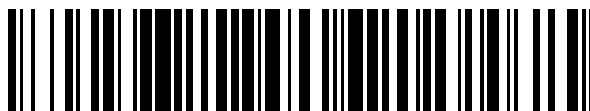


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 383 592**

51 Int. Cl.:
C07K 14/22 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **02726389 .6**
96 Fecha de presentación: **28.03.2002**
97 Número de publicación de la solicitud: **1379551**
97 Fecha de publicación de la solicitud: **14.01.2004**

54 Título: **Toxinas bacterianas ADP-ribosilantes**

30 Prioridad:
30.03.2001 GB 0108024

45 Fecha de publicación de la mención BOPI:
22.06.2012

45 Fecha de la publicación del folleto de la patente:
22.06.2012

73 Titular/es:
Novartis Vaccines and Diagnostics S.r.l.
Via Fiorentina 1
53100 Siena (SI), IT

72 Inventor/es:
MASIGNANI, Vega;
PIZZA, Mariagrazia y
RAPPUOLI, Rino

74 Agente/Representante:
Carpintero López, Mario

ES 2 383 592 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

Toxinas bacterianas ADP-ribosilantes

Campo técnico

La presente invención está en el campo de las toxinas bacterianas ADP-ribosilantes y sus usos.

5 Técnica anterior

Las exotoxinas bacterianas ADP-ribosilantes son ampliamente conocidas. Ejemplos incluyen toxina diftérica (*Corynebacterium diphtheriae*), exotoxina A (*Pseudomonas aeruginosa*), toxina colérica (CT; *Vibrio cholerae*), enterotoxina lábil al calor (LT; *E. coli*) y toxina pertussis (PT).

10 Las toxinas catalizan la transferencia de una unidad de ADP-ribosa de NAD⁺ a una proteína diana. CT, por ejemplo, transfiere ADP-ribosa a una cadena lateral de arginina específica de una subunidad de G_S, que bloquea la capacidad de G_S para hidrolizar GTP a GDP. Esto cierra la proteína en su forma 'activa', por lo que la actividad de la adenilato ciclasa está permanentemente activada. Los niveles de cAMP celular aumentan, conduciendo al transporte activo de iones de la célula y la pérdida de agua en el intestino [1].

15 Las toxinas se dividen normalmente en dos dominios funcionalmente distintos - A y B. La subunidad A es responsable de la actividad enzimática tóxica, mientras que la subunidad B es responsable de la unión celular. Las subunidades podrían ser dominios en la misma cadena de polipéptidos, o podrían ser cadenas de polipéptidos separadas. Las subunidades pueden por sí mismas ser oligómeros, por ejemplo, la subunidad A de CT consiste en A₁ y A₂ que están ligadas por un enlace disulfuro, y su subunidad B es un homopentámero. Normalmente, el contacto inicial con una célula diana está mediado por la subunidad B y luego la subunidad A sola entra en la célula.

20 Se conocen estructuras cristalinas [2] para LT [3], CT [4] y PT [5].

Las toxinas son normalmente inmunogénicas, y se han propuesto para su uso en vacunas acelulares. Un problema, sin embargo, es que las proteínas retienen su actividad tóxica en las vacunas. Para evitar este problema, la mutagénesis dirigida a sitio de residuos de sitio activo claves se ha usado para eliminar actividad enzimática tóxica, a la vez que se retiene la inmunogenicidad [por ejemplo refs. 6 (CT y LT), 7 (PT), 8 etc.]. Las actuales vacunas
25 contra la tos ferina acelulares incluyen una forma de toxina pertussis con dos sustituciones de aminoácidos (Arg⁹→Lys y Glu¹²⁹→Gly; 'PT-9K/129G' [9]).

Además de sus propiedades inmunogénicas, las toxinas se han usado como adyuvantes. La adyuvancia parenteral se observó por primera vez en 1972 [10] y la adyuvancia mucosa en 1984 [11]. Se encontró sorprendentemente en 1993 que las formas desintoxicadas de las toxinas retienen la adyuvancia [12].

30 Es un objeto de la invención proporcionar más toxinas bacterianas ADP-ribosilantes.

Divulgación de la invención

Las secuencias de aminoácidos de seis toxinas ADP-ribosilantes diferentes de bacterias Gram-positivas y Gram-negativas se facilitan en SEC ID 1, 3, 4, 5, 6 y 7. Estas toxinas son de *Neisseria meningitidis*, *Streptomyces coelicolor*, *Mycoplasma pneumoniae*, *Salmonella typhimurium*, *Salmonella paratyphi* y *Streptococcus pyogenes*. La
35 existencia de toxinas ADP-ribosilantes en estas especies bacterianas no se ha sugerido previamente y, además, sólo hay un bajo nivel de identidad de secuencias entre estas toxinas y toxinas tales como CT, LT y PT.

Toxinas mutantes de la invención

Toxinas mutantes de la invención

40 La invención proporciona una proteína que consiste en una secuencia de aminoácidos de SEC ID 1, excepto que la secuencia de aminoácidos contiene entre una y veinte mutación (mutaciones), en la que cada mutación implica un único aminoácido y en la que la mutación (mutaciones) reduce(n) o elimina(n) actividad ADP-ribosilante de la proteína.

45 Las mutaciones pueden ser cada una independientemente una sustitución, una inserción o una deleción. Las secuencias de aminoácidos contienen menos de veinte mutaciones (por ejemplo, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 ó 1).

Las mutaciones preferidas son sustituciones de un único aminoácido (por ejemplo, SEC ID 10, 12, 14 y 16).

50 La invención también proporciona un procedimiento para disminuir la actividad ADP-ribosilante enzimática de una toxina de la invención, que comprende mutar uno o más residuos de aminoácidos de dicha toxina. Esto puede lograrse convenientemente realizando mutagénesis dirigida a sitio en el ácido nucleico que codifica la toxina. La invención proporciona además una proteína que puede obtenerse por este procedimiento.

Las mutaciones también pueden introducirse para mejorar la estabilidad, por ejemplo, la inserción de enlaces disulfuro [13].

Las proteínas definidas anteriormente se denominan más adelante en este documento "toxinas mutantes de la invención" o "toxoides de la invención".

- 5 Los sitios preferidos para la mutación se facilitan en la Tabla 1, junto con mutaciones preferidas en esos sitios.

Proteínas de la invención

La invención proporciona una proteína que comprende la secuencia de aminoácidos de una toxina mutante de la invención.

- 10 La divulgación también proporciona una proteína que comprende una secuencia de aminoácidos que tiene identidad de secuencias con la secuencia de aminoácidos de una toxina mutante de la invención. El grado de identidad de secuencias es preferentemente superior al 50% (por ejemplo, el 60%, 70%, 80%, 90%, 95%, el 99% o más). Estas proteínas incluyen homólogos, ortólogos, variantes alélicas y mutantes funcionales. Normalmente, el 50% de identidad o más entre dos proteínas se considera que es una indicación de equivalencia funcional. La identidad entre proteínas se determina preferentemente por el algoritmo de búsqueda de homología de Smith-Waterman como se
15 implementa en el programa MPSRCH (Oxford Molecular) usando una búsqueda de huecos afin con los parámetros *penalidad abierta por huecos=12* y *penalidad por extensión de huecos=1*.

Las toxinas y proteínas mutantes definidas anteriormente se denominan en conjunto más adelante en este documento las "proteínas de la invención".

- 20 Las proteínas de la invención pueden, por supuesto, prepararse por diversos medios (por ejemplo, expresión recombinante, purificación a partir de huésped nativo, síntesis química, etc.) y en diversas formas (por ejemplo, nativa, fusiones etc.). Se preparan preferentemente en forma sustancialmente pura (es decir, sustancialmente libre de proteínas de célula huésped).

La invención también proporciona las proteínas de la invención (particularmente las toxinas mutantes) para su uso como adyuvantes y, en particular, como adyuvantes de la mucosa.

- 25 La invención también proporciona el uso de las proteínas de la invención en la preparación de un medicamento para producir una respuesta inmunitaria en un animal. El medicamento es preferentemente una composición inmunogénica (por ejemplo, una vacuna) y comprenderá, además de una proteína de la invención, un antígeno contra el que va a producirse una respuesta inmunitaria. El medicamento se administra preferentemente mucosalmente, por ejemplo, por vía oral o intranasalmente.

- 30 La invención también proporciona composiciones inmunogénicas (por ejemplo, una vacuna) que comprenden una proteína de la invención en mezcla con un segundo antígeno. También proporciona un kit que comprende una proteína de la invención y un segundo antígeno para administración simultánea, separada o secuencial. El segundo antígeno es preferentemente uno de las proteínas de *N. meningitidis* desveladas en las referencias 14 a 20. La composición puede comprender un tercer antígeno, un cuarto antígeno, un quinto antígeno, etc., uno o más de los
35 cuales puede seleccionarse de las proteínas de *N. meningitidis* desveladas en estas siete referencias.

Según otro aspecto, la invención proporciona anticuerpo que se une a una proteína de la invención. Éstos pueden ser policlonales o monoclonales y pueden producirse por cualquier medio adecuado. El anticuerpo puede incluir una marca detectable.

- 40 Según otro aspecto, la invención proporciona ácido nucleico que codifica las proteínas de la invención. Éste puede usarse en reacciones de hibridación (por ejemplo, transferencias Northern o Southern, o en micromatrices de ácido nucleico o 'chips de genes') y reacciones de amplificación (por ejemplo, PCR, SDA, SSSR, LCR, NASBA, TMA), etc.

La invención también proporciona ácido nucleico que comprende una o más de SEC ID 11, 13, 15, 17, 18, 20, 21, 22, 23, 24, 25 y/o 26.

- 45 También debe apreciarse que la divulgación proporciona ácido nucleico que comprende complementariedad de secuencias con aquellas descritas anteriormente (por ejemplo, para antisentido o sondaje, o para su uso como cebadores).

El ácido nucleico según la invención puede, por supuesto, prepararse de muchas formas (por ejemplo, por síntesis química, a partir de bibliotecas genómicas o de ADNc, a partir del propio organismo, etc.) y puede tomar diversas formas (por ejemplo, monocatenario, bicatenario, vectores, cebadores, sondas, etc.).

- 50 El ácido nucleico según la invención puede marcarse, por ejemplo, con una marca radiactiva o fluorescente. Esto es particularmente útil cuando el ácido nucleico va a usarse como cebador o sonda, por ejemplo, en PCR, LCR o TMA.

Además, el término "ácido nucleico" incluye ADN y ARN, y también sus análogos, tales como aquellos que contienen esqueletos modificados, y también ácidos nucleicos peptídicos (PNA), etc.

Según otro aspecto, la invención proporciona vectores que comprenden ácido nucleico de la invención (por ejemplo, vectores de clonación o de expresión) y células huésped transformadas con tales vectores.

- 5 Según otro aspecto, la invención proporciona composiciones que comprenden proteína, anticuerpo y/o ácido nucleico según la invención. Estas composiciones pueden ser adecuadas como composiciones inmunogénicas, por ejemplo, o como reactivos de diagnóstico, o como vacunas.

10 La invención también proporciona ácido nucleico, proteína, o anticuerpo según la invención para su uso como medicamentos (por ejemplo, vacunas) o reactivos de diagnóstico. La divulgación también proporciona el uso de ácido nucleico, proteína o anticuerpo de la invención en la preparación de: (i) un medicamento para tratar o prevenir infección bacteriana; (ii) un reactivo de diagnóstico para detectar la presencia de bacterias o de anticuerpos producidos contra bacterias; y/o (iii) un reactivo que pueden producir anticuerpos contra bacterias. Dichas bacterias son preferentemente *Neisseria meningitidis*.

15 Sigue un resumen de técnicas y procedimientos convencionales que pueden emplearse para realizar la invención (por ejemplo, para utilizar las secuencias desveladas para la vacunación o fines de diagnóstico). Este resumen no es una limitación de la invención sino que, más bien, da ejemplos que pueden usarse, pero no se requieren.

General

20 La práctica de la presente invención empleará, a menos que se indique lo contrario, técnicas convencionales de biología molecular, microbiología, ADN recombinante e inmunología, que están dentro de la habilidad de la materia. Tales técnicas se explican completamente en la bibliografía, por ejemplo, Sambrook Molecular Cloning: A Laboratory Manual, tercera edición (2001); DNA Cloning, volúmenes I y II (D.N Glover ed. 1985); Oligonucleotide Synthesis (M.J. Gait ed, 1984); Nucleic Acid Hybridization (B.D. Hames & S.J. Higgins eds. 1984); Transcription and Translation (B.D. Hames & S.J. Higgins eds. 1984); Animal Cell Culture (R.I. Freshney ed. 1986); Immobilized Cells and Enzymes (IRL Press, 1986); B. Perbal, A Practical Guide to Molecular Cloning (1984); the Methods in Enzymology series (Academic Press, Inc.), especialmente los volúmenes 154 y 155; Gene Transfer Vectors for Mammalian Cells (J.H. Miller y M.P. Calos eds. 1987, Cold Spring Harbor Laboratory); Mayer y Walker, eds. (1987), Immunochemical Methods in Cell and Molecular Biology (Academic Press, London); Scopes, (1987) Protein Purification: Principles and Practice, segunda edición (Springer-Verlag. N.Y.) y Handbook of Experimental Immunology, volúmenes I-IV (D.M. Weir and C. C. Blackwell eds 1986). En esta memoria descriptiva se usan abreviaturas estándar para nucleótidos y aminoácidos.

Definiciones

Una composición que contiene X está "sustancialmente libre de" Y cuando al menos el 85% en peso del X+Y total en la composición es X. Preferentemente, X comprende al menos aproximadamente el 90% en peso del total de X+Y en la composición, más preferentemente al menos aproximadamente el 95% o incluso el 99% en peso.

- 35 El término "que comprende" significa "que incluye", además de "que consiste", por ejemplo, una composición "que comprende" X pueden consistir exclusivamente en X o puede incluir algo adicional, por ejemplo, X + Y.

40 El término "heterólogo" se refiere a dos componentes biológicos que no se encuentran juntos en la naturaleza. Los componentes pueden ser células huésped, genes o regiones reguladoras tales como promotores. Aunque los componentes heterólogos no se encuentran juntos en la naturaleza, pueden funcionar juntos, como cuando un promotor heterólogo para un gen está operativamente ligado al gen. Otro ejemplo es cuando una secuencia de *Staphylococcus* es heteróloga a una célula huésped de ratón. Otro ejemplo serían dos epítopes de proteínas iguales o diferentes que se han ensamblado en una única proteína en una disposición no encontrada en la naturaleza.

45 Un "origen de replicación" es una secuencia de polinucleótidos que inicia y regula la replicación de polinucleótidos, tal como un vector de expresión. El origen de replicación se comporta como una unidad autónoma de replicación de polinucleótidos dentro de una célula, que puede ser replicación bajo su propio control. Un vector puede necesitar un origen de replicación para replicarse en una célula huésped particular. Con ciertos orígenes de replicación, un vector de expresión puede reproducirse a un alto número de copias en presencia de proteínas apropiadas dentro de la célula. Ejemplos de orígenes son las secuencias autónomamente replicantes que son eficaces en levadura; y el antígeno T vírico, eficaz en células COS-7.

50 Como se usa en este documento, una "variante alélica" de una molécula de ácido nucleico, o región, para la que la secuencia de ácidos nucleicos se proporciona en este documento es una molécula de ácido nucleico, o región, que se produce esencialmente en el mismo locus en el genoma de otra cepa aislada o segunda cepa aislada y que, debido a la variación natural producida por, por ejemplo, mutación o recombinación, tiene una secuencia de ácidos nucleicos similar, pero no idéntica. Una variante alélica de la región codificante normalmente codifica una proteína que tiene actividad similar a la de la proteína codificada por el gen con la que se está comparando. Una variante alélica también puede comprender una alteración en las regiones sin traducir 5' o 3' del gen, tal como en regiones de

control reguladoras (por ejemplo, véase la patente de EE.UU. 5.753.235).

Sistemas de expresión

Las secuencias de nucleótidos pueden expresarse en una variedad de diferentes sistemas de expresión; por ejemplo, aquellas usadas con células de mamífero, baculovirus, plantas, bacterias y levadura.

5 i. Sistemas de mamífero

Los sistemas de expresión de mamífero se conocen en la técnica. Un promotor de mamífero es cualquier secuencia de ADN que puede unirse a ARN polimerasa de mamífero e iniciar la transcripción en la dirección 3' de una secuencia codificante (por ejemplo, gen estructural) en ARNm. Un promotor tendrá una región de iniciación de la transcripción, que está normalmente situada proximal al extremo 5' de la secuencia codificante, y una caja TATA, normalmente localizada 25-30 pares de bases (pb) en la dirección 5' del sitio de iniciación de la transcripción. Se cree que la caja TATA dirige la ARN polimerasa II para empezar la síntesis de ARN en el sitio correcto. Un promotor de mamífero también contendrá un elemento promotor en la dirección 5', normalmente localizado dentro de 100-200 pb en la dirección 5' de la caja TATA. Un elemento promotor en la dirección 5' determina la tasa a la que se inicia la transcripción y pueden actuar en cualquier orientación [Sambrook y col.].

Los genes víricos de mamífero se expresan frecuentemente altamente y tienen una amplia gama de huéspedes; por tanto, las secuencias que codifican los genes víricos de mamífero proporcionan secuencias promotoras particularmente útiles. Ejemplos incluyen el promotor temprano del SV40, promotor del LTR del virus de tumor mamario de ratón, promotor tardío principal del adenovirus (Ad MLP) y promotor del virus del herpes simple. Además, las secuencias derivadas de genes no víricos (por ejemplo, el gen de la metalotioneína murina) también proporcionan secuencias promotoras útiles. La expresión puede ser tanto constitutiva como regulada (inducible), dependiendo del promotor puede inducirse con glucocorticoide en células sensibles a hormonas.

La presencia de un elemento potenciador (potenciador), combinado con los elementos promotores descritos anteriormente, normalmente aumentará los niveles de expresión. Un potenciador es una secuencia de ADN reguladora que puede estimular la transcripción hasta 1000 veces cuando se liga a promotores homólogos o heterólogos, empezando la síntesis en el sitio de iniciación del ARN normal. Los potenciadores también son activos cuando se sitúan en la dirección 5' o en la dirección 3' del sitio de iniciación de la transcripción, en orientación tanto normal como volcada, o a una distancia superior a 1000 nucleótidos desde el promotor [Maniatis y col. (1987) Science 236:1237; Alberts y col. (1989) Molecular Biology of the Cell, 2ª ed.]. Los elementos potenciadores derivados de virus pueden ser particularmente útiles, ya que normalmente tienen una gama de huéspedes más amplia. Ejemplos incluyen el potenciador del gen temprano del SV40 [Dijkema y col. (1985) EMBO J. 4:761] y el potenciador/promotores derivados de la repetición terminal larga (LTR) del virus del sarcoma de Rous [Gorman y col. (1982b) PNAS USA 79:6777] y del citomegalovirus humano [Boshart y col. (1985) Cell 41:521]. Adicionalmente, algunos potenciadores son regulables y se vuelven activos sólo en presencia de un inductor tal como una hormona o ión metálico [Sassone-Corsi y Borelli (1986) Trends Genet. 2:215; Maniatis y col. (1987) Science 236:1237].

Una molécula de ADN puede expresarse intracelularmente en células de mamífero. Una secuencia promotora puede ligarse directamente a la molécula de ADN, en cuyo caso el primer aminoácido en el extremo N de la proteína recombinante siempre será una metionina, que está codificada por el codón de iniciación ATG. Si se desea, el extremo N puede escindirse de la proteína por incubación *in vitro* con bromuro de cianógeno.

Alternativamente, proteínas extrañas también pueden secretarse de la célula en los medios de crecimiento creando moléculas de ADN quiméricas que codifican una proteína de fusión que comprende un fragmento de la secuencia conductora que proporciona la secreción de la proteína extraña en células de mamífero. Preferentemente, hay sitios de procesamiento codificados entre el fragmento conductor y el gen extraño que pueden escindirse tanto *in vivo* como *in vitro*. El fragmento de la secuencia conductora normalmente codifica un péptido señal que comprende aminoácidos hidrófobos que dirigen la secreción de la proteína de la célula. El conductor tripartito de adenovirus es un ejemplo de una secuencia conductora que proporciona secreción de una proteína extraña en células de mamífero.

Normalmente, las secuencias de terminación de la transcripción y de poliadenilación reconocidas por células de mamífero son regiones reguladoras localizadas 3' al codón de terminación de la traducción y, por tanto, junto con los elementos promotores, flanquean la secuencia codificante. El extremo 3' del ARNm maduro se forma por escisión postraduccional específica de sitio y poliadenilación [Birnstiel y col. (1985) Cell 41:349; Proudfoot & Whitelaw (1988) "Termination and 3' end processing of eukaryotic RNA. En Transcription and splicing (ed. B.D. Hames y D.M. Glover); Proudfoot (1989) Trends Biochem. Sci. 14:105]. Estas secuencias dirigen la transcripción de un ARNm que puede traducirse en el polipéptido codificado por el ADN. Ejemplos de señales terminadoras de la transcripción/de poliadenilación incluyen aquellas derivadas del SV40 [Sambrook y col.].

Normalmente, los componentes anteriormente descritos, que comprenden una secuencia promotora, señal de la poliadenilación y de terminación de la transcripción se ponen juntos en construcciones de expresión. Potenciadores, intrones con sitios de donante y aceptor de corte y empalme funcionales y secuencias conductoras también pueden incluirse en una construcción de expresión, si se desea. Las construcciones de expresión se mantienen

frecuentemente en un replicón, tal como un elemento extracromosómico (por ejemplo, plásmidos) que pueden mantenerse establemente en un huésped, tal como células de mamífero o bacterias. Los sistemas de replicación de mamífero incluyen aquellos derivados de virus de animales que requieren factores transactivantes para replicarse. Por ejemplo, plásmidos que contienen los sistemas de replicación de papovavirus tales como SV40 [Gluzman (1981) Cell 23:175] o virus del polioma se replican en número de copias extremadamente alto en presencia del antígeno T vírico apropiado. Ejemplos adicionales de replicones de mamífero incluyen aquellos derivados de virus del papiloma bovino y virus de Epstein-Barr. Adicionalmente, el replicón puede tener dos sistemas de replicón, permitiendo así que se mantengan, por ejemplo, en células de mamífero para la expresión y en un huésped procariota para la clonación y amplificación. Ejemplos de tales vectores lanzadera de bacterias de mamífero incluyen pMT2 [Kaufman y col. (1989) Mol. Cell. Biol. 9:946] y pHEBO [Shimizu y col. (1986) Mol. Cell. Biol. 6:1074].

Los procedimientos de transformación usados dependen del huésped que va a transformarse. Los procedimientos para la introducción de polinucleótidos heterólogos en células de mamífero se conocen en la técnica e incluyen transfección mediada por dextrano, precipitación con fosfato de calcio, transfección mediada por polibreno, fusión de protoplastos, electroporación, encapsulación de polinucleótido(s) en liposomas y microinyección directa de ADN en núcleos.

Las líneas celulares de mamífero disponibles como huéspedes para la expresión se conocen en la técnica e incluyen muchas líneas celulares inmortalizadas disponibles de la Colección americana de cultivos tipo (ATCC) que incluyen, pero no se limitan a, células de ovario de hámster chino (CHO), células HeLa, células de riñón de hámster bebé (BHK), células de riñón de mono (COS), células de carcinoma hepatocelular humano (por ejemplo, Hep G2), y varias otras líneas celulares.

ii. Sistemas de baculovirus

El polinucleótido que codifica la proteína también puede insertarse en un vector de expresión de insecto adecuado, y está operativamente ligado a los elementos de control dentro de ese vector. La construcción de vectores emplea técnicas que se conocen en la técnica. Generalmente, los componentes del sistema de expresión incluyen un vector de transferencia, normalmente un plásmido bacteriano, que contiene tanto un fragmento del genoma de baculovirus como un sitio de restricción conveniente para la inserción del gen o genes heterólogos que van a expresarse; un baculovirus natural con una secuencia homóloga al fragmento de baculovirus específico en el vector de transferencia (esto permite la recombinación homóloga del gen heterólogo en el genoma de baculovirus); y células huésped de insecto apropiadas y medios de crecimiento.

Después de insertarse la secuencia de ADN que codifica la proteína en el vector de transferencia, el vector y el genoma vírico natural se transfieren en una célula huésped de insecto en la que se permite que el vector y el genoma vírico se recombinen. El virus recombinante encapsidado se expresa y las placas recombinantes se identifican y se purifican. Los materiales y procedimientos para los sistemas de expresión de células de baculovirus/insecto están comercialmente disponibles en forma de kit de, entre otros, In vitrogen, San Diego CA (kit "MaxBac"). Estas técnicas son generalmente conocidas para aquellos expertos en la materia y se describen completamente en Summers & Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987) (denominado en lo sucesivo "Summers & Smith").

Antes de insertarse la secuencia de ADN que codifica la proteína en el genoma de baculovirus, los componentes anteriormente descritos, que comprenden un promotor, conductor (si se desea), secuencia codificante y secuencia de terminación de la transcripción, se ensamblan normalmente en una construcción de sustitución intermedia (vector de transferencia). Ésta puede contener un único gen y elementos reguladores operativamente ligados; múltiples genes, cada uno con su conjunto propio de elementos reguladores operativamente ligados; o múltiples genes, regulados por el mismo conjunto de elementos reguladores. Las construcciones de sustitución intermedia se mantienen frecuentemente en un replicón, tal como un elemento extracromosómico (por ejemplo, plásmidos), que pueden mantenerse establemente en un huésped, tal como una bacteria. El replicón tendrá un sistema de replicación, permitiéndose así que se mantenga en un huésped adecuado para la clonación y amplificación.

Actualmente, el vector de transferencia más comúnmente usado para la introducción de genes extraños en AcNPV es pAc373. También se han diseñado muchos otros vectores, conocidos para aquellos expertos en la materia. Éstos incluyen, por ejemplo, pVL985, que altera el codón de iniciación de la polihedrina de ATG a ATT, y que introduce un sitio de clonación BamHI de 32 pb en la dirección 3' desde ATT; véase Luckow & Summers, Virology (1989) 17:31.

El plásmido normalmente también contiene la señal de poliadenilación de la polihedrina (Miller y col. (1988) Ann. Rev. Microbiol., 42:177) y un gen de resistencia a ampicilina procariota (*amp*) y origen de replicación para la selección y propagación en *E. coli*.

Los vectores de transferencia de baculovirus normalmente contienen un promotor de baculovirus. Un promotor de baculovirus es cualquier secuencia de ADN que pueda unirse a ARN polimerasa de baculovirus e iniciar la transcripción en la dirección 3' (5' a 3') de una secuencia codificante (por ejemplo, gen estructural) en ARNm. Un promotor tendrá una región de iniciación de la transcripción que está normalmente situada proximal al extremo 5' de la secuencia codificante. Esta región de iniciación de la transcripción normalmente incluye un sitio de unión a ARN

polimerasa y un sitio de iniciación de la transcripción. Un vector de transferencia de baculovirus puede también tener un segundo dominio llamado un potenciador que, si está presente, está normalmente distal al gen estructural. La expresión puede ser tanto regulada como constitutiva.

5 Los genes estructurales, abundantemente transcritos en momentos posteriores en un ciclo de infección vírica, proporcionan secuencias promotoras particularmente útiles. Ejemplos incluyen secuencias derivadas del gen que codifica la proteína de la polihedrina vírica, Friesen y col., (1986) "The Regulation of Baculovirus Gene Expression" en: *The Molecular Biology of Baculoviruses* (ed. Walteran Doerfler); EPO Publ. Nos. 127 839 y 155 476; y el gen que codifica la proteína p10, Vlax y col., (1988). *J. Gen. Virol.* 69:765.

10 El ADN que codifica secuencias señal adecuadas puede derivarse de genes para proteínas de insecto o baculovirus secretadas tales como el gen de la polihedrina de baculovirus (Carbonell y col. (1988) *Gene*, 73:409). Alternativamente, como las señales para las modificaciones postraduccionales de células de mamíferos (tales como señal de escisión de péptidos, escisión proteolítica y fosforilación) parecen ser reconocidas por células de insecto, y las señales requeridas para la secreción y acumulación nuclear también parece que se conservan entre las células invertebradas y las células vertebradas, los conductores de origen no insecto, tales como aquellos derivados de genes que codifican interferón α humano, Maeda y col., (1985), *Nature* 315:592; péptido liberador de gastrina humana, Lebacq-Verheyden y col., (1988), *Molec. Cell. Biol.* 8:3129; IL-2 humana, Smith y col., (1985) *Proc. Nat'l Acad. Sci. USA*, 82:8404; IL-3 de ratón, (Miyajima y col., (1987) *Gene* 58:273); y glucocerebrosidasa humana, Martin y col. (1988) *DNA*, 7:99, también pueden usarse para proporcionar secreción en insectos.

20 Un polipéptido o poliproteína recombinante pueden expresarse intracelularmente o, si se expresan con las secuencias reguladoras apropiadas, pueden secretarse. La buena expresión intracelular de proteínas extrañas no infundidas normalmente requiere genes heterólogos que idealmente tienen una secuencia conductora corta que contiene señales de iniciación de la traducción adecuadas que preceden a una señal de iniciación de ATG. Si se desea, la metionina en el extremo N puede escindirse de la proteína madura por incubación *in vitro* con bromuro de cianógeno.

25 Alternativamente, las poliproteínas o proteínas recombinantes que no son naturalmente secretadas pueden secretarse de la célula de insecto creando moléculas de ADN quiméricas que codifican una proteína de fusión que comprende un fragmento de la secuencia conductora que proporciona la secreción de la proteína extraña en insectos. El fragmento de la secuencia conductora normalmente codifica un péptido señal que comprende aminoácidos hidrófobos que dirigen la translocalización de la proteína en el retículo endoplásmico.

30 Después de la inserción de la secuencia de ADN y/o el gen que codifica el producto de expresión precursor de la proteína, una célula huésped de insecto se co-transforma con el ADN heterólogo del vector de transferencia y el ADN genómico de baculovirus natural - normalmente por co-transfección. El promotor y la secuencia de terminación de la transcripción de la construcción normalmente comprenderán una sección de 2-5 kb del genoma de baculovirus. Los procedimientos para introducir ADN heterólogo en el sitio deseado en el virus de baculovirus se conocen en la técnica (véanse Summers & Smith, arriba; Ju y col. (1987); Smith y col., *Mol. Cell. Biol.* (1983) 3:2156; y Luckow y Summers (1989)). Por ejemplo, la inserción puede ser en un gen tal como el gen de la polihedrina, por recombinación cruzada doble homóloga; la inserción también puede ser en un sitio de enzima de restricción manipulado en el gen de baculovirus deseado. Miller y col., (1989), *Bioessays* 4:91. La secuencia de ADN, cuando se clona en lugar del gen de la polihedrina en el vector de expresión, está flanqueada tanto en 5' como en 3' por secuencias específicas de la polihedrina y está posicionada en la dirección 3' del promotor de polihedrina.

35 El vector de expresión de baculovirus recientemente formado se encapsida posteriormente en un baculovirus recombinante infeccioso. La recombinación homóloga se produce a baja frecuencia (entre aproximadamente el 1% y aproximadamente el 5%); por tanto, la mayoría del virus producido después de la cotransfección es todavía virus natural. Por tanto, es necesario un procedimiento para identificar virus recombinantes. Una ventaja del sistema de expresión es un cribado visual que permita distinguir virus recombinantes. La proteína de la polihedrina, que se produce por el virus nativo, se produce a niveles muy altos en los núcleos de células infectadas en momentos posteriores después de la infección vírica. La proteína de la polihedrina acumulada forma cuerpos de oclusión que también contienen partículas incorporadas. Estos cuerpos de oclusión, de hasta 15 μm de tamaño, son altamente refractivos, haciendo que tengan un aspecto lustroso brillante que es fácilmente visualizado bajo el microscopio óptico. Células infectadas con virus recombinantes carecen de cuerpos de oclusión. Para distinguir virus recombinante de virus natural, el sobrenadante de transfección se siembra en placa sobre una monocapa de células de insecto por técnicas conocidas para aquellos expertos en la materia. Concretamente, las placas se criban bajo el microscopio óptico para la presencia (indicativa de virus natural) o ausencia (indicativa de virus recombinante) de cuerpos de oclusión. "Current Protocols in Microbiology" vol. 2 (Ausubel y col. eds) en 16.8 (sup. 10, 1990); Summers & Smith, arriba; Miller y col. (1989).

55 Los vectores de expresión de baculovirus recombinantes se han desarrollado para la infección en varias células de insecto. Por ejemplo, los baculovirus recombinantes se han desarrollado para, entre otros: *Aedes aegypti*, *Autographa californica*, *Bombyx mori*, *Drosophila melanogaster*, *Spodoptera frugiperda* y *Trichoplusia ni* (documento WO 89/046699; Carbonell y col., (1985) *J. Virol.* 56:153; Wright (1986) *Nature* 321:718; Smith y col., (1983) *Mol. Cell. Biol.* 3:2156; y véase generalmente, Fraser y col. (1989) *In Vitro Cell. Dev. Biol.* 25:225).

Las células y los medios de cultivo celular están comercialmente disponibles para tanto expresión directa como de fusión de polipéptidos heterólogos en un baculovirus/sistema de expresión; la tecnología de cultivo celular es generalmente conocida para aquellos expertos en la materia. Véase, por ejemplo, Summers & Smith, arriba.

- 5 Las células de insecto modificadas pueden entonces cultivarse en un medio nutritivo apropiado que permita el mantenimiento estable del (de los) plásmido(s) presente(s) en el huésped de insecto modificado. Si el gen del producto de expresión está bajo el control inducible, el huésped puede cultivarse a alta densidad, e inducirse la expresión. Alternativamente, si la expresión es constitutiva, el producto se expresará continuamente en el medio y el medio nutritivo deberá circularse continuamente, a la vez que se extrae el producto de interés y se aumentan los nutrientes agotados. El producto puede purificarse por técnicas tales como cromatografía, por ejemplo, HPLC, 10 cromatografía de afinidad, cromatografía de intercambio iónico, etc.; electroforesis; centrifugación en gradiente de densidad; extracción en disolvente, etc. Como es apropiado, el producto puede purificarse adicionalmente, según se requiera, de manera que se elimine sustancialmente cualquier proteína de insecto que también está presente en el medio, de manera que se proporcione un producto que está al menos sustancialmente libre de residuos huésped, por ejemplo, proteínas, lípidos y polisacáridos.
- 15 Con el fin de obtener la expresión de proteínas, células huésped recombinantes derivadas de los transformantes se incuban en condiciones que permitan la expresión de la secuencia codificante de la proteína recombinante. Estas condiciones variarán, dependiendo de la célula huésped seleccionada. Sin embargo, las condiciones pueden determinarse fácilmente por aquellos expertos en la materia, basándose en lo que se sabe en la técnica.

iii. Sistemas de plantas

- 20 Hay muchos sistemas de expresión genéticos de cultivos de células de plantas y de plantas completas conocidos en la técnica. Sistemas de expresión genéticos celulares de plantas a modo de ejemplo incluyen aquellos descritos en patentes tales como: US 5.693.506; US 5.659.122; y US 5.608.143. Ejemplos adicionales de expresión genética en cultivo de células de plantas se han descrito por Zenk, *Phytochemistry* 30:3861-3863 (1991). Las descripciones de péptidos señal de proteínas de plantas pueden encontrarse, además de en las referencias descritas anteriormente, 25 en Vaulcombe y col., *Mol. Gen. Genet.* 209:33-40 (1987); Chandler y col., *Plant Molecular Biology* 3:407-418 (1984); Rogers, *J. Biol. Chem.* 260:3731-3738 (1985); Rothstein y col., *Gene* 55:353-356 (1987); Whittier y col., *Nucleic Acids Research* 15:2515-2535 (1987); Wirsal y col., *Molecular Microbiology* 3:3-14 (1989); Yu y col., *Gene* 122:247-253 (1992). Una descripción de la regulación de la expresión génica de plantas por la fitohormona, ácido giberélico y enzimas secretadas inducidas por ácido giberélico puede encontrarse en Jones & MacMillin, páginas 21-52 de *Advanced Plant Physiology*, Malcolm B. Wilkins, ed., 1984 Pitman Publishing Limited. Referencias que describen otros genes metabólicamente regulados: Sheen, *Plant Cell*, 2:1027-1038(1990); Maas y col., *EMBO J.* 9:3447-3452 (1990); Benkel & Hickey, *PNAS USA.* 84:1337-1339 (1987).

- Normalmente, usando técnicas conocidas en la técnica, una secuencia de polinucleótidos deseada se inserta en un casete de expresión que comprende elementos reguladores genéticos diseñados para el funcionamiento en plantas. 35 El casete de expresión se inserta en un vector de expresión deseado con secuencias de acompañamiento en la dirección 5' y en la dirección 3' del casete de expresión adecuado para la expresión en un huésped de planta. Las secuencias de acompañamiento serán de plásmido u origen vírico y proporcionan características necesarias al vector para permitir que los vectores muevan ADN desde un huésped de clonación original, tal como bacterias, hasta el huésped de planta deseado. La construcción del vector bacteriano/de planta básico preferentemente proporcionará un origen de replicación procariota de amplio intervalo de huéspedes; un marcador de selección procariota; y, para transformaciones en *Agrobacterium*, secuencias de T-ADN para la transferencia mediada por *Agrobacterium* a cromosomas de planta. Si el gen heterólogo no es fácilmente aceptado para la detección, la construcción también tendrá preferiblemente un gen del marcador de selección adecuado para determinar si una célula ha sido transformada. Una revisión general de marcadores adecuados, por ejemplo, para los miembros de la 45 familia de los pastos, se encuentra en Wilkink & Dons, 1993, *Plant Mol. Biol. Repr*, 11(2):165-185.

- También se recomiendan secuencias adecuadas para permitir la integración de la secuencia heteróloga en el genoma de la planta. Éstas podrían incluir secuencias de transposón y similares para la recombinación homóloga, además de secuencias de Ti que permitan la inserción al azar de un casete de expresión heterólogo en un genoma de la planta. Marcadores de selección procariotas adecuados incluyen resistencia hacia antibióticos tales como ampicilina o tetraciclina. Otras secuencias de ADN que codifican funciones adicionales también pueden estar presentes en el vector, como se conoce en la técnica. 50

- Las moléculas de ácidos nucleicos de la invención objeto puede incluirse en un casete de expresión para la expresión de la(s) proteína(s) de interés. Normalmente, sólo tendrán un casete de expresión, aunque sean viables dos o más. El casete de expresión recombinante contendrá, además de la secuencia codificante de proteínas heterólogas, los siguientes elementos, una región promotora, secuencias sin traducir en 5' de la planta, codón de iniciación que depende de si el gen estructural viene equipado o no con uno, y una secuencia de terminación de la transcripción y de la traducción. Sitios de enzimas de restricción únicos en los extremos 5' y 3' del casete permiten la fácil inserción en un vector preexistente. 55

Una secuencia codificante heteróloga puede ser para cualquier proteína relacionada con la presente invención. La secuencia que codifica la proteína de interés codificará un péptido señal que permite el procesamiento y la translocalización de la proteína, según sea apropiado, y normalmente carecerá de cualquier secuencia que pueda producir la unión de la proteína deseada de la invención a una membrana. Como, en general, la región de iniciación de la transcripción será para un gen que se expresa y se translocaliza durante la germinación, empleando el péptido señal que proporciona la translocalización, también puede proporcionarse una para la translocalización de la proteína de interés. De esta forma, la(s) proteína(s) de interés se translocalizará(n) a partir de las células en las que se expresan y pueden recogerse eficientemente. Normalmente, la secreción en semillas es a través de la capa de aleurona o de epitelio escutelar en el endosperma de la semilla. Aunque no se requiere que la proteína se secrete de las células en las que se produce la proteína, esto facilita el aislamiento y la purificación de la proteína recombinante.

Como la expresión final del producto génico deseado será en una célula eucariota, se desea determinar si cualquier porción del gen clonado contiene secuencias que se procesarán fuera como intrones por la maquinaria de esplicomas huésped. Si es así, la mutagénesis dirigida a sitio de la región de "intrón" puede realizarse para evitar la pérdida de una parte del mensajero genético como código de intrón falso, Reed & Maniatis, Cell 41:95-105, 1985.

El vector puede microinyectarse directamente en células vegetales usando micropipetas para transferir mecánicamente el ADN recombinante. Crossway, Mol. Gen. Genet, 202:179-185, 1985. El material genético también puede transferirse a la célula de planta usando polietilenglicol, Krens y col., Nature, 296, 72-74, 1982. Otro procedimiento de introducción de segmentos de ácido nucleico es la penetración balística a alta velocidad por pequeñas partículas con el ácido nucleico tanto dentro de la matriz de perlas o partículas pequeñas como sobre la superficie, Klein y col., Nature, 327, 70-73, 1987 y Knudsen y Muller, 1991, Planta, 185:330-336, que enseña que el bombardeo con partículas de endosperma de cebada crea cebada transgénica. Todavía otro procedimiento de introducción sería la fusión de protoplastos con otras entidades, tanto minicélulas, células, lisosomas como otros cuerpos con superficie de lípido fusibles, Fraley y col., PNAS USA, 79, 1859-1863, 1982.

El vector también puede introducirse en las células vegetales por electroporación (Fromm y col., PNAS USA 82:5824, 1985). En esta técnica, protoplastos de planta se electroporan en presencia de plásmidos que contienen la construcción de gen. Impulsos eléctricos de intensidad de campo alto permeabilizan reversiblemente las biomembranas permitiendo la introducción de los plásmidos. Los protoplastos de planta electroporados vuelven a formar la pared celular, dividen y forman callo de planta. Todas las plantas de las que pueden aislarse protoplastos y cultivarse para dar plantas regeneradas completas pueden transformarse por la presente invención de manera que se recuperen plantas completas que contienen el gen transferido. Se sabe que prácticamente todas las plantas pueden regenerarse a partir de células cultivadas o tejidos, que incluyen, pero no se limitan a, todas las especies importantes de caña de azúcar, remolacha azucarera, algodón, frutos y otros árboles, legumbres y verduras. Algunas plantas adecuadas incluyen, por ejemplo, especies de los géneros *Fragaria*, *Lotus*, *Medicago*, *Onobrychis*, *Trifolium*, *Trigonella*, *Vigna*, *Citrus*, *Linum*, *Geranium*, *Manihot*, *Daucus*, *Arabidopsis*, *Brassica*, *Raphanus*, *Sinapis*, *Atropa*, *Capsicum*, *Datura*, *Hyoscyamus*, *Lycopersion*, *Nicotiana*, *Solanum*, *Petunia*, *Digitalis*, *Majorana*, *Cichorium*, *Helianthus*, *Lactuca*, *Bromus*, *Asparagus*, *Antirrhinum*, *Hererocallis*, *Nemesia*, *Pelargonium*, *Panicum*, *Pennisetum*, *Ranunculus*, *Senecio*, *Salpiglossis*, *Cucumis*, *Browaalia*, *Glycine*, *Lolium*, *Zea*, *Triticum*, *Sorghum* y *Datura*.

Los medios para la regeneración varían entre especies de plantas, pero generalmente primero se proporciona una suspensión de protoplastos transformados que contienen copias del gen heterólogo. Se forma tejido de callo y pueden inducirse brotes de callo y posteriormente echar raíces. Alternativamente, la formación de embriones puede inducirse a partir de la suspensión de protoplastos. Estos embriones germinan como embriones naturales para formar plantas. Los medios de cultivo contendrán generalmente diversos aminoácidos y hormonas tales como auxina y citocininas. También es ventajoso añadir ácido glutámico y prolina al medio, especialmente para especies tales como trigo y alfalfa. Los brotes y raíces normalmente se desarrollan simultáneamente. La regeneración eficiente dependerá del medio, del genotipo y de la historia del cultivo. Si estas tres variables están controladas, entonces la regeneración es completamente reproducible y repetible.

En algunos sistemas de cultivo de células de planta, la proteína deseada de la invención puede secretarse o, alternativamente, la proteína puede extraerse de la planta completa. Si la proteína deseada de la invención es secretada en el medio, puede recogerse. Alternativamente, los embriones y las semisemillas sin embrión u otro tejido de planta pueden romperse mecánicamente para que libere cualquier proteína secretada entre células y tejidos. La mezcla puede suspenderse en una disolución de tampón para recuperar proteínas solubles. Entonces, el aislamiento convencional de proteínas y procedimientos de purificación se usará para purificar la proteína recombinante. Parámetros de tiempo, temperatura, pH, oxígeno y volúmenes se ajustarán mediante procedimientos rutinarios para optimizar la expresión y recuperación de proteína heteróloga.

iv. Sistemas bacterianos

Las técnicas de expresión bacteriana se conocen en la técnica. Un promotor bacteriano es cualquier secuencia de ADN que pueda unirse a ARN polimerasa bacteriana e iniciar la transcripción en la dirección 3' de una secuencia codificante (por ejemplo, gen estructural) en ARNm. Un promotor tendrá una región de iniciación de la transcripción que está normalmente situada proximal al extremo 5' de la secuencia codificante. Esta región de iniciación de la transcripción normalmente incluye un sitio de unión a ARN polimerasa y un sitio de iniciación de la transcripción. Un

promotor bacteriano pueden también tener un segundo dominio llamado un operador, que puede solapar un sitio de unión a ARN polimerasa adyacente en el que empieza la síntesis de ARN. El operador permite la transcripción regulada (inducible) negativa, ya que una proteína represora de genes puede unir el operador y así inhibir la transcripción de un gen específico. La expresión constitutiva puede producirse en ausencia de elementos reguladores negativos tales como el operador. Además, la regulación positiva puede lograrse por una secuencia de unión a proteína activadora de genes que, si está presente, está normalmente proximal (5') a la secuencia de unión a ARN polimerasa. Un ejemplo de una proteína activadora de genes es la proteína activadora de catabolitos (CAP), que ayuda a iniciar la transcripción del operón *lac* en *E. coli* [Raibaud y col. (1984) Annu. Rev. Genet. 18:173]. Por tanto, la expresión regulada puede ser tanto positiva como negativa, tanto potenciando como reduciendo así la transcripción.

Las secuencias que codifican las enzimas de la vía metabólica proporcionan secuencias promotoras particularmente útiles. Ejemplos incluyen secuencias promotoras derivadas de enzimas metabolizantes de azúcares tales como galactosa, lactosa (*lac*) [Chang y col. (1977) Nature 198:1056] y maltosa. Ejemplos adicionales incluyen secuencias promotoras derivadas de enzimas biosintéticas tales como triptófano (*trp*) [Goeddel y col. (1980) Nuc. Acids Res. 8:4057; Yelverton y col. (1981) Nucl. Acids Res. 9:731; patente de EE.UU. 4.738.921; documentos EP-A-0036776 y EP-A-0121775]. El sistema promotor de g-laotamasa (*bla*) [Weissmann (1981) "The cloning of interferon and other mistakes" en Interferon 3 (ed. Gresser)], bacteriófago lambda PL [Shimatake y col. (1981) Nature 292:128] y los sistemas promotores de T5 [patente de EE.UU. 4.689.406] también proporcionan secuencias promotoras útiles.

Además, los promotores sintéticos que no se producen en la naturaleza también actúan de promotores bacterianos. Por ejemplo, las secuencias de activación de la transcripción de un promotor bacteriano o bacteriófago pueden unirse con las secuencias del operón de otro promotor bacteriano o bacteriófago, creando un promotor híbrido sintético [patente de EE.UU. 4.551.433]. Por ejemplo, el promotor *tac* es un promotor híbrido *trp-lac* que comprende tanto el promotor *trp* como las secuencias del operón *lac* que está regulado por el represor *lac* [Amann y col. (1983) Gene 25:167; de Boer y col. (1983) PNAS USA 80:21]. Además, un promotor bacteriano puede incluir promotores que se producen naturalmente de origen no bacteriano que tienen la capacidad de unirse a ARN polimerasa bacteriana e iniciar la transcripción. Un promotor que se produce naturalmente de origen no bacteriano también puede acoplarse a una ARN polimerasa compatible para producir altos niveles de expresión de algunos genes en procariontes. El sistema de ARN polimerasa T7/promotor de bacteriófago es un ejemplo de un sistema de promotor acoplado [Studier y col. (1986) J. Mol. Biol. 189:113; Tabor y col. (1985) PNAS USA 82:1074]. Además, un promotor híbrido también puede comprender un promotor de bacteriófago y una región de operador de *E. coli* (documento EP-A-0 267 851).

Además de una secuencia promotora en funcionamiento, un sitio de unión a ribosoma eficiente también es útil para la expresión de genes extraños en procariontes. En *E. coli*, el sitio de unión a ribosoma se llama la secuencia de Shine-Dalgarno (SD) e incluye un codón de iniciación (ATG) y una secuencia de 3-9 nucleótidos de longitud localizada 3-11 nucleótidos en la dirección 5' del codón de iniciación [Shine y col. (1975) Nature 254:34]. Se cree que la secuencia de SD promueve la unión de ARNm al ribosoma por el apareamiento de bases entre la secuencia de SD y 3' y de ARNr de *E. coli* 16S [Steitz y col. (1979) "Genetic signals and nucleotide sequences in messenger RNA". En Biological regulation and Development: Gene Expression (ed. R.F. Goldberger)].

Una molécula de ADN puede expresarse intracelularmente. Una secuencia promotora puede ligarse directamente con la molécula de ADN, en cuyo caso el primer aminoácido en el extremo N nunca será una metionina, que está codificada por el codón de iniciación ATG. Si se desea, la metionina en el extremo N pueden escindirse de la proteína por incubación *in vitro* con bromuro de cianógeno o por tanto incubación *in vivo* como *in vitro* con una peptidasa del extremo N de metionina bacteriana (documento EPO-A-0 219 237).

Las proteínas de fusión proporcionan una alternativa a la expresión directa. Normalmente, una secuencia de ADN que codifica la porción del extremo N de una proteína bacteriana endógena, u otra proteína estable, se fusiona con el extremo 5' de secuencias codificantes heterólogas. Tras la expresión, esta construcción proporcionará una fusión de las dos secuencias de aminoácidos. Por ejemplo, el gen de la célula de bacteriófago lambda puede ligarse en el extremo 5' de un gen extraño y expresarse en bacterias. La proteína de fusión resultante retiene preferentemente un sitio para una enzima de procesamiento (factor Xa) para escindir la proteína de bacteriófago del gen extraño [Nagai y col. (1984) Nature 309:810]. Las proteínas de fusión también pueden prepararse con secuencias de los genes *lacZ* [Jia y col. (1987) Gene 60:197], *trpE* [Allen y col. (1987) 1. Biotechnol. 5:93; Makoff y col. (1989) J. Gen. Microbiol. 135:11] y Chey [documento EP-A-0 324 647]. La secuencia de ADN en la unión de las dos secuencias de aminoácidos puede o puede no codificar un sitio escindible. Otro ejemplo es una proteína de fusión de ubiquitina. Una proteína de fusión tal se prepara con la región de ubiquitina que preferentemente retiene un sitio para una enzima de procesamiento (por ejemplo, proteasa de procesamiento específico de ubiquitina) para escindir la ubiquitina de la proteína extraña. En todo este procedimiento, la proteína extraña nativa puede aislarse [Miller y col. (1989) Bio/Technology 7:698]. Alternativamente, las proteínas extrañas también pueden secretarse de la célula creando moléculas de ADN quiméricas que codifican una proteína de fusión que comprende un fragmento de la secuencia de péptidos señal que proporciona la secreción de la proteína extraña en bacterias [patente de EE.UU. 4.336.336]. El fragmento de la secuencia señal normalmente codifica un péptido señal que comprende aminoácidos hidrófobos que dirigen la secreción de la proteína de la célula. La proteína es tanto secretada en los medios de crecimiento (bacterias Gram-positivas) como en el espacio periplásmico localizado entre la membrana interna y

externa de la célula (bacterias Gram-negativas). Preferentemente, hay sitios de procesamiento que pueden escindirse tanto *in vivo* como *in vitro* codificados entre el fragmento del péptido señal y el gen extraño.

El ADN que codifica secuencias señal adecuadas puede derivarse de genes para proteínas bacterianas secretadas tales como el gen de la proteína de la membrana externa de *E. coli* (*ompA*) [Masui y col. (1983), en: Experimental Manipulation of Gene Expression; Ghayeb y col. (1984) EMBO J. 3:2437] y la secuencia señal de la fosfatasa alcalina de *E. coli* (*phoA*) [Oka y col. (1985) PNAS USA 82:72121]. Como ejemplo adicional, la secuencia señal del gen de alfa-amilasa de diversas cepas de *Bacillus* puede usarse para secretar proteínas heterólogas de *B. subtilis* [Palva y col. (1982) PNAS USA 79:5582; documento EP-A-0 244 042].

Normalmente, las secuencias de terminación de la transcripción reconocidas por bacterias son regiones reguladoras localizadas 3' con respecto al codón de terminación de la traducción y, por tanto, junto con el promotor flanquean la secuencia codificante. Estas secuencias dirigen la transcripción de un ARNm que puede traducirse en el polipéptido codificado por el ADN. Las secuencias de terminación de la transcripción frecuentemente incluyen secuencias de ADN de aproximadamente 50 nucleótidos que pueden formar estructuras de tallo-bucle que ayudan en la terminación de la transcripción. Ejemplos incluyen secuencias de terminación de la transcripción derivadas de genes con promotores fuertes tales como el gen *trp* en *E. coli*, además de otros genes biosintéticos.

Normalmente, los componentes anteriormente descritos, que comprenden un promotor, secuencia señal (si se desea), secuencia codificante de interés y secuencia de terminación de la transcripción, se ponen juntos en construcciones de expresión. Las construcciones de expresión se mantienen frecuentemente en un replicón, tal como un elemento extracromosómico (por ejemplo, plásmidos), que puede mantenerse establemente en un huésped, tal como bacterias. El replicón tendrá un sistema de replicación, permitiendo así que se mantenga en un huésped procarionta tanto para la expresión como para la clonación y amplificación. Además, un replicón puede ser un plásmido tanto de alto como de bajo número de copias. Un plásmido de alto número de copias tendrá generalmente un número de copias que oscila de aproximadamente 5 a aproximadamente 200, y normalmente de aproximadamente 10 a aproximadamente 150. Un huésped que contiene un plásmido de alto número de copias contendrá preferentemente al menos aproximadamente 10, y más preferentemente al menos aproximadamente 20 plásmidos. Puede seleccionarse un vector con tanto alto como bajo número de copias, dependiendo del efecto del vector y la proteína extraña sobre el huésped.

Alternativamente, las construcciones de expresión pueden integrarse en el genoma bacteriano con un vector integrante. Los vectores integrantes normalmente contienen al menos una secuencia homóloga al cromosoma bacteriano que permite que el vector se integre. Las integraciones parecen resultar de recombinaciones entre ADN homólogo en el vector y el cromosoma bacteriano. Por ejemplo, los vectores integrantes construidos con ADN de diversas cepas de *Bacillus* se integran en el cromosoma de *Bacillus* (documento EP-A-0 127 328). Los vectores integrantes también pueden comprender secuencias de bacteriófago o de transposón.

Normalmente, las construcciones de expresión extracromosómicas e integrantes pueden contener marcadores de selección para permitir la selección de cepas bacterianas que se han transformado. Los marcadores de selección pueden expresarse en el huésped bacteriano y pueden incluir genes que hacen que las bacterias sean resistentes a fármacos tales como ampicilina, cloranfenicol, eritromicina, kanamicina (neomicina) y tetraciclina [Davies y col. (1978) Annu. Rev. Microbiol. 32:469]. Los marcadores de selección también pueden incluir genes biosintéticos tales como aquellos en las vías biosintéticas de histidina, triptófano y leucina.

Alternativamente, algunos de los componentes anteriormente descritos pueden ponerse juntos en vectores de transformación. Los vectores de transformación comprenden normalmente un marcador de selección que tanto se mantiene en un replicón como se desarrolla en un vector integrante, como se ha descrito anteriormente.

Los vectores de expresión y de transformación, tanto replicones extra-cromosómicos como vectores integrantes, se han desarrollado para la transformación en muchas bacterias. Por ejemplo, se han desarrollado vectores de expresión para, entre otras, las siguientes bacterias: *Bacillus subtilis* [Palva y col. (1982) PNAS USA 79:5582; documentos EP-A-0 036 259 y EP-A-0 063 953; documento WO 84/04541], *Escherichia coli* [Shimatake y col. (1981) Nature 292:128; Amann y col. (1985) Gene 40:183; Studier y col. (1986) J. Mol. Biol. 189:113; documentos EP-A-0 036 776, EP-A-0 136 829 y EP-A-0 136 907], *Streptococcus cremoris* [Powell y col. (1988) Appl. Environ. Microbiol. 54:655]; *Streptococcus lividans* [Powell y col. (1988) Appl. Environ. Microbiol. 54:655], *Streptomyces lividans* [patente de EE.UU. 4.745.056].

Los procedimientos de introducción de ADN exógeno en huéspedes bacterianos son muy conocidos en la técnica, y normalmente incluyen tanto la transformación de bacterias tratadas con CaCl₂ como otros agentes, tales como cationes divalentes y DMSO. El ADN también puede introducirse en células bacterianas por electroporación. Los procedimientos de transformación normalmente varían con las especies bacterianas que van a transformarse. Véase, por ejemplo, [Masson y col. (1989) FEMS Microbiol. Lett. 60:273; Palva y col. (1982) PNAS USA 79:5582; documentos EP-A-0 036 259 y EP-A-0 063 953; documento WO 84/04541, *Bacillus*], [Miller y col. (1988) PNAS USA. 85:856; Wang y col. (1990) J. Bacteriol. 172:949, *Campylobacter*], [Cohen y col. (1973) PNAS USA 69:2110; Dower y col. (1988) Nucleic Acids Res. 16:6127; Kushner (1978) "An improved method for transformation of *Escherichia coli* with ColEI-derived plasmids. En Genetic Engineering: Proceedings of the International Symposium on Genetic

Engineering (eds. Boyer & Nicosia); Mandel y col. (1970) J. Mol. Biol. 53:159; Taketo (1988) Biochim. Biophys. Acta 949:318; *Escherichia*, [Chassy y col. (1987) FEMS Microbiol. Lett. 44:173 *Lactobacillus*]; [Fiedler y col. (1988) Anal. Biochem 170:38, *Pseudomonas*]; [Augustin y col. (1990) FEMS Microbiol. Lett. 66:203, *Staphylococcus*], [Barany y col. (1980) J. Bacteriol. 144:698; Harlander (1987) "Transformation of *Streptococcus lactis* by electroporation, en: Streptococcal Genetics (ed. J. Ferretti y R. Curtiss III); Perry y col. (1981) Infect. Immun. 32:1295; Powell y col. (1988) Appl. Environ. Microbiol. 54:655; Somkuti y col. (1987) Proc. 4th Evr. Congo Biotechnology 1:412, *Streptococcus*].

v. Expresión en levadura

Los sistemas de expresión en levadura también son conocidos para un experto en la materia. Un promotor de levadura es cualquier secuencia de ADN que pueda unirse a ARN polimerasa de levadura e iniciar la transcripción en la dirección 3' de una secuencia codificante (por ejemplo, gen estructural) en ARN. Un promotor tendrá una región de iniciación de la transcripción que está normalmente situada proximal al extremo 5' de la secuencia codificante. Esta región de iniciación de la transcripción normalmente incluye un sitio de unión a ARN polimerasa (la "caja TATA") y un sitio de iniciación de la transcripción. Un promotor de levadura también puede tener un segundo dominio llamado una secuencia activadora en la dirección 5' (UAS) que, si está presente, es normalmente distal al gen estructural. La UAS permite la expresión regulada (inducible). La expresión constitutiva se produce en ausencia de una UAS. La expresión regulada puede ser tanto positiva como negativa, tanto potenciando como reduciendo así la transcripción.

La levadura es un organismo fermentante con una vía metabólica activa, por lo que secuencias que codifican enzimas en la vía metabólica proporcionan secuencias promotoras particularmente útiles. Ejemplos incluyen alcohol deshidrogenasa (documento EP-A-0 284 044), enolasa, glucocinasa, glucosa-6-fosfato isomerasa, gliceraldehído-3-fosfato-deshidrogenasa (GAP o GAPDH), hexocinasa, fosfofructocinasa, 3-fosfoglicerato mutasa y piruvato cinasa (PyK) (documento EPO-A-0 329 203). El gen *PH05* de la levadura, que codifica fosfatasa ácida, también proporciona secuencias promotoras útiles [Myanohara y col. (1983) PNAS USA 80:1].

Además, los promotores sintéticos que no se producen en la naturaleza también funcionan de promotores de levadura. Por ejemplo, las secuencias UAS de un promotor de levadura pueden unirse con la región de activación de la transcripción de otro promotor de levadura, creando un promotor híbrido sintético. Ejemplos de tales promotores híbridos incluyen la secuencia reguladora de ADH ligada a la región de activación de la transcripción de GAP (patentes de EE.UU. nº 4.876.197 y 4.880.734). Otros ejemplos de promotores híbridos incluyen promotores que consisten en las secuencias reguladoras de tanto los genes *ADH2*, *GAL4*, *GAL10* como *PH05*, combinados con la región de activación de la transcripción de un gen de enzima glicolítica tal como GAP o PyK (documento EP-A-0 164 556). Además, un promotor de levadura puede incluir promotores que se producen naturalmente de origen de no levadura que tienen la capacidad de unirse a ARN polimerasa de levadura e iniciar la transcripción. Ejemplos de tales promotores incluyen, entre otros [Cohen y col. (1980) PNAS USA 77:1078; Henikoff y col. (1981) Nature 283:835; Hollenberg y col. (1981) Curr. Topics Microbiol. Immunol. 96:119; Hollenberg y col. (1979) "The expression of Bacterial Antibiotic Resistance Genes in the Yeast *Saccharomyces cerevisiae*," en: Plasmids of Medical, Environmental and Commercial Importance (eds. K.N. Timmis y A. Puhler); Mercerau-Puigalon y col. (1980) Gene 11:163; Panthier y col. (1980) Curr. Genet. 2:109;].

Una molécula de ADN puede expresarse intracelularmente en levadura. Una secuencia promotora puede unirse directamente a la molécula de ADN, en cuyo caso el primer aminoácido en el extremo N de la proteína recombinante siempre será una metionina, que está codificada por el codón de iniciación ATG. Si se desea, la metionina en el extremo N pueden escindirse de la proteína por incubación *in vitro* con bromuro de cianógeno.

Las proteínas de fusión proporcionan una alternativa a los sistemas de expresión de levadura, además de en sistemas de expresión de mamífero, baculovirus y bacterianos. Normalmente, una secuencia de ADN que codifica la porción del extremo N de una proteína de levadura endógena, u otra proteína estable, está fusionada con el extremo 5' de secuencias codificantes heterólogas. Tras la expresión, esta construcción proporcionará una fusión de las dos secuencias de aminoácidos. Por ejemplo, el gen de levadura o de superóxido dismutasa (SOD) humana puede unirse en el extremo 5' de un gen extraño y expresarse en levadura. La secuencia de ADN en la unión de las dos secuencias de aminoácidos puede o puede no codificar un sitio escindible. Véase, por ejemplo, el documento EP-A-0 196 056. Otro ejemplo es una proteína de fusión de ubiquitina. Una proteína de fusión tal se prepara con la región de ubiquitina que preferentemente retiene un sitio para una enzima de procesamiento (por ejemplo, proteasa de procesamiento específica de ubiquitina) para escindir la ubiquitina de la proteína extraña. Por tanto, en este procedimiento puede aislarse proteína extraña nativa (por ejemplo, el documento WO 88/024066).

Alternativamente, también pueden secretarse proteínas extrañas de la célula en los medios de crecimiento creando moléculas de ADN quiméricas que codifican una proteína de fusión que comprende un fragmento de secuencia conductora que proporciona la secreción en levadura de la proteína extraña. Preferentemente, hay sitios de procesamiento codificados entre el fragmento conductor y el gen extraño que pueden escindirse tanto *in vivo* como *in vitro*. El fragmento de la secuencia conductora normalmente codifica un péptido señal que comprende aminoácidos hidrófobos que dirigen la secreción de la proteína de la célula. El ADN que codifica secuencias señal adecuadas puede derivarse de genes para proteínas de levadura secretada tales como los genes de levadura

invertasa (documentos EP-A-0012873; JPO 62.096.086) y de factor A (patente de EE.UU. 4.588.684). Alternativamente, también existen conductores de origen de no levadura, tales como un conductor de interferón, que proporcionan secreción en levadura (documento EP-A-0060057).

5 Una clase preferida de conductores de la secreción son aquellos que emplean un fragmento del gen del factor alfa de levadura que contiene tanto una secuencia señal "pre" como una región "pro". Los tipos de fragmentos de factor alfa que pueden emplearse incluyen el conductor del factor alfa de pre-pro de longitud completa (aproximadamente 83 residuos de aminoácidos), además de conductores del factor alfa truncados (normalmente aproximadamente 25 a aproximadamente 50 residuos de aminoácidos) (patentes de EE.UU. 4.546.083 y 4.870.008; documento EP-A-0 324 274). Conductores adicionales que emplean un fragmento del factor alfa que proporciona la secreción incluyen conductores del factor alfa híbridos preparados con una presecuencia de una primera levadura, pero una pro-región de un segundo factor alfa de levadura (por ejemplo, véase el documento WO 89/02463).

10 Normalmente, las secuencias de terminación de la transcripción reconocidas por levadura son regiones reguladoras localizadas en 3' con respecto al codón de terminación de la traducción y, por tanto, junto con el promotor flanquean la secuencia codificante. Estas secuencias dirigen la transcripción de un ARNm que puede traducirse en el polipéptido codificado por el ADN. Ejemplos de secuencia terminadora de la transcripción y otras secuencias de terminación reconocidas por levadura, tales como aquellas que codifican enzimas glicolíticas.

15 Normalmente, los componentes anteriormente descritos, que comprenden un promotor, conductor (si se desea), secuencia codificante de interés y secuencia de terminación de la transcripción, se ponen juntos en construcciones de expresión. Las construcciones de expresión se mantienen frecuentemente en un replicón, tal como un elemento extracromosómico (por ejemplo, plásmidos), que puede mantenerse establemente en un huésped, tal como levadura o bacterias. El replicón puede tener dos sistemas de replicación, permitiendo así que se mantenga, por ejemplo, en levadura para la expresión y en un huésped procariota para la clonación y amplificación. Ejemplos de tales vectores lanzadera de bacterias para levadura incluyen YEp24 [Botstein y col. (1979) Gene 8:17-241, pCI/1 [Brake y col. (1984) PNAS USA 81:4642-46461] y YRp17 [Stinchcomb y col. (1982) J. Mol. Biol. 158:157]. Además, un replicón puede ser un plásmido tanto de alto como de bajo número de copias. Un plásmido de alto número de copias tendrá generalmente un número de copias que oscila de aproximadamente 5 a aproximadamente 200, y normalmente -10 a -150. Un huésped que contiene un plásmido de alto número de copias tendrá preferentemente al menos aproximadamente 10, y más preferentemente al menos ~20. Puede seleccionarse un vector tanto con alto como con bajo número de copias, dependiendo del efecto del vector y la proteína extraña sobre el huésped. Véase, por ejemplo, Brake y col., arriba.

20 Alternativamente, las construcciones de expresión pueden integrarse en el genoma de levadura con un vector integrante. Los vectores integrantes normalmente contienen al menos una secuencia homóloga al cromosoma de levadura que permite que el vector se integre, y preferentemente contienen dos secuencias homólogas que flanquean la construcción de expresión. Las integraciones parecen resultar de recombinaciones entre ADN homólogo en el vector y el cromosoma de levadura [Orr-Weaver y col. (1983) Methods in Enzymol. 101:228-245]. Un vector integrante puede dirigirse a un locus específico en levadura seleccionando la secuencia homóloga apropiada para la inclusión en el vector. Véase Orr-Weaver y col., arriba. Una o más construcciones de expresión pueden integrarse, afectando posiblemente los niveles de proteína recombinante producida [Rine y col. (1983) PNAS USA 80:6750]. Las secuencias cromosómicas incluidas en el vector pueden producirse tanto como un único segmento en el vector, que produce la integración del vector entero, como dos segmentos homólogos a segmentos adyacentes en el cromosoma y que flanquean la construcción de expresión en el vector; que pueden producir la integración estable de sólo la construcción de expresión.

25 Normalmente, las construcciones de expresión extracromosómicas e integrantes pueden contener marcadores de selección para permitir la selección de cepas de levadura que se han transformado. Los marcadores de selección pueden incluir genes biosintéticos que pueden expresarse en la levadura huésped tales como *ADE2*, *HIS4*, *LEU2*, *TRP* y *ALG7*, y el gen de resistencia a G418, que confiere resistencia en células de levadura a tunicamicina y G418, respectivamente. Además, un marcador de selección adecuado también puede proporcionar levadura con la capacidad de cultivarse en presencia de compuestos tóxicos tales como metal. Por ejemplo, la presencia de *CUP1* permite que la levadura se cultive en presencia de iones cobre [Butt y col. (1987) Microbiol. Rev. 51:351].

30 Alternativamente, algunos de los componentes anteriormente descritos pueden ponerse juntos en vectores de transformación. Los vectores de transformación comprenden normalmente un marcador de selección que tanto se mantiene en un replicón como se desarrolla en un vector integrante, como se ha descrito anteriormente.

35 Los vectores de expresión y de transformación, tanto replicones extracromosómicos como vectores integrantes, se han desarrollado para la transformación en muchas levaduras. Por ejemplo, se han desarrollado vectores de expresión para, entre otras, las siguientes levaduras: *Candida albicans* [Kurtz y col. (1986) Mol. Cell. Biol. 6:142], *Candida maltosa* [Kunze y col. (1985) J. Basic Microbiol. 25:141], *Hansenula polymorpha* [Gleeson y col. (1986) J. Gen. Microbiol. 132:3459; Roggenkamp y col. (1986) Mol. Gen. Genet. 202:302], *Kluyveromyces fragilis* [Das y col. (1984) J. Bacteriol. 158:1165], *Kluyveromyces lactis* [De Louvencourt y col. (1983) J. Bacteriol. 154:737; Van den Berg y col. (1990) Biotechnology 8:135], *Pichia guilliermondii* [Kunze y col. (1985) J. Basic Microbiol. 25:141], *Pichia pastoris* [Cregg y col. (1985) Mol. Cell. Biol. 5: 3376; patente de EE.UU. nº 4.837.148 y 4.929.555], *Saccharomyces*

cerevisiae [Hinnen y col. (1978) PNAS USA 75:1929; Ito y col. (1983) J. Bacteriol. 153:163], *Schizosaccharomyces pombe* [Beach y Nurse (1981) Nature 300:706] y *Yarrowia lipolytica* [Davidow y col. (1985) Curr. Genet. 10:380471 Gaillardin y col. (1985) Curr. Genet. 10:49].

5 Los procedimientos de introducción de ADN exógeno en huéspedes de levadura son muy conocidos en la técnica, y normalmente incluyen tanto la transformación de esferoplastos como de células de levadura intactas tratadas con cationes alcalinos. Los procedimientos de transformación normalmente varían con las especies de levadura que van a transformarse. Véase, por ejemplo, [Kurtz y col. (1986) Mol. Cell. Biol. 6:142; Kunze y col. (1985) J. Basic Microbiol. 25:141; *Candida*]; [Gleeson y col. (1986) J. Gen. Microbiol. 132:3459; Roggenkamp y col. (1986) Mol. Gen. Genet. 202:302; *Hansenula*]; [Das y col. (1984) J. Bacteriol. 158:1165; De Louvencourt y col. (1983) J. Bacteriol. 154:1165; Van den Berg y col. (1990) Bio/Technology 8:135; *Kluyveromyces*]; [Cregg y col. (1985) Mol. Cell. Biol. 5:3376; Kunze y col. (1985) J. Basic Microbiol. 25:141; patentes de EE.UU. nº 4.837.148 y 4.929.555; *Pichia*]; [Hinnen y col. (1978) PNAS USA 75:1929; Ito y col. (1983) J. Bacteriol. 153:163 *Saccharomyces*]; [Beach y Nurse (1981) Nature 300:706; *Schizosaccharomyces*]; [Davidow y col. (1985) Curr. Genet. 10:39; Gaillardin y col. (1985) Curr. Genet. 10:49; *Yarrowia*].

15 Anticuerpos

Como se usa en este documento, el término “anticuerpo” se refiere a un polipéptido o grupo de polipéptidos compuesto por al menos un sitio de combinación de anticuerpos. Un “sitio de combinación de anticuerpos” es el espacio de unión tridimensional con una forma de la superficie interna y distribución de carga complementaria a las características de un epítipo de un antígeno, que permite una unión del anticuerpo con el antígeno. “Anticuerpo” incluye, por ejemplo, anticuerpos de vertebrados, anticuerpos híbridos, anticuerpos quiméricos, anticuerpos humanizados, anticuerpos alterados, anticuerpos univalentes, proteínas de Fab y anticuerpos de un único dominio.

Los anticuerpos contra las proteínas de la invención son útiles para cromatografía de afinidad, inmunoensayos y distinguir/identificar proteínas de la invención.

25 Los anticuerpos para las proteínas de la invención, tanto policlonales como monoclonales, pueden prepararse mediante procedimientos convencionales. En general, la proteína se usa primero para inmunizar un animal adecuado, preferentemente un ratón, rata, conejo o cabra. Se prefieren conejos y cabras para la preparación de sueros policlonales debido al volumen de suero obtenible y la disponibilidad de anticuerpos anti-conejo y anti-cabra marcados. La inmunización se realiza generalmente mezclando o emulsionando la proteína en solución salina, preferentemente en un adyuvante tal como adyuvante completo de Freund, e inyectando la mezcla o emulsión parenteralmente (generalmente subcutáneamente o intramuscularmente). Una dosis de 50-200 µg/inyección es normalmente suficiente. La inmunización se refuerza generalmente 2-6 semanas después con una o más inyecciones de la proteína en solución salina, preferentemente usando adyuvante incompleto de Freund. Alternativamente pueden generarse anticuerpos por inmunización *in vitro* usando procedimientos conocidos en la técnica, que para los fines de la presente invención se consideran equivalentes a la inmunización *in vivo*. Los antisueros policlonales se obtienen sangrando el animal inmunizado en un recipiente de vidrio o plástico, incubando la sangre a 25°C durante una hora, seguido de incubación a 4°C durante 2-18 horas. El suero se recupera por centrifugación (por ejemplo, 1.000 g durante 10 minutos). Pueden obtenerse aproximadamente 20-50 ml por sangrado de conejos.

40 Los anticuerpos monoclonales se preparan usando el procedimiento convencional de Kohler & Milstein [Nature (1975) 256:495-96], o una modificación del mismo. Normalmente, un ratón o rata se inmuniza como se ha descrito anteriormente. Sin embargo, en vez de sangrar al animal para extraer el suero, el bazo (y opcionalmente varios ganglios linfáticos grandes) se extrae y se disocia en células individuales. Si se desea, las células del bazo pueden cribarse (después de eliminar las células no específicamente adherentes) aplicando una suspensión de células a una placa o pocillo recubierto con el antígeno de proteína. Los linfocitos B que expresan inmunoglobulina unida a membrana específica para el antígeno se unen a la placa, y no se aclaran con el resto de la suspensión. Los linfocitos B resultantes, o todas las células del bazo disociadas, son luego inducidas a fusionarse con células de mieloma para formar hibridomas, y se cultivan en un medio selectivo (por ejemplo, medio hipoxantina, aminopterina, timidina, “HAT”). Los hibridomas resultantes se siembran en placa por dilución limitante y se ensayan para la producción de anticuerpos que se unen específicamente al antígeno inmunizante (y que no se unen a antígenos sin relacionar). Los hibridomas que secretan MAb seleccionado se cultivan luego tanto *in vitro* (por ejemplo, en botellas de cultivo de tejido o reactores de fibra hueca) como *in vivo* (como ascitis en ratones).

Si se desea, los anticuerpos (tanto policlonales como monoclonales) pueden marcarse usando técnicas convencionales. Marcas adecuadas incluyen fluoróforos, cromóforos, átomos radiactivos (particularmente ³²P y ¹²⁵I), reactivos densos en electrones, enzimas y ligandos que tienen componentes de unión específica. Las enzimas se detectan normalmente por su actividad. Por ejemplo, la peroxidasa de rábano picante se detecta normalmente por su capacidad para convertir 3,3',5,5'-tetrametilbencidina (TMB) en un pigmento azul cuantificable con un espectrofotómetro. “Componente de unión específica” se refiere a una proteína que puede unirse a una molécula de ligando con alta especificidad como, por ejemplo, en el caso de un antígeno y un anticuerpo monoclonal específico para el mismo. Otros componentes de unión específica incluyen biotina y avidina o estreptavidina, IgG y proteína A, y las numerosas parejas de receptor-ligando conocidas en la técnica. Debe entenderse que la descripción anterior

no pretende clasificar las diversas marcas en distintas clases, ya que la misma marca puede servir en varios modos diferentes. Por ejemplo, ¹²⁵I puede servir de marca radiactiva o como reactivo denso en electrones. La HRP puede servir de enzima o de antígeno para un MAb. Además, pueden combinarse diversas marcas para el efecto deseado. Por ejemplo, los MAb y la avidina también requieren marcas en la práctica de la presente invención: por tanto, uno podría marcar un MAb con biotina, y detectar su presencia con avidina marcada con ¹²⁵I, o con un Mab anti-biotina marcado con HRP. Otras permutaciones y posibilidades serán rápidamente evidentes para aquellos expertos en la materia.

Composiciones farmacéuticas

Las composiciones farmacéuticas pueden comprender tanto polipéptidos, anticuerpos como ácido nucleico de la invención. Las composiciones farmacéuticas comprenderán una cantidad terapéuticamente eficaz de tanto polipéptidos, anticuerpos como polinucleótidos de la invención reivindicada.

El término "cantidad terapéuticamente eficaz" como se usa en este documento se refiere a una cantidad de un agente terapéutico para tratar, mejorar o prevenir una enfermedad o afección deseada, o para presentar un efecto terapéutico o preventivo detectable. El efecto puede detectarse, por ejemplo, por marcadores químicos o niveles de antígeno. Los efectos terapéuticos también incluyen la reducción en los síntomas físicos tales como la disminución de la temperatura corporal. La cantidad eficaz precisa para un sujeto dependerá de su tamaño y salud, la naturaleza y grado de la afección, y los agentes terapéuticos o combinación de agentes terapéuticos seleccionados para administración. Por tanto, no es útil especificar una cantidad eficaz exacta de antemano. Sin embargo, la cantidad eficaz para una situación dada puede determinarse por experimentación rutinaria y está dentro del criterio del profesional clínico.

Para los fines de la presente invención, una dosis eficaz será de aproximadamente 0,01 mg/kg a 50 mg/kg o 0,05 mg/kg a aproximadamente 10 mg/kg de las construcciones de ADN en el individuo al que se administra.

Una composición farmacéutica también puede contener un vehículo farmacéuticamente aceptable. El término "vehículo farmacéuticamente aceptable" se refiere a un vehículo para la administración de un agente terapéutico, tal como anticuerpos o un polipéptido, genes y otros agentes terapéuticos. El término se refiere a cualquier vehículo farmacéutico que por sí mismo no induce la producción de anticuerpos perjudiciales para el individuo que recibe la composición, y que puede administrarse sin excesiva toxicidad. Vehículos adecuados pueden ser macromoléculas grandes lentamente metabolizadas tales como proteínas, polisacáridos, ácidos polilácticos, ácido poliglicólicos, aminoácidos poliméricos, copolímeros de aminoácidos y partículas de virus inactivos. Tales vehículos son muy conocidos para aquellos expertos en la materia.

En esto pueden usarse sales farmacéuticamente aceptables, por ejemplo, sales de ácidos minerales tales como clorhidratos, bromhidratos, fosfatos, sulfatos y similares; y las sales de ácidos orgánicos tales como acetatos, propionatos, malonatos, benzoatos y similares. Una discusión meticulosa de excipientes farmacéuticamente aceptables está disponible en Remington's Pharmaceutical Sciences (Mack Pub. Co., N.J. 1991).

Los vehículos farmacéuticamente aceptables en composiciones terapéuticas pueden contener líquidos tales como agua, solución salina, glicerol y etanol. Adicionalmente, en tales vehículos pueden estar presentes sustancias auxiliares, tales como agentes humectantes o emulsionantes, sustancias de tamponamiento del pH y similares. Normalmente, las composiciones terapéuticas se preparan como inyectables, tanto como disoluciones líquidas como suspensiones; también puede prepararse formas sólidas adecuadas para disolución en, o suspensión en, vehículos líquidos antes de la inyección. Los liposomas están incluidos dentro de la definición de un vehículo farmacéuticamente aceptable.

Procedimientos de administración

Una vez formuladas, las composiciones de la invención pueden administrarse directamente al sujeto. Los sujetos que van a tratarse pueden ser animales; en particular, pueden tratarse sujetos humanos.

La administración directa de las composiciones se llevará generalmente a cabo mediante inyección, tanto subcutáneamente, intraperitonealmente, intravenosamente como intramuscularmente, o se administrarán al espacio intersticial de un tejido. Las composiciones también pueden administrarse a una lesión. Otros modos de administración incluyen administración oral y pulmonar, supositorios, y aplicaciones transdérmicas o transcutáneas (por ejemplo, véase el documento WO98/20734), agujas y pistolas de genes o hiposprays. El tratamiento de dosificación puede ser un programa de dosis única o un programa de múltiples dosis.

Vacunas

Las vacunas según la invención pueden ser tanto profilácticas (es decir, para prevenir la infección) como terapéuticas (es decir, para tratar la enfermedad después de la infección).

Tales vacunas comprenden inmunizar antígeno(s), inmunógeno(s), polipéptido(s), proteína(s) o ácido nucleico, normalmente en combinación con "vehículos farmacéuticamente aceptables," que incluyen cualquier vehículo que no

induzca por sí mismo la producción de anticuerpos perjudiciales para el individuo que recibe la composición. Vehículos adecuados son macromoléculas grandes lentamente metabolizadas tales como proteínas, polisacáridos, ácidos polilácticos, ácidos poliglicólicos, aminoácidos poliméricos, copolímeros de aminoácidos, agregados lipídicos (tales como gotitas de aceite o liposomas) y partículas de virus inactivos. Tales vehículos son muy conocidos para aquellos expertos en la materia. Adicionalmente, estos vehículos pueden actuar de agentes inmunoestimulantes (“adyuvantes”). Además, el antígeno o inmunógeno puede conjugarse con un toxoide bacteriano, tal como un toxoide de difteria, tétanos, cólera, patógenos de *H. pylori*, etc.

Adyuvantes preferidos para potenciar la eficacia de la composición incluyen, pero no se limitan a: (1) sales de aluminio (alumbre) tales como hidróxido de aluminio, fosfato de aluminio, sulfato de aluminio, etc.; (2) formulaciones de emulsión de aceite en agua (con o sin otros agentes inmunoestimulantes específicos tales como muramilpéptidos (véase más adelante) o componentes de la pared celular bacteriana) tales como, por ejemplo, (a) MF59™ (documento WO 90/14837; Capítulo 10 en *Vaccine design: the subunit and adjuvant approach*, eds. Powell & Newman, Plenum Press 1995), que contiene 5% de escualeno, 0,5% de Tween 80 y 0,5% de Span 85 (que opcionalmente contiene diversas cantidades de MTP-PE (véase más adelante), aunque no se requieren) formulado en partículas submicrométricas usando un microfluidizador tal como el microfluidizador modelo 110Y (Microfluidics, Newton, MA), (b) SAF, que contiene 10% de escualeno, 0,4% de Tween 80, 5% de polímero L121 bloqueado con plurónico y thr-MDP (véase más adelante) tanto microfluidizado en una emulsión submicrométrica como agitado con vórtex para generar una emulsión de mayor tamaño de partícula, y (c) sistema de adyuvante Ribit™ (RAS), (Ribi Immunochem, Hamilton, MT) que contiene 2% de escualeno, 0,2% de Tween 80 y uno o más componentes de la pared celular bacteriana del grupo que consiste en monofosforil lípido A (MPL), dimicolato de trehalosa (TDM) y esqueleto de la pared celular (CWS), preferentemente MPL + CWS (Detox™); (3) pueden usarse adyuvantes de saponina tales como Stimulon™ (Cambridge Bioscience, Worcester, MA) o partículas generadas a partir de los mismos tales como ISCOM (complejos inmunoestimulantes); (4) adyuvante completo de Freund (CFA) y adyuvante incompleto de Freund (IFA); (5) citocinas tales como interleucinas (por ejemplo, IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12, etc.), interferones (por ejemplo, interferón γ), factor estimulante de colonias de macrófagos (M-CSF), factor de necrosis tumoral (TNF), etc.; y (6) otras sustancias que actúan de agentes inmunoestimulantes para potenciar la eficacia de la composición. Se prefieren alumbre y MF59™.

Como se menciona anteriormente, los muramilpéptidos incluyen, pero no se limitan a, N-acetil-muramil-L-treonil-D-isoglutamina (thr-MDP), N-acetil-normuramil-L-alanil-D-isoglutamina (nor-MDP), N-acetilmuramil-L-alanil-D-isoglutaminil-L-alanina-2-(1'-2'-dipalmitoil-sn-glicero-3-hidroxifosforilo)-etilamina (MTP-PE), etc.

Las composiciones inmunogénicas (por ejemplo, el antígeno inmunizante/inmunógeno/polipéptido/proteína/ácido nucleico, vehículo farmacéuticamente aceptable y adyuvante) normalmente contendrán diluyentes tales como agua, solución salina, glicerol, etanol, etc. Adicionalmente, sustancias auxiliares tales como agentes humectantes o emulsionantes, sustancias de tamponamiento del pH y similares pueden estar presentes en tales vehículos.

Normalmente, las composiciones inmunogénicas se preparan como inyectables, tanto como disoluciones líquidas como suspensiones; también pueden prepararse formas sólidas adecuadas para disolución en, o suspensión en, vehículos líquidos antes de la inyección. La preparación también puede emulsionarse o encapsularse en liposomas para potenciar el efecto adyuvante, como se trata anteriormente bajo vehículos farmacéuticamente aceptables.

Las composiciones inmunogénicas usadas como vacunas comprenden una cantidad inmunológicamente eficaz de los polipéptidos antigénicos o inmunogénicos, además de cualquier otro de los componentes anteriormente mencionados, según se necesite. Por “cantidad inmunológicamente eficaz” se indica que la administración de esa cantidad a un individuo, tanto en una dosis única como parte de una serie, es eficaz para el tratamiento o la prevención. Esta cantidad varía dependiendo de la salud y condición física del individuo que va a tratarse, su grupo taxonómico (por ejemplo, primate no humano, primate, etc.), la capacidad de su sistema inmunitario para sintetizar anticuerpos, el grado de protección deseado, la formulación de la vacuna, la evaluación del médico práctico en la situación médica y otros factores relevantes. Se espera que la cantidad se encuentre en un intervalo relativamente amplio que puede determinarse por ensayos rutinarios.

Las composiciones inmunogénicas se administran convencionalmente parenteralmente, por ejemplo, mediante inyección, tanto subcutáneamente, intramuscularmente como transdérmicamente/transcutáneamente (por ejemplo, el documento WO98/20734). Formulaciones adicionales adecuadas para otros modos de administración incluyen formulaciones orales y pulmonares, supositorios y aplicaciones transdérmicas. El tratamiento de dosificación puede ser un programa de dosis única o un programa de múltiples dosis. La vacuna puede administrarse conjuntamente con otros agentes inmunorreguladores.

Como una alternativa a las vacunas basadas en proteínas puede usarse vacunación con ADN [por ejemplo, Robinson & Torres (1997) *Seminars in Immunol* 9:211-283; Donnelly y col. (1997) *Annu Rev Immunol* 15:617-648].

Vehículos de administración de genes

Los vehículos de terapia génica para la administración de construcciones que incluyen una secuencia codificante de un agente terapéutico de la invención que va a administrarse al mamífero para la expresión en el mamífero pueden

administrarse tanto localmente como sistémicamente. Estas construcciones pueden utilizar enfoques de vector vírico o no vírico en modalidad *in vivo* o *ex vivo*. La expresión de la secuencia codificante puede inducirse usando promotores endógenos o heterólogos de mamífero. La expresión de la secuencia codificante *in vivo* puede ser tanto constitutiva como regulada.

- 5 La invención incluye vehículos de administración de genes que pueden expresar las secuencias de ácidos nucleicos contempladas. El vehículo de administración de genes es preferentemente un vector vírico, y más preferentemente un vector retrovívico, adenovívico, vírico adenoasociado (AAV), de virus del herpes o alfavívico. El vector vírico también puede ser un vector vírico de astrovirus, coronavirus, ortomixovirus, papovavirus, paramixovirus, parvovirus, picornavirus, poxvirus o togavirus. Véanse generalmente Jolly (1994) *Cancer Gene Therapy* 1:51-64; Kimura (1994) *Human Gene Therapy* 5:845-852; Connelly (1995) *Human Gene Therapy* 6:185-193; y Kaplitt (1994) *Nature Genetics* 6:148-153. Los vectores retrovívicos son muy conocidos en la técnica y los presentes inventores contemplan que cualquier vector de terapia génica retrovívico puede emplearse en la invención, que incluye los retrovirus de tipo B, C y D, retrovirus xenotrópicos (por ejemplo, NZB-X1, NZB-X2 y NZB9-1 (véase O'Neill (1985) *J. Virol.* 53:160), retrovirus politrópicos, por ejemplo, MCF y MCF-MLV (véase Kelly (1983) *J. Virol.* 45:291),
10
15 espumavirus y lentivirus. Véase *RNA Tumor Viruses*, segunda edición, Cold Spring Harbor Laboratory, 1985.

Las porciones del vector de terapia génica retrovívico pueden derivarse de diferentes retrovirus. Por ejemplo, las LTR de retrovectores pueden derivarse de un virus del sarcoma murino, un sitio de unión a ARNt de un virus del sarcoma de Rous, una señal de encapsidación de un virus de la leucemia murina y un origen de la síntesis de la segunda cadena de un virus de la leucosis aviar.

- 20 Estos vectores retrovívicos recombinantes pueden usarse para generar partículas de vectores retrovívicos competentes para la transducción introduciéndolos en líneas celulares de encapsidación apropiadas (véase la patente de EE.UU. 5.591.624). Los vectores de retrovirus pueden construirse para integración específica de sitio en ADN de células huésped por incorporación de una enzima integrasa quimérica en la partícula retrovívica (véase el documento WO96/37626). Es preferible que el vector vírico recombinante sea un virus recombinante defectuoso en la replicación.
25

Líneas celulares de encapsidación adecuadas para su uso con los vectores de retrovirus anteriormente descritos son muy conocidas en la técnica, se preparan fácilmente (véanse los documentos WO95/30763 y WO92/05266) y pueden usarse para crear líneas celulares productoras (también llamadas líneas celulares de vector o "VCL") para la producción de partículas de vectores recombinantes. Preferentemente, las líneas celulares de encapsidación se preparan a partir de células parentales humanas (por ejemplo, células HT1080) o líneas celulares parentales de visón, que elimina la inactivación en suero humano.
30

- Los retrovirus preferidos para la construcción de vectores de terapia génica retrovívica incluyen virus de la leucosis aviar, virus de la leucemia bovina, virus de la leucemia murina, virus inductor de focos en célula de visón, virus del sarcoma murino, virus de la reticuloendoteliosis y virus del sarcoma de Rous. Virus de la leucemia murina particularmente preferidos incluyen 4070A y 1504A (Hartley y Rowe (1976) *J Virol* 19:19-25), Abelson (ATCC nº VR-999), Friend (ATCC nº VR-245), Graffi, Gross (ATCC nº VR-590), Kirsten, virus del sarcoma de Harvey y virus de la leucemia murina de Rauscher (ATCC nº VR-998) y Moloney (ATCC nº VR-190). Tales retrovirus pueden obtenerse a partir de repositorios o colecciones tales como la Colección americana de cultivos tipo ("ATCC") en Rockville, Mariland, o aislarse de fuentes conocidas usando técnicas comúnmente disponibles.
35

- 40 Vectores de terapia génica retrovívicos conocidos a modo de ejemplo empleables en la presente invención incluyen aquellos descritos en las solicitudes de patente GB2200651, EP0415131, EP0345242, EP0334301, WO89/02468; WO89/05349, WO89/09271, WO90/02806, WO90/07936, WO94/03622, WO93/25698, WO93/25234, WO93/0230, WO93/10218, WO91/02805, WO91/02825, WO95/07994, US 5.219.740, US 4.405.712, US 4.861.719, US 4.980.289, US 4.717.127, US 5.591.624. Véanse también Vile (1993) *Cancer Res* 53:3860-3864; Vile (1993) *Cancer Res* 53:962-967; Ram (1993) *Cancer Res* 53 (1993) 83-88; Takamiya (1992) *J Neurosci Res* 33:493-503; Baba (1993) *J Neurosurg* 79:729-735; Mann (1983) *Cell* 33:153; Cane (1984) *PNAS USA* 81:6349; y Miller (1990) *Human Gene Therapy* 1.
45

- Los vectores de terapia génica adenovívicos humanos también se conocen en la técnica y son empleables en la presente invención. Véanse, por ejemplo, Berkner (1988) *Biotechniques* 6:616 y Rosenfeld (1991) *Science* 252:431, y los documentos WO93/07283, WO93/06223 y WO93/07282. Vectores de terapia génica adenovívicos conocidos a modo de ejemplo empleables en la presente invención incluyen aquellos descritos en los documentos anteriormente referenciados y en los documentos WO94/12649, WO93/03769, WO93/19191, WO94/28938, WO95/11984, WO95/00655, WO95/27071, WO95/29993, WO95/3467, WO96/05320, WO94/08026, WO94/11506, WO93/06223, WO94/24299, WO95/14102, WO95/24297, WO95/02697, WO94/28152, WO94/24299, WO95/09241, WO95/25807, WO95/05835, WO94/18922 y WO95/09654. Alternativamente, puede emplearse la administración de ADN ligado a adenovirus muerto como se describe en Curiel (1992) *Hum. Gene Ther.* 3:147-154. Los vehículos de administración de genes de la invención también incluyen vectores de virus asociados a adenovirus (AAV). Ejemplos importantes y preferidos de tales vectores para su uso en la presente invención son los vectores basados en AAV-2 desvelados en Srivastava, documento WO93/09239. Los vectores de AAV más preferidos comprenden las dos repeticiones invertidas de AAV en las que las D-secuencias nativas se modifican por sustitución de nucleótidos, de forma que al
50
55
60

menos 5 nucleótidos nativos y hasta 18 nucleótidos nativos, preferentemente al menos 10 nucleótidos nativos hasta 18 nucleótidos nativos, lo más preferentemente 10 nucleótidos nativos, son retenidos y el resto de los nucleótidos de la D-secuencia están deletados o sustituidos con nucleótidos no nativos. Las D-secuencias nativas de las repeticiones terminales invertidas de AAV son secuencias de 20 nucleótidos consecutivos en cada repetición terminal invertida de AAV (es decir, hay una secuencia en cada extremo) que no participan en la formación de HP. El nucleótido de sustitución no nativa puede ser cualquier nucleótido distinto del nucleótido encontrado en la D-secuencia nativa en la misma posición. Otros vectores de AAV a modo de ejemplo empleables son pWP-19, pW N-1, ambos de los cuales se desvelan en Nahreini (1993) Gene 124:257-262. Otro ejemplo de un vector de AAV tal es psub201 (véase Samulski (1987) J. Virol. 61:3096). Otro vector de AAV a modo de ejemplo es el vector de ITR de D doble. La construcción del vector de ITR de D doble se desvela en la patente de EE.UU. 5.478.745. Todavía otros vectores son los desvelados en la patente de EE.UU. de Carter 4.797.368 y la patente de EE.UU. de Muzyczka 5.139.941, la patente de EE.UU. de Chartejee 5.474.935 y el documento de Kotin WO94/288157. Todavía otro ejemplo de un vector de AAV empleable en la presente invención es SSV9AFABTKneo, que contiene el potenciador de AFP y el promotor de albúmina y dirige la expresión predominantemente en el hígado. Su estructura y construcción se desvelan en Su (1996) Human Gene Therapy 7:463-470. Vectores de terapia génica de AAV adicionales se describen en los documentos US 5.354.678, US 5.173.414, US 5.139.941 y US 5.252.479.

Los vectores de terapia génica de la invención también incluyen vectores de herpes. Ejemplos importantes y preferidos son vectores del virus del herpes simple que contienen una secuencia que codifica un polipéptido de timidina cinasa tal como los desvelados en los documentos US 5.288.641 y EP0176170 (Roizman). Vectores del VHS a modo de ejemplo adicionales incluyen HFEM/ICP6-LacZ desvelado en el documento WO95/04139 (Wistar Institute), pHSVlac descrito en Geller (1988) Science 241:1667-1669 y en los documentos WO90/09441 y WO92/07945, VHS Us3::pgC-lacZ descrito en Fink (1992) Human Gene Therapy 3:11-19 y VHS 7134, 2 RH 105 y GAL4 descritos en el documento EP 0453242 (Breakefield), y aquellos depositados en la ATCC con los números de acceso VR-977 y VR-260.

También se contemplan vectores de terapia génica del alfavirus que pueden emplearse en la presente invención. Vectores del alfavirus preferidos son vectores del virus Sindbis. Togavirus, virus del bosque de Semliki (ATCC VR-67; ATCC VR-1247), virus de Middleberg (ATCC VR-370), virus de río Ross (ATCC VR-373; ATCC VR-1246), virus de la encefalitis equina venezolana (ATCC VR923; ATCC VR-1250; ATCC VR-1249; ATCC VR-532), y aquellos descritos en las patentes de EE.UU. 5.091.309, 5.217.879 y WO92/10578. Más particularmente, son empleables aquellos vectores de alfavirus descritos en el documento de EE.UU. nº de serie 08/405.627 presentada el 15 de marzo de 1995, los documentos WO94/21792, WO92/10518, WO95/07994, US 5.091.309 y US 5.217.879. Tales alfavirus pueden obtenerse de repositorios o colecciones tales como la ATCC en Rockville, Maryland, o aislarse de fuentes conocidas usando técnicas comúnmente disponibles. Preferentemente se usan vectores de alfavirus con citotoxicidad reducida (véase el documento USSN 08/679640).

Los sistemas de vector de ADN tales como los sistemas de expresión en capas eucariotas también son útiles para expresar los ácidos nucleicos de la invención. Véase el documento WO95/07994 para una descripción detallada de sistemas de expresión en capas eucariotas. Preferentemente, los sistemas de expresión en capas eucariotas de la invención se derivan de vectores de alfavirus y lo más preferentemente de vectores víricos de Sindbis.

Otros vectores víricos adecuados para su uso en la presente invención incluyen aquellos derivados del virus de la poliomielitis, por ejemplo, ATCC VR-58 y aquellos descritos en Evans, Nature 339 (1989) 385 y Sabin (1973) J. Biol. Standardization 1:115; rinovirus, por ejemplo, ATCC VR-1110 y aquellos descritos en Arnold (1990) J Cell Biochem L401; virus de la viruela tales como el virus de la viruela del canario o el virus de la variolovacuna, por ejemplo, ATCC VR-111 y ATCC VR-2010 y aquellos descritos en Fisher-Hoch (1989) PNAS USA 86:317; Flexner (1989) Ann NY Acad Sci 569:86; Flexner (1990) Vaccine 8:17; en los documentos US 4.603.112 y US 4.769.330 y WO89/01973; virus SV40, por ejemplo, ATCC VR-305 y aquellos descritos en Mulligan (1979) Nature 277:108 y Madzak (1992) J Gen Virol 73:1533; virus de la gripe, por ejemplo, ATCC VR-797 y virus recombinantes de la gripe producidos empleando técnicas de genética inversa como se describen en el documento US 5.166.057 y en Enami (1990) PNAS USA 87:3802-3805; Enami & Palese (1991) J Virol 65:2711-2113 y Luytjes (1989) Cell 59:110, (véanse también McMichael (1983) NEJ Med 309:13 y Yap (1978) Nature 273:238 y Nature (1979) 277:108); virus de la inmunodeficiencia humana como se describe en el documento EP-0386882 y en Buchschacher (1992) J. Virol. 66:2731; virus del sarampión, por ejemplo, ATCC VR-67 y VR-1247 y aquellos descritos en el documento EP-0440219; virus de Aura, por ejemplo, ATCC VR-368; virus de Bebaru, por ejemplo, ATCC VR-600 y ATCC VR-1240; virus de Cabassou, por ejemplo, ATCC VR-922; virus de Chikungunya, por ejemplo, ATCC VR-64 y ATCC VR-1241; virus de Fort Morgan, por ejemplo, ATCC VR-924; virus de Getah, por ejemplo, ATCC VR-369 y ATCC VR-1243; virus de Kyzilagach, por ejemplo, ATCC VR-927; virus de Mayaro, por ejemplo, ATCC VR-66; virus de Mucambo, por ejemplo, ATCC VR-580 y ATCC VR-1244; virus de Ndumu, por ejemplo, ATCC VR-371; virus de Pixuna, por ejemplo, ATCC VR-372 y ATCC VR-1245; virus de Tonate, por ejemplo, ATCC VR-925; virus de Trinita, por ejemplo, ATCC VR-469; virus de Una, por ejemplo, ATCC VR-374; virus de Whataroa, por ejemplo, ATCC VR-926; virus Y-62-33, por ejemplo, ATCC VR-375; virus de O'Nyong, virus de la encefalitis oriental, por ejemplo, ATCC VR-65 y ATCC VR-1242; virus de la encefalitis occidental, por ejemplo, ATCC VR-70, ATCC VR-1251, ATCC VR-622 y ATCC VR-1252; y coronavirus, por ejemplo, ATCC VR-740 y aquellos descritos en Hamre (1966) Proc Soc Exp Biol Med 121:190.

La administración de las composiciones de la presente invención en células no se limita a los vectores víricos anteriormente mencionados. Pueden emplearse otros procedimientos y medios de administración tales como, por ejemplo, vectores de expresión de ácidos nucleico, ADN condensado policatiónico ligado o sin ligar a adenovirus muerto solo, por ejemplo, véase el documento USSN 08/366.787 presentada el 30 de diciembre de 1994 y Curiel (1992) Hum Gene Ther 3:147-154, ADN ligado de ligando, por ejemplo, véase Wu (1989) J Biol Chem 264:16985-16987, vehículos de células de administración de células eucariotas, por ejemplo, véase el documento USSN 08/240.030 presentada el 9 de mayo de 1994 y el documento USSN 08/404.796, deposición de materiales de hidrogel fotopolimerizados, pistola de partículas de transferencia de genes de mano como se describe en la patente de EE.UU. nº 5.149.655, radiación ionizante como se describe en los documentos US5.206.152 y en WO92/11033, neutralización de carga nucleica o fusión con membranas celulares. Enfoques adicionales se describen en Philip (1994) Mol Cell Biol 14:2411-2418 y en Woffendin (1994) PNAS USA 91:1581-1585.

La transferencia de genes mediada por partículas puede emplearse, por ejemplo, véase el documento de EE.UU. nº de serie 60/023.867. Brevemente, la secuencia puede insertarse en vectores convencionales que contienen secuencias de control convencionales para la expresión de alto nivel, y luego incubarse con moléculas de transferencia de genes sintéticas tales como cationes de unión a ADN polimérico como polilisina, protamina y albúmina, ligados a ligandos que eligen células como diana tales como asialo-orosomucoide, como se describe en Wu & Wu (1987) J. Biol. Chem. 262:4429-4432, insulina como se describe en Hucked (1990) Biochem Pharmacol 40:253-263, galactosa como se describe en Plank (1992) Bioconjugate Chem 3:533-539, lactosa o transferrina. También puede emplearse ADN desnudo. Los procedimientos de introducción de ADN desnudo a modo de ejemplo se describen en los documentos WO 90/11092 y US 5.580.859. La eficiencia de captación puede mejorarse usando perlas de látex biodegradables. Las perlas de látex recubiertas de ADN se transportan eficientemente en células después de la iniciación de la endocitosis por las perlas. El procedimiento puede mejorarse adicionalmente por tratamiento de las perlas para aumentar la hidrofobia y así facilitar la rotura del endosoma y liberación del ADN en el citoplasma.

Los liposomas que pueden actuar de vehículos de administración de genes se describen en los documentos US 5.422.120, WO95/13796, WO94/23697, WO91/14445 y EP-524.968. Como se describe en el documento USSN 60/023.867, en la administración no vírica, las secuencias de ácidos nucleicos que codifican un polipéptido pueden insertarse en vectores convencionales que contienen secuencias de control convencionales para la expresión de alto nivel, y luego se incuban con moléculas de transferencia de genes sintéticas tales como cationes de unión a ADN poliméricos como polilisina, protamina y albúmina, ligados a ligandos que eligen células como diana tales como asialo-orosomucoide, insulina, galactosa, lactosa o transferrina. Otros sistemas de administración incluyen el uso de liposomas para encapsular ADN que comprende el gen bajo el control de una variedad de promotores específicos de tejido o ubicuamente activos. Otra administración no vírica adecuada para su uso incluye sistemas de administración mecánica tales como el enfoque descrito en Woffendin y col. (1994) PNAS USA 91(24):11581-11585. Además, la secuencia codificante y el producto de expresión de tal pueden administrarse mediante la deposición de materiales de hidrogel fotopolimerizados. Otros procedimientos convencionales para la administración de genes que pueden usarse para la administración de la secuencia codificante incluyen, por ejemplo, uso de pistola de partículas de transferencia de genes de mano, como se describe en el documento US 5.149.655; uso de radiación ionizante para activar el gen transferido, como se describe en los documentos US 5.206.152 y WO92/11033.

Los vehículos de administración de genes de liposomas y policatiónicos a modo de ejemplo son aquellos descritos en los documentos US 5.422.120 y 4.762.915; en los documentos WO 95/13796; WO94/23697; y WO91/14445; en el documento EP0524968; y en Stryer, Biochemistry, páginas 236-240 (1975) W.H. Freeman, San Francisco; Szoka (1980) Biochem Biophys Acta 600:1; Bayer (1979) Biochem Biophys Acta 550:464; Rivnay (1987) Meth Enzymol 149:119; Wang (1987) PNAS USA 84:7851; Plant (1989) Anal Biochem 176:420.

Una composición de polinucleótido puede comprender una cantidad terapéuticamente eficaz de un vehículo de terapia génica, como el término se ha definido anteriormente. Para los fines de la presente invención, una dosis eficaz será de aproximadamente 0,01 mg/kg a 50 mg/kg o 0,05 mg/kg a aproximadamente 10 mg/kg de las construcciones de ADN en el individuo al que se administra.

Procedimientos de administración

Una vez formuladas, las composiciones de polinucleótido de la invención pueden administrarse (1) directamente al sujeto; (2) administrarse *ex vivo*, a células derivadas del sujeto; o (3) *in vitro* para la expresión de proteínas recombinantes. Los sujetos que van a tratarse pueden ser aves o mamíferos (incluyendo seres humanos).

La administración directa de las composiciones se llevará generalmente a cabo mediante inyección, tanto subcutáneamente, intraperitonealmente, intravenosamente como intramuscularmente, o se administrarán al espacio intersticial de un tejido. Las composiciones también pueden administrarse a una lesión. Otros modos de administración incluyen administración oral y pulmonar, supositorios y aplicaciones transdérmicas o transcutáneas (por ejemplo, véase el documento WO98/20734), agujas y pistolas de genes o hipoesprays. El tratamiento de dosificación puede ser un programa de dosis única o un programa de múltiples dosis. Los procedimientos para la administración *ex vivo* y la reimplantación de células transformadas en un sujeto se conocen en la técnica y se describen en, por ejemplo, el documento WO93/14778. Ejemplos de células útiles en aplicaciones *ex vivo* incluyen,

por ejemplo, citoblastos, particularmente hematopoyéticos, linfocitos, macrófagos, células dendríticas o células tumorales.

Generalmente, la administración de ácidos nucleicos para aplicaciones tanto *ex vivo* como *in vitro* puede llevarse a cabo por los siguientes procedimientos, por ejemplo, transfección mediada por dextrano, precipitación de fosfato de calcio, transfección mediada por polibreno, fusión de protoplastos, electroporación, encapsulación de polinucleótido(s) en liposomas y microinyección directa del ADN en núcleos, todos muy conocidos en la técnica.

Composiciones farmacéuticas de polinucleótidos y polipéptidos

Además de los vehículos y sales farmacéuticamente aceptables descritos anteriormente, los siguientes agentes adicionales pueden usarse con composiciones de polinucleótidos y/o polipéptidos.

A. Polipéptidos

Un ejemplo son polipéptidos que incluyen, sin limitación: asialo-orosomucoide (ASOR); transferrina; asialoglicoproteínas; anticuerpos; fragmentos de anticuerpos; ferritina; interleucinas; interferones, factor estimulante de colonias de granulocitos y macrófagos (GM-CSF), factor estimulante de colonias de granulocitos (G-CSF), factor estimulante de colonias de macrófagos (M-CSF), factor de citoblastos y eritropoyetina. También pueden usarse antígenos víricos, tales como proteínas de la envuelta. Por tanto, proteínas de otros organismos invasivos tales como el péptido de 17 aminoácidos de la proteína circumsporozoito de *Plasmodium falciparum* conocida como RII.

B. Hormonas, vitaminas, etc.

Otros grupos que pueden incluirse son, por ejemplo: hormonas, esteroides, andrógenos, estrógenos, hormona tiroidea o vitaminas, ácido fólico.

C. Polialquilenos, polisacáridos, etc.

Por tanto, el polialquilenglicol puede incluirse con los polinucleótidos/polipéptidos deseados. En una realización preferida, el polialquilenglicol es polietilenglicol. Además, pueden incluirse mono-, di- o polisacáridos. En una realización preferida de este aspecto, el polisacárido es dextrano o DEAE-dextrano. Por tanto, quitosano y poli(lactida-co-glicolida).

D. Lípidos y liposomas

El polinucleótido/polipéptido deseado también puede encapsularse en lípidos o encapsidarse en liposomas antes de la administración al sujeto o a células derivadas del mismo.

La encapsulación de lípidos se realiza generalmente usando liposomas que pueden unirse establemente o ser atrapados y retener el ácido nucleico. La relación de polinucleótido condensado con respecto a preparación de lípido puede variar, pero generalmente será aproximadamente 1:1 (mg de ADN:micromoles de lípido), o más de lípido. Para una revisión del uso de liposomas como vehículos para la administración de ácidos nucleicos véase Hug y Sleight (1991) *Biochim. Biophys. Acta.* 1097:1-17; Straubinger (1983) *Meth. Enzymol.* 101:512-527.

Las preparaciones liposómicas para su uso en la presente invención incluyen preparaciones catiónicas (positivamente cargadas), aniónicas (negativamente cargadas) y neutras. Se ha mostrado que los liposomas catiónicos median en la administración intracelular de ADN de plásmido (Felgner (1987) *PNAS USA* 84:7413-7416); ARNm (Malone (1989) *PNAS USA* 86:6077-6081); y factores de transcripción purificados (Debs (1990) *J. Biol. Chem.* 265:10189-10192), en forma funcional.

Los liposomas catiónicos están fácilmente disponibles. Por ejemplo, liposomas de N[1,2,3-dioleiloxi]propil]-N,N,N-trietilamonio (DOTMA) están disponibles bajo la marca registrada Lipofectin, de GIBCO BRL, Grand Island, NY, (véase, también, Felgner, arriba). Otros liposomas comercialmente disponibles incluyen transfectace (DDAB/DOPE) y DOTAP/DOPE (Boehringer). Pueden prepararse otros liposomas catiónicos a partir de materiales fácilmente disponibles usando técnicas muy conocidas en la técnica. Véase, por ejemplo, Szoka (1978) *PNAS USA* 75:4194-4198; el documento WO90/11092 para una descripción de la síntesis de liposomas de DOTAP (1,2-bis(oleoiloxi)-3-(trimetilamonio)propano).

Similarmente, los liposomas aniónicos y neutros están fácilmente disponibles, tales como de Avanti Polar Lipids (Birmingham, AL), o pueden prepararse fácilmente usando materiales fácilmente disponibles. Tales materiales incluyen fosfatidilcolina, colesterol, fosfatidiletanolamina, dioleoilfosfatidilcolina (DOPC), dioleoilfosfatidilglicerol (DOPG), dioleoilfosfatidiletanolamina (DOPE), entre otros. Estos materiales también pueden mezclarse con los materiales de partida de DOTMA y DOTAP en relaciones apropiadas. Los procedimientos para preparar liposomas usando estos materiales son muy conocidos en la técnica.

Los liposomas pueden comprender vesículas multilaminares (MLV), vesículas unilaminares pequeñas (SUV) o vesículas unilaminares grandes (LUV). Los diversos complejos de liposoma-ácido nucleico se preparan usando procedimientos conocidos en la técnica. Véase, por ejemplo, Straubinger (1983) *Meth. Immunol.* 101:512-527; Szoka

(1978) PNAS USA 75:4194-4198; Papahadjopoulos (1975) Biochim. Biophys. Acta 394:483; Wilson (1979) Cell 17:77; Deamer & Bangham (1976) Biochim. Biophys. Acta 443:629; Ostro (1977) Biochem. Biophys. Res. Commun. 76:836; Fraley (1979) PNAS USA 76:3348; Enoch & Strittmatter (1979) PNAS USA 76:145; Fraley (1980) J. Biol. Chem. (1980) 255:10431; Szoka & Papahadjopoulos (1978) PNAS USA 75:145; y Schaefer-Ridder (1982) Science 215:166.

E. Lipoproteínas

Además, las lipoproteínas pueden incluirse con el polinucleótido/polipéptido que va a administrarse. Ejemplos de lipoproteínas que van a utilizarse incluyen: quilomicrones, HDL, IDL, LDL y VLDL. También pueden usarse mutantes, fragmentos o fusiones de estas proteínas. Por tanto, pueden usarse modificaciones de lipoproteínas que se producen naturalmente, tales como LDL acetilado. Estas lipoproteínas pueden elegir como diana la administración de polinucleótidos a células que expresan receptores de lipoproteínas. Preferentemente, si las lipoproteínas están incluidas con el polinucleótido que va a administrarse, en la composición no se incluye otro ligando que elige diana.

Las lipoproteínas que se producen naturalmente comprenden un lípido y una porción de proteína. La porción de proteína se conoce como apoproteínas. En la presente, las apoproteínas A, B, C, D y E han sido aisladas e identificadas. Al menos dos de éstas contienen varias proteínas, designadas por número romanos, AI, AII, AIV; CI, CII, CIII.

Una lipoproteína puede comprender más de una apoproteína. Por ejemplo, los quilomicrones que se producen naturalmente comprenden A, B, C y E, con el tiempo estas lipoproteínas pierden A y adquieren C y E. VLDL comprende apoproteínas A, B, C y E, LDL comprende apoproteína B; y HDL comprende apoproteínas A, C y E.

El aminoácido de estas apoproteínas es conocido y se describe en, por ejemplo, Breslow (1985) Annu Rev. Biochem 54:699; Law (1986) Adv. Exp Med. Biol. 151:162; Chen (1986) J Biol Chem 261:12918; Kane (1980) PNAS USA 77:2465; y Utermann (1984) Hum Genet 65:232.

Las lipoproteínas contienen una variedad de lípidos que incluyen triglicéridos, colesterol (libre y ésteres) y fosfolípidos. La composición de los lípidos varía en las lipoproteínas que se producen naturalmente. Por ejemplo, los quilomicrones comprenden principalmente triglicéridos. Una descripción más detallada del contenido de lípidos de lipoproteínas que se producen naturalmente puede encontrarse, por ejemplo, en Meth. Enzymol. 128 (1986). La composición de los lípidos se elige para ayudar en la conformación de la apoproteína para la actividad de unión a receptores. La composición de los lípidos también puede elegirse para facilitar la interacción hidrófoba y la asociación con la molécula de unión a polinucleótidos.

Las lipoproteínas que se producen naturalmente pueden aislarse de suero, por ejemplo, por ultracentrifugación. Tales procedimientos se describen en Meth. Enzymol. (arriba); Pitas (1980) J. Biochem. 255:5454-5460 y Mahey (1979) J Clin. Invest 64:743-750. Las lipoproteínas también pueden producirse por procedimientos *in vitro* o recombinantes por expresión de los genes de apoproteína en una célula huésped deseada. Véase, por ejemplo, Atkinson (1986) Annu Rev Biophys Chem 15:403 y Radding (1958) Biochim Biophys Acta 30: 443. Las lipoproteínas también pueden comprarse de proveedores comerciales, tales como Biomedical Technologies, Inc., Stoughton, MA, EE.UU. Otra descripción de lipoproteínas puede encontrarse en el documento WO98/06437.

F. Agentes policatiónicos

Los agentes policatiónicos pueden incluirse, con o sin lipoproteína, en una composición con el polinucleótido/polipéptido deseado que va a administrarse.

Los agentes policatiónicos presentan normalmente una carga positiva neta a pH fisiológico relevante y pueden neutralizar la carga eléctrica de ácidos nucleicos para facilitar la administración a una localización deseada. Estos agentes tienen aplicaciones tanto *in vitro*, *ex vivo* como *in vivo*. Los agentes policatiónicos pueden usarse para administrar ácidos nucleicos a un sujeto vivo tanto intramuscularmente, subcutáneamente, etc.

Lo siguiente son ejemplos de polipéptidos útiles como agentes policatiónicos: polilisina, poliarginina, poliornitina y protamina. Otros ejemplos incluyen histonas, prolaminas, albúmina de suero humano, proteínas de unión a ADN, proteínas cromosómicas de no histona, proteínas de la envuelta de virus de ADN tales como ϕ X174, los factores de transcripción también contienen dominios que se unen ADN y, por tanto, pueden ser útiles como agentes de condensación de ácidos nucleicos. Brevemente, factores de transcripción tales como C/CEBP, c-jun, c-fos, AP-I, AP-2, AP-3, CPF, Prot-1, Sp-1, Oct-1, Oct-2, CREP y TFIID contienen dominios básicos que se unen a secuencias de ADN.

Agentes policatiónicos orgánicos incluyen: espermina, espermidina y putrescina.

Las dimensiones y las propiedades físicas de un agente policatiónico pueden extrapolarse de la lista anterior para la construcción de otros agentes policatiónicos de polipéptido o para producir agentes policatiónicos sintéticos.

Agentes policationicos sintéticos que son útiles incluyen, por ejemplo, DEAE-dextrano, polibreno. Lipofectin™ y lipofectAMINE™ son monómeros que forman complejos policationicos cuando se combinan con polinucleótidos/polipéptidos.

Ensayos de inmunodiagnóstico

5 Las proteínas de la invención pueden usarse en inmunoensayos para detectar niveles de anticuerpos (o, en cambio, los anticuerpos de la invención pueden usarse para detectar niveles de proteínas). Los inmunoensayos basados en antígenos recombinantes bien definidos pueden desarrollarse para sustituir procedimientos de diagnóstico invasivos. Pueden detectarse anticuerpos para proteínas dentro de muestras biológicas que incluyen, por ejemplo, muestras de sangre o suero. El diseño de los inmunoensayos está sometido a una gran variación, y una variedad de estos se conocen en la técnica. Los protocolos para el inmunoensayo pueden basarse, por ejemplo, en la competencia o reacción directa, o en ensayos de tipo sándwich. Por tanto, los protocolos pueden usar, por ejemplo, soportes sólidos, o pueden ser por inmunoprecipitación. La mayoría de los ensayos implican el uso de anticuerpo marcado o polipéptido; las marcas pueden ser, por ejemplo, fluorescentes, quimioluminiscentes, radiactivas, o moléculas de colorante. También se conocen ensayos que amplifican las señales de la sonda; ejemplos de los cuales son ensayos que utilizan biotina y avidina, e inmunoensayos marcados y mediados por enzima, tales como ensayos de ELISA.

10 Kits adecuados para inmunodiagnóstico y que contienen los reactivos marcados apropiados se construyen envasando los materiales apropiados, que incluyen las composiciones de la invención, en recipientes adecuados, junto con los reactivos y materiales restantes (por ejemplo, tampones adecuados, disoluciones de sal, etc.) requeridos para la realización del ensayo, además del conjunto adecuado de instrucciones de ensayo.

Hibridación de ácidos nucleicos

La "hibridación" se refiere a la asociación de dos secuencias de ácidos nucleicos entre sí por enlaces de hidrógeno. Normalmente, una secuencia se fijará a un soporte sólido y la otra estará libre en disolución. Entonces, las dos secuencias se pondrán en contacto entre sí en condiciones que favorezcan los enlaces de hidrógeno. Factores que afectan estos enlaces incluyen: el tipo y el volumen de disolvente; temperatura de reacción; tiempo de hibridación; agitación; agentes para bloquear la unión no específica de la secuencia en fase líquida al soporte sólido (reactivo de Denhardt o BLOTTO); concentración de las secuencias; uso de compuestos para aumentar la tasa de asociación de secuencias (sulfato de dextrano o polietilenglicol); y la rigurosidad de las condiciones de lavado tras la hibridación. Véase Sambrook y col. [arriba] volumen 2, capítulo 9, páginas 9.47 a 9.57.

30 La "rigurosidad" se refiere a las condiciones en una reacción de hibridación que favorecen la asociación de secuencias muy similares con respecto a secuencias que se diferencian. Por ejemplo, la combinación de temperatura y concentración de sales debería elegirse de aproximadamente 120 a 200°C por debajo de la T_m calculada del híbrido en estudio. Las condiciones de temperatura y sal pueden determinarse frecuentemente empíricamente en experimentos preliminares en los que muestras de ADN genómico inmovilizadas sobre filtros se hibridan con la secuencia de interés y luego se lavan en condiciones de diferentes rigurosidades. Véase Sambrook y col. en la página 9.50.

Las variables a considerar cuando se realiza, por ejemplo, una transferencia Southern son (1) la complejidad del ADN que se transfiere y (2) la homología entre la sonda y las secuencias que se detectan. La cantidad total del (de los) fragmento(s) que va(n) a estudiarse puede variar una magnitud de 10, de 0,1 a 1 µg para un plásmido o digestión de fago a 10⁻⁹ a 10⁻⁸ g para un gen de una única copia en un genoma eucariota altamente complejo. Para reducir la complejidad de polinucleótidos pueden usarse tiempos de transferencia, hibridación y exposición sustancialmente más cortos, una cantidad más pequeña de polinucleótidos de partida y menor actividad específica de sondas. Por ejemplo, un gen de levadura de una única copia puede detectarse con un tiempo de exposición de sólo 1 hora empezando con 1 µg de ADN de levadura, transfiriendo durante dos horas e hibridando durante 4-8 horas con una sonda de 10⁸ cpm/µg. Para un gen de mamífero de una única copia, un enfoque conservativo empezaría con 10 µg de ADN, transferencia durante la noche e hibridación durante la noche en presencia de 10% de sulfato de dextrano usando una sonda de más de 10⁸ cpm/µg, produciendo un tiempo de exposición de -24 horas.

50 Varios factores pueden afectar la temperatura de fusión (T_m) de un híbrido de ADN-ADN entre la sonda y el fragmento de interés y, por consiguiente, las condiciones apropiadas para hibridación y lavado. En muchos casos, la sonda no es el 100% homóloga al fragmento. Otras variables comúnmente encontradas incluyen la longitud y contenido de G+C total de las secuencias de hibridación y la fuerza iónica y contenido de formamida del tampón de hibridación. Los efectos de todos estos factores pueden aproximarse por una única ecuación:

$$T_m = 81 + 16,6(\log_{10}C_i) + 0,4[\% \text{ de (G+C)}] - 0,6(\% \text{ de formamida}) - 600/n - 1,5(\% \text{ de desapareamiento})$$

55 en la que C_i es la concentración de sales (iones monovalentes) y n es la longitud del híbrido en pares de bases (ligeramente modificadas a partir de Meinkoth & Wahl (1984) Anal. Biochem. 138: 267-284).

En el diseño de un experimento de hibridación, algunos factores que afectan la hibridación de ácidos nucleicos pueden alterarse convenientemente. La temperatura de la hibridación y lavados y la concentración de sales durante los lavados son lo más fácil de ajustar. Como la temperatura de hibridación aumenta (es decir, la rigurosidad), es menos probable que la hibridación se produzca entre cadenas que son no homólogas y, como resultado, disminuya el ruido. Si una sonda radiomarcada no es completamente homóloga al fragmento inmovilizado (como es frecuentemente el caso en la familia del gen y en experimentos de hibridación entre especies), la temperatura de hibridación debe reducirse, y aumentará el ruido. La temperatura de los lavados afecta la intensidad de la banda de hibridación y el grado de ruido de un modo similar. La rigurosidad de los lavados también aumenta al disminuir las concentraciones de sales.

- 5
- 10 En general, temperaturas de hibridación convenientes en presencia de 50% de formamida son 42°C para una sonda con 95% al 100% de homología con el fragmento diana, 37°C para el 90% al 95% de homología, y 32°C para el 85% al 90% de homología. Para menores homologías, el contenido de formamida debería reducirse y consecuentemente ajustarse la temperatura usando la ecuación anterior. Si la homología entre la sonda y el fragmento diana no es conocida, el enfoque más simple es empezar con condiciones de hibridación como de lavado que son no rigurosas.
- 15 Si no se observan bandas específicas o alto ruido después de la autorradiografía, el filtro puede lavarse a alta rigurosidad y volverse a exponer. Si el tiempo requerido para la exposición hace que este enfoque sea poco práctico, deberían probarse en paralelo varias rigurosidades de hibridación y/o lavado.

Ensayos de sondas de ácido nucleico

- 20 Procedimientos tales como PCR, ensayos de sondas de ADN ramificadas o técnicas de transferencia que utilizan sondas de ácido nucleico según la invención pueden determinar la presencia de ADNc o ARNm. Se dice que una sonda se "hibrida" con una secuencia de la invención si puede formar un dúplex o complejo bicatenario que es suficientemente estable para ser detectado. Las sondas de ácido nucleico se hibridarán con el ácido nucleico de la invención (hebras codificantes y/o no codificantes). Aunque muchas secuencias de nucleótidos diferentes codificarán la secuencia de aminoácidos, se prefiere la secuencia natural debido a que es la secuencia real presente en células.
- 25 El ARNm representa una secuencia codificante y, por tanto, una sonda debería ser complementaria a la secuencia codificante; el ADNc monocatenario es complementario al ARNm, y, por tanto, una sonda de ADNc debería ser complementaria a la secuencia no codificante.

- 30 La secuencia de la sonda no necesita ser idéntica a una secuencia (o su complemento) - alguna variación en la secuencia y longitud puede conducir a un aumento de la sensibilidad del ensayo si la sonda de ácido nucleico puede formar un dúplex con nucleótidos diana, que pueden detectarse. Por tanto, la sonda de ácido nucleico puede incluir nucleótidos adicionales para estabilizar el dúplex formado. La secuencia adicional también puede ser útil como marca para detectar el dúplex formado. Por ejemplo, una secuencia de nucleótidos no complementaria puede unirse al extremo 5' de la sonda, siendo el resto de la secuencia de la sonda complementaria a una secuencia bacteriana.
- 35 Alternativamente, bases no complementarias o secuencias más largas pueden estar intercaladas en la sonda, siempre que la secuencia de la sonda tenga complementariedad suficiente con una secuencia bacteriana con el fin de hibridarse con la misma y así formar un dúplex que pueda detectarse.

- 40 La longitud y la secuencia exactas de la sonda dependerán de las condiciones de hibridación (por ejemplo, temperatura, condición de sales, etc.). Por ejemplo, para aplicaciones de diagnóstico, dependiendo de la complejidad de la secuencia del analito, la sonda de ácido nucleico normalmente contiene al menos 10-20 nucleótidos, preferentemente 15-25, y más preferentemente al menos 30 nucleótidos, aunque puede ser más corta que esto. Cebadores cortos generalmente requieren temperaturas más frías para formar complejos híbridos suficientemente estables con el molde.

- 45 Las sondas pueden producirse mediante procedimientos de síntesis tales como el procedimiento de triéster de Matteucci y col. [J. Am. Chem. Soc. (1981) 103:3185], o según Urdea y col. [PNAS. USA (1983) 80: 7461], o usando sintetizadores de oligonucleótidos automatizados comercialmente disponibles.

- 50 La naturaleza química de la sonda puede seleccionarse según preferencia. Para ciertas aplicaciones, ADN o ARN son apropiados. Para otras aplicaciones, las modificaciones pueden incorporar, por ejemplo, modificaciones del esqueleto tales como fosforioatos o metilfosfonatos, pueden usarse para aumentar la semivida *in vivo*, alterar la afinidad del ARN, aumentar la resistencia a nucleasa, etc. [por ejemplo, Agrawal & Iyer (1995) Curr. Opin. Biotechnol 6:12-19; Agrawal (1996) TIBTECH 14:376-3871; también pueden usarse análogos tales como PNA [por ejemplo, véase Corey (1997) TIBTECH 15:224-229; Buchardt y col. (1993) TIBTECH 11:384-386].

- 55 Alternativamente, la reacción en cadena de la polimerasa (PCR) es otro medio muy conocido para detectar pequeñas cantidades de ácido nucleico diana. El ensayo se describe en Mullis y col. [Meth. Enzymol. (1987) 155:335-350] y las patentes de EE.UU. 4.683.195 y 4.683.202. Dos nucleótidos del "cebador" se hibridan con los ácidos nucleicos diana y se usan para cebar la reacción. Los cebadores pueden comprender secuencias que no se hibridan con la secuencia de la amplificación diana (o su complemento) para ayudar con la estabilidad del dúplex o, por ejemplo, para incorporar un sitio de restricción conveniente. Normalmente, tal secuencia flanqueará la secuencia bacteriana deseada.

5 Una polimerasa termoestable crea copias de ácidos nucleicos diana de los cebadores usando los ácidos nucleicos diana originales como molde. Después de generarse una cantidad umbral de ácidos nucleicos diana por la polimerasa, pueden detectarse por procedimientos más tradicionales tales como transferencias Southern. Si se usa el procedimiento de transferencia Southern, la sonda marcada se hibridará con la secuencia bacteriana (o su complemento).

10 Por tanto, el ARNm o ADNc puede detectarse por técnicas de transferencia tradicionales descritas en Sambrook y col. [arriba]. ARNm, o ADNc generado a partir de ARNm usando una enzima polimerasa, pueden purificarse y separarse usando electroforesis en gel. Entonces, los ácidos nucleicos sobre el gel se transfieren sobre un soporte sólido, tal como nitrocelulosa. El soporte sólido se expone a una sonda marcada y luego se lava para eliminar cualquier sonda sin hibridar. A continuación, los dúplex que contienen la sonda marcada se detectan. Normalmente, la sonda se marca con un resto radiactivo.

Breve descripción de los dibujos

15 La Figura 1 muestra un alineamiento de dominios catalíticos de diversas toxinas bacterianas que incluyen la toxina de *N. meningitidis* de la invención (NMB1343). Residuos importantes para la actividad catalítica se muestran agrandados.

La Figura 2 muestra un alineamiento múltiple de secuencias de regiones conservadas de LT y toxinas de la invención. Residuos importantes para la actividad catalítica se muestran subrayados. Residuos importantes para la conservación de la estructura tienen un fondo sombreado. Otros residuos conservados se indican en negrita.

20 La Figura 3 muestra la incorporación de NAD radiomarcada en la toxina de *N. meningitidis*. En la Figura 3B, los carriles son: (L1) 95°C; (L2) novobiocina, 5 mMol; (L3) GTP, 10 mMol; (L4) ATP, 10 mMol; (L5) ADP-ribosa, 10 mMol; (L6) nicotinamida (NAM), 10 mMol; (L7) control. La flecha muestra la posición de la toxina.

Modos para llevar a cabo la invención

Gen de la toxina de Neisseria meningitidis, serogrupo B

25 Una proteína con la secuencia de aminoácidos <SEC ID 1> se identificó en *N. meningitidis* (serogrupo B, cepa MC58). Ésta está codificada por un gen que tiene la secuencia de nucleótidos <SEC ID 2>.

La proteína ('NMB1343') muestra el 27% de identidad con CT-A [CAA41592] sobre el solapamiento de 127 aa


```

SEC ID 3      PACPQFDDRTKAAADRGVDVDRITPEPVWRTTCGTLVRSDSRGPQVV-----
                : :                               |||:| | | :
ct . pep      MVKII FVFFIFLSSFSYANDDKLYRADSRRPDEIKQSGGLMPRGQS

SEC ID 3      -----FEEGFHAKDVQNGQYDVEKYVLVNQPSPYVSTSYD----HDLYKTWYKSG
                || |: : : ||: : |: | : : : ||| | | : | : : | :
ct . pep      EYFDRGTQMNINLYDHARGTQTG-----FVRHDDGYVSTSI SLRSAHLVGGTILSGH

SEC ID 3      YNYYV----DAPGGIDVNKTIGDTHKWADQVEVAFPGGIQRKYIIGVCPVDRQTKTEIMS
                :||: ||: :|| :| :| | : ||: : | | |
ct . pep      STYYIYVIATAPNMFVNDVLGAYSHPHPDEQEVSA LGGI PYSQIYGWYRVHFGVLDEQLH

SEC ID 3      DCESNPHYQPWH

ct . pep      RNRGYRDRYYSNLDIAPAADGYGLAGFPPEHRAWREEPWIHHAPPGCGNAPRSSMSNTCD
    
```

Por tanto, como para *N. meningitidis*, sólo hay un bajo nivel de identidad global con toxinas conocidas, pero los residuos catalíticos clave se conservan. La anotación de la base de datos de 'proteína secretada putativa' no sugiere actividad de ADP-ribosiltransferasa.

5 **Gen de la toxina de *Mycoplasma pneumoniae* M129 [22]**

Una proteína con la secuencia de aminoácidos <SEC ID 4> [P75409] se identificó en *M. pneumoniae*:

```

1 MPNPVRFVYR VDLRSPEEIF EHGPFSTLGDV RNFPEEILST NFGRSYFIST SETPTAAIRP
61 FGSWLREYVP EHPRRAYLYE IRADQHPYNA RATGENLLDL MRQRQVVFDS GDREMAQMG I
121 RALRTSFAYQ REWFTDGP IAAANVRSAWLV DAVPVEPGHA HHPAGRVVET TRINEPEMHN
181 PHYQELQTA NDQFWLPTG IATPVHLSIP QAASVADVSE GTSASLSFAC PDWSPSSNG
241 ENPLDKCIAE KIDNYNLQSL PQYASSVKEL EDTPVYLRGI KTQKTFMLQA DPQNNVFLV
301 EVNPKQKSSF PQTIFFDVY QRICLKDLTG AQISLSLTA TQYAGQLKV HLSVSAVNAV
361 NQWKMTPOD IAITQFRVSS ELLGQTE NGLFWNTKSGGSQ HDLYVCPLKN PPSDLEELQI
421 IVDECTTHAQ FVTMRAASTF FVDVOLGWYW RGYYPQLS GWSYQMKTPD GQIFYDLKTS
481 KIFFVQDNQN VFPLHNKLNK QTGYSWDWVE WLKHD MNEDK DENFKWYFSR DDLTIPSVEG
541 LNFHRIRCYA DNQQLKVIIS GSRWGGWYST YDKVESNVED KILVKDGFDR F
    
```

Esta proteína muestra el 29% de identidad con PT sobre el solapamiento de 243 aa:

```

SEC ID 4      MPNPVRFVYR VDLRSPEEIF EHGPFSTLGDVRN
                : | || | | | : : : | : |
pt . pep      AIRQTARTGWLTLWLA IAVTAPVTS PAWADDPATVYRYDSRPPEDV FQNGFTAWGNNDN

SEC ID 4      FFEHIL--STNFGRSY--FISTSETPTAAIRFFGSWLREYVP-EHPRRA-----YLYEI
                : : | : | | | : : : : : : : | : | : | : | :
pt . pep      VLDHLTGRSCQV GSSNSAFVSTSSRRRYTEVYLEHRMQEAVEAERAGRGTGHFIGYIYEV

SEC ID 4      RADQHPYNARATGENLLDLMRQRQVVFDSGDREMAQMGIRALRTSFAYQREWFTDGP IAA
                |||:| |:| : : : | : : : | : | | | | | : : |
pt . pep      RADNNPYGAASSYFEYVDTYG-----DNAGRILAG----ALAT---YSEYLAHRRIPP

SEC ID 4      ANVRSAWLVDAVPVEPGHAHHPAGRVVETTRINEPEMHNPHYQELQTA NDQFWLPTPGI
                | : | : | : | : | | | | : | : | | : | : | : |
pt . pep      ENIRRVTRV-----YH-NGITGETTTT---EYSNARYVSQQTRANPNFYTSRRRSV

SEC ID 4      ATPVHLSIPQAASVADVSEGTSASLSFACPDWSPSSNGENPLDKCIAEKIDNYNLQSLP
                | : | : : | | : : : | | | | |
pt . pep      ASIVG-TLVRMAPVIGACMARQAESSEAMA AWSE RAGEAMVLVYYESIAYSF

SEC ID 4      QYASSVKELEDTPVYLRGIKTQKTFMLQADPQNNVFLVEVNPQKSSFPQTIFFDVYQ
    
```

Por tanto, como para *N. meningitidis* y *S. coelicolor*, sólo hay un bajo nivel de identidad global con una toxina conocida, pero los residuos catalíticos clave se conservan. La anotación de la base de datos de 'proteína hipotética'

no sugiere actividad de ADP-ribosiltransferasa.

Gen de la toxina de Salmonella typhimurium LT2 (cepa SGSC1412)

Una proteína con la secuencia de aminoácidos <SEC ID 5> se identificó en *S. typhimurium*:

```

1  MKKLIFLTL  IVSFNNYAVD  FVYRVVDSTPP  DVIPRDGFSL  LGYNRNQOF
51  ISGRSCSGGS  SDSRYIATTS  SVNQTYAIAR  AYYSRSTFKG  NLYRYQIRAD
101 NNFYSLLPSI  TYLETQGGHF  NAYEKTMMRL  QREYVSTLSI  LPENIQKAVA
151 LVYDSATGLV  KDGVSTMNAS  YLGLSTTSP  GVIPFLPEPQ  TYTQORIXAF
201 GPLISSCFSI  GSVCHSHRQ  RADVYNMSFY  DARPVIELIL  SK
    
```

5 Esta proteína muestra el 29% de identidad con PT sobre el solapamiento de 232 aa:

```

SEC ID 5          MKKLIFLTL  IVSFNNYAVD  FVYRVVDSTPP  DVIPRDGFSL  LGYNRNQOFI
pt . pep          ARTGWLTLA  IAVTAPVTS  PAWADDPPAT  VYRYDSRPP  EDVVFQNGFT  AWGNNDNVLD  HLDHL
                  20          30          40          50          60          70

SEC ID 5          SGRSCSGGS  SDSRYIATTS  -----  VNQTYAIAR  AYYSRSTFKG  NLYRYQIRAD
pt . pep          TGRSCQVGS  SNSAFVST  SSSRRYTEV  YLEHRMQEA  VEAEERAGR  GTGHFIG--  YIYEVRAD

SEC ID 5          NNFYSLLPS  -ITYLETQGG  HFN-AYEKT  MMRLQREYV  STLSILPEN  IQKAVLVYD  SATG
pt . pep          NNFYGAASS  YFEYVDTY  GDNAGRILA  GALATYQSE  YLAHRRIPP  ENIRRVT  RVYHNGITG

SEC ID 5          LVKDGVSTM  NASYLGLST  TSNPGVI  PFLPEPQ  TYTQORIXA  FGPLISSCF  SIGSVCHSHR
pt . pep          ETTTTEYS  -NARYVSQ  QTRANPN-  -----  PYTSRRSV  ASIVGTLV  RMAPVIGA-  CMARQ

SEC ID 5          GQRADVYN  MSFYDARPV  IELILSK
pt . pep          AESSEA-  -MAAWSER  AGEAMVLV  VYESIAYSF
    
```

Por tanto, como se ha descrito anteriormente, sólo hay un bajo nivel de identidad global con una toxina conocida, pero los residuos catalíticos clave se conservan.

10 Además, un gen en la dirección 5' de la proteína (SEC ID 9) en *S. typhimurium* muestra homología con la subunidad S2 de la toxina pertussis:

```

Puntuación = 34,0 bits (77), Esperado = 0,91
Identities = 31/101 (30%), Positivos = 42/101 (40%), Huecos = 7/101 (6%)

Búsqueda: 30  TNAYYSDEVI  SELHVGQID  TSPYFCIK  TVKANGSG  TTPVV-AC  AVSKQSI  WAPSPKEL  LDQ 88
T+ YYS+ + L T+ C V+ SG PV+ AC + + L
Objeto: 98  TDHYYSNV  TATRLLS-  -STNSRLC  AVFVR--  -SGQPVI  GACTSPY  DGKYWSM  YSRLRKM  151

Búsqueda: 89  ARYFYSTG  OSVRIHVQ  KNIWTYPL  FVNTFSA  NALVGLS  SCS 129
Y G SVR+HV K Y TF AL G+S C+
Objeto: 152  LYLIYVAG  ISVRVHVS  KEEQYDY  EDATFET  YALTGIS  ICN 192
    
```

Gen de la toxina de Salmonella paratyphi A (cepa ATCC 9150)

Una proteína con la secuencia de aminoácidos <SEC ID 6> se identificó en *S. paratyphi*. Ésta muestra buena homología con la secuencia de *S. typhimurium* mostrada anteriormente:

Puntuación = 1231 (438,4 bits), Esperado = 1,6e-125, P = 1,6e-125
 Identidades = 241/242 (99%), Positivos = 241/242 (99%)

Typhi: 1 MKKLIPLTSLIVSFNNYAVDFVYRVDSTPPDVIFRDGFSLLGYNRNFQOFISGRSCSGGS 60
 MKKLIPLTSLIVSFNNYAVDFVYRVDSTPPDVIFRDGFSLLGYNRNFQOFISGRSCSGGS

Parat: 1 MKKLIPLTSLIVSFNNYAVDFVYRVDSTPPDVIFRDGFSLLGYNRNFQOFISGRSCSGGS 60

Búsqueda 61 SDSRYIATTSSVNQTYAIARAYYSRSTFKGNLYRYQIRADNNFYSLLPSTITYLETQGGHF 120
 SDSRYIATTSSVNQTYAIARAYYSRSTFKGNLYRYQIRADNNFYSLLPSTITYLETQGGHF

Objeto 61 SDSRYIATTSSVNQTYAIARAYYSRSTFKGNLYRYQIRADNNFYSLLPSTITYLETQGGHF 120

Búsqueda 121 NAYEKTMMRLQREYVSTLSILPENIQKAVLVYDSATGLVKDGVSTMNASYLGLSTTSNP 180
 NAYEKTMMRLQREYVSTLSILPENIQKAVLVYDSATGLVKDGVSTMNASYLGLSTTSNP

Objeto 121 NAYEKTMMRLQREYVSTLSILPENIQKAVLVYDSATGLVKDGVSTMNASYLGLSTTSNP 180

Búsqueda 181 GVIPFLPEPQTYTQQRIFAAGPLISSCPSIGSVCHSHRGQRADVYNMSFYDARPVIELIL 240
 GVIPFLPEPQTYTQQRIFAAGPLISSCPSIGSVCHSHRGQRADVYNMSFYDARPVIELIL

Objeto 181 GVIPFLPEPQTYTQQRIFAAGPLISSCPSIGSVCHSHRGQRADVYNMSFYDARPVIELIL 240

Búsqueda 241 SK* 242
 SK*
 241 SK* 242

De nuevo, esta proteína muestra sólo un bajo nivel de identidad global con una toxina conocida, pero los residuos catalíticos clave se conservan.

Como para *S. typhimurium*, hay un homólogo de PT-S2 en la dirección 5', pero está desplazado en el marco:

Puntuación = 387 (141,3 bits), Esperado = 4,9e-36, P = 4,9e-36
 Identidades = 73/73 (100%), Positivos = 73/73 (100%), Marco = +3

Búsqueda: 65 TPVVACAVSKQSIWAPSFKELLDQARYFYSTGQSVRIHVQKNIWYPLFVNTFSPANALVG 124
 TPVVACAVSKQSIWAPSFKELLDQARYFYSTGQSVRIHVQKNIWYPLFVNTFSPANALVG

Objeto: 14802 TPVVACAVSKQSIWAPSFKELLDQARYFYSTGQSVRIHVQKNIWYPLFVNTFSPANALVG 14981

Búsqueda: 125 LSSCSATQCFGPK 137
 LSSCSATQCFGPK

Objeto: 14982 LSSCSATQCFGPK 15020

Puntuación = 327 (120.2 bits), Esperado = 1.1e-29, P = 1.1e-29
 Identidades = 65/96 (67%), Positivos = 73/96 (76%), Marco = +1

Búsqueda: 1 MYMSKYVPVYTLILLIYSFNASAWEWTDNTNAYYSDEVISELHVQIDTSPYFCIKTVKA 60
 MY++K+VPVYTLILLIYSFNASAWEWTDNTNAYYSDEVISELHVQIDTSPYFCIKTVKA

Objeto: 14611 MYINKFVPVYTLILLIYSFNASAWEWTDNTNAYYSDEVISELHVQIDTSPYFCIKTVKA 14790

Búsqueda: 61 NGSQTPVVACAVSKQSIWAPSFKELLDQARYFYSTG 96
 NGS ++ ++ P K L + P G

Objeto: 14791 NGSVHQLLHVRYQSRAYGRPPLKNFLIRQDIPTVQG 14898

5

Gen de la toxina de *Streptococcus pyogenes*

Una proteína con la secuencia de aminoácidos <SEC ID 7> se identificó en *S. pyogenes*. Ésta está codificada por un gen que tiene la secuencia de nucleótidos <SEC ID 8>.

Esta proteína muestra el 24% de identidad con la toxina C3 de *Clostridium limosum*:

```

SEC ID 7      MLKKRYQLAIVLLLSCFSLIWQTEGLVELFVCEHYERAVCEGTP---AYFTFSQKGA
              | : | : | | | | : : | : : : | | : : | : : |
exoc3_cloli. MNKLTERVLCVGVSGILIFSVAALVQGTKKCYANPVRNRAASRVKPYADSPKEFTNIDEA

SEC ID 7      ETLIKKRWGKGLIYPRAEQEAMAAYTCQQAGPINTSLDKAKGELSQLTPELRDQVAQLDA
              :: |::| : : :|::| | | :|::| | | :|::| : | :| :| :|
exoc3_cloli. RAWGDKQFAKYKL-SSSEKNALTIYT-RNAARINGPLRANQGNLNPADIRKEVEQIDK

SEC ID 7      ATHRLVIPWNIVVYRYVYETPLRDIQVSHADL--TSYRNHQFDPHILCKIKL---GTRY
              : : : | |::| | | | : | | : : : : | |
exoc3_cloli. SFTKMQTPENIILFRG-----DDPGYLGPDFENTILNRDGTINKAVFEQVKLRFKGRDR

SEC ID 7      TKHSFMSTTALKNGAMTHRPVEVRICVKKGAKAAFPVSAVPSEVELLFPFGCQLEVVG
              ::::| | : : : | : | : : : | | : : : | : | : : :
exoc3_cloli. KEYGYISTSLVNGSAFAGRPIITKFKVLDGSKAGYIEPISTFKGQLEVLPRSTYTISD

SEC ID 7      AYVSQDQKKLHIEAYFKGSL
              : : : | : | : |
exoc3_cloli. MQIAPNNKQIIITALLKR
    
```

También muestra el 29% de identidad con la transferasa EDIN [M63917; ref. 23] de *S. aureus*:

```

Identidades = 58/195 (29%), Positivos = 106/195 (53%), Huecos = 13/195 (6%)

Búsqueda : 67  RWGKGLI----YPRAEQEAMAAYTCQQAGPINTSLDKAKGELSQLTPELRDQVAQLDAAT 122
              +WG LI Y ++ A+ YT + + IN L A G+++L +D+V +LD++
Objeto : 49  KWGNKLIKQAKYSSDDKIALYEYT-KDSSKINGPLRLAGGDINKLDSTTQDKVRRLDSSI 107

Búsqueda : 123 HRLVIPWNIVVYRYVYETPLRDI-GVSHADLTSYR--NHQFDPHILCKIK--LGTR-YT 176
              + P ++ VYR + +L I G ++ DL + N Q+D +++ K+ + +R Y
Objeto : 108 SKSTTPESVYVYRLLNLDYLTIVGFTNEDLYKLOQTNGQYDENLVRKLNVMNSRIYR 167

Búsqueda : 177 KHSFMSTTALKNGAMTHRPVEVRICVKKGAKAAAFV--EPYSAVPSEVELLFPFGCQLEVV 234
              + + ST + A+ RP+E+R+ + KG KAA++ + +A + E+L PRG + V
Objeto : 168 EDGYSSTQLVSGAAVGGRPIELRLELPGTKAAYLNSKOLTAYYQQEVLLPRGTEYAVG 227

Búsqueda : 235 GAYVSQDQKKLHIEA 249
              +S D+KK+ I A
Objeto : 228 SVELSNDKKKIIITA 242
    
```

Estudios enzimáticos

5 La proteína de *N. meningitidis* se expresó y se purificó en *E. coli* como un producto marcado con His. Los anticuerpos policlonales de ratón producidos contra la proteína recombinante se usaron en análisis de transferencia Western y mostraron una banda repentina a 20kDa en un lisado celular de la cepa MC58 de *N. meningitidis*. Una preparación de vesículas de membrana externa no mostró tal banda.

10 Como ensayo preliminar para la actividad de NAD-glicohidrolasa se usó agmatina como aceptor de ADP-ribosa. La proteína purificada se incubó en presencia de agmatina 0 mM, 20 mM o 75 mM en fosfato de potasio 50 mM, [carbonil-¹⁴C]NAD 0,01 mM (0,05 µCi), pH 7,5, en un volumen total de 0,3 ml. Después de las incubaciones a 30°C durante 18 h, las muestras (100 µl) se aplicaron a una columna de 1 ml de Dowex AG 1-X2. La [¹⁴C]nicotinamida se eluyó con 5 ml de H₂O para radioensayo.

La actividad enzimática fue del siguiente modo:

Concentración de agmatina (mM)	Actividad enzimática (pmoles/h)
0	34
20	62
75	72

Se realizaron otros estudios de actividad de NAD-glicohidrolasa y ADP-ribosiltransferasa usando agmatina como aceptor. El ensayo de NAD-glicohidrolasa usó [carbonil-¹⁴C]NAD 0,1 M, que se sustituyó con [adenina-U-¹⁴C]NAD para el ensayo de ADP-ribosiltransferasa. Después de la incubación a 30°C durante 1 hora, las muestras se probaron como antes. Los resultados fueron del siguiente modo:

[Agmatina] (mM)	ADP-ribosilagmatina formada (nmol/hora)	Nicotinamida liberada (nmol/hora)
0	-	8,9
20	6,3	15,4
75	18,1	25,9

5

Se probaron otros aminoácidos como aceptores de ADP-ribosa a la concentración 20 mM. Los resultados fueron:

Aminoácido	Nicotinamida liberada (nmol/hora)
Control	8,2
Agmatina	15,6
Arginina	14,6
Glicina	8,5
Cisteína	11,7
Serina	8,6
Lisina	8,9
Histidina	8,8
Prolina	8,6

Auto ADP-ribosilación de toxina de *N. meningitidis*

10 NMB1343 purificado (5,7 µg) se incubó en fosfato de potasio 50 mM (pH 7,5) con [adenilato-³²P]NAD 10 µM (10 µCi por ensayo) en un volumen total de 50 µl durante 1 h a 30°C. La proteína se precipitó con la adición de 50 µl de ácido tricloroacético frío en hielo (concentración final 25%) y, tras una incubación durante la noche a 4°C, se recogió por centrifugación (10.000 x g durante 30 min). El sedimento se suspendió en LDS y se calentó a 70°C durante 5 min. Entonces, las muestras se sometieron a electroforesis en (4-12% o 10%) de geles NuPAGE usando MES como tampón de electroforesis y se electrotransferieron a membranas de nitrocelulosa que se expusieron a la película X-Omat durante 5 días a temperatura ambiente.

15 Durante la incubación con [³²P]NAD, la radiomarca de NMB1343 aumenta de un modo dependiente de la concentración (Figura 3A). Para confirmar que la modificación de proteína era enzimática (es decir, no participaba esa adición química de ADP-ribosa reactiva), los experimentos de marcado se llevaron a cabo en presencia de Nam, un inhibidor de NADasa conocido, y en presencia de un gran exceso de ADP-ribosa fría. Ambos experimentos dieron una cantidad idéntica de proteína radiomarcada (Figura 3B, L5 y L6). La incorporación enzimática también fue soportada por el hallazgo de que la pre-incubación de NMB 1343 a 95°C durante 5 minutos abole completamente el marcado (L1) y que el marcado no se produce en presencia de novobiocina (L2), un conocido inhibidor de ADPRT.

Toxinas mutantes

25 Basándose en homología con toxinas conocidas y la predicción de residuos catalíticos se generaron cuatro mutantes de SEC ID 1, conteniendo cada uno una sustitución de un único aminoácido. La arginina en la posición 7 y los ácidos glutámicos en las posiciones 109, 111 y 120 se sustituyeron con residuos de lisina y glicina, respectivamente, usando mutagénesis dirigida a sitio (SDM) basada en PCR. Se diseñaron cebadores internos que contienen un cambio de codón de Arg a Lys, y de Glu a Gly:

30

Cebador	Secuencia	{SEC ID}	Cambio de codón
WT-directo	CGCGGATCCCATATGGGAAATTTCTTATATAGAGGCATTAGTTGC	{18}	
WT-inverso	CCCCTCGAGGTTAATTTCTATCAACTCTTTAGCAAT	{19}	
R7K-directo	CGCGGATCCCATATGGGAAATTTCTTATAT <u>Aa</u> AGGCATTAGTTGC	{20}	AGA → AaA
E109G-directo	ATTTTTGAATAT <u>Gg</u> GGTTGAACATCCAGAAAAC	{21}	GAG → GgG
E109G-inverso	TTCTGGATGTTCAACCcCATAITCAAAAATAGA	{22}	
E111G-directo	TATGAGGTT <u>Gg</u> ACATCCAGAAAACCCA	{23}	GAA → GgA
E111G-inverso	GTTTTCTGGATGTcCAACCTCATATTC	{24}	
E120G-directo	CCAAATGAGAA <u>Gg</u> AGTAACAATCAGAG	{25}	GAA → GgA
E120G-inverso	GATTGTTACTcCCTTCTCATTTGGGTT	{26}	

Los nucleótidos subrayados codifican lisina o glicina, y los nucleótidos mutados están en minúscula.

Para generar el mutante de R7K se realizó una única etapa de PCR. El molde fue 20 ng de ADN que codifica NMB 1343 marcado con His. Los cebadores fueron R7K-directo y WT-inverso.

5 Para generar los restantes mutantes, la PCR se realizó usando 20 ng de ADN de pET 21b+ como molde, y los siguientes pares de cebadores: (1) WT-directo / E109G-inverso; (2) E109G-directo / WT-inverso; (3) WT-directo / E111G-inverso; (4) E111G-directo / WT-inverso; (5) WT-directo / E120G-inverso; (6) E120G-directo / WT-inverso.

La segunda ronda de PCR se realizó usando el producto de PCR 1-2, 3-4 ó 5-6 como molde, y WT-directo y WT-inverso como cebadores directo e inverso, respectivamente.

10 Los fragmentos de PCR que contienen cada mutación se procesaron siguiendo el procedimiento convencional, se digirieron con enzimas de restricción *NdeI* y *XhoI* y se clonaron en el vector pET-21b+. La presencia de cada mutación se confirmó por análisis de secuencias.

15 Después de clonar cada gen en el vector de expresión, los plásmidos recombinantes se transformaron en cepas de *E. coli* adecuadas para la expresión de la proteína recombinante como marca de His. Se usaron 1,5 µl de cada construcción para transformar BL21-DE3 de *E. coli*. Las colonias recombinantes individuales se inocularon en 4 ml de LB+Amp (100 µg/ml), se incubaron a 37°C durante la noche, luego se diluyeron 1:30 en 20 ml de LB+Amp (100 µg/ml) en matraces de 125 ml dando una DO₆₀₀ entre 0,1 y 0,2. Los matraces se incubaron a 37°C en un agitador con baño de agua giratorio hasta que la DO₆₀₀ indicó el crecimiento exponencial adecuado para la inducción de expresión (DO 0,4-0,8). La expresión de proteínas se indujo mediante la adición de IPTG 1,0 mM. Después de 3 horas de incubación a 37°C, la DO₆₀₀ se midió y se examinó la expresión. 1,0 ml de cada muestra se centrifugó en una microcentrífuga, el sedimento se resuspendió en PBS y se analizó por SDS-PAGE y tinción con azul de Coomassie. Todos los mutantes se expresaron tan eficientemente como los naturales, y se purificaron como formas solubles.

20 Las actividades de ADP-ribosiltransferasa y NAD-glicohidrolasa se determinaron como se ha descrito anteriormente. Las mutaciones de Lys y Gly se estudiaron por separado. Los resultados fueron del siguiente modo:

Mutante	Actividad de ADP-ribosiltransferasa (pmol/hora)	Actividad de NAD-glicohidrolasa (nmol/hora)
Natural	1,8	2,6
R7K	0,1	0,08
Natural	305	680
E109 g	30	315
E111 g	116	289
E120 g	36	100

Por tanto, la actividad enzimática de la toxina de *N. meningitidis* puede eliminarse racionalmente y eficientemente por mutagénesis dirigida a sitio.

Se entenderá que la invención sólo se ha descrito a modo de ejemplo y que pueden hacerse modificaciones mientras que estén dentro del alcance y espíritu de la invención.

5

TABLA 1 - Mutaciones preferidas (Δ = deleción del residuo)

SEC ID	Sitio para mutación (mutaciones)	Residuo de sustitución
1	Arg-7	Lys, Δ
	Lys-24	Gly, Trp
	Tyr-34	Trp, Ala, His
	His-57	Asn, Tyr, Gln, Val, Ser, Pro
	Glu-60	Ala
	Thr-61	Gly
	Tyr-68	Met, Glu
	Ser-70	Phe
	Thr-72	Lys, Tyr
	Ala-82	Arg
	Gly-86	Lys
	Tyr-103	Lys, Asp, Ser
	Glu-107	Asp, Ser, Δ
	Glu-109	Ala, Gly, Lys, Asp, Ser, Δ
	Glu-111	Ala, Gly, Lys, Asp, Gln, Δ
	Glu-118	Asp, Ser, Δ
Glu-120	Ala, Gly, Lys, Asp, Gln, Δ	

ES 2 383 592 T3

(continuación)

SEC ID	Sitio para mutación (mutaciones)	Residuo de sustitución
3	Asp-64	Glu, Tyr
	Arg-81	Lys, Gly, Trp, Δ
	Asp-83	Glu, Tyr, Ser
	Arg-85	Lys, His, Leu
	Gly-86	Arg, Asp
	His-96	Asn, Tyr, Gln, Val, Ser, Pro
	Gln-101	Ala
	Tyr-110	Met
	Val-113	Asp, Glu, Tyr
	Tyr-119	Glu
	Ser-121	Phe
	Ser-123	Lys, Tyr
	His-126	Asn, Tyr, Gln, Val
	Gly-136	Lys
	Val-149	Lys, Tyr
	His-157	Glu
Gln-162	Asp, Ser, Δ	
Glu-164	Ala, Gly, Lys, Asp, Gln, Δ	
4	Arg-10	Lys, Δ
	Asp-12	Glu, Tyr, Ser
	Arg-14	Lys, His, Leu
	Ile-19	Ala
	His-36	Asn, Tyr, Gln, Val, Ser, Pro
	Thr-40	Gly
	Arg-44	Ala, Lys
	Phe-47	Met, Glu
	Ser-49	Phe
	Ser-51	Lys, Tyr
	Ala-57	Arg
	Gln-130	Asp, Ser, Δ
	Glu-132	Ala, Gly, Lys, Asp, Gln, Δ
	Trp-133	Gly

(continuación)

SEC ID	Sitio para mutación (mutaciones)	Residuo de sustitución
5/6	Arg-24	Lys, Δ
	Asp-26	Glu, Tyr, Ser
	Ile-33	Ala
	Tyr-43	Trp, Ala, His
	Arg-46	Ala, Lys
	Cys-56	Δ
	Ser-61	Phe
	Ser-63	Lys, Tyr
	Tyr-65	Met, Glu
	Thr-68	Phe
	Thr-69	Phe
	Ser-70	Lys, Tyr
	Ser-84	Phe
	Ser-86	Lys, Tyr
	Gln-131	Asp, Ser, Δ
Glu-133	Gly, Ala, Lys, Asp, Gln, Δ	
Tyr-134	Gly	
7	Arg-130	Lys, Δ
	Ser-177	Phe
	Thr-179	Lys, Tyr
	Cys-196	Ser
	Glu-215	Ser, Asp, Δ
	Glu-217	Gly, Ala, Lys, Asp, Gln, Δ

REFERENCIAS (cuyo contenido completo se incorpora en el presente documento)

5 [1] Rappuoli & Pizza (1991) Capítulo 1 de Sourcebook of Bacterial Protein Toxins (Alouf & Freer, eds). ISBN 0-12-053078-3.

[2] Bazan & Koch-Nolte (1997) Adv. Exp. Med. Biol. 419:99-107.

[3] Sixma y col. (1991) Nature 351:371-377.

[4] Zhang y col. (1995) J. Mol. Biol. 251:563-573.

10 [5] Stein y col. (1994) Structure 2:45-57.

[6] Solicitud de patente internacional WO93/13202.

[7] Solicitudes de patente europea 0306618, 0322533 y 0322115.

[8] Del Guidice & Rappuoli (1999) Vaccine 1999 17 supl. 2:544-52

[9] Patente europea 0396964.

15 [10] Northrup & Fauci (1972) J. Infect. Dis. 125:672ff.

[11] Elson & Ealding (1984) J. Immunol. 133:2892ff y 132:2736ff.

[12] Solicitud de patente internacional W095/17211.

[13] van den Akker y col. (1997) Protein Sci 6:2644-2649.

[14] Solicitud de patente internacional WO99/24578.

20 [15] Solicitud de patente internacional WO99/36544.

[16] Solicitud de patente internacional WO99/57280.

[17] Tettelin y col. (2000) Science 287:1809-1815.
 [18] Pizza y col. (2000) Science 287:1816-1820.
 [19] Solicitud de patente internacional WO01/64920.
 [20] Solicitud de patente internacional WO01/64922.
 5 [21] Parkhill y col. (2000) Nature 404:502-506.
 [22] Himmelreich y col. (1996) Nucleic Acids Res. 24:4420-4449.
 [23] Wilde y col. (2001) J. Biol. Chem. 276:9537-9542.

Lista de secuencias

- <110> CHIRON SpA
 10 <120> TOXINAS BACTERIANAS ADP-RIBOSILANTES
 <130> P02696SWO
 <150> GB-0108024.1
 <151> 30/03/2001
 <160> 26
 15 <170> SeqWin99, versión 1.02
 <210> 1
 <211> 145
 <212> PRT
 <213> Neisseria meningitidis
 20 <400> 1

```

Met Gly Asn Phe Leu Tyr Arg Gly Ile Ser Cys Gln Gln Asp Glu Gln
 1           5                10                15

Asn Asn Gly Gln Leu Lys Pro Lys Gly Asn Lys Ala Glu Val Ala Ile
          20                25                30

Arg Tyr Asp Gly Lys Phe Lys Tyr Asp Gly Lys Ala Thr His Gly Pro
          35                40                45

Ser Val Lys Asn Ala Val Tyr Ala His Gln Ile Glu Thr Gly Leu Tyr
 50                55                60

Asp Gly Cys Tyr Ile Ser Thr Thr Thr Asp Lys Glu Ile Ala Lys Lys
65                70                75                80

Phe Ala Thr Ser Ser Gly Ile Glu Asn Gly Tyr Ile Tyr Val Leu Asn
          85                90                95

Arg Asp Leu Phe Gly Gln Tyr Ser Ile Phe Glu Tyr Glu Val Glu His
          100                105                110

Pro Glu Asn Pro Asn Glu Lys Glu Val Thr Ile Arg Ala Glu Asp Cys
          115                120                125

Gly Cys Ile Pro Glu Glu Val Ile Ile Ala Lys Glu Leu Ile Glu Ile
130                135                140

Asn
145
    
```

ES 2 383 592 T3

<210>2
 <211>435
 <212> ADN
 <213> Neisseria meningitidis

5 <400> 2

```

    .atgggaaatt tcttatatag aggcattagt tgccaacaag atgagcaaaa taatggacag    60
      ttaaaaccta aaggtaataa agctgaagtt gcaattcgtt atgatggtaa gtttaaatat    120
      gatggtaaag ctacacatgg tccaagtgtg aagaatgcag tttacgcca tcaaattgaa    180
      acaggtctat atgacggatg ttatatatct acgacaacag acaaggaaat tgccaagaaa    240
    
```

```

      tttgcaaaa gttccggcat cgaaaatggc tatatatatg ttttaaatag ggatttgttt    300
      ggtcaatatt ctatthttga atatgaggtt gaacatccag aaaaccctaaa tgagaaggaa    360
      gtaacaatca gagctgaaga ttgtggctgt attcctgaag aagtgattat tgctaaagag    420
      ttgatagaaa ttaac                                                    435
    
```

<210>3
 <211>204
 <212> PRT
 <213> Streptomyces coelicolor

10

<400> 3

ES 2 383 592 T3

Met Ile Thr Thr Ser Leu Arg Arg Arg Thr Ala Ala Ala Val Leu Ser
 1 5 10 15

Leu Ser Ala Val Leu Ala Thr Thr Ala Ala Thr Ala Pro Gly Ala Ala
 20 25 30

Pro Ala Pro Ser Ala Ala Pro Ala Lys Ala Ala Pro Ala Cys Pro Gln
 35 40 45

Phe Asp Asp Arg Thr Lys Ala Ala Ala Asp Arg Gly Val Asp Val Asp
 50 55 60

Arg Ile Thr Pro Glu Pro Val Trp Arg Thr Thr Cys Gly Thr Leu Tyr
 65 70 75 80

Arg Ser Asp Ser Arg Gly Pro Gln Val Val Phe Glu Glu Gly Phe His
 85 90 95

Ala Lys Asp Val Gln Asn Gly Gln Tyr Asp Val Glu Lys Tyr Val Leu
 100 105 110

Val Asn Gln Pro Ser Pro Tyr Val Ser Thr Ser Tyr Asp His Asp Leu
 115 120 125

Tyr Lys Thr Trp Tyr Lys Ser Gly Tyr Asn Tyr Tyr Val Asp Ala Pro
 130 135 140

Gly Gly Ile Asp Val Asn Lys Thr Ile Gly Asp Thr His Lys Trp Ala
 145 150 155 160

Asp Gln Val Glu Val Ala Phe Pro Gly Gly Ile Gln Arg Lys Tyr Ile
 165 170 175

Ile Gly Val Cys Pro Val Asp Arg Gln Thr Lys Thr Glu Ile Met Ser
 180 185 190

Asp Cys Glu Ser Asn Pro His Tyr Gln Pro Trp His
 195 200

<210>4

<211> 591

<212> PRT

5 <213> Mycoplasma pneumoniae

<400> 4

Met Pro Asn Pro Val Arg Phe Val Tyr Arg Val Asp Leu Arg Ser Pro
 1 5 10 15

Glu Glu Ile Phe Glu His Gly Phe Ser Thr Leu Gly Asp Val Arg Asn
 20 25 30

ES 2 383 592 T3

Phe Phe Glu His Ile Leu Ser Thr Asn Phe Gly Arg Ser Tyr Phe Ile
 .. 35 40 45
 Ser Thr Ser Glu Thr Pro Thr Ala Ala Ile Arg Phe Phe Gly Ser Trp
 50 55 60
 Leu Arg Glu Tyr Val Pro Glu His Pro Arg Arg Ala Tyr Leu Tyr Glu
 65 70 75 80
 Ile Arg Ala Asp Gln His Phe Tyr Asn Ala Arg Ala Thr Gly Glu Asn
 85 90 95
 Leu Leu Asp Leu Met Arg Gln Arg Gln Val Val Phe Asp Ser Gly Asp
 100 105 110
 Arg Glu Met Ala Gln Met Gly Ile Arg Ala Leu Arg Thr Ser Phe Ala
 115 120 125
 Tyr Gln Arg Glu Trp Phe Thr Asp Gly Pro Ile Ala Ala Ala Asn Val
 130 135 140
 Arg Ser Ala Trp Leu Val Asp Ala Val Pro Val Glu Pro Gly His Ala
 145 150 155 160
 His His Pro Ala Gly Arg Val Val Glu Thr Thr Arg Ile Asn Glu Pro
 165 170 175
 Glu Met His Asn Pro His Tyr Gln Glu Leu Gln Thr Gln Ala Asn Asp
 180 185 190
 Gln Pro Trp Leu Pro Thr Pro Gly Ile Ala Thr Pro Val His Leu Ser
 195 200 205
 Ile Pro Gln Ala Ala Ser Val Ala Asp Val Ser Glu Gly Thr Ser Ala
 210 215 220
 Ser Leu Ser Phe Ala Cys Pro Asp Trp Ser Pro Pro Ser Ser Asn Gly
 225 230 235 240
 Glu Asn Pro Leu Asp Lys Cys Ile Ala Glu Lys Ile Asp Asn Tyr Asn
 245 250 255
 Leu Gln Ser Leu Pro Gln Tyr Ala Ser Ser Val Lys Glu Leu Glu Asp
 260 265 270
 Thr Pro Val Tyr Leu Arg Gly Ile Lys Thr Gln Lys Thr Phe Met Leu
 275 280 285
 Gln Ala Asp Pro Gln Asn Asn Asn Val Phe Leu Val Glu Val Asn Pro
 290 295 300
 Lys Gln Lys Ser Ser Phe Pro Gln Thr Ile Phe Phe Trp Asp Val Tyr
 305 310 315 320
 Gln Arg Ile Cys Leu Lys Asp Leu Thr Gly Ala Gln Ile Ser Leu Ser
 325 330 335
 Leu Thr Ala Phe Thr Thr Gln Tyr Ala Gly Gln Leu Lys Val His Leu
 340 345 350

ES 2 383 592 T3

Ser Val Ser Ala Val Asn Ala Val Asn Gln Lys Trp Lys Met Thr Pro
 355 360 365

Gln Asp Ile Ala Ile Thr Gln Phe Arg Val Ser Ser Glu Leu Leu Gly
 370 375 380

Gln Thr Glu Asn Gly Leu Phe Trp Asn Thr Lys Ser Gly Gly Ser Gln
 385 390 395 400

His Asp Leu Tyr Val Cys Pro Leu Lys Asn Pro Pro Ser Asp Leu Glu
 405 410 415

Glu Leu Gln Ile Ile Val Asp Glu Cys Thr Thr His Ala Gln Phe Val
 420 425 430

Thr Met Arg Ala Ala Ser Thr Phe Phe Val Asp Val Gln Leu Gly Trp
 435 440 445

Tyr Trp Arg Gly Tyr Tyr Tyr Thr Pro Gln Leu Ser Gly Trp Ser Tyr
 450 455 460

Gln Met Lys Thr Pro Asp Gly Gln Ile Phe Tyr Asp Leu Lys Thr Ser
 465 470 475 480

Lys Ile Phe Phe Val Gln Asp Asn Gln Asn Val Phe Phe Leu His Asn
 485 490 495

Lys Leu Asn Lys Gln Thr Gly Tyr Ser Trp Asp Trp Val Glu Trp Leu
 500 505 510

Lys His Asp Met Asn Glu Asp Lys Asp Glu Asn Phe Lys Trp Tyr Phe
 515 520 525

Ser Arg Asp Asp Leu Thr Ile Pro Ser Val Glu Gly Leu Asn Phe Arg
 530 535 540

His Ile Arg Cys Tyr Ala Asp Asn Gln Gln Leu Lys Val Ile Ile Ser
 545 550 555 560

Gly Ser Arg Trp Gly Gly Trp Tyr Ser Thr Tyr Asp Lys Val Glu Ser
 565 570 575

Asn Val Glu Asp Lys Ile Leu Val Lys Asp Gly Phe Asp Arg Phe
 580 585 590

<210>5

<211>242

<212> PRT

5 <213> Salmonella typhimurium

<400> 5

ES 2 383 592 T3

Met Lys Lys Leu Ile Phe Leu Thr Leu Ser Ile Val Ser Phe Asn Asn
 1 5 10 15
 Tyr Ala Val Asp Phe Val Tyr Arg Val Asp Ser Thr Pro Pro Asp Val
 20 25 30
 Ile Phe Arg Asp Gly Phe Ser Leu Leu Gly Tyr Asn Arg Asn Phe Gln
 35 40 45
 Gln Phe Ile Ser Gly Arg Ser Cys Ser Gly Gly Ser Ser Asp Ser Arg
 50 55 60
 Tyr Ile Ala Thr Thr Ser Ser Val Asn Gln Thr Tyr Ala Ile Ala Arg
 65 70 75 80
 Ala Tyr Tyr Ser Arg Ser Thr Phe Lys Gly Asn Leu Tyr Arg Tyr Gln
 85 90 95
 Ile Arg Ala Asp Asn Asn Phe Tyr Ser Leu Leu Pro Ser Ile Thr Tyr
 100 105 110
 Leu Glu Thr Gln Gly Gly His Phe Asn Ala Tyr Glu Lys Thr Met Met
 115 120 125
 Arg Leu Gln Arg Glu Tyr Val Ser Thr Leu Ser Ile Leu Pro Glu Asn
 130 135 140
 Ile Gln Lys Ala Val Ala Leu Val Tyr Asp Ser Ala Thr Gly Leu Val
 145 150 155 160
 Lys Asp Gly Val Ser Thr Met Asn Ala Ser Tyr Leu Gly Leu Ser Thr
 165 170 175
 Thr Ser Asn Pro Gly Val Ile Pro Phe Leu Pro Glu Pro Gln Thr Tyr
 180 185 190
 Thr Gln Gln Arg Ile Xaa Ala Phe Gly Pro Leu Ile Ser Ser Cys Phe
 195 200 205
 Ser Ile Gly Ser Val Cys His Ser His Arg Gly Gln Arg Ala Asp Val
 210 215 220
 Tyr Asn Met Ser Phe Tyr Asp Ala Arg Pro Val Ile Glu Leu Ile Leu
 225 230 235 240
 Ser Lys

<210>6

<211>242

<212> PRT

5 <213> Salmonella paratyphi

<400> 6

ES 2 383 592 T3

Met Lys Lys Leu Ile Phe Leu Thr Leu Ser Ile Val Ser Phe Asn Asn
 1 5 10 15

Tyr Ala Val Asp Phe Val Tyr Arg Val Asp Ser Thr Pro Pro Asp Val
 20 25 30

Ile Phe Arg Asp Gly Phe Ser Leu Leu Gly Tyr Asn Arg Asn Phe Gln
 35 40 45

Gln Phe Ile Ser Gly Arg Ser Cys Ser Gly Gly Ser Ser Asp Ser Arg
 50 55 60

Tyr Ile Ala Thr Thr Ser Ser Val Asn Gln Thr Tyr Ala Ile Ala Arg
 65 70 75 80

Ala Tyr Tyr Ser Arg Ser Thr Phe Lys Gly Asn Leu Tyr Arg Tyr Gln
 85 90 95

Ile Arg Ala Asp Asn Asn Phe Tyr Ser Leu Leu Pro Ser Ile Thr Tyr
 100 105 110

Leu Glu Thr Gln Gly Gly His Phe Asn Ala Tyr Glu Lys Thr Met Met
 115 120 125

Arg Leu Gln Arg Glu Tyr Val Ser Thr Leu Ser Ile Leu Pro Glu Asn
 130 135 140

Ile Gln Lys Ala Val Ala Leu Val Tyr Asp Ser Ala Thr Gly Leu Val
 145 150 155 160

Lys Asp Gly Val Ser Thr Met Asn Ala Ser Tyr Leu Gly Leu Ser Thr
 165 170 175

Thr Ser Asn Pro Gly Val Ile Pro Phe Leu Pro Glu Pro Gln Thr Tyr
 180 185 190

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

Ser Ile Gly Ser Val Cys Gln Ser His Arg Gly Gln Arg Ala Asp Val
 210 215 220

Tyr Asn Met Ser Phe Tyr Asp Ala Arg Pro Val Ile Glu Leu Ile Leu
 225 230 235 240

Ser Lys

<210>7

<211>250

<212> PRT

5 <213> Streptococcus pyogenes

<400> 7

ES 2 383 592 T3

Met Leu Lys Lys Arg Tyr Gln Leu Ala Ile Val Leu Leu Leu Ser Cys
 1 5 10 15

Phe Ser Leu Ile Trp Gln Thr Glu Gly Leu Val Glu Leu Phe Val Cys
 20 25 30

Glu His Tyr Glu Arg Ala Val Cys Glu Gly Thr Pro Ala Tyr Phe Thr
 35 40 45

Phe Ser Asp Gln Lys Gly Ala Glu Thr Leu Ile Lys Lys Arg Trp Gly
 50 55 60

Lys Gly Leu Ile Tyr Pro Arg Ala Glu Gln Glu Ala Met Ala Ala Tyr
 65 70 75 80

Thr Cys Gln Gln Ala Gly Pro Ile Asn Thr Ser Leu Asp Lys Ala Lys
 85 90 95

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

Leu Asp Ala Ala Thr His Arg Leu Val Ile Pro Trp Asn Ile Val Val
 115 120 125

Tyr Arg Tyr Val Tyr Glu Thr Phe Leu Arg Asp Ile Gly Val Ser His
 130 135 140

Ala Asp Leu Thr Ser Tyr Tyr Arg Asn His Gln Phe Asp Pro His Ile
 145 150 155 160

Leu Cys Lys Ile Lys Leu Gly Thr Arg Tyr Thr Lys His Ser Phe Met
 165 170 175

Ser Thr Thr Ala Leu Lys Asn Gly Ala Met Thr His Arg Pro Val Glu
 180 185 190

Val Arg Ile Cys Val Lys Lys Gly Ala Lys Ala Ala Phe Val Glu Pro
 195 200 205

Tyr Ser Ala Val Pro Ser Glu Val Glu Leu Leu Phe Pro Arg Gly Cys
 210 215 220

Gln Leu Glu Val Val Gly Ala Tyr Val Ser Gln Asp Gln Lys Lys Leu
 225 230 235 240

His Ile Glu Ala Tyr Phe Lys Gly Ser Leu
 245 250

<210>8

<211>765

<212> ADN

5 <213> Streptococcus pyogenes

<400> 8

ES 2 383 592 T3

```

gtgtcagggg gaactatgct aaaaaagcgc tatcaactgg ctattgtcct tcttcttagc 60
tgtttttagtc tgatctggca aactgagggc ttggtcgagc tttttgtctg tgagcactat 120
gagcgggcgg tttgtgaggg gacgcctgct tattttacct tttcggatca aaagggcgct 180
gagacactga ttaaaaagcg atggggcaag ggtctcatct acccaagggc tgagcaagag 240
gcgatggctg cttatacctg tcagcaggca ggcctatca acaccagcct agacaaagcc 300
aaaggtgagc tcagccaact cacgcctgag ctaaggggatc aggtggccca gctcgatgct 360
gcgactcacc ggctagtcac cccgtggaac attgtagtat accgctatgt atacgagacg 420
tttttgctg atatecgggtg ttcacatgct gatctcacgt cttactaccg taaccatcag 480
tttgaccctc atatcctttg taagatcaag cttggtacac gctacaccaa gcacagtttt 540
atgagcacga cagcettgaa aaacggcgcc atgacccatc gaccgggtgga ggtgcgcatc 600
tgtgtcaaaa aaggggcca ggcagccttt gtcgagcctt attcggctgt gccttcagag 660
gttgagctct tgtttccaag aggctgtcag ctggaggctg ttggagctta cgtgtcacag 720
gaccaaaaaa agctccacat agaagcgtat ttcaagggca gtttg 765

```

<210>9

<211> 137

<212> PRT

5 <213> Salmonella typhimurium

<400> 9

```

Met Tyr Met Ser Lys Tyr Val Pro Val Tyr Thr Leu Leu Ile Leu Ile
1           5           10           15

Tyr Ser Phe Asn Ala Ser Ala Glu Trp Thr Gly Asp Asn Thr Asn Ala
                20           25           30

Tyr Tyr Ser Asp Glu Val Ile Ser Glu Leu His Val Gly Gln Ile Asp
                35           40           45

Thr Ser Pro Tyr Phe Cys Ile Lys Thr Val Lys Ala Asn Gly Ser Gly
50           55           60

Thr Pro Val Val Ala Cys Ala Val Ser Lys Gln Ser Ile Trp Ala Pro
65           70           75           80

Ser Phe Lys Glu Leu Leu Asp Gln Ala Arg Tyr Phe Tyr Ser Thr Gly
                85           90           95

Gln Ser Val Arg Ile His Val Gln Lys Asn Ile Trp Thr Tyr Pro Leu
                100          105          110

Phe Val Asn Thr Phe Ser Ala Asn Ala Leu Val Gly Leu Ser Ser Cys
                115          120          125

Ser Ala Thr Gln Cys Phe Gly Pro Lys
130           135

```

<210> 10

<211>145

<212> PRT

10 <213> Neisseria meningitidis

<400> 10

ES 2 383 592 T3

Met Gly Asn Phe Leu Tyr Lys Gly Ile Ser Cys Gln Gln Asp Glu Gln
 1 5 10 15
 Asn Asn Gly Gln Leu Lys Pro Lys Gly Asn Lys Ala Glu Val Ala Ile
 20 25 30
 Arg Tyr Asp Gly Lys Phe Lys Tyr Asp Gly Lys Ala Thr His Gly Pro
 35 40 45
 Ser Val Lys Asn Ala Val Tyr Ala His Gln Ile Glu Thr Gly Leu Tyr
 50 55 60
 Asp Gly Cys Tyr Ile Ser Thr Thr Thr Asp Lys Glu Ile Ala Lys Lys
 65 70 75 80
 Phe Ala Thr Ser Ser Gly Ile Glu Asn Gly Tyr Ile Tyr Val Leu Asn
 85 90 95
 Arg Asp Leu Phe Gly Gln Tyr Ser Ile Phe Glu Tyr Glu Val Glu His
 100 105 110
 Pro Glu Asn Pro Asn Glu Lys Glu Val Thr Ile Arg Ala Glu Asp Cys
 115 120 125
 Gly Cys Ile Pro Glu Glu Val Ile Ile Ala Lys Glu Leu Ile Glu Ile
 130 135 140
 Asn
 145

<210> 11
 <211>435
 <212> ADN
 5 <213> Neisseria meningitidis
 <400> 11

atgggaaatt tcttatataa aggcattagt tgccaacaag atgagcaaaa taatggacag 60
 ttaaaaccta aaggtaataa agctgaagtt gcaattcgtt atgatggtaa gtttaaataat 120
 gatggtaaag ctacacatgg tccaagtgtg aagaatgcag tttacgcca tcaaattgaa 180

 acaggtctat atgacggatg ttatatatct acgacaacag acaaggaaat tgccaagaaa 240
 tttgcaacaa gttccggcat cgaaaatggc tatatatatg ttttaaatag ggatttgttt 300
 ggtcaatatt ctatTTTTga atatgagggt gaacatccag aaaacccaaa tgagaaggaa 360
 gtaacaatca gagctgaaga ttgtggctgt attcctgaag aagtgattat tgctaaagag 420
 ttgatagaaa ttaac 435

<210> 12
 <211>145
 10 <212> PRT
 <213> Neisseria meningitidis
 <400> 12

ES 2 383 592 T3

Met Gly Asn Phe Leu Tyr Arg Gly Ile Ser Cys Gln Gln Asp Glu Gln
 1 5 10 15
 Asn Asn Gly Gln Leu Lys Pro Lys Gly Asn Lys Ala Glu Val Ala Ile
 20 25 30
 Arg Tyr Asp Gly Lys Phe Lys Tyr Asp Gly Lys Ala Thr His Gly Pro
 35 40 45
 Ser Val Lys Asn Ala Val Tyr Ala His Gln Ile Glu Thr Gly Leu Tyr
 50 55 60
 Asp Gly Cys Tyr Ile Ser Thr Thr Thr Asp Lys Glu Ile Ala Lys Lys
 65 70 75 80
 Phe Ala Thr Ser Ser Gly Ile Glu Asn Gly Tyr Ile Tyr Val Leu Asn
 85 90 95
 Arg Asp Leu Phe Gly Gln Tyr Ser Ile Phe Glu Tyr Gly Val Glu His
 100 105 110
 Pro Glu Asn Pro Asn Glu Lys Glu Val Thr Ile Arg Ala Glu Asp Cys
 115 120 125
 Gly Cys Ile Pro Glu Glu Val Ile Ile Ala Lys Glu Leu Ile Glu Ile
 130 135 140
 Asn
 145

<210> 13
 <211>435
 <212> ADN
 5 <213> Neisseria meningitidis
 <400> 13

atgggaaatt tcttatatag aggcattagt tgccaacaag atgagcaaaa taatggacag 60
 ttaaaaccta aaggtaataa agctgaagtt gcaattcgtt atgatggtaa gtttaaatat 120
 gatggtaaag ctacacatgg tccaagtgtg aagaatgcag tttacgcca tcaaattgaa 180
 acaggtctat atgacggatg ttatatatct acgacaacag acaaggaaat tgccaagaaa 240
 ttgcaacaa gttccggcat cgaaaatggc tatatatatg ttttaaatag ggatttgttt 300
 ggtcaatatt ctatTTTTga atatgggggt gaacatccag aaaacccaaa tgagaaggaa 360
 gtaacaatca gagctgaaga ttgtggctgt attcctgaag aagtgattat tgctaaagag 420
 ttgatagaaa ttaac 435

<210> 14
 <211> 145
 10 <212> PRT
 <213> Neisseria meningitidis
 <400> 14

ES 2 383 592 T3

Met Gly Asn Phe Leu Tyr Arg Gly Ile Ser Cys Gln Gln Asp Glu Gln
 1 5 10 15
 Asn Asn Gly Gln Leu Lys Pro Lys Gly Asn Lys Ala Glu Val Ala Ile
 20 25 30
 Arg Tyr Asp Gly Lys Phe Lys Tyr Asp Gly Lys Ala Thr His Gly Pro
 35 40 45
 Ser Val Lys Asn Ala Val Tyr Ala His Gln Ile Glu Thr Gly Leu Tyr
 50 55 60
 Asp Gly Cys Tyr Ile Ser Thr Thr Thr Asp Lys Glu Ile Ala Lys Lys
 65 70 75 80
 Phe Ala Thr Ser Ser Gly Ile Glu Asn Gly Tyr Ile Tyr Val Leu Asn
 85 90 95
 Arg Asp Leu Phe Gly Gln Tyr Ser Ile Phe Glu Tyr Glu Val Gly His
 100 105 110
 Pro Glu Asn Pro Asn Glu Lys Glu Val Thr Ile Arg Ala Glu Asp Cys
 115 120 125
 Gly Cys Ile Pro Glu Glu Val Ile Ile Ala Lys Glu Leu Ile Glu Ile
 130 135 140
 Asn
 145

<210> 15
 <211>435
 <212> ADN
 5 <213> Neisseria meningitidis
 <400> 15

atgggaaatt tcttatatag aggcattagt tgccaacaag atgagcaaaa taatggacag 60
 ttaaaaccta aaggtaataa agctgaagtt gcaattcggt atgatggtaa gtttaaatat 120
 gatggtaaag ctacacatgg tccaagtgtg aagaatgcag tttacgcca tcaaattgaa 180
 acaggtctat atgacggatg ttatatatct acgacaacag acaaggaaat tgccaagaaa 240
 tttgcaacaa gttccggcat cgaaaatggc tatatatatg ttttaaatag ggatttgttt 300
 ggtcaatatt ctatttttga atatgaggtt ggacatccag aaaacccaaa tgagaaggaa 360
 gtaacaatca gagctgaaga ttgtggctgt attcctgaag aagtgattat tgctaaagag 420
 ttgatagaaa ttaac 435

<210> 16
 <211>145
 10 <212> PRT
 <213> Neisseria meningitidis
 <400> 16

ES 2 383 592 T3

Met Gly Asn Phe Leu Tyr Arg Gly Ile Ser Cys Gln Gln Asp Glu Gln
 1 5 10 15
 Asn Asn Gly Gln Leu Lys Pro Lys Gly Asn Lys Ala Glu Val Ala Ile
 20 25 30
 Arg Tyr Asp Gly Lys Phe Lys Tyr Asp Gly Lys Ala Thr His Gly Pro
 35 40 45
 Ser Val Lys Asn Ala Val Tyr Ala His Gln Ile Glu Thr Gly Leu Tyr
 50 55 60
 Asp Gly Cys Tyr Ile Ser Thr Thr Thr Asp Lys Glu Ile Ala Lys Lys
 65 70 75 80
 Phe Ala Thr Ser Ser Gly Ile Glu Asn Gly Tyr Ile Tyr Val Leu Asn
 85 90 95
 Arg Asp Leu Phe Gly Gln Tyr Ser Ile Phe Glu Tyr Glu Val Glu His
 100 105 110
 Pro Glu Asn Pro Asn Glu Lys Gly Val Thr Ile Arg Ala Glu Asp Cys
 115 120 125
 Gly Cys Ile Pro Glu Glu Val Ile Ile Ala Lys Glu Leu Ile Glu Ile
 130 135 140
 Asn
 145

<210> 17
 <211>435
 <212> ADN
 5 <213> Neisseria meningitidis
 <400> 17

atgggaaatt tcttatatag aggcattagt tgccaacaag atgagcaaaa taatggacag 60
 ttaaaaccta aaggtaataa agctgaagtt gcaattcgtt atgatggtaa gtttaaatat 120
 gatggtaaag ctacacatgg tccaagtgtg aagaatgcag tttacgcca tcaaattgaa 180
 acaggtctat atgacggatg ttatatatct acgacaacag acaaggaaat tgccaagaaa 240
 tttgcaacaa gttccggcat cgaaaatggc tatatatatg ttttaaatag ggatttgttt 300
 ggtcaatatt ctatTTTTga atatgaggtt gaacatccag aaaacccaaa tgagaagggg 360
 gtaacaatca gagctgaaga ttgtggctgt attcctgaag aagtgattat tgctaaagag 420
 ttgatagaaa ttaac 435

<210> 18
 <211> 45
 10 <212> ADN
 <213> PCR cebador

<220> 37
 <223> PCR cebador

<400> 18
 15 cgcgatccc atatgggaaa tttctatat aaaggcatta gttgc 45

<210> 19
 <211> 37
 <212> ADN

ES 2 383 592 T3

<213> PCR cebador
 <400> 19
 cccgctcgag gttaatttct atcaactctt tagcaat 37

 5 <210> 20
 <211> 45
 <212> ADN
 <213> PCR cebador
 <400> 20
 cgcgatccc atatgggaaa ttcttatat aaaggcatta gttgc 45

 10 <210> 21
 <211> 33
 <212> ADN
 <213> PCR cebador
 <400> 21
 attttgaat atggggtga acatccagaa aac 33
 15
 <210> 22
 <211> 33
 <212> ADN
 <213> PCR cebador
 20 <400> 22
 ttctgatgt tcaaccccat atfcaaaaat aga 33
 <210> 23
 <211> 27
 <212> ADN
 25 <213> PCR cebador
 <400> 23
 tatgaggtg gacatccaga aaaccca 27
 <210> 24
 <211> 27
 30 <212> ADN
 <213> PCR cebador
 <400> 24
 gtttctgga tgtccaacct catattc 27
 35 <210> 25
 <211> 28
 <212> ADN
 <213> PCR cebador
 <400> 25
 ccaaatgaga agggagtaac aatcagag 28
 40 <210> 26
 <211> 27
 <212> ADN
 <213> PCR cebador
 <400> 26
 45 gtttctgga tgtccaacct catattc 27

REIVINDICACIONES

1. Una proteína que consiste en una secuencia de aminoácidos de SEC ID N°: 1, excepto que la secuencia de aminoácidos contiene entre una y veinte mutaciones, en la que cada mutación implica un único aminoácido y en la que la mutación o mutaciones reducen o eliminan la actividad ADP-ribosilante de la proteína.
- 5 2. La proteína de la reivindicación 1, en la que cada mutación es independientemente una sustitución, una inserción o una deleción.
3. La proteína de la reivindicación 2, en la que cada mutación es una sustitución o una deleción.
4. La proteína de una cualquiera de las reivindicaciones 1 a 3, en la que la proteína comprende una o más de las mutaciones enumeradas en la Tabla 1.
- 10 5. La proteína de la reivindicación 4, que consiste en SEC ID N°: 10, 12, 14 ó 16.
6. Una proteína que comprende la secuencia de aminoácidos de una proteína según una cualquiera de las reivindicaciones precedentes.
7. La proteína de una cualquiera de las reivindicaciones precedentes para su uso como un adyuvante.
8. La proteína de la reivindicación 7 para su uso como un adyuvante de la mucosa.
- 15 9. Una composición inmunogénica que comprende la proteína de una cualquiera de las reivindicaciones 1 a 6 en mezcla con un antígeno.
10. Anticuerpo que se une a la proteína de la reivindicación 3.
11. Ácido nucleico que codifica la proteína de una cualquiera de las reivindicaciones 1 a 6.
12. Un vector que comprende una secuencia de ácidos nucleicos del ácido nucleico de la reivindicación 11.
- 20 13. Una célula huésped transformada con el vector de la reivindicación 12.
14. Una composición que comprende la proteína de una cualquiera de las reivindicaciones 1 a 6, el anticuerpo de la reivindicación 10 y/o el ácido nucleico de la reivindicación 11.
15. Una composición que comprende la proteína de una cualquiera de las reivindicaciones 1 a 6 y/o el ácido nucleico de la reivindicación 11 para su uso como un adyuvante.
- 25 16. Uso de la composición de la reivindicación 15 en la preparación de un medicamento para tratar o prevenir infección bacteriana.
17. El uso de la reivindicación 16, en el que la infección bacteriana es producida por *Neisseria meningitidis*.
18. Una proteína que comprende una secuencia de aminoácidos de SEC ID N°: 1 para su uso como un adyuvante.
19. Un anticuerpo que se une a la proteína de SEC ID N°: 1 para su uso como un medicamento.
- 30 20. Una composición que comprende la proteína de la reivindicación 18 y/o el anticuerpo de la reivindicación 19 para su uso como un adyuvante.
21. Una composición inmunogénica que comprende una proteína que comprende una secuencia de aminoácidos de SEC ID N°: 1 en mezcla con un antígeno.
- 35 22. Ácido nucleico que comprende una secuencia de nucleótidos seleccionada de SEC ID 11, 13, 15, 17, 18, 20, 21, 22, 23, 24, 25 y 26.

FIGURA 2

<i>E. coli</i> LT	KLYRVDSRPPD	LYDHARGTQ	YDDG W STSL S RS A	SPHP E Q E V S ALGGI
<i>N. meningitidis</i>	FLYRGI S COOD	VYAHQIETG	YDGCYI S T T DK E IA	PENP N E K E V T T RAED
<i>S. coelicolor</i>	FLYRSDSRGPQ	VVFEEGFHAKDVQNGQYDVEKYVLVNQPSFYVSTSYDHDLY		HKWADQ V E V A F PPGGI
<i>M. pneumoniae</i>	FVYRVD S RSPE	FFEHLSTN	GRSY F I S T S E T P T AA	T S FAY Q R E W T TDGPI
<i>S. typhi</i>	FVYRVD S T P PD		SCSGGSDSR V YAT T	T M R R U Q R E Y V S T L S I
<i>S. paratyphi</i>	FVYRVD S T P PD		SCSGGSDSR V YAT T	T M R R U Q R E Y V S T L S I
<i>S. pyogenes</i>	VYR V Y E T F L		TKHSEM S T T AK K NG A	SAV P S E V E L L PRGC

FIGURA 3B

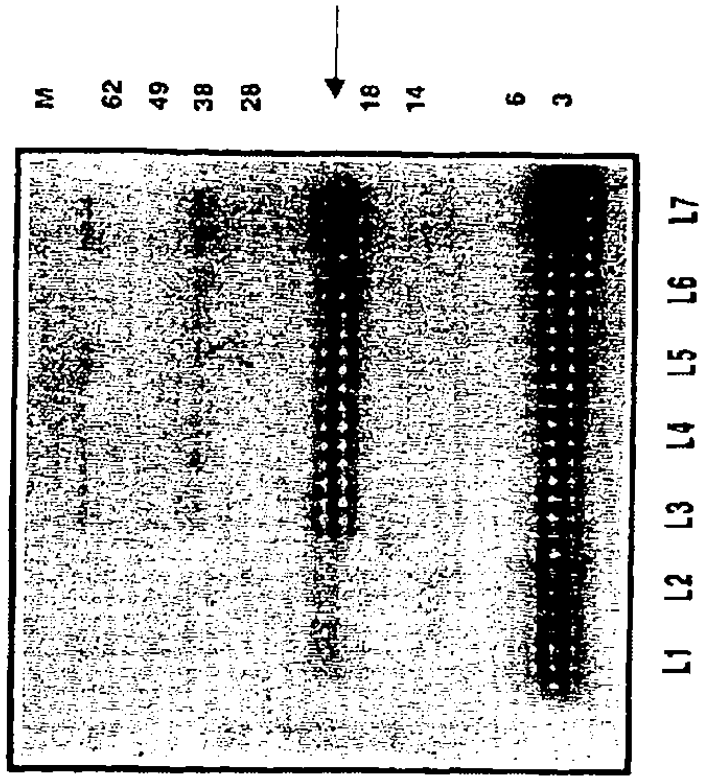


FIGURA 3A

