoglobulina de la clase IgA, IgE, IgM o IgG y se selecciona entre los ácidos nucleicos de SEQ ID NO: 17, 18, 19, 22, 23, 30 o 31. En el presente documento se divulga un ácido nucleico que codifica una parte del dominio constante C-terminal de una cadena pesada de inmunoglobulina de la clase IgA, IgE, IgM o IgG y se selecciona entre los ácidos nucleicos de SEQ ID NO: 17, 18, 19, 22 o 23.

En el presente documento se divulga un ácido nucleico que codifica el dominio constante C-terminal de una inmunoglobulina de la clase IgA, IgE, IgM o IgG, y comprende un ácido nucleico seleccionado entre los ácidos nucleicos de SEQ ID NO: 17, 18, 19, 20, 21, 22 o 23 o 30 o 31, que codifica una parte del dominio C-terminal de la cadena pesada de inmunoglobulina. En el presente documento se divulga un ácido nucleico que codifica el dominio constante C-terminal de una inmunoglobulina de la clase IgA, IgE, IgM o IgG, y comprende un ácido nucleico seleccionado entre los ácidos nucleicos de SEQ ID NO: 17, 18, 19, 22, 23, 30 o 31, que codifica una parte del dominio C-terminal de la cadena pesada de inmunoglobulina. En el presente documento se divulga un ácido nucleico que codifica el dominio constante C-terminal de una inmunoglobulina de la clase IgA, IgE, IgM o IgG, y comprende un ácido nucleico seleccionado entre los ácidos nucleicos de SEQ ID NO: 17, 18, 19, 22 o 23, que codifica una parte del dominio C-terminal de la cadena pesada de inmunoglobulina.

Tabla 3: Secuencias de ácido nucleico que codifican una parte de la secuencia de aminoácidos de dominio constante C-terminal de una cadena pesada de inmunoglobulina de diferentes clases.

Clase de inmunoglobulina	Secuencia de ácido nucleico que codifica la secuencia de aminoácidos C-terminal	SEQ ID NO:
IgA	ggcaaacc ca cccatgtcaa tgtgtctgtt gtcattggcgg aggtggacgg cacctgctac	17
IgE	catgaggcag cgagcccctc acagaccgtc cagcgagcgg tgtctgtaaa tcccggcaaa	18
IgM	ggcaaacc ca cctgtacaa cgtgtccctg gtcattgtccg acacagctgg cacctgctac	19
Variante 1 de IgG1	atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggcaaa	20
Variante 2 de IgG1	atgcatgagg ctctgcacaa ccactacaca cagaagagcc tctccctgtc tccgggcaaa	30
Variante 1 de IgG2	atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggcaaa	21
Variante 2 de IgG2	atctccgtgg agtgggagag caatgggcag ccggagaaca actacaagac cacacctccc atgctggact ccgacggctc cttcttctc tacagcaagc tcaccgtgga caagagcagg tggcagcagg ggaacgtctt ctcatgctcc gtgatgcatg aggctctgca caaccactac acgcagaaga gcctctcct gtctccgggc aaa	31
IgG3	atgcatgagg ctctgcacaa ccgcttcacg cagaagagcc tctccctgtc tccgggcaaa	22
IgG4	atgcatgagg ctctgcacaa ccactacaca cagaagagcc tctccctgtc tctgggcaaa	23

- 5 El término "dominio constante C-terminal" denota el dominio CH3 de una cadena pesada de inmunoglobulina de la clase IgA o IgG, o el dominio CH4 de una cadena pesada de inmunoglobulina de la clase IgE o IgM. La expresión "una parte de" denota una fracción C-terminal del dominio constante C-terminal de una cadena pesada de inmunoglobulina de la clase IgA, IgE, IgM o IgG, de al menos 20 aminoácidos consecutivos o de al menos 50 aminoácidos consecutivos o de al menos 100 aminoácidos consecutivos de la secuencia de aminoácidos primaria contados desde el extremo C en dirección al extremo N de la cadena pesada de inmunoglobulina.

15 La producción recombinante de inmunoglobulinas se conoce bien en el estado de la técnica y se describe, por ejemplo, en los artículos de revisión de Makrides, S.C., *Protein Expr. Purif.* 17 (1999) 183-202; Geisse, S., *et al.*, *Protein Expr. Purif.* 8 (1996) 271-282; Kaufman, R.J., *Mol. Biotechnol.* 16 (2000) 151-160 Werner, R.G., *Arzneimittelforschung - Drug Research* 48 (1998) 870-880.

20 Para la producción de una inmunoglobulina que comprende una secuencia de aminoácidos codificada por un ácido nucleico de acuerdo con la invención, la invención comprende adicionalmente un procedimiento para la producción de una inmunoglobulina en una célula de mamífero que comprende las siguientes etapas:

- a) transfectar la célula de mamífero con un ácido nucleico que codifica una cadena pesada de inmunoglobulina de acuerdo con la invención,
- b) cultivar la célula de mamífero transfectada en condiciones adecuadas para la expresión de la inmunoglobulina,
- c) recuperar la inmunoglobulina del cultivo o de la célula.

El término "en condiciones adecuadas para la expresión de la inmunoglobulina" denota las condiciones que se usan para el cultivo de una célula de mamífero que expresa una inmunoglobulina y que un experto en la técnica conoce o puede determinar fácilmente. Un experto en la técnica también conoce que estas condiciones pueden variar dependiendo del tipo de célula de mamífero cultivada y del tipo de inmunoglobulina expresada. En general, la célula de mamífero se cultiva a una temperatura, por ejemplo, de entre 20 °C y 40 °C, y durante un período de tiempo suficiente para permitir la producción eficaz de proteína de la inmunoglobulina, por ejemplo, durante 4 a 28 días.

En un modo de realización, la transfección de la célula de mamífero se realiza con un ácido nucleico que codifica una cadena pesada de inmunoglobulina, en la que un ácido nucleico seleccionado entre los ácidos nucleicos de SEQ ID NO: 23 o 30 codifica una parte del dominio C-terminal de la cadena pesada de inmunoglobulina. En un modo de realización adicional, la transfección de la célula de mamífero se realiza con un ácido nucleico que codifica una cadena pesada de inmunoglobulina, en la que un ácido nucleico de SEQ ID NO: 23 codifica una parte del dominio C-terminal de la cadena pesada de inmunoglobulina. En un modo de realización, la célula de mamífero se transfecta con un, es decir, uno, ácido nucleico que comprende un casete de expresión que codifica una cadena pesada de inmunoglobulina y un casete de expresión que codifica una cadena ligera de inmunoglobulina. En otro modo de realización, la célula de mamífero se transfecta con dos ácidos nucleicos, uno que comprende un casete de expresión que codifica una cadena ligera de inmunoglobulina, y uno que comprende un casete de expresión que codifica una cadena pesada de inmunoglobulina.

La inmunoglobulina producida con el procedimiento de acuerdo con la invención es preferentemente una inmunoglobulina heteróloga. El término "inmunoglobulina heteróloga" denota una inmunoglobulina que no es producida naturalmente por dicha célula de mamífero. La inmunoglobulina producida de acuerdo con el procedimiento de la invención se produce por medios recombinantes. Dichos procedimientos son ampliamente conocidos en el estado de la técnica y comprenden la expresión de proteínas en células procariotas y eucariotas con la subsiguiente recuperación y aislamiento de la inmunoglobulina heteróloga, y habitualmente, la purificación hasta una pureza farmacéuticamente aceptable. Para la producción, es decir, la expresión, de una inmunoglobulina, se insertan un ácido nucleico que codifica la cadena ligera y un ácido nucleico que codifica la cadena pesada, que comprende el ácido nucleico de acuerdo con la invención, cada uno en un casete de expresión mediante procedimientos estándar. Los ácidos nucleicos que codifican inmunoglobulinas se aíslan y secuencian fácilmente usando procedimientos convencionales. Las células de hibridoma pueden servir como fuente de dichos ácidos nucleicos. Los casetes de expresión se pueden insertar en un plásmido o plásmidos de expresión, que luego se transfectan en células huéspedes, que de lo contrario no producen inmunoglobulinas. La expresión se realiza en células huéspedes procariotas o eucariotas apropiadas y la inmunoglobulina se recupera de las células después de la lisis o del sobrenadante del cultivo.

Diferentes procedimientos están bien establecidos y se usan ampliamente para la recuperación y purificación de proteínas, tales como cromatografía de afinidad con proteínas microbianas (por ejemplo, cromatografía de afinidad con proteína A o proteína G), cromatografía de intercambio iónico (por ejemplo, intercambio catiónico (resinas de carboximetilo), intercambio aniónico (resinas de aminoetilo) e intercambio mixto), adsorción tiorfílica (por ejemplo, con beta-mercaptoetanol y otros ligandos SH), cromatografía de interacción hidrófoba o adsorción aromática (por ejemplo, con fenil-sefarsa, resinas aza-arenofílicas o ácido m-aminofenilborónico), cromatografía de afinidad por quelatos metálicos (por ejemplo, con material de afinidad por Ni(II) y Cu(II)), cromatografía de exclusión por tamaño y procedimientos electroforéticos (tales como electroforesis en gel, electroforesis capilar) (Vijayalakshmi, M.A. Appl. Biochem. Biotech. 75 (1998) 93-102).

La presente invención comprende adicionalmente un ácido nucleico que codifica una cadena pesada de inmunoglobulina que comprende el ácido nucleico divulgado en el presente documento. Además, la invención comprende un plásmido que comprende el ácido nucleico de acuerdo con la invención y una célula que comprende este plásmido.

En el presente documento se divulga un procedimiento para mejorar la expresión de una inmunoglobulina en una célula de mamífero, en el que el ácido nucleico que codifica la cadena pesada de inmunoglobulina comprende el ácido nucleico ggaaaa o el ácido nucleico ggcaaa o el ácido nucleico gggaaa o el ácido nucleico ggaag o el ácido nucleico ggcaag o el ácido nucleico gggaag, que codifica el dipéptido glicina-lisina contenido en el dominio C_H3 o C_H4. Con este ácido nucleico se pueden reducir o suprimir los sucesos de corte y empalme indeseables.

En un modo de realización de la invención, la cadena pesada de inmunoglobulina es una cadena pesada de inmunoglobulina de un anticuerpo humano de la subclase IgG4 o una cadena pesada de inmunoglobulina de un anticuerpo humano de la subclase IgG1, IgG2 o IgG3. En un modo de realización, la cadena pesada de inmunoglobulina es una cadena pesada de inmunoglobulina humana y, preferentemente, de la subclase IgG4 humana o una cadena pesada de inmunoglobulina mutada de la subclase IgG1 humana. En otro modo de realización, la cadena pesada de inmunoglobulina es de la subclase IgG1 humana con las mutaciones L234A y L235A. En un modo de realización adicional, la cadena pesada de inmunoglobulina es una cadena pesada de inmunoglobulina IgG4 humana con la mutación S228P. En un modo de realización, la cadena pesada de inmunoglobulina es de la subclase IgG4 o de la subclase IgG1 o IgG2, con una mutación en L234, L235 y/o D265, y/o contiene la mutación PVA236. En otros modos de realización, las mutaciones son S228P, L234A, L235A, L235E y/o PVA236 (PVA236 significa que la secuencia de aminoácidos ELLG (dada en un código de aminoácidos de una letra) desde la posición de aminoácidos 233 a 236 de la IgG1 o EFLG de la IgG4 se reemplaza por PVA). Son preferentes las mutaciones S228P de IgG4 y L234A y L235A de IgG1 (numeración de acuerdo con el índice UE de Kabat).

Los "elementos reguladores", como se usan en el presente documento, se refieren a secuencias nucleotídicas presentes en cis, necesarias para la transcripción y/o traducción de la secuencia del ácido nucleico que codifica un polipéptido de interés. Los elementos reguladores transcripcionales comprenden normalmente un promotor en dirección 5' de la secuencia del ácido nucleico que se va a expresar, lugares de iniciación y terminación transcripcional y una secuencia señal de poliadenilación. El término "lugar de iniciación transcripcional" se refiere a la base de ácido nucleico en el ácido nucleico correspondiente al primer ácido nucleico incorporado en el transcrito primario, es decir, el precursor del ARNm; el lugar de iniciación transcripcional se puede solapar con la secuencia promotora. El término "lugar de terminación transcripcional" se refiere a una secuencia nucleotídica normalmente representada en el extremo 3' de un gen de interés que se va a transcribir, que provoca que la ARN polimerasa termine la transcripción. La secuencia señal de poliadenilación, o señal de adición de poli-A proporciona la señal para la escisión en un lugar específico en el extremo 3' del ARNm eucariota y la adición postranscripcional en el núcleo de una secuencia de aproximadamente 100-200 nucleótidos de adenina (cola de poliA) al extremo 3' escindido. La secuencia señal de poliadenilación puede incluir la secuencia de consenso AATAAA localizada a aproximadamente 10-30 nucleótidos en dirección 5' del lugar de escisión.

Para producir un polipéptido secretado, el ácido nucleico de interés incluye un segmento de ADN que codifica una secuencia señal / péptido líder. La secuencia señal dirige el polipéptido recién sintetizado hacia y a través de la membrana ER donde el polipéptido se puede encaminar para la secreción. La secuencia señal se escinde mediante una peptidasa señal durante el cruce de la proteína de la membrana ER. En cuanto a la función de la secuencia señal, es esencial el reconocimiento por parte del mecanismo de secreción de la célula huésped. Por lo tanto, la secuencia señal usada debe ser reconocida por las proteínas y enzimas de la célula huésped del mecanismo de secreción.

Los elementos reguladores de la traducción incluyen un codón de iniciación (AUG) y un codón de terminación de la traducción (TAA, TAG o TGA). En algunas construcciones se puede incluir un lugar de entrada de ribosomas interno (IRES).

Un "promotor" se refiere a una secuencia polinucleotídica que controla la transcripción de un gen / gen estructural o secuencia de ácido nucleico a la que está unida funcionalmente. Un promotor incluye señales para la fijación de la ARN polimerasa y la iniciación de la transcripción. Los promotores usados serán funcionales en el tipo celular de la célula huésped en la que se contempla la expresión de la secuencia seleccionada. En la técnica se conocen bien un gran número de promotores, incluyendo promotores constitutivos, inducibles y reprimibles de una variedad de fuentes diferentes (y están identificados en bases de datos tales como GenBank) y están disponibles como o dentro de polinucleótidos clonados (a partir de, por ejemplo, depositarios tales como ATCC, así como otras fuentes comerciales o individuales). Un "promotor" comprende una secuencia nucleotídica que dirige la transcripción de un gen estructural. Típicamente, un promotor se localiza en la región no codificadora o no traducida 5' de un gen, próximo al lugar de inicio transcripcional de un gen estructural. Los elementos de secuencia dentro de los promotores que actúan en la iniciación de la transcripción a menudo se caracterizan por secuencias nucleotídicas consenso. Estos elementos promotores incluyen lugares de fijación de la ARN polimerasa, secuencias TATA, secuencias CAAT, elementos específicos de diferenciación (DSE; McGehee, R.E., *et al.*, Mol. Endocrinol. 7 (1993) 551-560), elementos de respuesta al AMP cíclico (CRE), elementos de respuesta al suero (SRE, Treisman, R., Seminars in Cancer Biol. 1 (1990) 47-58), elementos de respuesta a glucocorticoides (GRE) y lugares de fijación para otros factores de transcripción, tales como CRE/ATF (Treisman, R., Seminars in Cancer Biol. 1 (1990) 47-58; O'Reilly, M.A., *et al.*, J. Biol. Chem. 267 (1992) 19938), AP2 (Ye, J., *et al.*, J. Biol. Chem. 269 (1994) 25728), SP1, proteína de fijación al elemento de respuesta al AMPc (CREB; Loeken, M.R., Gene Expr. 3 (1993) 253) y factores octámeros (véase, en general, Watson *et al.*, eds., Molecular Biology of the Gene, 4.^a ed. (The Benjamin/Cummings Publishing Company, Inc. 1987) y Lemaigre, F.P. y Rousseau, G.G., Biochem. J. 303 (1994) 1-14). Si un promotor es un promotor inducible, entonces la tasa de transcripción aumenta en respuesta a un agente inductor. Por el contrario, la

tasa de transcripción no está regulada por un agente inductor si el promotor es un promotor constitutivo. También se conocen los promotores reprimibles. Por ejemplo, el promotor c-fos se activa específicamente tras la fijación de la hormona del crecimiento a su receptor en la superficie celular. La expresión regulada por tetraciclina (tet) se puede conseguir mediante promotores híbridos artificiales que consisten, por ejemplo, en un promotor de CMV seguido de dos lugares del operador de Tet. El represor de Tet se fija a los dos lugares del operador de Tet y bloquea la transcripción. Tras la adición de la tetraciclina inductora, se libera el represor de Tet de los lugares del operador de Tet y prosigue la transcripción (Gossen, M. y Bujard, H. PNAS 89 (1992) 5547-5551). Para otros promotores inducibles que incluyen metalotioneína y promotores de choque térmico, véanse, por ejemplo, Sambrook *et al.* (*supra*) y Gossen, M., Curr. Opin. Biotech. 5 (1994) 516-520. Entre los promotores eucariotas que se han identificado como promotores potentes para expresión de alto nivel están el promotor temprano de SV40, el promotor tardío principal de adenovirus, el promotor de metalotioneína I murina, la repetición terminal larga del virus del sarcoma de Rous, el factor 1 alfa de elongación de hámster chino (CHEF-1, véase, por ejemplo, el documento US 5.888.809), el EF-1 alfa humano, la ubicuitina y el promotor temprano inmediato de citomegalovirus humano (CMV IE).

El "promotor" puede ser constitutivo o inducible. Un potenciador (es decir, un elemento de ADN que actúa en cis que actúa sobre un promotor para aumentar la transcripción) puede ser necesario para actuar en conjunto con el promotor para aumentar el nivel de expresión obtenido con un promotor solo, y se puede incluir como elemento regulador transcripcional. A menudo, el segmento de polinucleótidos que contiene el promotor incluirá también secuencias potenciadoras (por ejemplo, CMV o SV40).

Un "potenciador", como se usa en el presente documento, se refiere a una secuencia polinucleotídica que potencia la transcripción de un gen o secuencia codificadora a la que está unida funcionalmente. A diferencia de los promotores, los potenciadores son relativamente independientes de la orientación y la posición y se han encontrado en 5' o 3' (Lusky, M., *et al.*, Mol. Cell Bio., 3 (1983) 1108-1122) con respecto a la unidad de transcripción, dentro de un intrón (Banerji, J., *et al.*, Cell, 33 (1983) 729-740), así como dentro de la propia secuencia codificadora (Osborne, T.F., *et al.*, Mol. Cell Bio., 4 (1984) 1293-1305). Por lo tanto, los potenciadores se pueden colocar en dirección 5' o 3' del lugar de iniciación de la transcripción o a distancias considerables del promotor, aunque en la práctica los potenciadores se pueden superponer física y funcionalmente con los promotores. En la técnica se conocen bien un gran número de potenciadores de una variedad de fuentes diferentes (y están identificados en bases de datos tales como GenBank) y están disponibles como o dentro de secuencias polinucleotídicas clonadas (a partir de, por ejemplo, depositarios tales como ATCC, así como otras fuentes comerciales o individuales). Numerosos polinucleótidos que comprenden secuencias promotoras (tales como el promotor de CMV comúnmente usado) también comprenden secuencias potenciadoras. Por ejemplo, todos los promotores potentes enumerados anteriormente también pueden contener potenciadores potentes (véase, por ejemplo, Bendig, M.M., Genetic Engineering, 7 (Academic Press, 1988) 91-127).

Un "lugar de entrada al ribosoma interno" o "IRES" describe una secuencia que promueve funcionalmente la iniciación de la traducción independiente del gen 5' del IRES y permite que dos cistrones (marcos abiertos de lectura) se traduzcan a partir de un único transcrito en una célula animal. El IRES proporciona un lugar de entrada al ribosoma independiente para la traducción del marco abierto de lectura situado inmediatamente en dirección 3' (en dirección 3' se usa indistintamente con 3' en el presente documento) de este. A diferencia del ARNm bacteriano que puede ser policistrónico, es decir, que codifica varios polipéptidos diferentes que se traducen secuencialmente de los ARNm, la mayoría de los ARNm de las células animales son monocistrónicos y codifican la síntesis de solo una proteína. Con un transcrito monocistrónico en una célula eucariota, la traducción se iniciaría desde el lugar de iniciación de traducción situado más en 5', terminaría en el primer codón de terminación y el transcrito se liberaría del ribosoma, dando como resultado la traducción de solo el primer polipéptido codificado en el ARNm. En una célula eucariota, un transcrito policistrónico que tiene un IRES unido funcionalmente al segundo o subsiguiente marco abierto de lectura en el transcrito permite la traducción secuencial de ese marco abierto de lectura en dirección 3' para producir los dos o más polipéptidos codificados por el mismo transcrito. El uso de elementos IRES en la construcción de vectores se ha descrito previamente, véase, por ejemplo, Pelletier, J., *et al.*, Nature 334 (1988) 320-325; Jang, S.K., *et al.*, J. Virol. 63 (1989) 1651-1660; Davies, M.V., *et al.*, J. Virol. 66 (1992) 1924-1932; Adam, M.A., *et al.*, J. Virol. 65 (1991) 4985-4990; Morgan, R.A., *et al.*, Nucl. Acids Res. 20 (1992) 1293-1299; Sugimoto, Y., *et al.*, Biotechnology 12 (1994) 694-698; Ramesh, N., *et al.*, Nucl. Acids Res. 24 (1996) 2697-2700; y Mosser, D.D. *et al.*, BioTechniques 22 (1997) 150-161).

"Unido funcionalmente" se refiere a una yuxtaposición de dos o más componentes, en la que los componentes así descritos están en una relación que les permite funcionar de su manera pretendida. Por ejemplo, un promotor y/o potenciador están unidos funcionalmente a una secuencia codificadora si actúa en cis para controlar o modular la transcripción de la secuencia unida. En general, pero no necesariamente, las secuencias de ADN que están "unidas funcionalmente" son contiguas y, si fuera necesario, unen dos regiones codificadoras de proteína tales como un líder secretor y un polipéptido, contiguos y en marco (de lectura). Sin embargo, aunque un promotor unido funcionalmente se localiza en general en dirección 5' de la secuencia codificadora, no es necesariamente contiguo a ella. Los potenciadores no tienen que ser contiguos. Un potenciador está unido funcionalmente a una secuencia codificadora si el potenciador aumenta la transcripción de la secuencia

codificadora. Los potenciadores unidos funcionalmente se pueden localizar en dirección 5', dentro o en dirección 3' de las secuencias codificadoras y a una distancia considerable del promotor. Un lugar de poliadenilación está unido funcionalmente a una secuencia codificadora si se localiza en el extremo en dirección 3' de la secuencia codificadora de manera que la transcripción prosigue a través de la secuencia codificadora en la secuencia de poliadenilación. Un codón de terminación de la traducción está unido funcionalmente a una secuencia de ácido nucleico exónica si se localiza en el extremo en dirección 3' (extremo 3') de la secuencia codificadora de manera que la traducción prosigue a través de la secuencia codificadora hasta el codón de terminación y termina allí. La unión se logra mediante procedimientos recombinantes conocidos en la técnica, por ejemplo, usando metodología de PCR y/o mediante unión en lugares de restricción convenientes. Si no existen lugares de restricción convenientes, entonces se usan los adaptadores o conectores oligonucleotídicos sintéticos de acuerdo con la práctica convencional.

"ADN heterólogo" o "polipéptido heterólogo" se refiere a una molécula de ADN o a un polipéptido, o a una población de moléculas de ADN o a una población de polipéptidos, que no existen de manera natural dentro de una célula huésped dada. Las moléculas de ADN heterólogas con respecto a una célula huésped particular pueden contener ADN derivado de la especie de la célula huésped (es decir, ADN endógeno) siempre que el ADN huésped se combine con ADN no huésped (es decir, ADN exógeno). Por ejemplo, se considera que una molécula de ADN que contiene un segmento de ADN no huésped que codifica un polipéptido unido funcionalmente a un segmento de ADN huésped que comprende un promotor es una molécula de ADN heteróloga. Contrariamente, una molécula de ADN heteróloga puede comprender un gen estructural endógeno unido funcionalmente a un promotor exógeno.

Un péptido o polipéptido codificado por una molécula de ADN no huésped es un péptido o polipéptido "heterólogo".

Un "plásmido de expresión" es una molécula de ácido nucleico que codifica una proteína que se va a expresar en una célula huésped. Típicamente, un plásmido de expresión comprende una unidad de propagación de plásmidos procariotas, por ejemplo, para *E. coli*, que comprende un origen de replicación, y un marcador de selección, un marcador de selección eucariota, y uno o más casetes de expresión para la expresión del gen o genes estructurales de interés, comprendiendo cada uno un promotor, un gen estructural y un terminador de la transcripción que incluye una señal de poliadenilación. La expresión génica se coloca habitualmente bajo el control de un promotor, y dicho gen estructural se dice que está "unido funcionalmente al" promotor. De forma similar, un elemento regulador y un promotor básico están unidos funcionalmente si el elemento regulador modula la actividad del promotor básico.

Un "polipéptido aislado" es un polipéptido que está esencialmente libre de componentes celulares contaminantes, tales como carbohidratos, lípidos u otras impurezas proteínicas asociadas con el polipéptido en la naturaleza. Típicamente, una preparación de polipéptido aislado contiene el polipéptido en una forma altamente purificada, es decir, al menos aproximadamente un 80 % de pureza, al menos aproximadamente un 90 % de pureza, al menos aproximadamente un 95 % de pureza, superior a un 95 % de pureza o superior a un 99 % de pureza. Una forma de mostrar que una preparación de proteína particular contiene un polipéptido aislado es mediante la aparición de una única banda después de la electroforesis en gel con dodecilsulfato sódico (SDS)-poliacrilamida de la preparación de proteína y la tinción con azul de Coomassie del gel. Sin embargo, el término "aislado" no excluye la presencia del mismo polipéptido en formas físicas alternativas, tales como dímeros o, de forma alternativa, en formas glucosiladas o derivadas.

El "terminador de la transcripción", como se denota en la presente solicitud, es una secuencia de ADN de 50-750 pares de bases de longitud que da a la ARN polimerasa la señal para la terminación de la síntesis de ARNm. Son aconsejables terminadores muy eficaces (potentes) en el extremo 3' de un casete de expresión para prevenir la lectura íntegra de la ARN polimerasa, particularmente cuando se usan promotores potentes. Los terminadores de transcripción ineficaces pueden dar lugar a la formación de un ARNm de tipo operón que puede ser la causa de una expresión génica indeseada, por ejemplo, codificada por un plásmido.

Los siguientes ejemplos, listado de secuencias y figuras se proporcionan para ayudar a la comprensión de la presente invención, cuyo verdadero alcance se expone en las reivindicaciones adjuntas.

Descripción de las secuencias

60	SEQ ID NO: 01	Parte C-terminal de la secuencia de aminoácidos del dominio constante de cadena pesada (C _H 3) de una inmunoglobulina de clase IgA humana
	SEQ ID NO: 02	Parte C-terminal de la secuencia de aminoácidos del dominio constante de cadena pesada (C _H 3) de una inmunoglobulina de clase IgD humana
65	SEQ ID NO: 03	Parte C-terminal de la secuencia de aminoácidos del dominio constante de cadena pesada (C _H 4) de una inmunoglobulina de clase IgE humana

	SEQ ID NO: 04	Parte C-terminal de la secuencia de aminoácidos del dominio constante de cadena pesada (C _H 4) de una inmunoglobulina de clase IgM humana
5	SEQ ID NO: 05	Parte C-terminal de la secuencia de aminoácidos del dominio constante de cadena pesada (C _H 3) de una inmunoglobulina de clase IgG1 humana
	SEQ ID NO: 06	Parte C-terminal de la secuencia de aminoácidos del dominio constante de cadena pesada (C _H 3) de una inmunoglobulina de clase IgG2 humana
10	SEQ ID NO: 07	Parte C-terminal de la secuencia de aminoácidos del dominio constante de cadena pesada (C _H 3) de una inmunoglobulina de clase IgG3 humana
15	SEQ ID NO: 08	Parte C-terminal de la secuencia de aminoácidos del dominio constante de cadena pesada (C _H 3) de una inmunoglobulina de clase IgG4 humana
	SEQ ID NO: 09	Secuencia de ácido nucleico que codifica la parte C-terminal de la secuencia de aminoácidos del dominio constante de cadena pesada (C _H 3) de una inmunoglobulina de clase IgA humana
20	SEQ ID NO: 10	Secuencia de ácido nucleico que codifica la parte C-terminal de la secuencia de aminoácidos del dominio constante de cadena pesada (C _H 3) de una inmunoglobulina de clase IgD humana
25	SEQ ID NO: 11	Secuencia de ácido nucleico que codifica la parte C-terminal de la secuencia de aminoácidos del dominio constante de cadena pesada (C _H 4) de una inmunoglobulina de clase IgE humana
30	SEQ ID NO: 12	Secuencia de ácido nucleico que codifica la parte C-terminal de la secuencia de aminoácidos del dominio constante de cadena pesada (C _H 4) de una inmunoglobulina de clase IgM humana
	SEQ ID NO: 13, 28	Secuencia de ácido nucleico que codifica la parte C-terminal de la secuencia de aminoácidos del dominio constante de cadena pesada (C _H 3) de una inmunoglobulina de clase IgG1 humana (variante 1 y 2)
35		
	SEQ ID NO: 14, 29	Secuencia de ácido nucleico que codifica la parte C-terminal de la secuencia de aminoácidos del dominio constante de cadena pesada (C _H 3) de una inmunoglobulina de clase IgG2 humana (variante 1 y 2)
40		
	SEQ ID NO: 15	Secuencia de ácido nucleico que codifica la parte C-terminal de la secuencia de aminoácidos del dominio constante de cadena pesada (C _H 3) de una inmunoglobulina de clase IgG3 humana
45	SEQ ID NO: 16	Secuencia de ácido nucleico que codifica la parte C-terminal de la secuencia de aminoácidos del dominio constante de cadena pesada (C _H 3) de una inmunoglobulina de clase IgG4 humana
	SEQ ID NO: 17	Secuencia de ácido nucleico de acuerdo con la invención que codifica una parte de las secuencias de aminoácidos del dominio constante C-terminal (C _H 3) de una inmunoglobulina de clase IgA
50		
	SEQ ID NO: 18	Secuencia de ácido nucleico de acuerdo con la invención que codifica una parte de las secuencias de aminoácidos del dominio constante C-terminal (C _H 4) de una inmunoglobulina de clase IgE
55		
	SEQ ID NO: 19	Secuencia de ácido nucleico de acuerdo con la invención que codifica una parte de las secuencias de aminoácidos del dominio constante C-terminal (C _H 4) de una inmunoglobulina de clase IgM
60		
	SEQ ID NO: 20, 30	Secuencia de ácido nucleico de acuerdo con la invención que codifica una parte de las secuencias de aminoácidos del dominio constante C-terminal (C _H 3) de una inmunoglobulina de clase IgG1 (variante 1 y 2)

SEQ ID NO: 21, 31 Secuencia de ácido nucleico de acuerdo con la invención que codifica una parte de las secuencias de aminoácidos del dominio constante C-terminal (C_H3) de una inmunoglobulina de clase IgG2 (variante 1 y 2)

5 SEQ ID NO: 22 Secuencia de ácido nucleico de acuerdo con la invención que codifica una parte de las secuencias de aminoácidos del dominio constante C-terminal (C_H3) de una inmunoglobulina de clase IgG3

10 SEQ ID NO: 23 Secuencia de ácido nucleico de acuerdo con la invención que codifica una parte de las secuencias de aminoácidos del dominio constante C-terminal (C_H3) de una inmunoglobulina de clase IgG4

SEQ ID NO: 24 a 27 Cebadores de ácido nucleico usados en los ejemplos.

15 **Descripción de las figuras**

Figura 1 Mapa plasmídico de p4831 con anotaciones.

20 **Figura 2** Análisis de SDS-PAGE e inmunoelectrotransferencia de anticuerpos secretados por el clon n.º 23; a) tinción con Coomassie, b) análisis de inmunoelectrotransferencia con anticuerpo contra la cadena gamma de inmunoglobulina humana conjugado con peroxidasa, c) análisis de inmunoelectrotransferencia con anticuerpo contra la cadena ligera kappa de inmunoglobulina humana conjugado con peroxidasa. Carril 1: anticuerpo de referencia anti-IGF-1R humano; carril 2: sobrenadante del cultivo del clon n.º 23 de CHO-DG44 que comprende el anticuerpo del clon n.º 23.

Figura 3 Mapa plasmídico de p4817 con anotaciones.

30 **Figura 4** Análisis de inmunoelectrotransferencia de inmunoglobulinas secretadas por CHO-DXB11 transfectadas con p4831 o 4855. Carril 1: células CHO transfectadas con p4831, carril 2: células CHO transfectadas con p4855.

Figura 5 Mapa plasmídico de p5031 con anotaciones.

35 **Figura 6** Análisis de SDS-PAGE de anticuerpos secretados por células HEK-293-EBNA después de la transfección con el plásmido p5031 o p5032; carriles 1 + 2: p5031, carriles 3 + 4: p5032, carriles 5-8: anticuerpo anti-IGF-1R humano como anticuerpo de referencia; 5: 0,2 µg, 6: 0,7 µg, 7: 2 µg, 8: 6 µg.

40 **Figura 7** Células CHO-DG44 transfectadas con p5032 y seleccionadas para la integración estable del plásmido con metotrexato (MTX). Se purificaron los anticuerpos de seis clones y se analizaron mediante SDS-PAGE y tinción con Coomassie.

45 **Materiales y procedimientos**

Técnicas de ADN recombinante

Se usaron procedimientos estándar para manipular el ADN como se describe en Sambrook *et al.*, 1989 (*supra*). Todos los reactivos biológicos moleculares estaban disponibles comercialmente (si no se indica lo contrario) y se usaron de acuerdo con las instrucciones del fabricante.

Determinación de la secuencia de ácido nucleico

55 Las secuencias de ADN se determinaron mediante secuenciación bicatenaria realizada en MediGenomix GmbH (Martinsried, Alemania).

Análisis de secuencia de ADN y proteína y gestión de datos de secuencia

60 Se usaron el paquete de software de GCG (Genetics Computer Group, Madison, Wisconsin, EE. UU.), versión 10.2 y el paquete Vector NTI Advance, versión 8.0 de Infomax para la creación, mapeo, análisis, anotación e ilustración de secuencias.

Técnicas de cultivo celular

65 Se cultivaron células CHO-DXB11 en medio MEM alfa (Invitrogen Corp., Gibco®, n.º de cat.: 22571) con FCS al 10 % (suero bovino fetal obtenido de Hyclone, Thermo Fisher Scientific Inc., n.º de cat.: SH3007.03).

Se cultivaron células HEK-293-EBNA (n.º de ATCC: CRL-10852) en DMEM, enriquecido con glutamina 2 mM (Gibco®, n.º de cat.: 25030), aminoácidos no esenciales en MEM al 1 % (v/v) (Gibco®, n.º de cat.: 11140), FCS-IgG ultrabajo al 10 % (v/v) (Gibco®, n.º de cat.: 16250) y G418 250 µg/ml (Roche Applied Sciences, Roche Diagnostics GmbH, Alemania, n.º de cat.: 1464981).

El medio para el cultivo de células CHO-DG44 era medio MEM alfa (Gibco®, n.º de cat.: 22561) enriquecido con FCS dializado al 10 % (v/v) (Gibco®, n.º de cat.: 26400) y complemento de HT al 2 % (v/v) (Gibco®, n.º de cat.: 41065). Para la selección de líneas celulares CHO DG44 transfectadas establemente se omitió el complemento de HT y se añadió metotrexato (MTX) de 20 a 500 nM solo o en combinación con higromicina B 400 µg/ml (Roche Diagnostics GmbH, Roche Applied Sciences, Alemania, n.º de cat.: 10843555001).

Todas las líneas celulares se mantuvieron en incubadoras humidificadas a 37 °C con CO₂ al 5 %. La transfección de las células se realizó mediante nucleofección (Amaxa GmbH, Alemania) o mediante lipofección usando FuGENE 6 (Roche Diagnostics GmbH, Roche Applied Sciences, Alemania, n.º de cat.: 1815075).

Además, se aplicaron técnicas de cultivo celular estándar como se describe, por ejemplo, en Bonifacino, J. S., *et al.*, (Eds) (2000) *Current Protocols in Cell Biology*, John Wiley and Sons, Inc.

Precipitación con proteína A, SDS-PAGE e inmunoelectrotransferencia

Se precipitaron las inmunoglobulinas de los sobrenadantes del cultivo celular con microesferas de proteína A-sefarosa y luego se analizaron mediante electroforesis en gel de SDS/poliacrilamida (dodecilsulfato sódico, SDS-PAGE) e inmunoelectrotransferencia.

Para la precipitación de inmunoglobulinas, los sobrenadantes del cultivo celular que contenían hasta 7 µg de inmunoglobulina se diluyeron con tampón TBS (TRIS/HCl 50 mM, pH 7,5, enriquecido con NaCl 150 mM), Nonidet-P40 al 1 % (v/v) (Roche Diagnostics GmbH, Roche Applied Sciences, n.º de cat.: 1754599) hasta un volumen final de 1 ml, y después se incubaron durante una hora con 15 µl de microesferas de proteína A-sefarosa en volumen húmedo. Se recuperaron las microesferas mediante centrifugación y se lavaron con TBS con Nonidet P-40 al 1 % (v/v), seguidamente se lavaron 2 veces con PBS concentrado (solución salina tamponada con fosfato) y, finalmente, con tampón de citrato sódico 100 mM, pH 5. Después de la etapa de lavado final, se eliminó por completo el sobrenadante de las microesferas. Las inmunoglobulinas fijadas se eluyeron 2 veces con 20 µl de tampón de muestra LDS (dodecilsulfato de litio) concentrado (Invitrogen Corp.) que contenía DTT (ditiotreitól) 50 mM. Después de 5 minutos de incubación a 95 °C, se centrifugó la suspensión y se recuperó el sobrenadante para un análisis adicional.

SDS-PAGE: el SDS-PAGE se realizó usando el sistema de gel NuPAGE® (Invitrogen Corp.) de acuerdo con las recomendaciones del fabricante. Se cargaron las muestras en geles de NuPAGE® Novex Bis/TRIS al 10 % (Invitrogen Corp., n.º de cat.: NP0301) y se separaron las proteínas en tampón de carga reductor NuPAGE® MES SDS (ácido 4-morfolinoetananosulfónico/dodecilsulfato sódico). Típicamente, se cargaron de 2 a 3 µg de inmunoglobulina por carril para la tinción con Coomassie con Simply Blue Safe Stain® (Invitrogen Corp.) y de 0,4 a 0,6 µg para la inmunoelectrotransferencia.

Inmunoelectrotransferencia: para la electrotransferencia de proteínas a partir de geles de SDS/poliacrilamida se usaron membranas estándar de PVDF (difluoruro de polivinilideno) o nitrocelulosa. Después de la electrotransferencia, se lavaron las membranas en TBS (solución salina tamponada con tris, TRIS/HCl 50 mM, pH 7,5, NaCl 150 mM). Los lugares de fijación no específicos se bloquearon mediante incubación en TBS con reactivo de bloqueo de inmunoelectrotransferencia al 1 % (p/v) (Roche Diagnostics GmbH, Roche Applied Sciences, n.º de cat.: 11921673001). Las cadenas pesadas gamma y las cadenas ligeras kappa de inmunoglobulina humana se detectaron con anticuerpos de detección policlonal conjugados con peroxidasa (véase el siguiente párrafo para detalles adicionales) diluidos en TBS con reactivo de bloqueo de inmunoelectrotransferencia al 0,5 % (p/v). Después de tres etapas de lavado con TBS con Tween® 20 al 0,05 % (v/v) y de una etapa de lavado con TBS, los anticuerpos de detección conjugados con peroxidasa fijados se detectaron mediante quimioluminiscencia usando solución de sustrato LumiLightPlus (Roche Diagnostics GmbH, Roche Applied Sciences, n.º de cat.: 12015196001) y el analizador LUMI-Imager F1 (Roche Diagnostics GmbH, Roche Applied Sciences).

Las cadenas pesadas gamma (H) y las cadenas ligeras kappa (L) de inmunoglobulina humana se detectaron simultáneamente o por separado. Para la detección simultánea se usó anticuerpo caprino contra la IgG (H+L) humana conjugado con peroxidasa (Jackson ImmunoResearch Laboratories Inc., n.º de cat.: 109-035-088) a una dilución de 1:2500 (v/v). Para la detección consecutiva, en primer lugar se sondearon las membranas con anticuerpo de conejo contra la Ig gamma humana conjugado con peroxidasa (DAKO GmbH, Alemania, n.º de código P0214) a una dilución de 1:1000 (v/v) o con anticuerpo caprino contra el Fc gamma humano de F(ab')₂ conjugado con peroxidasa (Jackson ImmunoResearch Laboratories, n.º de cat.: 109-036-008) a una dilución de 1:7500 (v/v). Después de la detección de la cadena pesada gamma de inmunoglobulina, las membranas se

lavarón durante 30 minutos en TRIS/HCl 62,5 mM, pH 6,7, enriquecido con SDS al 2 % (p/v) y β -mercaptoetanol 100 mM, a 50 °C. Para la segunda detección, se volvieron a sondear las membranas con anticuerpo de conejo contra la Ig kappa humana conjugado con peroxidasa (DAKO GmbH, n.º de código P0129) a una dilución de 1:1000 (v/v).

5

Ejemplo 1

Preparación de un plásmido de expresión para una inmunoglobulina de clase IgG1

10 El plásmido 4831 (denotado como p4831 en lo sucesivo) es el plásmido de expresión para la expresión de un anticuerpo anti-IGF-1R (casete de expresión genómicamente organizado con organización exón-intrón retenida) en células eucariotas (para las secuencias, véase, por ejemplo, el documento US 2005/0008642 o el documento EP 1 646 720). Comprende los siguientes elementos funcionales:

15 – un origen de replicación derivado del vector pUC18 (origen de pUC),

– un gen de β (beta)-lactamasa que confiere resistencia a ampicilina en *E. coli* (Amp),

20 – un casete de expresión para la expresión de una cadena pesada gamma 1 que comprende los siguientes elementos:

– el promotor temprano inmediato principal y el potenciador del citomegalovirus humano (promotor hCMV IE1),

25 – un 5'UTR sintético que incluye una secuencia de Kozak,

– una secuencia señal de cadena pesada de inmunoglobulina murina que incluye el intrón de la secuencia señal (L1_Intron_L2),

30 – el ADNc para una región variable de cadena pesada (VH) dispuesta con un lugar donante de corte y empalme en el extremo 3',

– la región potenciadora de inmunoglobulina μ murina,

35 – un gen de cadena pesada gamma 1 de inmunoglobulina humana (IGHG1) que incluye los exones CH1, bisagra, CH2 y CH3, los intrones intermedios y el 3'UTR que lleva la secuencia señal de poliadenilación y que opcionalmente contiene mutaciones,

40 – un casete de expresión para la expresión de una cadena ligera kappa que comprende los siguientes elementos:

– el promotor temprano inmediato principal y el potenciador del citomegalovirus humano (promotor hCMV IE1),

45 – un 5'UTR sintético que incluye una secuencia de Kozak,

– una secuencia señal de cadena pesada de inmunoglobulina murina que incluye el intrón de la secuencia señal (L1_Intron_L2),

50 – el ADNc para una región variable de cadena ligera (VL) dispuesta con un lugar donante de corte y empalme en el extremo 3',

– la región potenciadora de Ig-kappa murina intrónica,

55 – el gen de inmunoglobulina kappa (IGK) humana que incluye el exón IGKC y el 3'UTR de IGK que lleva la secuencia señal de poliadenilación.

– una unidad de transcripción de higromicina B fosfotransferasa adecuada para la selección en células eucariotas, que incluye

60

– el promotor temprano y origen de SV40,

– la secuencia codificadora de la higromicina B fosfotransferasa (HygB),

65 – la señal de poliadenilación temprana de SV40

– un casete de expresión para la expresión de la dihidrofolato reductasa (DHFR) murina adecuado para la selección auxotrófica en células eucariotas, que incluye

– una versión acortada del promotor temprano y origen de SV40,

– la secuencia codificadora para la DHFR murina,

– la señal de poliadenilación temprana de SV40.

En la figura 1 se muestra un mapa plasmídico de p4831 con anotaciones.

Se transfectó p4831 en células CHO-DG44 y se aislaron líneas celulares estables después de la selección con higromicina B y metotrexato (MTX). Los anticuerpos secretados por el clon n.º 23 seleccionado se precipitaron con microesferas de proteína A-sefarosa y se analizaron mediante SDS-PAGE y tinción con Coomassie (figura 2 a)). Además de la cadena pesada gamma-1 de inmunoglobulina de 50 kDa y de la cadena ligera kappa de inmunoglobulina de 25 kDa esperadas, se detectaron cantidades considerables de una proteína subproducto de 80 kDa. Esta proteína fue reconocida por anticuerpos contra la cadena gamma de inmunoglobulina humana (figura 2 b)), así como por anticuerpos contra la cadena kappa de inmunoglobulina humana (figura 2 c)).

Ejemplo 2

Preparación de un plásmido de expresión para una inmunoglobulina de clase IgG1 con un dominio CH3 modificado

Con el fin de prevenir la generación de subproductos resultantes del pre-ARNm de gamma 1 empalmado de forma anómala, se destruyó el lugar de corte y empalme interno del exón CH3 de p4831 mutando la T en la posición 4573 a C. Al mismo tiempo, se reemplazó T4567 por C para la eliminación de un lugar de restricción BsmA I.

Se construyó p4855 de la siguiente manera. El p4817, un plásmido progenitor de p4831 con la misma unidad de transcripción de cadena pesada gamma 1, se compone de los siguientes elementos:

– un origen de replicación del vector pUC18 que permite la replicación de este plásmido en *E. coli* (origen de pUC)

– un gen de beta-lactamasa que confiere resistencia a la ampicilina en *E. coli* (Amp)

– un casete de expresión para la expresión de una cadena pesada gamma 1 que comprende los siguientes elementos:

– el promotor temprano inmediato principal y el potenciador del citomegalovirus humano (promotor hCMV IE1),

– un 5'UTR sintético que incluye una secuencia de Kozak,

– una secuencia señal de cadena pesada de inmunoglobulina murina que incluye el intrón de la secuencia señal (L1_Intron_L2),

– el ADNc para una región variable de cadena pesada (VH) dispuesta con un lugar donante de corte y empalme en el extremo 3',

– la región potenciadora de inmunoglobulina μ murina,

– un gen de cadena pesada gamma 1 de inmunoglobulina humana (IGHG1) que incluye los exones CH1, bisagra, CH2 y CH3, los intrones intermedios y el 3'UTR que lleva la secuencia señal de poliadenilación y que opcionalmente contiene mutaciones,

En la figura 3 se muestra el mapa plasmídico de p4817.

Se manipuló el p4817 mediante mutagénesis dirigida al lugar usando el kit de mutagénesis dirigida al lugar QuickChange (Stratagene, n.º de cat.: 200518) y los oligonucleótidos específicos de secuencia 1

agcctctccc tgtccccggg caaatgagtg cgacggccg SEQ ID NO: 24

y 2

cggccgctgc actcatttgc ccggggacag ggagaggct SEQ ID NO: 25.

5 Se escindió el fragmento SfiI/SgrAI de 839 pb de p4817 mutado y se unió al fragmento SgrAI/SfiI de 13133 pb de p4831 para formar p4855.

Ejemplo 3

10 **Expresión de ácidos nucleicos de acuerdo con el ejemplo 1 y 2, aislamiento de la inmunoglobulina producida y análisis de la inmunoglobulina producida**

Se transfectaron transitoriamente los plásmidos p4831 y p4855 en células CHO-DXB11. Se cultivaron las células en condiciones no selectivas. Después de tres días de cultivo, se recogieron los sobrenadantes del cultivo celular y se purificaron las inmunoglobulinas secretadas con microesferas de proteína A-sefarosa. El análisis de inmunoelectrotransferencia de las inmunoglobulinas con anticuerpos contra la IgG humana (H+L) mostró que el subproducto de 80 kDa había sido expresado por las células transfectadas con p4831, pero no por las células transfectadas con p4855 (figura 4).

Ejemplo 4

20 **Preparación de un plásmido de expresión para una inmunoglobulina de clase IgG4**

Se diseñó el plásmido 5031 para la expresión transitoria y estable de un anticuerpo contra P-selectina humana en células de cultivo de tejido eucariótico. Para los anticuerpos contra P-selectina ejemplares, véanse, por ejemplo, los documentos EP 1 737 891 o US 2005/0226876. El p5031 se compone de los siguientes elementos:

- un origen de replicación del vector pUC18 que permite la replicación de este plásmido en *E. coli* (origen de pUC),
- 30 – un gen de β -lactamasa que confiere resistencia a la ampicilina en *E. coli* (Amp),
- un casete de expresión para la expresión de una cadena pesada gamma 4 humana que comprende los siguientes elementos:
 - 35 – el promotor temprano inmediato principal y el potenciador del citomegalovirus humano (promotor hCMV IE1),
 - un 5'UTR sintético que incluye una secuencia de Kozak,
 - 40 – una secuencia señal de cadena pesada de inmunoglobulina murina que incluye el intrón de la secuencia señal (L1_Intron_L2),
 - el ADNc para una región variable de cadena pesada (VH) dispuesta con un lugar donante de corte y empalme en el extremo 3',
 - 45 – la región potenciadora de Ig μ murina,
 - un gen de cadena pesada gamma 4 de inmunoglobulina humana (IGHG4) que incluye los exones CH1, bisagra, CH2 y CH3, los intrones intermedios y el 3'UTR que lleva la secuencia señal de poliadenilación y que opcionalmente contiene mutaciones,
 - 50 – un casete de expresión para la expresión de la cadena ligera kappa humana que comprende los siguientes elementos:
 - 55 – el promotor temprano inmediato principal y el potenciador del citomegalovirus humano (promotor hCMV IE1),
 - un 5'UTR sintético que incluye una secuencia de Kozak,
 - 60 – una secuencia señal de cadena pesada de inmunoglobulina murina que incluye el intrón de la secuencia señal (L1_Intron_L2),
 - el ADNc para una región variable de cadena ligera (VL) dispuesta con un lugar donante de corte y empalme en el extremo 3',
 - 65

- la región potenciadora de Ig-kappa murina intrónica,
- un gen de inmunoglobulina kappa (IGK) humana que incluye el exón IGKC y el 3'UTR de IGK que lleva la secuencia señal de poliadenilación,
- una unidad de transcripción de higromicina B fosfotransferasa adecuada para la selección en células eucariotas que incluye
- el promotor temprano y origen de SV40,
- la secuencia codificadora de la higromicina B fosfotransferasa (HygB),
- la señal de poliadenilación temprana de SV40,
- un casete de expresión para la expresión de la dihidrofolato reductasa (DHFR) murina adecuado para la selección auxotrófica en células eucariotas, que incluye
- una versión acortada del promotor temprano y origen de SV40,
- la secuencia codificadora para la DHFR murina,
- la señal de poliadenilación temprana de SV40.

En la figura 5 se muestra un mapa plasmídico de p5031.

Cuando se transfectaron las células HEK-293-EBNA con el plásmido p5031, las células produjeron inmunoglobulinas que comprenden un subproducto de 80 kDa (figura 6, carril 1 y 2). Esta proteína se fijó mediante el anticuerpo contra la Ig gamma humana, así como mediante el anticuerpo contra la Ig kappa humana en análisis de inmunoelectrotransferencia.

Ejemplo 5

Preparación de un plásmido de expresión para una inmunoglobulina de clase IgG4 con un dominio CH3 modificado

La modificación se introdujo de acuerdo con el ejemplo 3. En resumen, se mutó T4565 a C junto con un segundo nucleótido, T4559, que también se intercambió por C para la eliminación de un lugar de restricción BsmAI. El oligonucleótido 3

gcctctccct gtcctgggc aaatgagtgc cagg SEQ ID NO: 26

y el oligonucleótido 4

cctggcactc atttgcccag ggacagggag aggc SEQ ID NO: 27

se usaron para la mutagénesis dirigida al lugar. El plásmido obtenido se denominó p5032.

Ejemplo 6

Expresión de ácidos nucleicos de acuerdo con el ejemplo 5, aislamiento de la inmunoglobulina producida y análisis de la inmunoglobulina producida

Se transfectaron transitoriamente los plásmidos p5031 y p5032 en células HEK-293-EBNA. Se cultivaron las células en condiciones no selectivas. Después de tres días de cultivo, se recogieron los sobrenadantes del cultivo celular y se purificaron las inmunoglobulinas secretadas con microesferas de proteína A-sefarosa. El análisis de inmunoelectrotransferencia de las inmunoglobulinas con anticuerpos contra la IgG humana (H+L) mostró que las células transfectadas con p5031 expresaban un subproducto de 80 kDa (figura 6, carriles 1+2), mientras que las células transfectadas con p5032 no habían expresado ningún subproducto (figura 6, carriles 3+4).

Se transfectó el plásmido p5032 en células HEK-293-EBNA, la proteína de 80 kDa no se expresó (figura 6, carril 3 y 4). Esto demuestra claramente que la proteína de fusión de 80 kDa es un resultado del corte y empalme anómalo de pre-ARNm y que la producción de dicha proteína indeseable durante la expresión transitoria se puede suprimir eficazmente mediante la mutación del lugar de corte y empalme de CH3 interno del gen de cadena pesada gamma 4 de inmunoglobulina (IGHG4).

5 La mutación del lugar de corte y empalme de C_H3 interno de IGHG4 previno la expresión de la proteína de 80 kDa también en líneas celulares estables. Se transfectaron las células CHO-DG44 con p5032 y se seleccionaron para la integración estable del plásmido con metotrexato (MTX). Se purificaron los anticuerpos de 6 clones y se analizaron mediante SDS-PAGE y tinción con Coomassie (figura 7). Ninguno de los anticuerpos contenía la subunidad de 80 kDa.

LISTADO DE SECUENCIAS

<110> F. Hoffmann-La Roche AG

5 <120> Producción de inmunoglobulinas

<130> 24355 FT

<150> EP 07012774.1

10 <151> 29-06-2007

<160> 31

<170> PatentIn versión 3.2

15 <210> 1

<211> 131

<212> PRT

<213> *Homo sapiens*

20 <400> 1

Gly Asn Thr Phe Arg Pro Glu Val His Leu Leu Pro Pro Pro Ser Glu
1 5 10 15

Glu Leu Ala Leu Asn Glu Leu Val Thr Leu Thr Cys Leu Ala Arg Gly
20 25 30

Phe Ser Pro Lys Asp Val Leu Val Arg Trp Leu Gln Gly Ser Gln Glu
35 40 45

Leu Pro Arg Glu Lys Tyr Leu Thr Trp Ala Ser Arg Gln Glu Pro Ser
50 55 60

Gln Gly Thr Thr Thr Phe Ala Val Thr Ser Ile Leu Arg Val Ala Ala
65 70 75 80

Glu Asp Trp Lys Lys Gly Asp Thr Phe Ser Cys Met Val Gly His Glu
85 90 95

Ala Leu Pro Leu Ala Phe Thr Gln Lys Thr Ile Asp Arg Leu Ala Gly
100 105 110

Lys Pro Thr His Val Asn Val Ser Val Val Met Ala Glu Val Asp Gly
115 120 125

Thr Cys Tyr
130

25 <210> 2

<211> 109

<212> PRT

ES 2 665 920 T3

<213> *Homo sapiens*

<400> 2

Ala Ala Gln Ala Pro Val Lys Leu Ser Leu Asn Leu Leu Ala Ser Ser
1 5 10 15

Asp Pro Pro Glu Ala Ala Ser Trp Leu Leu Cys Glu Val Ser Gly Phe
20 25 30

Ser Pro Pro Asn Ile Leu Leu Met Trp Leu Glu Asp Gln Arg Glu Val
35 40 45

Asn Thr Ser Gly Phe Ala Pro Ala Arg Pro Pro Pro Gln Pro Gly Ser
50 55 60

Thr Thr Phe Trp Ala Trp Ser Val Leu Arg Val Pro Ala Pro Pro Ser
65 70 75 80

Pro Gln Pro Ala Thr Tyr Thr Cys Val Val Ser His Glu Asp Ser Arg
85 90 95

Thr Leu Leu Asn Ala Ser Arg Ser Leu Glu Val Ser Tyr
100 105

5

<210> 3

<211> 110

<212> PRT

10 <213> *Homo sapiens*

<400> 3

ES 2 665 920 T3

Gly Pro Arg Ala Ala Pro Glu Val Tyr Ala Phe Ala Thr Pro Glu Trp
1 5 10 15

Pro Gly Ser Arg Asp Lys Arg Thr Leu Ala Cys Leu Ile Gln Asn Phe
20 25 30

Met Pro Glu Asp Ile Ser Val Gln Trp Leu His Asn Glu Val Gln Leu
35 40 45

Pro Asp Ala Arg His Ser Thr Thr Gln Pro Arg Lys Thr Lys Gly Ser
50 55 60

Gly Phe Phe Val Phe Ser Arg Leu Glu Val Thr Arg Ala Glu Trp Glu
65 70 75 80

Gln Lys Asp Glu Phe Ile Cys Arg Ala Val His Glu Ala Ala Ser Pro
85 90 95

Ser Gln Thr Val Gln Arg Ala Val Ser Val Asn Pro Gly Lys
100 105 110

<210> 4
<211> 131
5 <212> PRT
<213> *Homo sapiens*
<400> 4

ES 2 665 920 T3

Gly Val Ala Leu His Arg Pro Asp Val Tyr Leu Leu Pro Pro Ala Arg
1 5 10 15

Glu Gln Leu Asn Leu Arg Glu Ser Ala Thr Ile Thr Cys Leu Val Thr
20 25 30

Gly Phe Ser Pro Ala Asp Val Phe Val Gln Trp Met Gln Arg Gly Gln
35 40 45

Pro Leu Ser Pro Glu Lys Tyr Val Thr Ser Ala Pro Met Pro Glu Pro
50 55 60

Gln Ala Pro Gly Arg Tyr Phe Ala His Ser Ile Leu Thr Val Ser Glu
65 70 75 80

Glu Glu Trp Asn Thr Gly Glu Thr Tyr Thr Cys Val Val Ala His Glu
85 90 95

Ala Leu Pro Asn Arg Val Thr Glu Arg Thr Val Asp Lys Ser Thr Gly
100 105 110

Lys Pro Thr Leu Tyr Asn Val Ser Leu Val Met Ser Asp Thr Ala Gly
115 120 125

Thr Cys Tyr
130

<210> 5
<211> 107
<212> PRT
<213> *Homo sapiens*

<400> 5

5

ES 2 665 920 T3

Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp
1 5 10 15

Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe
20 25 30

Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu
35 40 45

Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe
50 55 60

Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly
65 70 75 80

Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr
85 90 95

Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
100 105

<210> 6
<211> 107
<212> PRT
<213> *Homo sapiens*

<400> 6

5

ES 2 665 920 T3

Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu
1 5 10 15

Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe
20 25 30

Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu
35 40 45

Asn Asn Tyr Lys Thr Thr Pro Pro Met Leu Asp Ser Asp Gly Ser Phe
50 55 60

Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly
65 70 75 80

Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr
85 90 95

Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
100 105

<210> 7
<211> 107
5 <212> PRT
<213> *Homo sapiens*
<400> 7

ES 2 665 920 T3

Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu
1 5 10 15

Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe
20 25 30

Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Ser Gly Gln Pro Glu
35 40 45

Asn Asn Tyr Asn Thr Thr Pro Pro Met Leu Asp Ser Asp Gly Ser Phe
50 55 60

Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly
65 70 75 80

Asn Ile Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn Arg Phe
85 90 95

Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
100 105

<210> 8
<211> 107
<212> PRT
<213> *Homo sapiens*

<400> 8

5

ES 2 665 920 T3

Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu
1 5 10 15

Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe
20 25 30

Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu
35 40 45

Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe
50 55 60

Phe Leu Tyr Ser Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly
65 70 75 80

Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr
85 90 95

Thr Gln Lys Ser Leu Ser Leu Ser Leu Gly Lys
100 105

<210> 9
<211> 300
5 <212> ADN
<213> *Homo sapiens*

<400> 9

ggcttcagcc ccaaggacgt gctggttcgc tggtgcagg ggtcacagga gctgccccgc 60

gagaagtacc tgacttgggc atcccggcag gagcccagcc agggcaccac caccttcgct 120

gtgaccagca tactgcgcgt ggcagccgag gactggaaga agggggacac cttctcctgc 180

atggtgggcc acgaggccct gccgctggcc ttcacacaga agaccatcga ccgcttggcg 240

10 ggtaaaccce cccatgtcaa tgtgtctgtt gtcattggcg aggtggacgg cacctgctac 300

<210> 10
<211> 300
<212> ADN
15 <213> *Homo sapiens*

<400> 10

ES 2 665 920 T3

	taccacccaa cgtccgtgac tgtcacctgg tacatgggga cacagagcca gccccagaga	60
	accttccttg agatacaaag acgggacagc tactacatga caagcagcca gctctccacc	120
	cccctccagc agtggcgcca aggcgagtag aaatgcgtgg tccagcacac cgccagcaag	180
	agtaagaagg agatcttccg ctggccaggt aggtcgcacc ggagatcacc cagaagggcc	240
	ccccaggacc ccagcacct tccactcagg gcctgaccac aaagacagaa gcaagggtg	300
5	<210> 11 <211> 300 <212> ADN <213> <i>Homo sapiens</i>	
	<400> 11	
	tttgcgacgc cggagtggcc ggggagccgg gacaagcgca ccctgcctg cctgatccag	60
	aacttcatgc ctgaggacat ctcggtgcag tggtgcaca acgaggtgca gctccggac	120
	gcccggcaca gcacgacgca gcccgcgaag accaagggtc ccggcttctt cgtcttcagc	180
	cgcttgagg tgaccagggc cgaatgggag cagaaagatg agttcatctg cgtgcagtc	240
10	catgaggcag cgagcccctc acagaccgtc cagcgagcgg tgtctgtaaa tcccggtaaa	300
15	<210> 12 <211> 300 <212> ADN <213> <i>Homo sapiens</i>	
	<400> 12	
	acgggcttct ctcccgcgga cgtcttcgtg cagtggatgc agagggggca gcccttgctc	60
	ccggagaagt atgtgaccag cgcaccaatg cctgagcccc agggcccagg ccggtacttc	120
	gcccacagca tcttgaccgt gtccgaagag gaatggaaca cgggggagac ctacacctgc	180
	gtggtggccc atgaggccct gccaacagg gtcaccgaga ggaccgtgga caagtccacc	240
20	ggtaaaccba ccctgtacaa cgtgtccctg gtcattgtccg acacagctgg cacctgctac	300
25	<210> 13 <211> 300 <212> ADN <213> <i>Homo sapiens</i>	
	<400> 13	
	gtgtacaccc tgccccatc ccgggatgag ctgaccaaga accaggctcag cctgacctgc	60
	ctggtcaaag gcttctatcc cagcgacatc gccgtggagt gggagagcaa tgggcagccg	120
	gagaacaact acaagaccac gcctcccgtg ctggactccg acggctcctt cttcctctac	180
	agcaagctca ccgtggacaa gagcaggtgg cagcagggga acgtcttctc atgctccgtg	240
	atgcatgagg ctctgcacaa ccactacag cagaagagcc tctccctgtc tccgggtaaa	300

ES 2 665 920 T3

<210> 14
 <211> 300
 <212> ADN
 5 <213> *Homo sapiens*

 <400> 14
 gtgtacaccc tgcccccatc ccgggaggag atgaccaaga accaggtcag cctgacctgc 60
 ctggtcaaag gcttctaccc cagcgacatc gccgtggagt gggagagcaa tgggcagccg 120
 gagaacaact acaagaccac acctcccatg ctggactccg acggctcctt cttcctctac 180
 agcaagctca ccgtggacaa gagcaggtgg cagcagggga acgtcttctc atgctccgtg 240
 atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggtaaa 300
 10
 <210> 15
 <211> 300
 <212> ADN
 15 <213> *Homo sapiens*

 <400> 15
 gtgtacaccc tgcccccatc ccgggaggag atgaccaaga accaggtcag cctgacctgc 60
 ctggtcaaag gcttctaccc cagcgacatc gccgtggagt gggagagcag cgggcagccg 120
 gagaacaact acaacaccac gcctcccatg ctggactccg acggctcctt cttcctctac 180
 agcaagctca ccgtggacaa gagcaggtgg cagcagggga acatcttctc atgctccgtg 240
 atgcatgagg ctctgcacaa ccgcttcacg cagaagagcc tctccctgtc tccgggtaaa 300
 20
 <210> 16
 <211> 300
 <212> ADN
 <213> *Homo sapiens*
 25 <400> 16
 gtgtacaccc tgcccccatc ccaggaggag atgaccaaga accaggtcag cctgacctgc 60
 ctggtcaaag gcttctaccc cagcgacatc gccgtggagt gggagagcaa tgggcagccg 120
 gagaacaact acaagaccac gcctcccgtg ctggactccg acggctcctt cttcctctac 180
 agcaggctaa ccgtggacaa gagcaggtgg caggagggga atgtcttctc atgctccgtg 240
 atgcatgagg ctctgcacaa ccactacaca cagaagagcc tctccctgtc tctgggtaaa 300
 30
 <210> 17
 <211> 60
 <212> ADN
 <213> Artificial
 35 <220>
 <223> IgA
 <400> 17

ES 2 665 920 T3

ggcaaacc ca cccatgtcaa tgtgtctgtt gtcattggcg aggtggacgg cacctgctac 60

<210> 18
<211> 60
5 <212> ADN
<213> Artificial

<220>
10 <223> IgE
<400> 18

catgaggcag cgagcccctc acagaccgtc cagcgagcgg tgtctgtaaa tcccggcaaa 60

15 <210> 19
<211> 60
<212> ADN
<213> Artificial

20 <220>
<223> IgM
<400> 19

25 ggcaaacc ca ccctgtacaa cgtgtccctg gtcattgtccg acacagctgg cacctgctac 60

<210> 20
<211> 60
<212> ADN
30 <213> Artificial

<220>
<223> IgG1 (variante 1)

35 <400> 20

atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggcaaa 60

40 <210> 21
<211> 60
<212> ADN
<213> Artificial

<220>
45 <223> IgG2 (variante 1)
<400> 21

atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggcaaa 60

50 <210> 22
<211> 60
<212> ADN
<213> Artificial

55 <220>
<223> IgG3
<400> 22

60 atgcatgagg ctctgcacaa ccgcttcacg cagaagagcc tctccctgtc tccgggcaaa 60

<210> 23
<211> 60

ES 2 665 920 T3

<212> ADN
 <213> Artificial
 <220>
 5 <223> IgG4
 <400> 23
 atgcatgagg ctctgcacaa ccactacaca cagaagagcc tctccctgtc tctgggcaaa 60
 10 <210> 24
 <211> 39
 <212> ADN
 <213> Artificial
 15 <220>
 <223> cebador de mutagénesis 1
 <400> 24
 20 agcctctccc tgtccccggg caaatgagtg cgacggccg 39
 <210> 25
 <211> 39
 25 <212> ADN
 <213> Artificial
 <220>
 <223> cebador de mutagénesis 2
 30 <400> 25
 cggccgtcgc actcatttgc ccggggacag ggagaggct 39
 35 <210> 26
 <211> 34
 <212> ADN
 <213> Artificial
 40 <220>
 <223> cebador de mutagénesis 3
 <400> 26
 45 gcctctccct gtccctgggc aaatgagtgc cagg 34
 <210> 27
 <211> 34
 <212> ADN
 50 <213> Artificial
 <220>
 <223> cebador de mutagénesis 4
 55 <400> 27
 cctggcactc atttgcccag ggacagggag aggc 34
 60 <210> 28
 <211> 321
 <212> ADN
 <213> *Homo sapiens*
 <400> 28

ES 2 665 920 T3

	gggcagcccc gagaaccaca ggtgtacacc ctgcccccat cccgggatga gctgaccaag	60
	aaccaggtca gcctgacctg cctgggtcaaa ggcttctatc ccagcgacat cgccgtggag	120
	tgggagagca atgggcagcc ggagaacaac tacaagacca cgctccccgt gctggactcc	180
	gacggctcct tcttctctta cagcaagctc accgtggaca agagcaggtg gcagcagggg	240
	aacgtcttct catgctccgt gatgcatgag gctctgcaca accactacac acagaagagc	300
	ctctccctgt ctccgggtaa a	321
5	<210> 29 <211> 321 <212> ADN <213> <i>Homo sapiens</i>	
10	<400> 29	
	gggcagcccc gagaaccaca ggtgtacacc ctgcccccat cccgggagga gatgaccaag	60
	aaccaggtca gcctgacctg cctgggtcaaa ggcttctacc ccagcgacat ctccgtggag	120
	tgggagagca atgggcagcc ggagaacaac tacaagacca cacctcccat gctggactcc	180
	gacggctcct tcttctctta cagcaagctc accgtggaca agagcaggtg gcagcagggg	240
	aacgtcttct catgctccgt gatgcatgag gctctgcaca accactacac gcagaagagc	300
	ctctccctgt ctccgggtaa a	321
15	<210> 30 <211> 60 <212> ADN <213> Artificial	
20	<220> <223> IgG1 (variante 2) <400> 30	
	atgcatgagg ctctgcacaa ccactacaca cagaagagcc tctccctgtc tccgggcaaa	60
25	<210> 31 <211> 213 <212> ADN <213> Artificial	
30	<220> <223> IgG2 (variante 2) <400> 31	
	atctccgtgg agtgggagag caatgggcag ccggagaaca actacaagac cacacctccc	60
	atgctggact ccgacggctc cttcttcttc tacagcaagc tcaccgtgga caagagcagg	120
	tggcagcagg ggaacgtctt ctcatgctcc gtgatgcatg aggctctgca caaccactac	180
35	acgcagaaga gcctctccct gtctccgggc aaa	213

REIVINDICACIONES

- 5 1. Un ácido nucleico que codifica una cadena pesada de inmunoglobulina con organización exón-intrón genómica retenida que comprende un ácido nucleico que codifica la secuencia de aminoácidos de la parte C-terminal del dominio C_H3 de una inmunoglobulina de la clase IgG, caracterizado por que el dipéptido glicina-lisina en el extremo C en dicha secuencia de aminoácidos de la parte C-terminal del dominio C_H3 es codificada por el ácido nucleico ggcaaa.
- 10 2. Un ácido nucleico de acuerdo con la reivindicación 1, caracterizado por que el ácido nucleico que codifica dicho dipéptido glicina-lisina va precedido por el nucleótido g o a.
- 15 3. Un ácido nucleico de acuerdo con la reivindicación 1, caracterizado por que dicho ácido nucleico que codifica la secuencia de aminoácidos de la parte C-terminal del dominio C_H3 de una inmunoglobulina de la clase IgG se selecciona entre los ácidos nucleicos de SEQ ID NO: 20, 23 o 30.
- 20 4. Un ácido nucleico de acuerdo con una cualquiera de las reivindicaciones previas, en el que la cadena pesada de inmunoglobulina es una cadena pesada de inmunoglobulina humana de la subclase IgG4 humana o una cadena pesada de inmunoglobulina mutada de la subclase IgG1 humana.
- 25 5. Un ácido nucleico de acuerdo con una cualquiera de las reivindicaciones previas, en el que la cadena pesada de inmunoglobulina es de la subclase IgG1 humana con las mutaciones L234A y L235A.
6. Un ácido nucleico de acuerdo con una cualquiera de las reivindicaciones 1 a 4, en el que la cadena pesada de inmunoglobulina es una cadena pesada de inmunoglobulina IgG4 humana con la mutación S228P.
7. Un plásmido que comprende el ácido nucleico de acuerdo con la reivindicación 1.
8. Una célula que comprende el ácido nucleico de acuerdo con la reivindicación 1.
- 30 9. Un procedimiento para la producción de una inmunoglobulina en una célula de mamífero que comprende las siguientes etapas:
 - 35 a) transfectar dicha célula de mamífero con un ácido nucleico que codifica una cadena pesada de inmunoglobulina de acuerdo con la reivindicación 3,
 - b) cultivar la célula de mamífero transfectada en condiciones adecuadas para la expresión de la inmunoglobulina,
 - 40 c) recuperar la inmunoglobulina del cultivo o de la célula.
10. Procedimiento de acuerdo con la reivindicación 9, caracterizado por que dicha célula de mamífero se transfecta con un ácido nucleico que comprende un casete de expresión que codifica una cadena pesada de inmunoglobulina y un casete de expresión que codifica una cadena ligera de inmunoglobulina.
- 45 11. Procedimiento de acuerdo con la reivindicación 9, caracterizado por que dicha célula de mamífero se transfecta con dos ácidos nucleicos, en el que el primer ácido nucleico comprende un casete de expresión que codifica una cadena ligera de inmunoglobulina, y el segundo ácido nucleico comprende un casete de expresión que codifica una cadena pesada de inmunoglobulina.

Figura 1

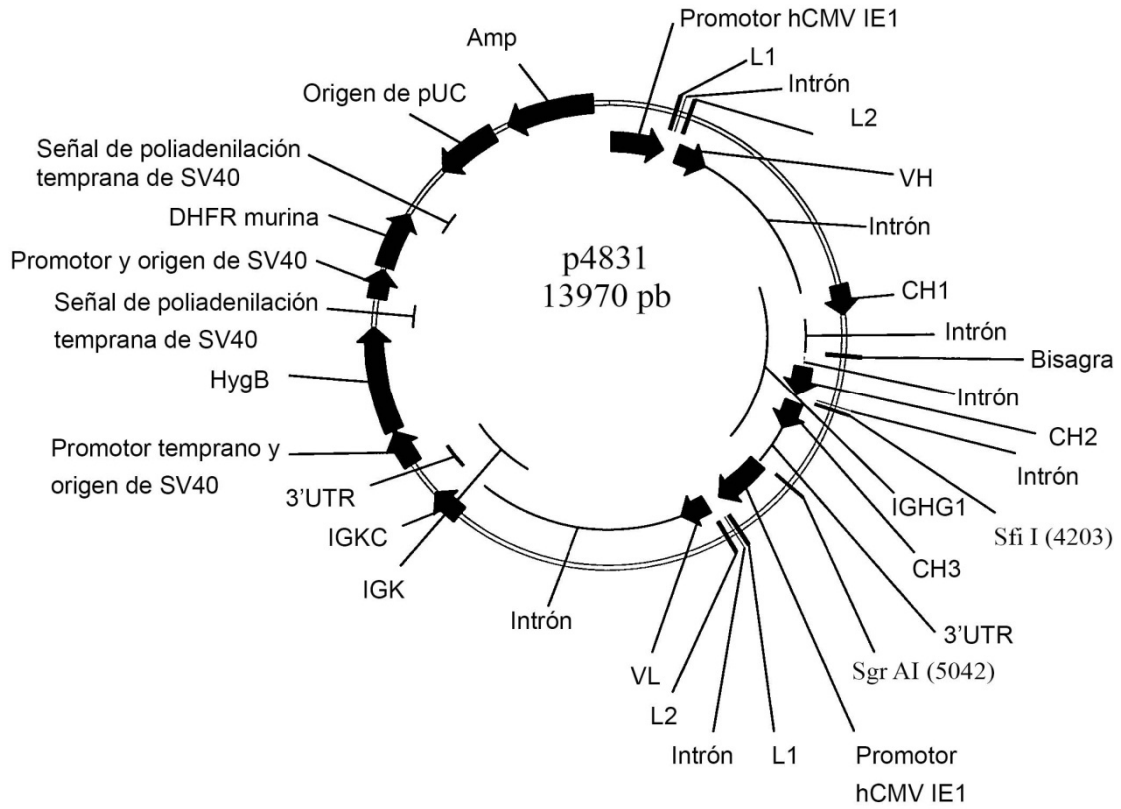


Figura 2

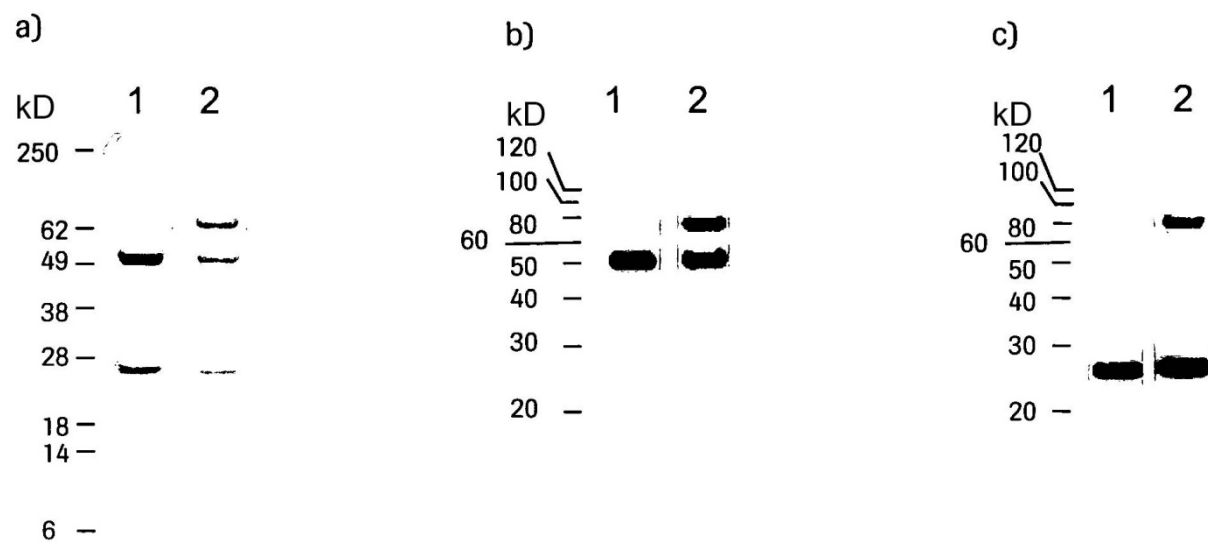


Figura 3

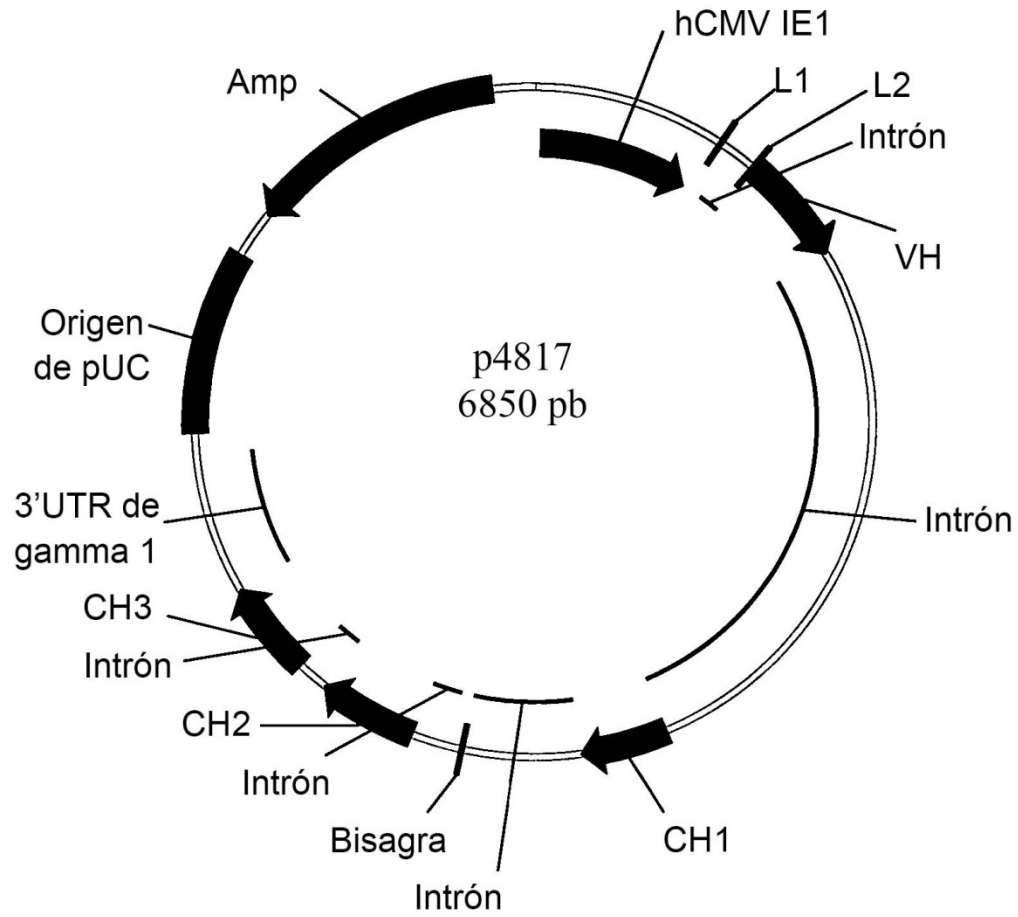


Figura 4

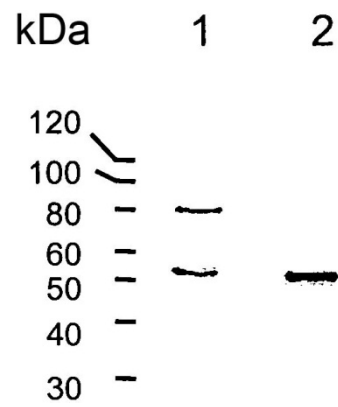


Figura 5

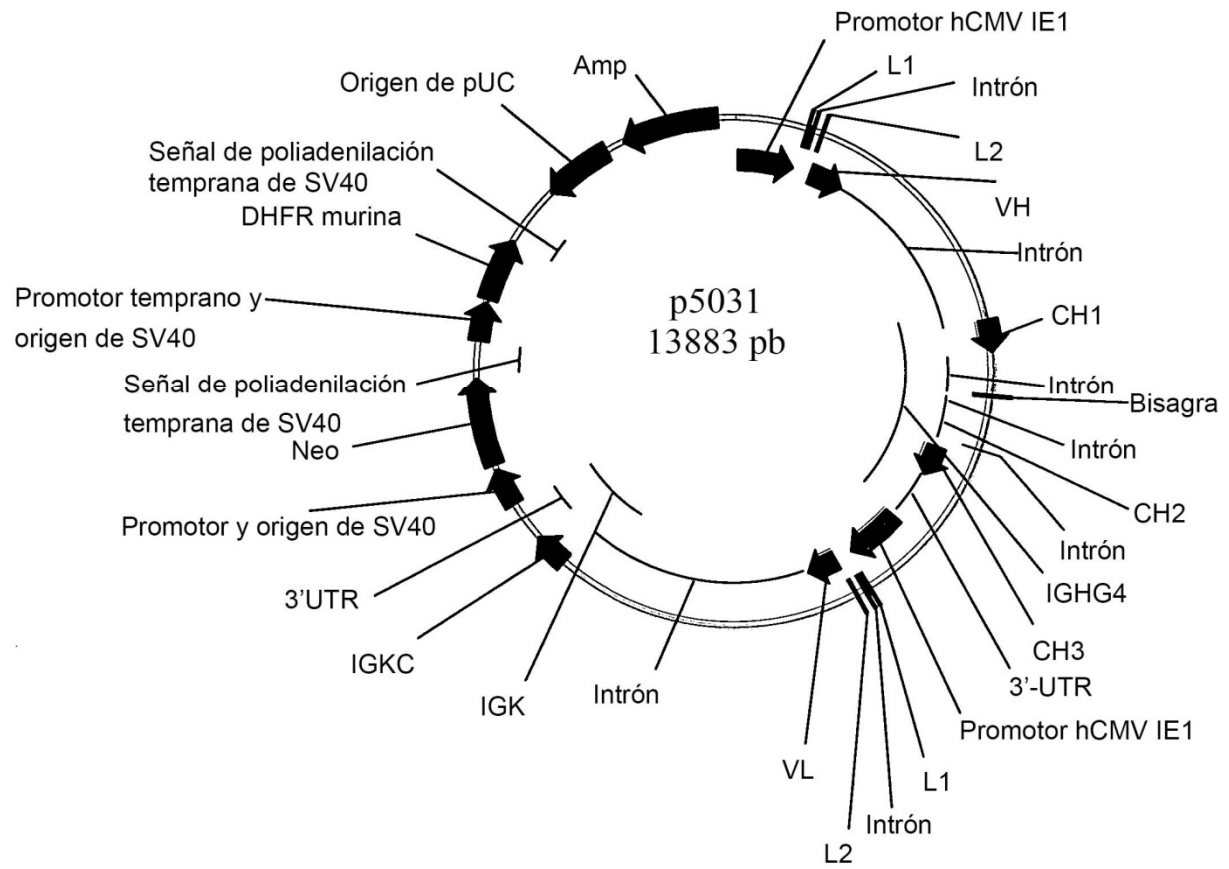


Figura 6

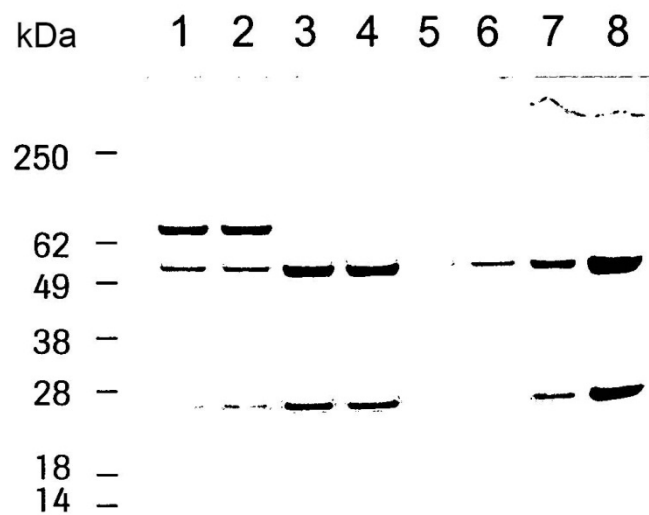


Figura 7

