



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 333 071**

51 Int. Cl.:

**C12N 15/31** (2006.01) **C07K 14/22** (2006.01)

**A61K 39/095** (2006.01) **G01N 33/53** (2006.01)

**C12Q 1/68** (2006.01) **C07K 16/12** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **99900583 .8**

96 Fecha de presentación : **14.01.1999**

97 Número de publicación de la solicitud: **1047784**

97 Fecha de publicación de la solicitud: **02.11.2000**

54 Título: **Antígenos de *Neisseria meningitidis*.**

30 Prioridad: **14.01.1998 GB 9800760**  
**01.09.1998 GB 9819015**  
**09.10.1998 GB 9822143**

45 Fecha de publicación de la mención BOPI:  
**16.02.2010**

45 Fecha de la publicación del folleto de la patente:  
**16.02.2010**

73 Titular/es:  
**Novartis Vaccines and Diagnostics S.R.L.**  
**Via Fiorentina 1**  
**53100 Siena, SI, IT**

72 Inventor/es: **Massignani, Vega;**  
**Rappuoli, Rino;**  
**Pizza, Mariagrazia;**  
**Scarlato, Vincenzo y**  
**Grandi, Guido**

74 Agente: **Carpintero López, Mario**

ES 2 333 071 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

Antígenos de *Neisseria meningitidis*.

5 La presente invención se refiere a antígenos de *Neisseria meningitidis*.

## Antecedentes

10 *Neisseria meningitidis* es un diplococo inmóvil, gram negativo patógeno del ser humano. Coloniza la faringe, causando meningitis y ocasionalmente septicemia en ausencia de meningitis. Está estrechamente relacionado con *N. gonorrhoeae*, aunque una característica que diferencia claramente a los meningococos de los gonococos es la presencia de una cápsula de polisacáridos que está presente en todos los meningococos patógenos.

15 *N. meningitidis* causa enfermedad endémica y epidémica. En los Estados Unidos el índice de infección es de 0,6-1 por cada 100.000 personas al año, y puede ser mucho mayor durante los brotes (véase Lieberman y col., (1996) Safety and Immunogenicity of a Serogroups A/C *Neisseria meningitidis* Oligosaccharide-Protein Conjugate Vaccine in Young Children. JAMA 275(19): 1499-1503; Schuchat y col., (1997) Bacterial Meningitis in the United States in 1995. N Engl J Med 337(14): 970-976). En los países en desarrollo, los índices de enfermedad endémica son mucho mayores y durante los casos de epidemia los índices pueden alcanzar hasta 500 casos por cada 100.000 personas al año. La mortalidad es extremadamente alta, del 10-20% en los Estados Unidos y mucho mayor en los países en desarrollo. Tras la introducción de la vacuna conjugada contra *Haemophilus influenzae*, *N. meningitidis* es la causa principal de meningitis bacteriana en todas las edades en los Estados Unidos (Schuchat y col. (1997) *supra*).

25 Basándose en el polisacárido capsular del organismo se han identificado 12 serogrupos de *N. meningitidis*. El grupo A es el patógeno implicado más a menudo en la enfermedad epidémica en el África Subsahariana. Los serogrupos B y C son responsables de la amplia mayoría de casos en los Estados Unidos y en la mayoría de los países desarrollados. Los serogrupos W135 e Y son responsables del resto de casos de los Estados Unidos y de los países desarrollados. La vacuna contra meningococos usada actualmente es una vacuna de polisacárido tetravalente constituida por serogrupos A, C, Y y W135. Aunque es eficaz en adolescentes y en adultos, induce una respuesta inmune pobre y una corta duración de la protección, y no puede usarse en niños [por ejemplo, Morbidity and Mortality weekly report, Vol. 46, N° RR-5 (1997)]. Esto se debe a que los polisacáridos son antígenos independientes de las células T que inducen una respuesta inmune débil que no puede reforzarse mediante la inmunización repetida. Después del éxito de la vacunación contra *H. influenzae*, se han desarrollado vacunas conjugadas contra serogrupos A y C y están en la etapa final del ensayo clínico (Zollinger WD "New and Improved Vaccines Against Meningococcal Disease" en: New Generation Vaccines, *supra*, págs. 469-488; Lieberman y col. (1996) *supra*; Costantino y col. (1992) Development and phase I clinical testing of a conjugate vaccine against meningococcus A and C. Vaccine 10: 691-698).

35 Sin embargo, el meningococo B sigue siendo un problema. Este serotipo es responsable actualmente de aproximadamente el 50% de todas las meningitis en los Estados Unidos, Europa, y Sudamérica. El enfoque de polisacáridos no puede usarse debido a que el polisacárido capsular de menB es un polímero de ácido N-acetil neuramínico con enlace  $\alpha(2-8)$  que también está presente en el tejido de los mamíferos. Esto da por resultado en la tolerancia al antígeno; de hecho, si se provocara una respuesta inmune, sería autoinmune, y por consiguiente no deseable. Para evitar la inducción de autoinmunidad e inducir una respuesta inmune protectora, por ejemplo, se ha modificado químicamente el polisacárido capsular sustituyendo los grupos N-acetilo con grupos N-propionilo, dejando inalterada la antigenicidad específica (Romero & Outschoorn (1994) Current status of Meningococcal group B vaccine candidates: capsular or non-capsular? Clin Microbiol Rev 7(4): 559-575).

45 Los enfoques alternativos para vacunas contra menB han usado mezclas complejas de proteínas de la membrana externa (OMP), que contenían o las OMP en solitario u OMP enriquecidas con porinas o con una delección en las OMP de clase 4 que se cree inducen anticuerpos que bloquean la actividad bactericida. Este enfoque produce vacunas que no están bien caracterizadas. Éstas son capaces de proteger contra la cepa homóloga, pero no son eficaces en general, ya que existen muchas variantes antigénicas de las proteínas de la membrana externa. Para superar la variabilidad antigénica, se han construido vacunas multivalentes que contienen hasta nueve porinas diferentes (por ejemplo, Poolman JT (1992) Development of a meningococcal vaccine. Infect. Agents Dis, 4: 13-28). Las proteínas adicionales a usarse en vacunas de la membrana externa han sido las proteínas opa y opc, pero ninguno de estos enfoques ha sido capaz de superar la variabilidad antigénica (por ejemplo, Ala'Aldeen & Borriello (1996) The meningococcal transferrin-binding proteins 1 and 2 are both surface exposed and generate bactericidal antibodies capable of killing homologous and heterologous strains. Vaccine 14(1): 49-53).

50 Para los genes y proteínas de los gonococos y meningococos está disponible una cierta cantidad de datos de secuencias (por ejemplo documentos EP-A-0467714, WO 96/29412), pero no están de ninguna manera completos. El suministro de secuencias adicionales podría proporcionar una oportunidad para identificar proteínas secretadas o expuestas en la superficie que sean presuntas dianas para el sistema inmune y que no sean antigénicamente variables. Por ejemplo, algunas de las proteínas identificadas podrían ser componentes de vacunas eficaces contra meningococos B, algunas podrían ser componentes de vacunas contra todos los serotipos de meningococos y otras podrían ser componentes de vacunas contra todos los serotipos de meningococos, y otros podrían ser componentes de vacunas contra todas las *Neisseriae* patogénicas.

El documento WO 99/3113 desvela un polipéptido de superficie de *Neisseria meningitidis* llamado "HiaNm". El polipéptido tiene la siguiente secuencia de aminoácidos (SEC. ID N°: 7 en el presente documento):

5 MNKIYRI IWN~~S~~ALN~~A~~WV~~V~~SELTRNHTKRASATVKTAVLATLLFATVQASANNERPR  
 KKDL~~Y~~LD~~P~~VQRTVAVLIVNSDKEGTGEKEKVEENS~~D~~WAVYFNEKGVLTAREITLKAG  
 DNLKIKQNGTINFTYSLK~~N~~DLTDLT~~S~~VGTEKLSFSANGN~~K~~VNITS~~D~~TKGLNFAKETAG  
 10 TNGD~~T~~TVHLNGIGSTLTOTLLNTGATTNVTND~~N~~VT~~D~~DEK~~K~~RAASVK~~D~~VLNAGWNI~~R~~G  
 VKPGT~~I~~AS~~D~~NVDFV~~R~~TYDTVEFLSADTKTTTVN~~V~~ESK~~D~~NGK~~K~~TEV~~K~~IGV~~K~~TSV~~I~~KEK  
 DGK~~L~~VTGK~~D~~KGENGSSTDEGEGLVIAKEVIDAVN~~K~~AGWRMKT~~T~~TANGQTGQAD~~K~~FET  
 VTS~~G~~TN~~V~~T~~F~~ASGK~~G~~TTATVSKDDQGNITVMYD~~V~~NVGDALN~~V~~NQLQNSGWN~~L~~DSKAVA  
 15 GSSG~~K~~VISGN~~V~~SPSKG~~M~~DETVNINAGNNIEITRNG~~K~~NI~~D~~IATSM~~T~~PQ~~F~~SSVSLGAG  
 ADAP~~T~~LSVDGDALN~~V~~GSK~~R~~DNK~~P~~VRI~~T~~INVA~~P~~GVKEGD~~V~~TNVAQLK~~G~~VAQ~~N~~LNNR~~I~~DN  
 VDG~~N~~ARAGIAQALATAGL~~V~~QAYLPGK~~S~~MM~~A~~IGGGTYRGEAGYAI~~G~~YSSISDGGN~~W~~II  
 KGTASGNSRGHFGASASVGYQH

20 Cierta contenido relacionado con HiaNm se da a conocer en el presente documento.

El Ejemplo 2 del documento WO 99/31132 informa que HiaNm es homólogo a *hia* y *hsf d* *Haemophilus influenzae*, como se dio a conocer en el documento WO 96/30519.

### 25 La invención

La invención proporciona proteínas que comprenden las secuencias de aminoácidos de *N. meningitidis* descritas en el ejemplo.

30 También proporciona proteínas que comprenden secuencias homólogas (es decir que tienen identidad de secuencias) a las secuencias de aminoácidos de *N. meningitidis* dadas a conocer en los ejemplos. El grado de identidad de secuencia es del 80% o más. Estas proteínas homólogas incluyen mutantes y variantes alélicas de las secuencias descritas en el ejemplo. Típicamente, se considera que 50% de identidad o más entre dos proteínas es una indicación de equivalencia funcional. La identidad entre proteínas se determina de preferencia mediante el algoritmo de búsqueda de homología de Smith-Waterman como está implementado en el programa MPSRCH (Oxford Molecular), usando una búsqueda de huecos afines con parámetros de penalización de hueco abierto = 12 y penalización de extensión de hueco = 1.

40 La invención proporciona adicionalmente proteínas que comprenden fragmentos de secuencias de aminoácidos de *N. meningitidis* descritos en el ejemplo. Los fragmentos deben comprender al menos 20 aminoácidos consecutivos de las secuencias y comprender un epítipo de la secuencia.

45 Las proteínas de la invención pueden, por supuesto, prepararse mediante diversos medios (por ejemplo expresión recombinante, purificación a partir de cultivo celular, síntesis química, etc.) y en diversas formas (por ejemplo nativa, condensaciones, etc.). Se preparan de preferencia en forma sustancialmente pura o aislada (es decir sustancialmente libre de otras *N. meningitidis* o proteínas celulares del huésped).

50 Según otro aspecto, la invención proporciona anticuerpos que se unen a estas proteínas. Éstos pueden ser policlonales o monoclonales y pueden producirse mediante cualquier medio adecuado.

Según otro aspecto, la invención proporciona ácidos nucleicos que codifican las proteínas y fragmentos de proteína de la invención. Según otro aspecto, la invención proporciona ácidos nucleicos que comprenden las secuencias de nucleótidos *N. meningitidis* dadas a conocer en el ejemplo.

55 Estos comprenderán al menos n nucleótidos consecutivos de secuencias de *N. meningitidis*, donde n es 10 ó más (por ejemplo 12, 14, 15, 18, 20, 25, 30, 35, 40 ó más).

60 Los ácido nucleicos pueden comprender secuencias complementarias a las descritas anteriormente (por ejemplo, para fines de antisentido o de sondeo).

65 Los ácidos nucleicos de acuerdo con la invención pueden, por supuesto, prepararse de muchas maneras (por ejemplo, mediante síntesis química, a partir de bibliotecas genómicas o de ADNc, a partir del propio organismo, etc.) y pueden tomar diversas formas (por ejemplo, de cadena sencilla, de doble cadena, vectores, sondas, etc.).

Además, la expresión "ácido nucleico" incluye ADN y ARN y también sus análogos, tales como los que contienen estructuras modificadas, y también ácidos nucleicos de péptidos (APN) etc.

## ES 2 333 071 T3

Según otro aspecto, la invención proporciona vectores que comprenden secuencias de nucleótidos de la invención (por ejemplo, vectores 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 de acuerdo con la invención. Estas composiciones pueden ser adecuadas como vacunas, por ejemplo, o como reactivos de diagnóstico o como composiciones inmunogénicas.

10 La invención también proporciona ácido nucleico, proteína o anticuerpo de acuerdo con la invención para usar como medicamentos (por ejemplo, como vacunas) o como reactivos de diagnóstico. También proporciona el uso de un ácido nucleico, proteína o anticuerpo de acuerdo con la invención en la fabricación de: (i) un medicamento para tratar o prevenir la infección por bacterias *Neisseria*, de preferencia *N. meningitidis*, especialmente cepa A, cepa B o cepa C.

Según otros aspectos, la invención proporciona diversos procedimientos.

15 Se proporciona un procedimiento para producir proteínas de la invención, que comprende la etapa de cultivar una célula huésped de acuerdo con la invención en condiciones que inducen la expresión de proteínas.

20 A diferencia de las secuencias dadas a conocer en el documento WO 99/24578, se cree que las secuencias dadas a conocer en la presente solicitud no tienen ningún homólogo significativo en *N. gonorrhoeae*. Por consiguiente, las secuencias de la presente invención también encuentran uso en la preparación de reactivos para distinguir entre *N. meningitidis* y *N. gonorrhoeae*.

25 A continuación se presenta un resumen de técnicas y procedimientos convencionales que pueden utilizarse para llevar a cabo la invención (por ejemplo, para utilizar las secuencias dadas a conocer con fines de vacunación o de diagnóstico). Este resumen no es una limitación de la invención sino, más bien, brinda ejemplos que pueden usarse, pero que no son necesarios.

### General

30 La puesta en práctica de la presente invención utilizará, a menos que se indique de otra manera, técnicas convencionales de biología molecular, microbiología, ADN recombinante e inmunología, que están dentro del conocimiento de la técnica. Tales técnicas se explican completamente en la bibliografía, por ejemplo, Sambrook Molecular Cloning; A Laboratory Manual, segunda edición (1989); 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 (IRI 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 and 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 convencionales para los nucleótidos y los aminoácidos.

### 45 Definiciones

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

50 El término “que comprende” significa “que incluye” así como “que está constituido” por ejemplo una composición “que comprende” X puede estar constituida exclusivamente por X o puede incluir algo además de X, tal como X + Y.

55 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 de forma conjunta, como cuando un promotor heterólogo a un gen se une de forma operativa al gen. Otro ejemplo es cuando una secuencia de *Neisseria* es heteróloga para una célula huésped de ratón. Otros ejemplos podrían ser dos epítopos de la misma o de distintas proteínas que se hayan estructurado en una única proteína en una organización no hallada en la naturaleza.

60 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 polinucleotídica en una célula, capaz de replicación bajo su propio control. Puede ser necesario un origen de replicación para que un vector se replique en una célula huésped particular. Con ciertos orígenes de replicación, un vector de expresión puede reproducirse a un elevado número de copias en presencia de las proteínas apropiadas dentro de la célula. Los ejemplos de orígenes son secuencias que se replican autónomamente, que son eficaces en levadura; y el antígeno T viral, eficaz en células COS-7.

Una secuencia “mutante” se define como una secuencia de ADN, ARN o aminoácidos que es diferente de pero que tiene una homología con la secuencia nativa o la secuencia dada a conocer. Dependiendo de la secuencia particular, el grado de homología (identidad de secuencia) entre la secuencia nativa o la secuencia descrita y la secuencia mutante es de preferencia mayor del 50% (por ejemplo, 60%, 70%, 80%, 90%, 95%, 99% o más, de lo que se calcula usando el algoritmo de Smith-Waterman como se describe anteriormente). Como se usa en el presente documento, una “variante alélica” de una molécula de ácido nucleico, o región de ácido nucleico, para la que se proporciona secuencia de ácido nucleico en el presente documento es una molécula de ácido nucleico, o región de ácido nucleico, que aparece esencialmente el mismo locus en el genoma de otro aislado o de un segundo aislado, y que, debido a la variación natural causada mediante, por ejemplo, mutación o recombinación, tiene una secuencia de ácido nucleico similar pero no idéntica. Una variante alélica de región codificante codifica típicamente una proteína que tiene actividad similar a aquella de la proteína codificada por el gen al que se está comparando. Una variante alélica también puede comprender una alteración en las regiones 5' o 3' no traducidas del gen, tal como en regiones de control reguladoras (véase por ejemplo, la Patente de EEUU N° 5.753.235).

## 15 *Sistemas de Expresión*

Las secuencias de nucleótidos de Neisseria pueden expresarse en una diversidad de sistemas de expresión diferentes; por ejemplo los utilizados con células de mamífero, baculovirus, plantas, bacterias y levaduras.

### 20 i. *Sistemas de Mamíferos*

En la técnica se conocen sistemas de expresión en mamíferos. Un promotor de mamífero es cualquier secuencia de ADN capaz de unir la ARN polimerasa de mamíferos e iniciar la transcripción secuencia abajo (3') de una secuencia codificadora (por ejemplo un gen estructural) en ARNm. Un promotor tendrá una región de iniciación de la transcripción, que se sitúa normalmente en posición proximal al extremo 5' de la secuencia codificadora y una caja TATA, normalmente situada 25-30 pares de bases (pb) secuencia arriba del sitio de iniciación de la transcripción. Se cree que la caja TATA dirige la ARN polimerasa II para que comience la síntesis de ARN en el sitio correcto. Un promotor de mamífero también contendrá un elemento promotor secuencia arriba, situado normalmente en 100 o 200 pb secuencia arriba de la caja TATA. Un elemento promotor secuencia arriba determina el índice con el que se inicia la transcripción y puede actuar en cualquier orientación [Sambrook y col. (1989) “Expression of Clone Genes in Mammalian Cells” In Molecular Cloning: A Laboratory Manual, 2ª ed.].

Los genes virales de mamífero se expresan a menudo en gran cantidad y tienen una amplia variedad de huéspedes; por consiguiente las secuencias que codifican genes virales de mamífero proporcionan secuencias promotoras particularmente útiles. Los ejemplos incluyen el promotor temprano de SV40, el promotor LTR del virus de tumor mamario de ratón, el promotor tardío principal de adenovirus (Ad MLP), y el promotor del virus herpes simplex. Además, las secuencias derivadas de genes no virales, tales como el gen de metalotioneína murina, también proporcionan secuencias promotoras útiles. La expresión puede ser constitutiva o regulada (inducible), dependiendo del promotor puede inducirse mediante glucocorticoides 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 1.000 veces cuando se une a promotores homólogos o heterólogos, con la síntesis comenzando en el sitio de iniciación del ARN normal. Los potenciadores también son activos cuando se sitúan secuencia arriba o secuencia abajo del sitio de iniciación de la transcripción, en orientación bien normal o bien invertida, o a una distancia de más de 1.000 nucleótidos del promotor [Maniatis y col. (1987) Science 236: 1237; Alberts y col. (1989) Molecular Biology of the Cell, 2ª ed.]. Los elementos potenciadores obtenidos a partir de virus pueden ser particularmente útiles, debido a que normalmente tienen un intervalo de huéspedes más amplio.

Los ejemplos incluyen el potenciador del gen temprano de SV40 [Dijkema y col. (1985) EMBO J. 4: 761] y los potenciadores/promotores obtenidos de la repetición terminal larga (LTR) del virus del sarcoma de Rous [Gorman y col. (1982b) Proc. Natl. Acad. Sci. 79: 6777] y del citomegalovirus humano [Boshart y col. (1985) Cell 41: 521]. Además, algunos potenciadores son regulables y llegan a ser activos solamente en presencia de un inductor, tal como una hormona o un ión metálico [Sassone-Corsi y Borelli (1986) Trends Genet. 2: 215; Maniatis y col. (1987) Science 236: 1237].

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

Como alternativa, las proteínas extrañas también pueden secretarse a partir de la célula al medio de cultivo creando moléculas de ADN quiméricas que codifican una proteína de fusión constituida por un fragmento de secuencia líder que hace posible la secreción de la proteína extraña en células de mamífero. De preferencia, existen sitios de procesamiento codificados entre el fragmento líder y el gen extraño que pueden escindirse *in vivo* o *in vitro*. El fragmento de secuencia líder codifica usualmente un péptido señal que comprende aminoácidos hidrófobos que dirigen la secreción de la proteína desde la célula. El líder tripartito de adenovirus es un ejemplo de una secuencia líder que posibilita la secreción de una proteína extraña en células de mamífero.

Usualmente, las secuencias de terminación de la transcripción y de poliadenilación reconocidas por las células de mamífero son regiones reguladoras situadas en posición 3' con respecto al codón de terminación de la traducción y por consiguiente, junto con los elementos promotores, flanquean a la secuencia codificadora. El extremo 3' del ARNm maduro se forma mediante escisión y poliadenilación postranscripcionales específicas de sitio [Birnstiel y col. (1985) Cell 41:349; Proudfoot y 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. Los ejemplos de señales de terminador de la transcripción/poliadenilación incluyen los obtenidos de SV40 [Sambrook y col. (1989) "Expression of cloned genes in cultured mammalian cells." En Molecular Cloning: A Laboratory Manual].

Usualmente, los componentes descritos anteriormente, que comprenden un promotor, señal de poliadenilación y secuencia de terminación de la transcripción se colocan conjuntamente en construcciones de expresión. Si se desea, en una construcción de expresión, también pueden incluirse, potenciadores intrones con sitios donadores y aceptores de corte y empalme funcionales y secuencias líderes. Las construcciones de expresión se mantienen a menudo en un replicón, tal como un elemento extracromosómico (por ejemplo, plásmidos) capaz del mantenimiento estable en un huésped, tal como células de mamífero o bacterias. Los sistemas de replicación en mamíferos incluyen los derivados de virus animales, que requieren factores que actúan en dirección trans para replicarse. Por ejemplo, los plásmidos que contienen el sistema de replicación de papovavirus, tales como SV40 [Gluzman (1981) Cell 23: 175] o polioma virus, se replican a un número de copias extremadamente alto en presencia del antígeno T viral apropiado. Los ejemplos adicionales de replicones de mamífero incluyen los obtenidos de papilomavirus bovino y del virus Epstein-Barr. Adicionalmente, el replicón puede tener dos sistemas de replicación, lo que le permite de esta manera mantenerse, por ejemplo, en células de mamífero para su expresión y en un huésped procariota para clonación y amplificación. Los ejemplos de dichos vectores lanzadera mamífero-bacteria incluyen pMT2 [Kaufman y col. (1989) Mol. Cell. Biol. 9: 946] y pHEBO [Shimizu y col. (1986) Mol. Cell. Biol. 6: 1074].

El procedimiento de transformación usado depende del huésped a transformarse. Los procedimientos para la introducción de los polinucleótidos heterólogos en las células de mamífero se conocen en la técnica e incluyen transfección mediada por dextrano, precipitación por fosfato de calcio, transfección mediada por polibreno, fusión de protoplastos, electroporación, encapsulamiento del/de los polinucleótido(s) en liposomas y microinyección directa del ADN en los 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 American Type Culture Collection (ATCC), incluyendo aunque sin limitarse a, células de ovario de hámster chino (CHO), células HeLa, células de riñón de cría de hámster (BHK), células de riñón de mono (COS), células de carcinoma hepatocelular humano (por ejemplo, Hep G2), y varias líneas celulares diferentes.

## ii. *Sistemas de Baculovirus*

El polinucleótido que codifica la proteína también puede insertarse en un vector de expresión en insectos adecuado, y se une de forma operativa a los elementos de control dentro de este 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, usualmente un plásmido bacteriano, que contiene tanto un fragmento del genoma del baculovirus como un sitio de restricción conveniente para la inserción del gen o genes heterólogos a expresarse; un baculovirus de tipo salvaje con una secuencia homóloga al fragmento específico de baculovirus en el vector de transferencia (esto permite la recombinación homóloga del gen heterólogo en el genoma del baculovirus); y células huésped de insecto y medios de cultivo apropiados.

Después de insertar la secuencia de ADN que codifica la proteína en el vector de transferencia, el vector y el genoma viral de tipo natural se transfectan dentro de una célula huésped de insecto donde se permite que el vector y el genoma viral se recombinen. El virus recombinante encapsulado se expresa y se identifican y se purifican placas recombinantes. Los materiales y los procedimientos para sistemas de expresión de baculovirus/células de insecto están disponibles comercialmente en forma de kit de, entre otros, Invitrogen, San Diego CA (kit "MaxBac"). Estas técnicas se conocen en general por los expertos en la técnica y se describen completamente en el documento de Summers y Smith, Texas Agricultural Experiment Station Bulletin N° 1555 (1987) (en lo sucesivo en el presente documento "Summers y Smith").

Antes de la inserción de la secuencia de ADN que codifica la proteína dentro del genoma del baculovirus, los componentes descritos anteriormente, que comprenden un promotor, una secuencia líder (si se desea), la secuencia codificadora de interés y la secuencia de terminación de la transcripción, se ensamblan usualmente en una construcción de translocación intermedia (vector de transferencia). Esta construcción puede contener un único gen y elementos reguladores unidos de forma operativa; múltiples genes, cada uno con su propio conjunto de elementos reguladores unidos de forma operativa; o múltiples genes, regulados por el mismo conjunto de elementos reguladores. Las construcciones de translocación intermedias se mantienen a menudo en un replicón, tal como un elemento extracromosómico (por ejemplo, plásmidos) capaz de mantenimiento estable en un huésped, tal como una bacteria. El replicón tendrá un sistema de replicación, que le permitirá mantenerse así en un huésped adecuado para la clonación y la amplificación.

## ES 2 333 071 T3

Actualmente, el vector de transferencia usado más comúnmente para introducir genes extraños dentro de AcNPV es pAc373. También se han diseñado, muchos otros vectores, conocidos por los expertos en la técnica. Estos incluyen, por ejemplo pVL985 (que altera el codón de iniciación polihedrina de ATG a ATT, y que introduce un sitio de clonación de BamHI 32 pares de bases secuencia abajo del ATT; véase Luckow y Summers, Virology (1989), 17: 31.

5

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

10

Los vectores de transferencia de baculovirus contienen usualmente un promotor de baculovirus. Un promotor de baculovirus es cualquier secuencia de ADN capaz de unir una ARN polimerasa de baculovirus e iniciar la transcripción secuencia abajo (5' a 3') de una secuencia codificadora (por ejemplo gen estructural) en ARNm. Un promotor tendrá una región de iniciación de la transcripción que se sitúa usualmente en posición proximal con respecto al extremo 5' de la secuencia codificadora. Esta región de iniciación de la transcripción usualmente incluye un sitio de unión a ARN polimerasa y un sitio de iniciación de la transcripción. Un vector de transferencia de baculovirus también puede tener un segundo dominio llamado potenciador, que, si está presente, está normalmente en posición distal con respecto al gen estructural. La expresión puede ser regulada o constitutiva.

15

20

Los genes estructurales, abundantemente transcritos en los últimos momentos en un ciclo de infección vírica, proporcionan secuencias promotoras particularmente útiles. Los ejemplos incluyen secuencias derivadas del gen que codifica la proteína poliédrica viral, Friesen y col., (1986) "The Regulation of Baculovirus Gene Expression" en: "The Molecular Biology of Baculoviruses" (ed. Walter Doerfler): Publicación EPO N° 127 839 y 155 476; y el gen que codifica la proteína p10, Vlak y col., (1988), J. Gen. Virol. 69: 765.

25

El ADN que codifica secuencias señal adecuadas puede derivarse de genes para proteínas de insecto o de baculovirus secretadas, tales como el gen de la polihedrina de baculovirus (Carbonell y col. (1988) Gene, 73: 409). Como alternativa, dado que las señales para las modificaciones post-traduccionales de células de mamífero (tales como la escisión de péptido señal, escisión proteolítica, y fosforilación) parecen reconocerse por las células de insecto, y las señales requeridas para la secreción y la acumulación nuclear también parecen conservarse entre las células de invertebrados y las células de vertebrados, líderes de origen no insecto, tales como los obtenidos de los genes que codifican interferón- $\alpha$  humano, Maeda y col., (1985), Nature 315: 592; péptido de liberación de gastrina humana, Lebacqz-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, pueden usarse también para proporcionar secreción en insectos.

30

35

Un polipéptido o poliproteína recombinante puede expresarse de forma intracelular o, si se expresa con las secuencias reguladoras apropiadas, puede secretarse. La buena expresión intracelular de proteínas extrañas sin condensar usualmente requiere genes heterólogos que idealmente tienen una secuencia líder corta que contiene señales de iniciación de la traducción adecuadas precediendo a una señal de iniciación ATG. Si se desea, la metionina en el extremo N puede escindirse de la proteína madura mediante incubación *in vitro* con bromuro de cianógeno.

40

45

Como alternativa, las poliproteínas o proteínas recombinantes que no se secretan de forma natural pueden secretarse a partir de la célula de insecto creando moléculas de ADN quiméricas que codifican una proteína de fusión constituida por un fragmento de secuencia líder que permite la secreción de las proteínas extrañas en insectos. El fragmento de secuencia líder codifica usualmente un péptido señal constituido por aminoácidos hidrófobos que dirigen la translocación de la proteína al retículo endoplasmático.

50

Después de la inserción de la secuencia de ADN y/o del gen que codifica el precursor del producto de expresión de la proteína, una célula huésped de insecto se cotransforma con el ADN heterólogo del vector de transferencia y el ADN genómico del baculovirus de tipo silvestre -usualmente mediante cotransfección. El promotor y la secuencia de terminación de la transcripción de la construcción estarán constituidos usualmente por una sección de 2-5 kb del genoma del baculovirus. Se conocen en la técnica los procedimientos para introducir ADN heterólogo en el sitio deseado en el virus baculovirus. (Véase Summers y Smith *supra*; Ju y col., (1987); Smith y col., Mol. Cell. Biol. (1983), 3: 2156; y Luckow y Summers (1989)). Por ejemplo, la inserción puede estar en un gen tal como el gen de polihedrina, mediante recombinación de doble entrecruzamiento homóloga; la inserción también puede estar en un sitio de enzima de restricción introducido 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 polihedrina en el vector de expresión, está flanqueada tanto en 5' como en 3' por secuencias específicas de polihedrina y está situado secuencia abajo del promotor de polihedrina.

60

65

El vector de expresión de baculovirus recién formado se empaqueta posteriormente en un baculovirus recombinante infeccioso. La recombinación homóloga se produce a baja frecuencia (entre aproximadamente el 1% y aproximadamente el 5%); por consiguiente, la mayoría del virus producido después de la cotransfección es aún virus de tipo salvaje. Por consiguiente, es necesario un procedimiento para identificar virus recombinantes. Una ventaja del sistema de expresión es una selección visual que permite distinguir virus recombinantes. La proteína polihedrina, que es producida por el virus nativo, se produce a niveles muy altos en los núcleos de las células infectadas en las etapas tardías después de la infección vírica. La proteína polihedrina acumulada forma cuerpos de oclusión que también contienen partículas incluidas. Estos cuerpos de oclusión, de hasta 15  $\mu$ m en tamaño, tienen gran capacidad de refracción, lo que

les proporciona una apariencia brillante que se visualiza fácilmente en el microscopio óptico. Las células infectadas con virus recombinantes carecen de cuerpos de oclusión. Para distinguir virus recombinante de virus de tipo salvaje, se plaquea el sobrenadante de transfección en una monocapa de células de insecto mediante técnicas conocidas por los expertos en la técnica. A saber, las placas se exploran en el microscopio óptico para detectar la presencia (indicativa de virus de tipo salvaje) o ausencia (indicativa de virus recombinantes) de cuerpos de oclusión. “Current Protocols in Microbiology” Vol. 2 (Ausubel y col., eds) a 16.8 (Supl. 10, 1990), Summers y Smith, *supra*, Miller y col., (1989).

Se han desarrollado vectores de expresión de baculovirus recombinantes para la infección de varias células de insecto. Por ejemplo, se han desarrollado baculovirus recombinantes 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 disponibles comercialmente para la expresión directa y de fusión de polipéptidos heterólogos en un sistema de expresión de baculovirus; la tecnología del cultivo celular se conoce generalmente por los expertos en la técnica. Véase por ejemplo Summers y Smith *supra*.

Las células de insecto modificadas pueden cultivarse a continuación en un medio nutriente adecuado, lo que permite el mantenimiento estable del(de los) plásmido(s) presente(s) en el huésped insecto modificado. Donde el gen producido por la expresión está bajo control inducible, el huésped puede cultivarse hasta alta densidad y puede inducirse la expresión. Como alternativa, cuando la expresión es constitutiva, el producto se expresará de forma continua en el medio y el medio nutriente debe circular de forma continua, mientras se retira el producto de interés y se aumentan los nutrientes agotados. El producto puede purificarse mediante técnicas tales como cromatografía, por ejemplo HPLC, cromatografía de afinidad, cromatografía de intercambio iónico, etc.; electroforesis; centrifugación por gradiente de densidad; extracción del disolvente, o similares. Según sea adecuado, el producto puede purificarse además, según necesidad, tal como para eliminar sustancialmente cualesquier proteína de insecto que se secreten en el medio o sean el resultado de la lisis de células de insectos, tal como para proporcionar un producto que esté al menos sustancialmente libre de restos del huésped, por ejemplo, proteínas, lípidos y polisacáridos.

Con el fin de obtener la expresión de las proteínas, se incuban las células huésped recombinantes obtenidas de los transformantes en condiciones que permitan la expresión de la secuencia que codifica la proteína recombinante. Estas condiciones variarán, dependiendo de la célula huésped seleccionada. Sin embargo, las condiciones son fácilmente determinables por los expertos en la técnica, en base a lo que se conoce en la técnica.

### iii. Sistemas de plantas

Hay muchos sistemas de expresión genética de cultivo de células vegetales y de plantas completas conocidos en la técnica. Los sistemas de expresión genética en células vegetales ejemplares incluyen los que se describen en patentes de EEUU N°: 5.693.506; 5.659.122; y 5.608.143. Otros ejemplos de expresión genética en cultivo de células vegetales se han descrito por Zenk, *Phytochemistry* 30: 3861-3863 (1991). Pueden encontrarse descripciones de péptidos señal de proteínas vegetales además de las referencias descritas anteriormente 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); Wirsel y col., *Molecular Microbiology* 3: 3-14 (1989); Yu y col., *Gene* 122: 247-253 (1992). Una descripción de la regulación de expresión de genes vegetales mediante la fitohormona, ácido giberélico y las enzimas secretadas inducidas por el ácido giberélico puede encontrarse en los documentos de R.L. Jones y J. MacMillin, *Gibberellins*: en: *Advanced Plant Physiology*, Malcolm B. Winkins, ed., 1984 Pitman Publishing Limited, London págs. 21-52. Referencias que describen otros genes regulados metabólicamente: Sheen, *Plant Cell*, 2: 1027-1038 (1990); Maas y col., *EMBO J.* 9: 3447-3452 (1990); Benkel y Hickey, *Proc. Natl. Acad. Sci.* 84: 1337-1339 (1987).

Típicamente, usando técnicas conocidas en la técnica, se inserta una secuencia de polinucleótidos deseada dentro de una casete de expresión que comprende elementos reguladores genéticos diseñados para el funcionamiento en plantas. La casete de expresión se inserta en un vector de expresión deseado con secuencias acompañantes secuencia arriba y cadena debajo de la casete de expresión adecuada para la expresión en un huésped vegetal. Las secuencias acompañantes serán de origen viral o plasmídico y proporcionan las características necesarias al vector para permitir que el vector mueva ADN de un huésped de clonación original, tal como una bacteria, al huésped vegetal deseado. La construcción básica bacteria/vector de planta proporcionará de preferencia un origen de replicación procariota de amplia variedad de huéspedes; un marcador seleccionable procariota; y, para transformaciones de *Agrobacterium*, secuencias de ADN T para la transferencia mediada por *Agrobacterium* a cromosomas vegetales. Donde el gen heterólogo no es susceptible de detección fácilmente, la construcción también tendrá de preferencia un gen marcador seleccionable adecuado para determinar si una célula vegetal se ha transformado. Una revisión general de los marcadores adecuados, por ejemplo para los miembros de la familia de las gramíneas, se encuentra en el documento Wilmlink y Dons, 1993, *Plant Mol. Biol. Reptr.* 11 (2): 165-185.

También se recomiendan las secuencias adecuadas para permitir la integración de la secuencia heteróloga en el genoma de la planta. Estas pueden incluir secuencias del transposón y similares para recombinación homóloga así como secuencias Ti que permiten la inserción aleatoria de una casete de expresión heteróloga en un genoma vegetal.

Los marcadores seleccionables procariotas adecuados incluyen resistencia a antibióticos tales como ampicilina o tetraciclina. También pueden estar presentes en el vector, otras secuencias de ADN que codifican funciones adicionales, como se conoce en la técnica.

- 5 Las moléculas de ácido nucleico de la invención sujeto pueden incluirse en una casete de expresión para la expresión de la(s) proteína(s) de interés. Usualmente, solamente habrá una casete de expresión, aunque sean factibles dos o más. La casete de expresión recombinante contendrá además de la secuencia que codifica la proteína heteróloga los siguientes elementos: una región promotora, secuencias sin traducir 5' vegetales, codón de iniciación dependiendo de si el gen estructural viene o no equipado con uno, y una secuencia de terminación de la transcripción y de la traducción.
- 10 Los sitios de restricción enzimática únicos en los extremos 5' y 3' de la casete permiten la fácil inserción en un vector preexistente.

Una secuencia codificadora heteróloga puede ser para cualquier proteína relacionada con la presente invención. La secuencia que codifica a la proteína de interés codificará un péptido señal que permitirá el procesamiento y la translocación de la proteína, según sea apropiado, y carecerá usualmente de cualquier secuencia que pudiera dar como resultado la unión de la proteína deseada de la invención a una membrana. Aunque, para la mayoría, la región de iniciación de la transcripción será la de un gen que se expresa y se transloca durante la germinación, utilizando el péptido señal que posibilita la translocación, también puede proporcionarse la translocación de la proteína de interés. De este modo, la(s) proteína(s) de interés se translocarán a partir de las células en las que se expresan y podrán recogerse de forma eficaz. Típicamente la secreción en las semillas es a través de la aleurona o a través de la capa de epitelio del escutelo dentro del endospermo de la semilla. Aunque no es necesario que la proteína secrete desde las células en las que la proteína se produce, esto facilita el aislamiento y la purificación de la proteína recombinante.

Dado que la expresión definitiva del producto génico deseado estará en una célula eucariota es deseable determinar si cualquier porción del gen clonado contiene secuencias que se procesarán como intrones por la maquinaria de corte y empalme del huésped. Si esto fuera así, puede realizarse mutagénesis dirigida de la región "intrón" para evitar la pérdida de una porción del mensaje genético en forma de falso código intrón, Reed y Maniatis, Cell 41: 95-105, 1985.

El vector puede microinyectarse directamente en células vegetales mediante el uso de micropipetas para transferir mecánicamente el ADN recombinante. Crossway, Mol. Gen. Genet, 202: 179-185, 1985. El material genético también puede transferirse a la célula vegetal usando polietilenoglicol, Krens y col., Nature, 296, 72-74, 1982. Otro procedimiento de introducción de segmentos de ácido nucleico es la penetración balística a alta velocidad mediante pequeñas partículas con el ácido nucleico dentro de la matriz de pequeñas perlas o partículas, o en la superficie, Klein y col., Nature, 327, 70-73, 1987 y Knudsen y Muller, 1991, Planta, 185: 330-336 que muestran el bombardeo de partículas del endospermo de cebada para crear cebada transgénica. Aún otro procedimiento de introducción sería la fusión de protoplastos con otras entidades, minicélulas, células, lisosomas u otros cuerpos con superficie lipídica que pueden condensarse, Fraley, y col., Proc. Natl. Acad. Sci. USA, 79, 1859-1863, 1982.

El vector también puede introducirse dentro de las células vegetales mediante electroporación (Fromm y col., Proc. Natl. Acad. Sci. USA 82: 5824, 1985). En esta técnica, los protoplastos vegetales se someten a electroporación en presencia de plásmidos que contienen la construcción génica. Los impulsos eléctricos de alto potencial iónico permeabilizan de forma reversible las membranas biológicas permitiendo la introducción de los plásmidos. Los protoplastos vegetales electroporados vuelven a formar la pared celular, se dividen y forman callos vegetales.

55 Todas las plantas a partir de las que pueden aislarse y cultivarse protoplastos para proporcionar plantas regeneradas completas pueden transformarse mediante la presente invención de modo que se recuperen plantas completas que contienen el gen transferido. Se sabe que prácticamente todas las plantas pueden regenerarse a partir de células o tejidos cultivados, incluidos pero no limitados a todas las especies principales de caña de azúcar, remolacha azucarera, algodón, fruta y otros árboles, legumbres y verduras. Algunas plantas adecuadas incluyen, por ejemplo, las 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*, *Browaalía*, *Glycine*, *Lolium*, *Zea*, *Triticum*, *Sorghum* y *Datura*.

Los medios de regeneración varían de especie a especie vegetal, pero generalmente se proporciona en primer lugar una suspensión de protoplastos transformados que contienen copias del gen heterólogo. Se forma tejido calloso y pueden inducirse brotes a partir de los callos, que posteriormente se pueden enraizar. Como alternativa, puede inducirse la formación de embriones a partir de la suspensión de protoplastos. Estos embriones germinan como los 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 maíz y alfalfa. Los brotes y las raíces se desarrollan normalmente de forma simultánea. La regeneración eficaz dependerá del medio, del genotipo, y del historial del cultivo. Si se controlan estas tres variables, entonces la regeneración es totalmente reproducible y repetible.

65 En algunos sistemas de cultivo de células vegetales, la proteína deseada de la invención puede excretarse o, como alternativa, la proteína puede extraerse a partir de la planta completa. En los casos en los que la proteína de la invención se secreta al medio, ésta puede recogerse. Como alternativa, los embriones y las semillas sin embrión u otro tejido

vegetal pueden romperse mecánicamente para liberar cualquier proteína secretada entre las células y los tejidos. La mezcla puede suspenderse en una disolución tampón para recuperar las proteínas solubles. Después se utilizarán procedimientos convencionales de aislamiento y purificación de proteínas para purificar la proteína recombinante. Los parámetros de tiempo, temperatura, pH, oxígeno y volumen se ajustarán mediante procedimientos rutinarios para optimizar la expresión y la recuperación de la proteína heteróloga.

#### iv. Sistemas Bacterianos

En la técnica se conocen técnicas de expresión bacteriana. Un promotor bacteriano es cualquier secuencia de ADN capaz de unirse a la ARN polimerasa bacteriana e iniciar la transcripción secuencia abajo (3') de una secuencia codificadora (por ejemplo gen estructural) en ARNm. Un promotor tendrá una región de iniciación de la transcripción que se sitúa normalmente en posición proximal con respecto al extremo 5' de la secuencia codificadora. Esta región de iniciación de la transcripción usualmente incluye un sitio de unión a la ARNm polimerasa y un sitio de iniciación de la transcripción. Un promotor bacteriano puede tener también un segundo dominio llamado un operador, que puede solaparse con un sitio de unión de ARN polimerasa adyacente en el que comienza la síntesis de ARN. El operador permite la transcripción regulada negativamente (inducible), ya que una proteína represora del gen puede unirse al operador y por consiguiente inhibir la transcripción de un gen específico. La expresión constitutiva puede producirse en ausencia de elementos reguladores negativos, tales como el operador. Además, puede conseguirse la regulación positiva mediante una secuencia de unión a la proteína activadora del gen, que, si está presente, está normalmente en posición proximal (5') con respecto a la secuencia de unión a la ARN polimerasa. Un ejemplo de una proteína activadora del gen es la proteína activadora de catabolitos (CAP), que ayuda a iniciar la transcripción del operón *lac* en *Escherichia coli* (*E. coli*) [Raibaud y col., (1984) Annu. Rev. Genet. 18: 173]. Por consiguiente la expresión regulada puede ser positiva o negativa, potenciando o reduciendo de este modo la transcripción.

Las secuencias que codifican enzimas de rutas metabólicas proporcionan secuencias promotoras particularmente útiles. Los ejemplos incluyen secuencias promotoras derivadas de enzimas que metabolizan azúcares, tales como galactosa, lactosa (*lac*) [Chang y col., (1977) Nature 198: 1056] y maltosa. Otros ejemplos incluyen secuencias promotoras derivadas de enzimas biosintéticas tales como triptofano (*trp*) [Goeddel y col., (1980) Nuc. Acids Res 8: 4057; Yelverton y col., (1981) Nucl. Acids Res. 9: 731, Patente de EEUU 4.738.921; Publicación EP-A-036776 y EP-A-0121775]. El sistema promotor de  $\beta$ -lactamasa (*bla*) [Weissmann (1981) "The cloning of interferon and other mistakes", en *Interferon 3* (ed. I. Gresser)], el bacteriófago lambda PL [Shimatake y col., (1981) Nature 292: 128] y los sistemas promotores T5 (Patente de EEUU 4.689.406) también proporcionan secuencias promotoras útiles.

Además, los promotores sintéticos que no se presentan en la naturaleza también funcionan como promotores bacterianos. Por ejemplo, las secuencias de activación de la transcripción de un promotor bacteriano o promotor de bacteriófago pueden unirse con las secuencias de operón de otro promotor bacteriano o promotor de bacteriófago, creando un promotor híbrido sintético [Patente de EEUU 4.551.433]. Por ejemplo, el promotor *tac* es un promotor *trp-lac* híbrido 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) Proc. Natl. Acad. Sci. 80: 21]. Además, un promotor bacteriano puede incluir promotores de origen natural de bacteriano que tengan la capacidad de unirse a ARN polimerasa bacteriana e iniciar la transcripción. Un promotor que se presenta en la naturaleza de origen no bacteriano puede acoplarse también con una ARN polimerasa compatible para producir altos niveles de expresión de algunos genes en procariontes. El sistema de ARN polimerasa/promotor del bacteriófago T7 es un ejemplo de un sistema promotor acoplado [Studier y col., (1986) J. Mol. Biol. 189: 113, Tabor y col., (1985) Proc Natl Acad. Sci. 82: 1074]. Además, un promotor híbrido puede también comprender un promotor de bacteriófago y una región operadora de *E. coli* (publicación EPO-A-0 267 851).

Además de la secuencia promotora en funcionamiento, también es útil un sitio de unión al ribosoma eficaz para la expresión de genes extraños en procariontes. En *E. coli*, el sitio de unión al ribosoma se denomina la secuencia Shine-Dalgarno (SD) e incluye un codón de iniciación (ATG) y una secuencia de 3-9 nucleótidos de longitud situada 3-11 nucleótidos secuencia arriba del codón de iniciación [Shine y col., (1975) Nature 254: 34]. Se cree que la secuencia SD promueve la unión de ARNm al ribosoma emparejando bases entre la secuencia SD y el extremo 3' del ARNr 16S de *E. coli* [Steitz y col., (1979) "Genetic signals and nucleotide sequences in Messenger RNA", en *Biological Regulation and Development: Gene Expression* (ed. R.F. Goldberger)]. Para expresar genes eucariotas y genes procariontes con un sitio de unión débil al ribosoma [Sambrook y col., (1989) "Expression of cloned genes in *Escherichia coli*", en *Molecular Cloning: A Laboratory Manual*].

Una molécula de ADN puede expresarse de manera intracelular. Una secuencia promotora puede unirse de forma directa con la molécula de ADN, caso en el que el primer aminoácido en el extremo N será siempre una metionina, que está codificada por el codón de iniciación ATG. Si se desea, la metionina en el extremo N puede escindirse de la proteína mediante incubación *in vitro* con bromuro de cianógeno o mediante incubación *in vivo* o *in vitro* con una peptidasa N terminal de metionina bacteriana (Publicación EPO-A-0 219 237).

Las proteínas de fusión proporcionan una alternativa a la expresión directa. Usualmente, una secuencia de ADN que codifica la porción N terminal de una proteína bacteriana endógena, u otra proteína estable, se fusiona con el extremo 5' de secuencias codificadoras heterólogas. Después de la expresión, esta construcción proporcionará una fusión de las dos secuencias de aminoácidos. Por ejemplo, el gen de la célula del bacteriófago lambda puede unirse en el extremo 5' de un gen extraño y puede expresarse en bacterias. La proteína de fusión resultante conserva de

preferencia un sitio para una enzima de procesamiento (factor Xa) para escindir la proteína del 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) *J. Biotechnol.* 5: 93; Makoff y col., (1989) *J. Gen. Microbiol.* 135: 11] y Chey [Publicación EP-A-0 324 647]. La secuencia de ADN en la unión de las dos secuencias de aminoácidos puede o no codificar un sitio escindible. Otro ejemplo es una proteína de fusión a ubiquitina. Una proteína de fusión tal se elabora con la región de ubiquitina que conserva de preferencia un sitio para una enzima de procesamiento (por ejemplo una proteasa de procesamiento específica de ubiquitina) para escindir la ubiquitina de la proteína extraña. A través de este procedimiento, puede aislarse una proteína extraña natural [Miller y col. (1989) *Biol Technology* 7: 698].

Como alternativa, también pueden secretarse de la célula proteínas extrañas creando moléculas de ADN quiméricas que codifican una proteína de fusión que comprende un fragmento de secuencia de péptido señal que permite la secreción de la proteína extraña en bacterias [Patente de EEUU 4.336.336]. El fragmento de secuencia señal usualmente codifica un péptido señal que comprende aminoácidos hidrófobos que dirigen la secreción de la proteína de la célula. La proteína se secreta en el medio de cultivo (bacterias gram positivas) o en el espacio periplasmático, situado entre la membrana interna y externa de la célula (bacterias gram negativas). De preferencia existen sitios de procesamiento, que puede escindirse *in vivo* o *in vitro* codificados entre el fragmento de péptido señal y el gen extraño.

El ADN que codifica secuencias de señal adecuadas puede obtenerse 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 fosfatasa alcalina de *E. coli* (*phoA*) [Oka y col. (1985) *Proc. Natl. Acad. Sci.* 82: 7212]. Como otro ejemplo, puede usarse la secuencia señal del gen de alfa amilasa de diversas cepas de *Bacillus* para secretar proteínas heterólogas de *B. subtilis* [Palva y col. (1982) *Proc. Natl. Acad. Sci. USA* 79: 5582; publicación EP-A-0 244 042].

Usualmente, las secuencias de terminación de la transcripción reconocidas por bacterias son regiones reguladoras situadas en posición 3' con respecto al codón de terminación de la traducción, y por consiguiente junto con el promotor flanquean a la secuencia codificadora. 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 incluyen frecuentemente secuencias de ADN de aproximadamente 50 nucleótidos capaces de formar estructuras en bucle que ayudan en la terminación de la transcripción. Los ejemplos incluyen secuencias de terminación de la transcripción obtenidas de genes con promotores fuertes, tales como el gen *trp* en *E. coli* así como otros genes biosintéticos.

Usualmente, los componentes descritos anteriormente, que comprenden un promotor, secuencia señal (si se desea), secuencia codificadora de interés, y secuencia de terminación de la transcripción, se colocan de forma conjunta en construcciones de expresión. Las construcciones de expresión se mantienen a menudo en un replicón, tal como un elemento extracromosómico (por ejemplo, plásmidos) capaz del mantenimiento estable en un huésped, tal como bacterias. El replicón tendrá un sistema de replicación, que le permite por consiguiente mantenerse en un huésped procarionota para expresión o para clonación y amplificación. Además, un replicón puede ser un plásmido de alto o de bajo número de copias. Un plásmido de alto número de copias tendrá generalmente un número de copias que varía entre aproximadamente 5 y aproximadamente 200 y normalmente entre aproximadamente 10 y aproximadamente 150. Un huésped que contiene un plásmido de alto número de copias contendrá de preferencia al menos aproximadamente 10 y de más de preferencia al menos aproximadamente 20 plásmidos. Puede seleccionarse un vector de alto o bajo número de copias, dependiendo del efecto del vector y de la proteína extraña sobre el huésped.

Como alternativa, las construcciones de expresión pueden integrarse en el genoma bacteriano con un vector de integración. Los vectores de integración contienen usualmente al menos una secuencia homóloga al cromosoma bacteriano que permite al vector integrarse. Las integraciones parecen ser el resultado de recombinaciones entre ADN homólogo en el vector y el cromosoma bacteriano. Por ejemplo, los vectores de integración contruidos con ADN de diversas cepas de *Bacillus* se integran en el cromosoma de *Bacillus* [publicación EP-A-0 127 328]. Los vectores de integración también pueden comprender secuencias de bacteriófagos o del transposón.

Usualmente, las construcciones de expresión extracromosómicas y de integración pueden contener marcadores seleccionables para permitir la selección de cepas bacterianas que se hayan transformado. Los marcadores seleccionables pueden expresarse en el huésped bacteriano y pueden incluir genes que hacen a las bacterias resistentes a fármacos tales como ampicilina, cloramfenicol, eritromicina, kanamicina (neomicina) y tetraciclina [Davies y col. (1978) *Annu. Rev. Microbiol.* 32: 469]. Los marcadores seleccionables también pueden incluir genes biosintéticos, tales como los de las rutas de histidina, triptófano y leucina.

Como alternativa, algunos de los componentes descritos anteriormente pueden colocarse conjuntamente en vectores de transformación. Los vectores de transformación comprenden usualmente un marcador seleccionable que se mantiene en un replicón o se desarrolla en un vector de integración, como se describió anteriormente.

Los vectores de expresión y transformación, ya sean replicones extracromosómicos o vectores de integración, 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) *Proc. Natl. Acad. Sci. USA* 79: 5582; publicaciones EP-A 0 036 259 y EP-A-0 063 953; Documento WO 84/04541], *Escherichia coli* (Shimatake y col., (1981) *Nature* 292: 128; Amman y col., (1985) *Gene* 40: 183; Studier y col. (1986) *J. Mol. Biol.* 189: 113; Publicación

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 EEUU 4.745.056].

5 Los procedimientos para introducir ADN exógeno en huéspedes bacterianos se conocen bien en la técnica, e incluyen usualmente la transformación de una bacteria tratada con  $\text{CaCl}_2$  u otros agentes, tales como cationes divalentes y DMSO. También puede introducirse ADN en células bacterianas por medio de electroporación. Los procedimientos de transformación varían usualmente con las especies bacterianas a transformar. Véase por ejemplo, [Masson y col. (1989) FEMS Microbiol. Lett. 60: 273, Palva y col., (1982) Proc. Natl. Acad. Sci. USA 79: 5582; Publicación EP-A-0 036 259 y EP-A-0 063 953; Documento WO 84/04541], [Miller y col. (1988) Proc. Natl. Acad. Sci. 85: 856, y Wang y col. (1990) J. Bacteriol. 172: 949, Campylobacter], [Cohen y col. (1973) Proc. Natl. Acad. Sci. 69: 2110; Dower y col. (1988) Nucleic Acids Res. 16: 6127; Kushner (1978) "An improved method for transformation of *Escherichia coli* with ColE1-derived plasmids, en Genetic Engineering: Proceedings of the International Symposium on Genetic Engineering (eds. H.W. Boyer y S. 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. Cong. Biotechnology 1:412, Streptococcus].

#### v. Expresión en levaduras

25 Los sistemas de expresión en levaduras también son conocidos por los expertos en la materia. Un promotor de levadura es cualquier secuencia de ADN capaz de unirse a la ARN polimerasa de levadura e iniciar la transcripción secuencia abajo (3') de una secuencia codificadora (por ejemplo, gen estructural) en ARNm. Un promotor tendrá una región de iniciación de la transcripción que se sitúa usualmente en posición proximal con respecto al extremo 5' de la secuencia codificadora. Esta región de iniciación de la transcripción usualmente incluye un sitio de unión a la ARN polimerasa (la "caja TATA") y un sitio de iniciación de la transcripción. Un promotor de levadura también puede tener un segundo dominio denominado secuencia activadora secuencia arriba (UAS), que, si está presente, está normalmente en posición distal con respecto 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 positiva o negativa, potenciando o reduciendo de este modo la transcripción.

35 La levadura es un organismo fermentador con una ruta metabólica activa, por consiguiente las secuencias que codifican enzimas en la ruta metabólica proporcionan secuencias promotoras particularmente útiles. Los ejemplos incluyen la alcohol deshidrogenasa (ADH) (Publicación EP-A-0 284 044), la enolasa, la glucocinasa, la glucosa-6-fosfato isomerasa, la gliceraldehído-3-fosfato-deshidrogenasa, (GAP o GAPDH), la hexocinasa, la fosfofructocinasa, la 3-fosfoglicerato mutasa y la piruvato cinasa (PyK) (Publicación EP-A-0 329 203). El gen *PHO5* de levadura, que codifica la fosfatasa ácida, también proporciona secuencias promotoras útiles [Myanohara y col. (1983) Proc. Natl. Acad. Sci. USA 80: 1].

45 Además, los promotores sintéticos que no se presentan en la naturaleza también funcionan como promotores de levaduras. 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. Los ejemplos de dichos promotores híbridos incluyen la secuencia reguladora ADH unida a la región de activación de la transcripción GAP (Patentes de EEUU N° 4.876.197 y 4.880.734). Otros ejemplos de promotores híbridos incluyen promotores que están constituidos por las secuencias reguladoras de los genes *ADH2*, *GALA*, *GAL10*, o *PHO5*, combinadas con la región de activación transcripcional de un gen de enzima glicolítica tal como GAP o PyK (Publicación EP-A-0 164 556). Además, un promotor de levadura puede incluir promotores que se presentan en la naturaleza de origen diferente a la levadura que tengan la capacidad de unirse a la ARN polimerasa de levadura e iniciar la transcripción. Los ejemplos de tales promotores incluyen, entre otros, [Cohen y col. (1980) Proc. Natl. Acad. Sci. 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].

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

65 Las proteínas de fusión proporcionan una alternativa para los sistemas de expresión en levadura, así como en sistemas de expresión en mamíferos, en baculovirus y bacterianos. Usualmente, una secuencia de ADN que codifica la porción N terminal de una proteína de levadura endógena, u otra proteína estable, se fusiona con el extremo 5' de las secuencias codificadoras heterólogas. Después de la expresión, esta construcción proporcionará una fusión de las dos secuencias de aminoácidos. Por ejemplo, el gen de la superóxido dismutasa (SOD) de levadura o humano, 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 no codificar un sitio escindible. Véase por ejemplo, Publicación EP-A-0 196 056. Otro ejemplo es una proteína de fusión a ubiquitina. Dicha proteína de fusión se elabora con la región de ubiquitina que de preferencia conserva un sitio para una enzima de procesamiento (por ejemplo, proteasa de procesamiento específica de ubiquitina) para escindir la ubiquitina de la proteína extraña. A través de este procedimiento, por consiguiente, puede aislarse una proteína extraña nativa (por ejemplo, documento WO 88/024066).

Como alternativa, también pueden secretarse proteínas extrañas desde la célula en los medios de cultivo creando moléculas de ADN quiméricas que codifiquen una proteína de fusión que comprenda un fragmento de secuencia líder que permite la secreción en levaduras de la proteína extraña. De preferencia, hay sitios de procesamiento codificados entre el fragmento líder y el gen extraño que pueden escindirse *in vivo* o *in vitro*. El fragmento de secuencia líder normalmente codifica un péptido señal que comprende aminoácidos hidrófobos que dirigen la secreción de la proteína a partir de la célula.

El ADN que codifica secuencias señal adecuadas puede obtenerse a partir de genes para proteínas de levaduras secretadas, tales como el gen de la invertasa de levadura (Publicación EP-A-0 012 873; Publicación JPO 62.096.086) y el gen del factor A (Patente de EEUU N° 4.588.684). Como alternativa, existen líderes de origen diferente de levadura, tal como un líder de interferón, que también permiten la secreción en levaduras (Publicación EP-A-0 060 057).

Una clase de preferencia de líderes de secreción son los que emplean un fragmento del gen del factor alfa de levadura, que contiene una secuencia señal "pre" y una región "pro". Los tipos de fragmentos de factor alfa que pueden utilizarse incluyen el líder de factor alfa pre-pro de longitud completa (de aproximadamente 83 residuos de aminoácidos) así como los líderes de factor alfa truncados (usualmente aproximadamente 25 a aproximadamente 50 residuos de aminoácidos) (Patentes de EEUU N° 4.546.083 y 4.870.008; Publicación EP-A-0 324 274). Otros líderes que emplean un fragmento líder de factor alfa que permite la secreción incluyen líderes de factor alfa híbridos elaborados con una presecuencia de una primera levadura, pero una región pro de un factor alfa de una segunda levadura. (Véase por ejemplo Documento WO 89/02463).

Usualmente, las secuencias de terminación de la transcripción reconocidas por levadura son secuencias reguladoras situadas en posición 3' con respecto al codón de terminación de la traducción, y por consiguiente junto con el promotor flanquean la secuencia codificadora. Estas secuencias dirigen la transcripción de un ARNm que puede traducirse al polipéptido codificado por el ADN. Los ejemplos de secuencia de terminación de la transcripción y otras secuencias de terminación reconocidas por la levadura, tales como las que codifican enzimas glicolíticas.

Usualmente, los componentes descritos anteriormente, que comprenden un promotor, líder (si se desea), secuencia codificadora de interés, y secuencia de terminación de la transcripción, se colocan conjuntamente en construcciones de expresión. Las construcciones de expresión se mantienen a menudo en un replicón, tal como un elemento extracromosómico (por ejemplo, plásmidos) capaz de mantenimiento estable en un huésped, tal como una levadura o una bacteria. El replicón puede tener dos sistemas de replicación, que le permiten por consiguiente mantenerse, por ejemplo, en levadura para expresión y en un huésped procariota para clonación y amplificación. Los ejemplos de dichos vectores lanzadera de levadura-bacteria incluyen Yep24 [Botstein y col. (1979) Gene 8: 17-24], pCI/1 [(Brake y col. (1984) Proc Natl. Acad. Sci. USA 81: 4642-4646] e YRp17 [Stinchcomb y col. (1982) J. Mol. Biol. 158: 157]. Además, un replicón puede ser un plásmido de alto número de copias o un plásmido de bajo número de copias. Un plásmido de alto número de copias tendrá generalmente un número de copias que varía desde aproximadamente 5 hasta aproximadamente 200, y usualmente aproximadamente 10 hasta aproximadamente 150. Un huésped que contiene un plásmido de alto número de copias de preferencia tendrá al menos aproximadamente 10, y de más preferencia al menos aproximadamente 20. Puede seleccionarse la introducción de un vector de alto o de bajo número de copias, dependiendo del efecto del vector y la proteína extraña sobre el huésped. Véase por ejemplo, Brake y col., *supra*.

Como alternativa, las construcciones de expresión pueden integrarse en el genoma de la levadura con un vector de integración. Los vectores de integración usualmente contienen al menos una secuencia homóloga a un cromosoma de levadura que permite que el vector se integre, y de preferencia contienen dos secuencias homólogas que flanquean a la construcción de expresión. Las integraciones parecen ser el resultado 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 de integración puede dirigirse a un locus específico en la levadura seleccionando la secuencia homóloga apropiada para la inclusión en el vector. Véase Orr-Weaver y col., *supra*. Pueden integrarse una o más construcciones de expresión, afectando posiblemente a los niveles de la proteína recombinante producida [Rine y col. (1983) Proc. Natl. Acad. Sci. USA 80: 6750]. Las secuencias cromosómicas incluidas en el vector pueden aparecer como un único segmento en el vector, que da como resultado la integración del vector completo, o como dos segmentos homólogos a segmentos adyacentes en el cromosoma y que flanquean la construcción de expresión en el vector, que puede dar como resultado la integración estable de solamente la construcción de expresión.

Usualmente, las construcciones de expresión extracromosómicas y de integración pueden contener marcadores seleccionables para permitir la selección de cepas de levaduras que se hayan transformado. Los marcadores seleccionables pueden incluir genes biosintéticos que pueden expresarse en el huésped de levadura, tales como *ADE2*, *HIS4*, *LEU2*, *TRP1* y *ALG7* y el gen de resistencia a *G418*, que confiere resistencia en células de levaduras a tunicamicina y *G418*, respectivamente. Además, un marcador seleccionable adecuado puede proporcionar también a la levadura la capacidad de crecer en presencia de compuestos tóxicos, tales como metal. Por ejemplo, la presencia de *CUP1* permite a la levadura crecer en presencia de iones de cobre [Butt y col. (1987) Microbiol. Rev. 51: 351].

## ES 2 333 071 T3

Como alternativa, algunos de los componentes descritos anteriormente pueden colocarse conjuntamente en vectores de transformación. Los vectores de transformación usualmente comprenden un marcador seleccionable que se mantiene en un replicón o se desarrolla dentro de un vector de integración, como se describió anteriormente.

5 Los vectores de expresión y transformación, replicones extracromosómicos o vectores de integración, 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), Bio/Technology 8: 135]; *Pichia guilliermondii* [Kunze y col., (1985) J. Basic Microbiol. 25: 141]; *Pichia pastoris* [Cregg, y col (1985), Mol. Cell. Biol. 5: 3376; Patentes de EEUU N° 4.837.148 y 4.929.555]; *Saccharomyces cerevisiae* [Hinnen y col. (1978) Proc. Natl. Acad. Sci. 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].

Los procedimientos para introducir ADN exógeno en los huéspedes de levadura son bien conocidos en la técnica, y usualmente incluyen la transformación de esferoplastos o la transformación de células de levadura intactas tratadas con cationes alcalinos. Los procedimientos de transformación varían usualmente con las especies de levadura a transformar. 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 EEUU N° 4.837.148 y 4.929.555; *Pichia*]; [Hinnen y col. (1978) Proc. Natl. Acad. Sci. 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*].

### Anticuerpos

30 Como se usa en el presente documento, el término “anticuerpo” se refiere a un polipéptido o a un 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 superficie interna y una distribución de cargas complementarias 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 Fab y anticuerpos de dominio único.

Los anticuerpos contra las proteínas de la invención son útiles para cromatografía de afinidad, inmunoensayos y distinción/identificación de proteínas de *Neisseria*.

Los anticuerpos obtenidos contra las proteínas de la invención, policlonales y monoclonales, pueden prepararse por medio de procedimientos convencionales. En general, la proteína se usa primero para inmunizar un animal adecuado, de preferencia un ratón, rata, conejo o cabra. Resultan de preferencia conejos y cabras para la preparación de sueros policlonales debido al volumen de suero obtenible, y a la disponibilidad de anticuerpos anticonejo y anticabra marcados. 45 La inmunización se realiza generalmente mezclando o emulsionando la proteína en disolución salina, de preferencia en un adyuvante tal como el adyuvante completo de Freund, e inyectando la mezcla o emulsión por vía parenteral (generalmente por vía subcutánea o intramuscular). Típicamente es suficiente una dosis de 50-200  $\mu\text{g}$ /inyección. La inmunización se refuerza generalmente 2-6 semanas después con una o más inyecciones de la proteína en disolución salina, de preferencia usando adyuvante incompleto de Freund. Pueden generarse como alternativa anticuerpos mediante inmunización *in vitro* usando procedimientos conocidos en la técnica, que para los fines de esta invención se considera equivalente a la inmunización *in vivo*. Los antisueros policlonales se obtienen extrayendo sangre de animales inmunizados en un recipiente de vidrio o de plástico, incubando la sangre a 25°C durante 1 hora, seguido por la incubación a 4°C durante 2-18 horas. El suero se recupera mediante centrifugado (por ejemplo, a 1000 g durante 10 minutos). Pueden obtenerse de aproximadamente 20-50 ml por extracción de sangre de conejos.

Los anticuerpos monoclonales se preparan usando el procedimiento convencional de Kohler y Milstein [Nature (1975) 256: 495-496], o una de sus modificaciones. Típicamente, se inmuniza un ratón o una rata como se describió anteriormente. Sin embargo, en lugar de extraer sangre del animal para extraer el suero, el bazo (y opcionalmente varios nódulos linfáticos grandes) se extirpa y se disocia en células únicas. Si se desea, las células del bazo pueden seleccionarse (después de retirar las células adherentes no específicas) aplicando una suspensión celular a una placa o a un pocillo recubierto con el antígeno de la proteína. Las células B que expresan inmunoglobulina unida a membrana específica para el antígeno se unen a la placa y no se eliminan mediante el aclarado con el resto de la suspensión. Las células B resultantes, o todas las células del bazo disociadas, se inducen después para que se fusionen con células de mieloma para formar hibridomas, y se cultivan en un medio selectivo (por ejemplo, medio de hipoxantina, aminopterina, timidina, “HAT”). Los hibridomas resultantes se plaquean mediante dilución limitante, y se ensayan para la producción de anticuerpos que se unen específicamente al antígeno inmunizador (y que no se unen a antígenos no relacionados). Los hibridomas que secretan MAB seleccionados se cultivan después *in vitro* (por ejemplo, en frascos de cultivo tisular o en reactores de fibra hueca) o *in vivo* (como ascitis en ratones).

Si se desea, los anticuerpos (policlonales o monoclonales) pueden marcarse usando técnicas convencionales. Los marcadores adecuados incluyen fluoróforos, cromóforos, átomos radioactivos (particularmente  $^{32}\text{P}$  y  $^{125}\text{I}$ ), reactivos densos a electrones, enzimas y ligandos que tienen parejas de unión específicas. Las enzimas se detectan normalmente mediante su actividad. Por ejemplo, la peroxidasa del rábano picante se detecta usualmente mediante su capacidad para convertir 3,3',5,5'-tetrametilbenzidina (TMB) a un pigmento azul, cuantificable con un espectrofotómetro. “Parejas de unión específica” se refiere a una proteína capaz de unirse a una molécula de ligando con alta especificidad, como por ejemplo en el caso de un antígeno y un anticuerpo monoclonal específico para éste. Otros parejas de unión específica incluyen biotina y avidina o estreptavidina, IgG y proteína A, y las numerosas parejas de receptor-ligando conocidas en la técnica. Debería entenderse que la descripción anterior no pretende categorizar los diversos marcadores en diferentes clases, ya que el mismo marcador puede servir en varios modos diferentes. Por ejemplo,  $^{125}\text{I}$  puede servir como un marcador radioactivo o como un reactivo denso en electrones. HRP puede servir como enzima o como antígeno para un MAb. Además, pueden combinarse diversos marcadores para el efecto deseado. Por ejemplo, los MAb y la avidina también requieren marcadores en la práctica de esta invención: por consiguiente, puede marcarse un MAb con biotina, y detectarse su presencia con avidina marcada con  $^{125}\text{I}$ , o con un MAb anti-biotina marcado con HRP. Otras permutaciones y posibilidades serán fácilmente evidentes para los expertos en la técnica, y se consideran como equivalentes dentro del alcance de la presente invención.

#### *Composiciones farmacéuticas*

Las composiciones farmacéuticas pueden comprender polipéptidos, anticuerpos o ácidos nucleicos de la invención. Las composiciones farmacéuticas comprenderán una cantidad terapéuticamente eficaz de polipéptidos, anticuerpos, o polinucleótidos de la invención reivindicada.

El término “cantidad terapéuticamente eficaz”, como se usa en el presente documento se refiere a una cantidad de un agente terapéutico para tratar, mejorar o evitar una enfermedad o afección deseada, o para mostrar un efecto terapéutico o preventivo detectable. El efecto puede detectarse mediante, por ejemplo, niveles de marcadores químicos o de antígenos. Los efectos terapéuticos también incluyen reducción en los síntomas físicos, tales como temperatura corporal disminuida. La cantidad eficaz exacta para un sujeto dependerá del tamaño y la salud del sujeto, de la naturaleza y el alcance de la afección, y de los agentes terapéuticos o combinación de agentes terapéuticos seleccionados para administración. Por consiguiente, no es útil especificar una cantidad eficaz exacta por adelantado. Sin embargo, la cantidad eficaz para una situación dada puede determinarse mediante experimentación rutinaria y está dentro del juicio del médico.

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

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 no induce por sí mismo la producción de anticuerpos perjudiciales para el individuo que recibe la composición, y que puede administrarse sin toxicidad indebida. Los vehículos adecuados pueden ser macromoléculas grandes, metabolizadas lentamente tales como proteínas, polisacáridos, ácidos polilácticos, ácidos poliglicólicos, aminoácidos poliméricos, copolímeros de aminoácidos y partículas de virus inactivos. Tales vehículos son bien conocidos por los expertos en la técnica.

Pueden usarse en los mismos 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 exhaustiva de los excipientes farmacéuticamente aceptables está disponible en el documento Remington Pharmaceutical Sciences (Mack Pub. Co. N.J. 1991).

Los vehículos farmacéuticamente aceptables en las composiciones terapéuticas pueden contener líquidos tales como agua, disolución salina, glicerol y etanol. Además, en tales vehículos pueden estar presentes sustancias auxiliares, tales como agentes humectantes o agentes emulsivos, sustancias reguladoras del pH, y similares. Típicamente, las composiciones terapéuticas se preparan como inyectables, como disoluciones líquidas o como suspensiones líquidas; también pueden prepararse formas sólidas adecuadas para disolución en, o suspensión en, vehículos líquidos antes de la inyección. Los liposomas se incluyen 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 a tratar pueden ser animales; en particular, pueden tratarse sujetos humanos.

La administración directa de las composiciones se realizará generalmente mediante inyección, por vía subcutánea, por vía intraperitoneal, intravenosa o intramuscular o se administrarán en el espacio intersticial de un tejido. Las composiciones también pueden administrarse en una lesión. Otros modos de administración incluyen administración oral y pulmonar, supositorios y aplicaciones transdérmicas y transcutáneas (por ejemplo véase el documento WO 98/20734), agujas, y pistolas génicas o inyectoras tipo hypospray. El tratamiento de dosificación puede ser un programa de monodosis o un programa de dosis múltiples.

## Vacunas

Las vacunas de acuerdo con la invención pueden ser profilácticas (es decir, para evitar la infección) o bien terapéuticas (es decir, para tratar la enfermedad después de la infección).

Tales vacunas comprenden antígeno(s), inmunógeno(s), polipéptido(s), proteína(s) o ácidos nucleicos inmunizadores, usualmente en combinación con “vehículos farmacéuticamente aceptables” que incluyen cualquier vehículo que no induce por sí mismo la producción de anticuerpos dañinos para el individuo que recibe la composición. Los vehículos adecuados son normalmente macromoléculas grandes, metabolizadas lentamente tales como proteínas, polisacáridos, ácidos polilácticos, ácidos poliglicólicos, aminoácidos poliméricos, copolímeros de aminoácidos, agregados lipídicos (tales como gotículas de aceite o liposomas) y partículas de virus inactivados. Tales vehículos son bien conocidos por los expertos en la técnica. Además, estos vehículos pueden funcionar como agentes inmunoestimuladores (“adyuvantes”). Además, el inmunógeno o antígeno puede conjugarse con un toxoide bacteriano, tal como un toxoide de difteria, tétanos, cólera, *H. pylori*, etc. patógenos.

Los adyuvantes de preferencia 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 emulsiones de aceite en agua (con o sin otros agentes inmunoestimuladores específicos tales como péptidos de muramilo (véase a continuación) o componentes de la pared celular bacteriana), tales como por ejemplo (a) MF59<sup>TM</sup> (documento WO 90/14837; Capítulo 10 en Vaccine design: the subunit and adjuvant approach, eds. Powell & Newman, Plenum Press 1995) que contienen escualeno al 5%, Tween 80 al 0,5% y Span 85 al 0,5% (que contienen opcionalmente diversas cantidades de MTP-PE (véase a continuación), aunque no se requieren) formuladas en partículas submicrométricas usando un microfluidizador tal como el microfluidizador modelo 110Y (Microfluidics, Newton, MA), (b) SAF, que contiene escualeno al 10%, Tween 80 al 0,4%; polímero L121 bloqueado con plurónico al 5%, y thr-MDP (véase a continuación) microfluidizado en una emulsión submicrométrica o agitado en vórtice para generar una emulsión de tamaño de partícula mayor y (c) el sistema adyuvante Ribit<sup>TM</sup> (RAS), (Ribi Immunochem, Hamilton, MT) que contiene escualeno al 2%, Tween 80 a 0,2%, y uno o más componentes de la pared celular bacteriana del grupo constituido por monofosforilípido A (MPL), dimicolato de trehalosa (TDM), y esqueleto de la pared celular (CWS), de preferencia MPL + CWS (Detox<sup>TM</sup>); (3) pueden usarse adyuvantes de saponina, tales como Stimulon<sup>TM</sup> (Cambridge Bioscience, Worcester, MA) o partículas generadas a partir de los mismos tales como ISCOM (complejos inmunoestimuladores); (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 gamma), factor estimulador de las colonias de macrófagos (M-CSF), factor de necrosis tumoral (TNF), etc.; y (6) otras sustancias que actúan como agentes inmunoestimuladores para aumentar la eficacia de la composición. Resultan de preferencia alumbre y MF59<sup>TM</sup>G.

Como se mencionó anteriormente, los péptidos de muramilo 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-isoglutamil-L-alanina-2-(1'-2'-dipalmitoil-sn-glicero-3-hidroxi-fosforiloxi)-etilamina (MTP-PE), etc.

Las composiciones inmunogénicas (por ejemplo, el antígeno/inmunógeno/polipéptido/proteína/ácido nucleico inmunizadores, el vehículo farmacéuticamente aceptable y el adyuvante) contendrán típicamente diluyentes, tales como agua, disolución salina, glicerol, etanol, etc. Además, pueden estar presentes en tales vehículos, sustancias auxiliares, tales como agentes humectantes o emulsivos, sustancias tamponadoras del pH, y similares.

Típicamente, las composiciones inmunogénicas se preparan como inyectables, en forma de soluciones líquidas o en forma de suspensiones líquidas; 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 un efecto adyuvante potenciado, como se describió anteriormente en vehículos farmacéuticamente aceptables.

Las composiciones inmunogénicas usadas como vacunas comprenden una cantidad inmunológicamente eficaz de los polipéptidos antigénicos o inmunogénicos, así como cualquier otro de los componentes mencionados anteriormente, según sea necesario. Por “cantidad inmunológicamente eficaz”, se entiende que la administración de esta cantidad a un individuo, en una monodosis o como parte de una serie, es eficaz para el tratamiento o prevención. Esta cantidad varía dependiendo de la salud y de la condición física del individuo a tratar, del grupo taxonómico del individuo a tratar (por ejemplo, primate no humano, primate, etc.), de la capacidad del sistema inmune del individuo para sintetizar anticuerpos, del grado de protección deseado, de la formulación de la vacuna, de la evaluación del médico tratante de la situación médica, y de otros factores pertinentes. Se espera que la cantidad esté dentro de un intervalo relativamente amplio que pueda determinarse a través de ensayos rutinarios.

Las composiciones inmunogénicas se administran convencionalmente por vía parenteral, por ejemplo, mediante inyección, por vía subcutánea, por vía intramuscular o por vía transdérmica/transcutánea (por ejemplo Documento WO 98/20734). Otras formulaciones adecuadas para otros modos de administración incluyen formulaciones orales y pulmonares, supositorios y aplicaciones transdérmicas. El tratamiento de la dosificación puede ser un programa de monodosis y un programa de dosis múltiples. La vacuna puede administrarse junto con otros agentes inmunorreguladores.

## ES 2 333 071 T3

Como una alternativa a las vacunas basadas en proteínas, puede utilizarse la vacunación con ADN [por ejemplo, Robinson & Torres (1997) *Seminars in Immunology* 9: 271-283; Donnelly y col. (1997) *Annu Rev Immunol* 15: 617-648]; véase a continuación en el presente documento].

### 5 Vehículos para la administración de genes

Los vehículos de terapia génica para la administración de construcciones, incluyendo una secuencia codificadora de un compuesto terapéutico de la invención, para administrar al mamífero para expresión en el mamífero, pueden administrarse localmente o sistémicamente. Estas construcciones pueden utilizar enfoques de vector viral o no viral en modalidad *in vivo* o *ex vivo*. La expresión de tal secuencia codificadora puede inducirse usando promotores de mamífero endógenos o promotores heterólogos. La expresión de la secuencia codificadora *in vivo* puede ser constitutiva o regulada.

La invención incluye vehículos de administración de genes capaces de expresar las secuencias de ácido nucleico contempladas. El vehículo de administración de genes es de preferencia un vector viral y, de más preferencia, un vector retroviral, adenoviral, viral adenoasociado (AAV), herpes viral o de alfavirus. El vector viral también puede ser un vector viral de astrovirus, coronavirus, ortomixovirus, papovavirus, paramixovirus, parvovirus, picornavirus, poxvirus o togavirus. Véase en general, 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 retrovirales se conocen bien en la técnica y contemplamos que es utilizable cualquier vector de terapia génica retroviral en la invención, incluyendo 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), espumavirus y lentivirus. Véase *RNA Tumor Viruses*, Segunda edición, Cold Spring Harbor Laboratory, 1985.

Las porciones del vector de terapia génica retroviral pueden obtenerse de diferentes retrovirus. Por ejemplo, pueden obtenerse LTR de retrovectores de un Virus de Sarcoma Murino, un sitio de unión a ARNt de un Virus de Sarcoma de Rous, una señal de empaquetado de un Virus de Leucemia Murina, y un origen de la síntesis de segunda hebra de un Virus de Leucosis Aviar.

Estos vectores retrovirales recombinantes pueden usarse para generar partículas de vector retroviral competentes para la transducción introduciéndolos en líneas celulares de empaquetado apropiadas (véase la patente de EEUU 5.591.624). Los vectores de retrovirus pueden construirse para integración específica de sitio en ADN de la célula huésped mediante incorporación de una enzima integrasa quimérica dentro de la partícula retroviral (véase el documento WO 96/37626). Resulta de preferencia que el vector viral recombinante sea un virus recombinante de replicación defectiva.

Las líneas celulares de empaquetado adecuadas para usar con los vectores de retrovirus descritos anteriormente se conocen bien en la técnica, se preparan fácilmente (véanse los documentos WO 95/30763 y WO 92/05266), y pueden usarse para crear líneas celulares productoras (también llamadas líneas celulares de vector o "VCL") para la producción de partículas de vectores recombinantes. De preferencia, las líneas celulares de empaquetado se preparan a partir de células parentales humanas (por ejemplo células HT1080) o líneas celulares parentales de visón, que eliminan la inactivación en suero humano.

Los retrovirus de preferencia para la construcción de vectores de terapia génica retrovirales incluyen Virus de la Leucosis Aviar, Virus de la Leucemia Bovina, Virus de la Leucemia Murina, Virus que Induce Focos en Células de Visón, Virus del Sarcoma Murino, Virus de Reticuloendoteliosis y Virus del Sarcoma de Rous. Los Virus de Leucemia Murina particularmente de preferencia incluyen 4070A y 1504A (Hartley y Rowe (1976) *J. Virol.* 19: 19-25), Abelson (ATCC N° VR-999), Friend (ATCC N° VR-245), Graffi, Gross (ATCC N° VR-590), Kirsten, Virus del Sarcoma de Harvey y Rauscher (ATCC N° VR-998) y Virus de la Leucemia Murina de Moloney (ATCC N° VR-190). Tales retrovirus pueden obtenerse de depósitos o colecciones tales como la American Type Culture Collection ("ATCC") en Rockville, Maryland o aislarse a partir de fuentes conocidas usando técnicas disponibles comúnmente.

Los vectores de terapia génica retroviral conocidos de ejemplo que pueden utilizarse en la presente invención incluyen los que se describen en las solicitudes de patente GB 2200651, EP 0415731, EP 0345242, EP 0334301, WO 89/02468; WO 89/05349, WO 89/09271, WO 90/02806, WO 90/07936, WO 94/03622, WO 93/25698, WO 93/25234, WO 93/11230, WO 93/10218, WO 91/02805, WO 91/02825, WO 95/07994, US 5.219.740, US 4.405.712, US 4.861.719, US 4.980.289, US 4.777.127, US 5.591.624. Véase 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) *Proc Natl Acad Sci* 81: 6349; y Miller (1990) *Human Gene Therapy* 1.

Los vectores de terapia génica de adenovirus humanos también se conocen en la técnica y pueden utilizarse en la presente invención. Véase, por ejemplo, Berkner (1988) *Biotechniques* 6: 616 y Rosenfeld (1991) *Science* 252: 431, y documentos WO 93/07283, WO 93/06223, y WO 93/07282. Los vectores de terapia génica de adenovirus co-

nocidos de ejemplo que pueden utilizarse en la presente invención incluyen los que se describen en los documentos mencionados anteriormente y en los documentos WO 94/12649, WO 93/03769, WO 93/19191, WO 94/28938, WO 95/11984, WO 95/00655, WO 95/27071, WO 95/29993, WO 95/34671, WO 96/05320, WO 94/08026, WO 94/11506, WO 93/06223, WO 94/24299, WO 95/14102, WO 95/24297, WO 95/02697, WO 94/28152, WO 94/24299, WO 95/09241, WO 95/25807, WO 95/05835, WO 94/18922 y WO 95/09654. Como alternativa, puede utilizarse la administración de ADN unido a adenovirus inactivados 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 principales y de preferencia de tales vectores para usar en la presente invención son los vectores basados en AAV-2 dados a conocer en Srivastava, WO 93/09239. Los vectores AAV de más preferencia, comprenden las dos repeticiones terminales invertidas de AAV en las que se modifican las secuencias D nativas mediante sustitución de nucleótidos, tal que se conservan al menos 5 nucleótidos nativos y hasta 18 nucleótidos nativos, de preferencia al menos 10 nucleótidos nativos y hasta 18 nucleótidos nativos, de más preferencia 10 nucleótidos nativos y los restantes nucleótidos de la secuencia D se delecionan o se sustituyen con nucleótidos no nativos. Las secuencias D 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 están implicadas en la formación de HP. El nucleótido de sustitución no nativo puede ser cualquier nucleótido diferente del nucleótido encontrado en la secuencia D nativa en la misma posición. Otros vectores de AAV ejemplares que pueden utilizarse son pWP-19, pWN-1 ambos dados a conocer en Nahreini (1993) Gene 124: 257-262. Otro ejemplo de tales vectores de AAV es psub201 (véase Samulski (1987) J. Virol. 61: 3096). Otro vector de AAV de ejemplo es el vector IRT Doble-D. La construcción del vector ITR Doble D se da a conocer en la Patente de EEUU N° 5.478.745. Aún otros vectores son los que se dan a conocer en la Patente de Estados Unidos N° 4.797.368 de Carter y en la Patente de EEUU 5.139.941 de Muzyczka, en la Patente de EEUU 5.474.935 de Chartejee y en el documento WO 94/282157 de Kotin. Aún otro ejemplo de un vector de AAV que puede utilizarse en la presente invención es SSV9AFABTKneo, que contiene el potenciador AFP y el promotor de albúmina y dirige la expresión de forma predominante en el hígado. Su estructura y su construcción se dan a conocer en Su (1996) Human Gene Therapy 7: 463-470. En los documentos US 5.354.678, US 5.173.414, US 5.139.941 y US 5.252.479 se describen vectores de terapia génica de AAV adicionales.

Los vectores de terapia génica de la invención también incluyen vectores de herpes. Los ejemplos principales y de preferencia son vectores del virus herpes simplex que contienen una secuencia que codifica un polipéptido de timidina cinasa tal como los descritos en los documentos US 5.288.641 y EP 0176170 (Roizman). Otros vectores de virus herpes simplex de ejemplo incluyen HFEM/ICP6-LacZ dado a conocer en el documento WO 95/04139 (Wistar Institute), pHSVlac descrito en Geller (1988) Science 241: 1667-1669 y en los documentos WO 90/09441 y WO 92/07945, HSV Us3::pgC-lacZ descrito en Fink (1992) Human Gene Therapy 3: 11-19 y HSV 7134, 2 RH 105 y GAL4 descritos en el documento EP 0453242 (Breakefield), y los depositados en la ATCC con los números de acceso ATCC VR-977 y ATCC VR-260.

También se contemplan vectores de terapia génica de virus alfa que pueden utilizarse en la presente invención. Los vectores de virus alfa de preferencia son los vectores de los virus Sindbis, Togavirus, virus Semliki Forest (ATCC VR-67, ATCC VR-1247), virus Middleberg (ATCC VR-370), virus Ross River (ATCC VR-373; ATCC VR-1246), virus de la encefalitis equina venezolana (ATCC VR923; ATCC VR-1250; ATCC VR-1249; ATCC VR-532), y los descritos en las Patente de EEUU 5.091.309, 5.217.879, y en el documento WO 92/10578. Más particularmente, los vectores de virus alfa descritos en los documentos WO 94/21792, WO 92/10578, WO 95/07994, US 5.091.309 y US 5.217.879 son utilizables. Tales virus alfa pueden obtenerse a partir de depósitos o de colecciones tales como la ATCC en Rockville, Maryland o aislarse a partir de fuentes conocidas usando técnicas disponibles comúnmente. De preferencia, se usan vectores de virus alfa con citotoxicidad reducida.

Los sistemas de vectores de ADN tales como sistemas de expresión en capas eucariotas también son útiles para expresar los ácidos nucleicos de la invención. Véase el documento WO 95/07994 para una descripción detallada de los sistemas de expresión en capas eucariotas. De preferencia, los sistemas de expresión en capas eucariotas de la invención se obtienen de vectores de virus alfa y lo de más preferencia de vectores virales de Sindbis.

Otros vectores virales adecuados para usar en la presente invención incluyen los derivados de poliovirus, por ejemplo ATCC VR-58 y los que se describen en Evans, Nature 339 (1989) 385 y Sabin (1973) J. Biol. Standardization 1: 115; rinovirus, por ejemplo ATCC VR-1110 y los que se describen en Arnold (1990) J Cell Biochem L401; poxvirus tales como poxvirus de canario o virus vacunal, por ejemplo ATCC VR-111 y ATCC VR-2010 y los que se describen en el documento Fisher-Hoch (1989) Proc. Natl Acad Sci 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 en el documento WO 89/01973; virus SV40, por ejemplo ATCC VR-305 y los descritos en los documentos Mulligan (1979) Nature 277: 108 y Madzak (1992) J Gen Virol 73: 1533; virus de la gripe, por ejemplo ATCC VR-797 y virus de la gripe recombinantes preparados utilizando técnicas genéticas inversas como se describen en el documento US 5.166.057 y en Enami (1990) Proc Natl Acad Sci 87: 3802-3805; Enami y Palese (1991) J Virol 65: 2711-2713 y Luytjes (1989) Cell 59: 110, (véase 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 los descritos en el documento EP-0440219; virus Aura, por ejemplo ATCC VR-368; virus Bebaru, por ejemplo ATCC VR-600 y ATCC VR-1240; virus Cabassou, por ejemplo ATCC VR-922; virus Chikungunya, por ejemplo ATCC VR-64 y ATCC VR-1241; virus Fort Morgan, por ejemplo ATCC VR-924; virus Getah, por ejemplo ATCC VR-369 y ATCC VR-1243; virus Kyzylgach, por ejemplo

ATCC VR-927; virus Mayaro, por ejemplo ATCC VR-66; virus Mucambo, por ejemplo ATCC VR-580 y ATCC VR-1244; virus Ndumu, por ejemplo ATCC VR-371; virus Pixuna, por ejemplo ATCC VR-372 y ATCC VR-1245; virus Tonate, por ejemplo ATCC VR-925; virus Trinita, por ejemplo ATCC VR-469; virus Una, por ejemplo ATCC VR-374; virus Whataroa, por ejemplo ATCC VR-926; virus Y-62-33, por ejemplo ATCC VR-375; virus 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 los que se describen en Hamre (1966) Proc Soc Exp Biol Med 121: 190.

La administración de las composiciones de esta invención dentro de las células no se limita a los vectores virales mencionados anteriormente. Pueden utilizarse otros procedimientos y medios de administración tales como, por ejemplo vectores de expresión de ácidos nucleicos, ADN condensado policatiónico unido o no unido a adenovirus inactivados en solitario, por ejemplo véase Curiel (1992) Hum Gene Ther 3: 147-154, ADN unido a ligando, por ejemplo véase Wu (1989) J Biol Chem 264: 16985-16987, células de vehículos para la administración de células eucariotas, por ejemplo véase la patente de EEUU 60/5686, depósito de materiales de hidrogel fotopolimerizado, pistola de partículas de transferencia génica de manejo manual, como se describe en la patente de EEUU 5.149.655, radiación ionizante como se describe en el documento US 5.206.152 y en el documento WO 92/11033, neutralización de carga nucleica o fusión con membranas celulares. Se describen otros enfoques en Philip (1994) Mol Cell Biol 14: 2411-2418 y en Woffendin (1994) Proc Natl Acad Sci 91: 1581-1585.

Puede utilizarse transferencia génica mediada por partículas. Brevemente, la secuencia puede insertarse dentro de vectores convencionales que contienen secuencias de control convencionales para alto nivel de expresión, y después incubarse con moléculas de transferencia génica sintética tales como cationes de unión a ADN polimérico tales como polilisina, protamina y albúmina, unidos a ligandos dirigidos a células tales como asialoorosomucoide, 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 utilizarse ADN desnudo. Los procedimientos de introducción de ADN desnudo de ejemplo se describen en el documento WO 90/11092 y en el documento US 5.580.859. La eficacia de la captación puede mejorarse usando perlas de látex biodegradables. Las perlas de látex recubiertas de ADN se transportan de manera eficaz dentro de células después de iniciación de endocitosis mediante las perlas. El procedimiento puede mejorarse además mediante el tratamiento de las perlas para aumentar la hidrofobia y facilitar de este modo la alteración del endosoma y la liberación del ADN al citoplasma.

Los liposomas que pueden actuar como vehículos de suministro de genes se describen en los documentos U.S. 5.422.120, WO 95/13796, WO 94/23697, WO 91/144 y EP-524.968. Las secuencias de ácido nucleico que codifican un polipéptido pueden insertarse en vectores convencionales que contienen secuencias de control convencionales para un alto nivel de expresión y después pueden incubarse con moléculas de transferencia de genes sintéticas tales como cationes de unión a ADN polimérico, tales como polilisina, protamina y albúmina, unidos a ligandos dirigidos a células tales como asialoorosomucoide, insulina, galactosa, lactosa o transferrina. Otros sistemas de suministro incluyen el uso de liposomas para encapsular ADN que comprende el gen bajo el control de diversos promotores específicos de tejido o activos de forma ubicua. Otra administración convencional no viral adecuada para usar incluye sistemas de suministro mecánico tales como el enfoque descrito en Woffendin y col. (1994) Proc. Natl. Acad. Sci. USA 91(24): 11581-11585. Además, la secuencia codificadora y el producto de la expresión de tal secuencia codificadora pueden administrarse a través del depósito de materiales de hidrogel fotopolimerizados. Otros procedimientos convencionales para la administración de genes que pueden usarse para administrar la secuencia codificadora incluyen, por ejemplo, el uso de pistola de partículas de transferencia de genes de uso manual, como se describe en el documento U.S. 5.149.655; uso de radiación ionizante para activar un gen transferido, como se describe en los documentos U.S. 5.206.152 y WO 92/11033.

Los vehículos de administración de genes policatiónicos y de liposomas de ejemplo son los descritos en los documentos US 5.422.120 y 4.762.915; en los documentos WO 95/13796; WO 94/23697 y WO 91/1444; en el documento EP-0524968; 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) Proc Natl Acad Sci 84: 7851; Plant (1989) Anal Biochem 176: 420.

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

#### *Procedimientos de Administración*

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

La administración directa de las composiciones se conseguirá generalmente mediante inyección, por vía subcutánea, intraperitoneal, intravenosa o intramuscular o se administrarán en el espacio intersticial de un tejido. Las composiciones también pueden administrarse dentro de una lesión. Otros modos de administración incluyen administración oral y pulmonar, supositorios, y aplicaciones transdérmicas o transcutáneas (véase por ejemplo el documento WO 98/20734), agujas, y pistolas génicas o inyectores de tipo hypospray. El tratamiento de la dosificación puede ser un programa de monodosis o un programa de dosis múltiples.

Los procedimientos para la administración *ex vivo* y la reimplantación de células transformadas en un sujeto se conocen en la técnica y se describen por ejemplo en el documento WO 93/14778. Los ejemplos de células útiles en aplicaciones *ex vivo* incluyen, por ejemplo, células madre, particularmente hematopoyéticas, células linfáticas, macrófagos, células dendríticas o células tumorales.

Generalmente, la administración de ácidos nucleicos para aplicaciones *ex vivo* e *in vitro* puede realizarse mediante los siguientes procedimientos, por ejemplo, transfección mediada por dextrano, precipitación con fosfato de calcio, transfección mediada por polibreno, fusión de protoplastos, electroporación, encapsulamiento de polinucleótido(s) en liposomas, y microinyección directa de ADN dentro de los núcleos, todos bien conocidos en la técnica.

#### *Composiciones farmacéuticas de polinucleótidos y polipéptidos*

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

##### *A. Polipéptidos*

Un ejemplo son polipéptidos que incluyen, sin limitación: asioloorosomucoide (ASOR); transferrina; asialoglicoproteínas; anticuerpos; fragmentos de anticuerpos; ferritina; interleucinas; interferones, factor estimulador de las colonias de granulocitos y macrófagos (GM-CSF), factor estimulador de las colonias de granulocitos (G-CSF), factor estimulador de colonias de macrófagos (M-CSF), factor de células madre y eritropoyetina. También pueden usarse antígenos virales, tales como proteínas de la envoltura. También, proteínas de otros organismos invasores, tales como el péptido de 17 aminoácidos de la proteína del circumporozoito 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.*

Además, puede incluirse polialquilenoglicol con los polinucleótidos o polipéptidos deseados. En una forma de realización de preferencia, el polialquilenoglicol es polietilenoglicol. Además, pueden incluirse mono-, di- o polisacáridos. En una forma de realización de preferencia de este aspecto, el polisacárido es dextrano o DEAE-dextrano. También, quitosano y poli(lactida-co-glicólido).

##### *D. Lípidos y Liposomas*

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

El encapsulamiento en lípidos se realiza generalmente usando liposomas que son capaces de unirse o atrapar de forma estable y conservar el ácido nucleico. La proporción de polinucleótido condensado para la preparación de lípidos puede variar pero generalmente será de 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 de liposomas para uso en la presente invención incluyen preparaciones catiónicas (cargadas positivamente), aniónicas (cargadas negativamente) y neutras. Se ha demostrado que los liposomas catiónicos median la administración intracelular de ADN de plásmidos (Felgner (1987) *Proc. Natl. Acad. Sci. USA* 84: 7413-7416); ARNm (Malone (1989) *Proc. Natl. Acad. Sci. 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, están disponibles liposomas de N[1-(2,3-dioleiloxi)propil]-N,N,N-trietilamonio (DOTMA) bajo la marca comercial Lipofectin, de GIBCO BRL, Grand Island, NY. (Véase también Felgner *supra*). Otros liposomas disponibles en el mercado incluyen transfectace (DDAB/DOPE) y DOTA/DOPE (Boehringer). Otros liposomas catiónicos pueden prepararse a partir de materiales fácilmente disponibles usando técnicas bien conocidas en la técnica. Véase, por ejemplo Szoka (1978) *Proc. Natl. Acad. Sci. USA* 75: 4194-4198; documento WO 901/1092 para una descripción de la síntesis de liposomas de DOTAP (1,2-bis(oleiloxi)-3-(trimetilamonio)propano).

De forma similar, 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, dioleilfosfatidilcolina (DOPC), dioleilfosfatidilglicerol (DOPG), dioleilfosfatidiletanolamina (DOPE), entre otros. Estos materiales también pueden mezclarse con los materiales de partida DOTMA y DOTAP en proporciones apropiadas. Los procedimientos para preparar liposomas usando estos materiales se conocen bien en la técnica.

Los liposomas pueden comprender vesículas multilaminares (MLV), pequeñas vesículas unilaminares (SUV), o grandes vesículas unilaminares (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) *Proc. Natl. Acad. Sci. 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) *Proc. Natl. Acad. Sci. USA* 76: 3348; Enoch & Strittmatter (1979) *Proc. Natl. Acad. Sci. USA* 76: 145; Fraley (1980) *J. Biol. Chem.* (1980) 255: 10431; Szoka & Papahadjopoulos (1978) *Proc. Natl. Acad. Sci. USA* 75: 145; y Schaefer-Ridder (1982) *Science* 215: 166.

#### E. Lipoproteínas

Además, pueden incluirse lipoproteínas con el polinucleótido o polipéptido a administrar. Los ejemplos de lipoproteínas a utilizar incluyen: quilomicrones, HDL, IDL, LDL y VLDL. También pueden usarse mutantes, fragmentos o condensaciones de estas proteínas. Además, pueden usarse modificaciones de lipoproteínas que se dan en la naturaleza, tales como LDL acetilada. Estas lipoproteínas pueden dirigir la administración de polinucleótidos a células que expresan receptores de lipoproteínas. De preferencia, si se incluyen lipoproteínas con el polinucleótido a administrarse, no se incluye ningún otro ligando objetivo en la composición.

Las lipoproteínas de origen natural comprenden un lípido y una porción proteica. La porción proteica se conoce como apoproteínas. Actualmente, se han aislado e identificado las apoproteínas A, B, C, D y E. Al menos dos de éstas contienen varias proteínas, denominadas mediante números romanos AI, AII, AIV; CI, CII, CIII.

Una lipoproteína puede comprender más de una apoproteína. Por ejemplo, los quilomicrones que se dan en la naturaleza comprenden A, B, C y E, a lo largo del tiempo estas lipoproteínas pierden A y adquieren apoproteínas C y E. VLDL comprende apoproteínas A, B, C y E, LDL comprende la apoproteína B; y HDL comprende las apoproteínas A, C y E.

Los aminoácidos de estas apoproteínas se conocen y se describen 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) *Proc Natl Acad Sci USA* 77: 2465; y Utermann (1984) *Hum Genet* 65: 232.

Las lipoproteínas contienen diversos lípidos incluyendo, triglicéridos, colesterol (libre y en forma de ésteres) y fosfolípidos. La composición de los lípidos varía en las lipoproteínas que se dan en la naturaleza. Por ejemplo, los quilomicrones comprenden principalmente triglicéridos. Una descripción más detallada del contenido de lípidos de las lipoproteínas que se dan en la naturaleza puede encontrarse, por ejemplo, en *Meth. Enzymol.* 128 (1986). La composición de los lípidos se selecciona para ayudar en la conformación de la apoproteína para la actividad de unión al receptor. La composición de lípidos también puede seleccionarse para facilitar la interacción hidrófoba y la asociación con la molécula de unión al polinucleótido.

Las lipoproteínas de origen natural pueden aislarse a partir del suero mediante, por ejemplo, ultracentrifugación. Tales procedimientos se describen en *Meth Enzymol (supra)*; Pitas (1980) *J. Biochem.* 255: 5454-5460 y Mahey (1979) *J Clin. Invest* 64: 743-750. Las lipoproteínas también pueden producirse mediante procedimientos *in vitro* o recombinantes mediante expresión de los genes de la 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 pueden además adquirirse de proveedores comerciales, tales como Biomedical Technologies, Inc., Stoughton, Massachusetts, EEUU. Puede encontrarse descripción adicional de las lipoproteínas en Zuckermann y col., documento WO 98/06437.

#### F. Agentes policatiónicos

Los agentes policatiónicos pueden incluirse, con o sin lipoproteína, en una composición con el polinucleótido o polipéptido a administrar deseado.

Los agentes policatiónicos, típicamente, muestran una carga neta positiva a un pH pertinente fisiológico y son capaces de neutralizar la carga eléctrica de ácidos nucleicos para facilitar la administración a una ubicación deseada. Estos agentes tienen tanto aplicaciones *in vitro*, *ex vivo* como *in vivo*. Los agentes policatiónicos pueden usarse para administrar ácidos nucleicos a un sujeto vivo por vía intramuscular, por vía subcutánea, etc.

Los siguientes son ejemplos de polipéptidos útiles como agentes policatiónicos: polilisina, poliarginina, poliornitina, y protamina. Otros ejemplos incluyen histonas, protaminas, albúmina de suero humano, proteínas de unión al ADN, proteínas cromosómicas no histonas, proteínas de envoltura a partir de virus de ADN tales como (X174, los

factores transcripcionales también contienen dominios que se unen al ADN y por consiguiente pueden ser útiles como agentes de fusión de ácidos nucleicos). Brevemente, los factores transcripcionales tales como C/CEBP, c-jun, c-fos, AP-1, 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.

5

Los agentes policatiónicos orgánicos incluyen: espermina, espermidina y putrescina.

Las dimensiones y las propiedades físicas de un agente policatiónico pueden extrapolarse a partir de la lista anterior para construir otros agentes policatiónicos de polipéptidos o para producir agentes policatiónicos sintéticos.

10

Los agentes policatiónicos sintéticos que son útiles incluyen, por ejemplo, DEAE-dextrano, polibreno, lipofectin™, y lipofectAMINE™ que son monómeros que forman complejos policatiónicos cuando se combinan con polinucleótidos o polipéptidos.

### 15 *Ensayos de Inmunodiagnóstico*

Los antígenos de Neisseria de la invención pueden usarse en inmunoensayos para detectar niveles de anticuerpos (o, por el contrario, pueden usarse anticuerpos anti-Neisseria para detectar niveles de antígenos). Los inmunoensayos basados en antígenos recombinantes, bien definidos, pueden desarrollarse para sustituir procedimientos de diagnóstico invasivos. Pueden detectarse anticuerpos para proteínas de Neisseria en muestras biológicas, incluyen por ejemplo, muestras de sangre o de suero. El diseño de los inmunoensayos está sometido a una gran cantidad de variación, y se conoce una variedad de estos en la técnica. Los protocolos para el inmunoensayo pueden basarse, por ejemplo, en competición, o en reacción directa, o en ensayos de tipo sándwich. Los protocolos pueden además, por ejemplo, usar soportes sólidos, o pueden ser mediante inmunoprecipitación. La mayoría de los ensayos implican el uso de un anticuerpo o polipéptido marcado; los marcadores pueden ser, por ejemplo, fluorescentes, quimioluminiscentes, radioactivos o moléculas de colorantes. También se conocen los ensayos que amplifican las señales a partir de la sonda; de los cuales son ejemplos los ensayos que utilizan biotina y avidina, e inmunoensayos marcados con enzimas, tales como ensayos ELISA.

Se construyen kits adecuados para inmunodiagnóstico y que contienen los reactivos marcados adecuados empaquetando los materiales adecuados, incluyendo las composiciones de la invención, en recipientes adecuados, junto con los reactivos y materiales restantes (por ejemplo, tampones adecuados, soluciones salinas, etc.) necesarios para realizar el ensayo, así como una serie adecuada de instrucciones para el ensayo.

### 35 *Hibridación de Ácidos Nucleicos*

“Hibridación” se refiere a la asociación de dos secuencias de ácidos nucleicos entre sí mediante enlaces de hidrógeno. Típicamente, una secuencia se fijará a un soporte sólido y la otra estará libre en disolución. A continuación, las dos secuencias se colocarán en contacto entre sí en condiciones que favorecen la formación de enlaces de hidrógeno. Los factores que afectan a esta formación de enlaces incluyen: el tipo y el volumen del disolvente, la temperatura de la reacción; el tiempo de hibridación; la agitación; agentes para bloquear la unión no específica de la secuencia de 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 después de la hibridación. Véase Sambrook y col. [*supra*] Volumen 2, capítulo 9, páginas 9.47 a 9.57.

“Rigurosidad” se refiere a condiciones en una reacción de hibridación que favorecen la asociación de secuencias muy similares por encima de la de secuencias que son diferentes. Por ejemplo, la combinación de temperatura y concentración de sales debería seleccionarse de modo que esté aproximadamente de 120 hasta 200°C por debajo de la  $T_f$  calculada para el híbrido estudiado. La temperatura y las condiciones de las sales pueden determinarse a menudo de forma empírica en experimentos previos en los que muestras del ADN genómico inmovilizadas en filtros se hibridan con la secuencia de interés y a continuación 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) a estudiar puede variar en una magnitud de 10, entre 0,1 y 1  $\mu\text{g}$  para un plásmido o fago digerido a  $10^{-9}$  a  $10^{-8}$  g a partir de una única copia del gen en un genoma eucariota altamente complejo. Para polinucleótidos de menor complejidad, pueden usarse tiempos de transferencia, de hibridación, y de exposición sustancialmente más cortos, una menor cantidad de polinucleótidos de partida, y actividad específica menor de las sondas. Por ejemplo, un gen de levadura de copia única puede detectarse con un tiempo de exposición de solamente 1 hora empezando con 1  $\mu\text{g}$  de ADN de levadura, transfiriendo durante dos horas, e hibridando durante 4-8 horas con una sonda de  $10^8$  cpm/ $\mu\text{g}$ . Para un gen de mamífero de copia única un enfoque conservador debería comenzar con 10  $\mu\text{g}$  de ADN, transferir durante la noche, e hibridar durante la noche en presencia de sulfato de dextrano al 10% usando una sonda de más de  $10^8$  cpm/ $\mu\text{g}$ , dando como resultado un tiempo de exposición de aproximadamente 24 horas.

## ES 2 333 071 T3

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

$$T_f = 81 + 16,6(\log_{10}Ci) + 0,4[\%(G + C)] - 0,6(\% \text{ de formamida}) - 600/n - 1,5(\% \text{ de emparejamiento erróneo}).$$

donde  $C_i$  es la concentración de sales (iones monovalentes) y  $n$  es la longitud del híbrido en pares de bases (modificada ligeramente de Meinkoth & Wahi (1984) Anal. Biochem. 138: 267-284).

Al diseñar un experimento de hibridación, pueden alterarse convenientemente algunos factores que afectan a la hibridación del ácido nucleico. La temperatura de la hibridación y de los lavados y la concentración de sales durante el lavado son lo más sencillo de ajustar. A medida que aumenta la temperatura de la hibridación (es decir, la rigurosidad), es menos probable que se produzca la hibridación entre cadenas que son no homólogas, y como resultado, el fondo disminuye. Si la sonda radiomarcada no es completamente homóloga con el fragmento inmovilizado (como es frecuentemente el caso en familias de genes y en experimentos de hibridación interespecies), debe reducirse la temperatura de hibridación y el fondo aumentará. La temperatura de los lavados afecta a la intensidad de la banda de hibridación y al grado de fondo de forma similar. La rigurosidad de los lavados también aumenta disminuyendo la concentración de sales.

En general, las temperaturas de hibridación adecuadas en presencia de formamida al 50% son de 42°C para una sonda que es homóloga del 95 al 100% con el fragmento diana, 37°C para homología del 90% al 95%, y 32°C para el 85 al 90% de homología. Para homologías inferiores, se rebajaría el contenido de formamida y la temperatura se ajustaría de acuerdo con ello, usando la ecuación anterior. Si la homología entre la sonda y el fragmento diana no se conoce, el enfoque más sencillo es comenzar tanto con condiciones de hibridación como con condiciones de lavado que sean no rigurosas. Si se observan bandas no específicas o un fondo alto después de la autorradiografía, el filtro puede lavarse con rigurosidad alta y volver a exponerse. Si el tiempo requerido para la exposición hace este enfoque poco práctico, deben ensayarse en paralelo varias rigurosidades de hibridación y/o de lavado.

### *Ensayos de Sonda de Ácido Nucleico*

Procedimientos tales como PCR, ensayos de sonda de ADN ramificado, o técnicas de transferencia utilizando sondas de ácidos nucleicos de acuerdo con la invención pueden determinar la presencia de ADNc o ARNm. Se dice que una sonda "hibrida" con una secuencia de la invención si puede formar un dúplex o complejo de doble cadena, que es suficientemente estable para detectarse.

Las sondas de ácido nucleico hibridarán con las secuencias de nucleótidos de Neisseria de la invención (incluyendo cadenas sentido y antisentido). Aunque muchas secuencias de nucleótidos diferentes codificarán la secuencia de aminoácidos, se prefiere la secuencia de Neisseria nativa ya que es la verdadera secuencia presente en las células. El ARNm representa una secuencia codificadora y de este modo una sonda sería complementaria con la secuencia codificadora; el ADNc de cadena sencilla es complementario al ARNm, y por tanto una sonda de ADNc sería complementaria con la secuencia no codificadora.

La secuencia de la sonda no necesita ser idéntica a la secuencia de Neisseria (o su complemento) - algunas variaciones en la secuencia y en la longitud pueden conducir a sensibilidad del ensayo incrementada si la sonda del ácido nucleico puede formar un dúplex con nucleótidos diana, que pueden detectarse. Además, la sonda de ácido nucleico puede incluir nucleótidos adicionales para estabilizar el dúplex formado. Otra secuencia de Neisseria también puede ser útil como un marcador para detectar el dúplex formado. Por ejemplo, una secuencia de nucleótidos no complementaria puede unirse en el extremo 5' de la sonda, con el resto de la sonda siendo complementario a una secuencia de Neisseria. Como alternativa, pueden introducirse dentro de la sonda bases no complementarias o secuencias más largas, siempre que la secuencia de la sonda tenga la suficiente complementariedad con la secuencia de Neisseria para hibridar con ella y formar de este modo un dúplex que pueda detectarse.

La longitud y secuencia exacta de la sonda dependerán de las condiciones de hibridación, tales como temperatura, condiciones salinas y similares. Por ejemplo, para aplicaciones de diagnóstico, dependiendo de la complejidad de la secuencia del analito, la sonda de ácido nucleico contiene normalmente al menos 10-20 nucleótidos, de preferencia 15-25, y de más preferencia al menos 30 nucleótidos, aunque puede ser más corta que esto. Los cebadores cortos requieren generalmente temperaturas más bajas para formar complejos híbridos suficientemente estables con el molde.

Pueden producirse sondas mediante procedimientos sintéticos, tales como el procedimiento del triéster de Matteucci y col. [J. Am. Chem. Soc. (1981) 103: 3185], o según Urdea y col. [Proc. Natl. Acad. Sci. USA (1983) 80: 7461], o usando sintetizadores de oligonucleótidos automatizados disponibles comercialmente.

La naturaleza química de la sonda puede seleccionarse según la preferencia. Para algunas aplicaciones, son apropiados ADN o ARN. Para otras aplicaciones, pueden incorporarse modificaciones por ejemplo modificaciones en la estructura, tales como fosforotioatos o metilfosfonatos, pueden usarse para aumentar la semivida *in vivo*, alterar la afinidad por ARN, aumentar la resistencia a nucleasa, etc. [por ejemplo véase Agrawal y Lyer (1995) *Curr Opin Biotecnol* 6: 12-19; Agrawal (1996) *TIBTECH* 14: 376-387]; también pueden usarse análogos tales como ácidos péptidonucleicos [por ejemplo véase Corey (1997) *TIBTECH* 15: 224-229, Buchardt y col. (1993) *TIBTECH* 11: 384-386].

Como alternativa, la reacción en cadena de la polimerasa (PCR) es otro procedimiento bien conocido para detectar pequeñas cantidades de ácidos nucleicos diana. El ensayo se describe en: Mullis y col [Meth. Enzymol. (1987) 155: 335-350]; en las Patentes de EEUU 4.683.195; y 4.683.202. Dos nucleótidos "cebadores" hibridan con los ácidos nucleicos diana y se usan para iniciar la reacción. Los cebadores pueden comprender secuencia que no hibrida con la secuencia de la diana de amplificación (o con su complemento) para ayudar a la estabilidad del dúplex o, por ejemplo, para incorporar un sitio de restricción adecuado. Típicamente, tal secuencia flanqueará a la secuencia de *Neisseria* deseada.

Una polimerasa termoestable crea copias de ácidos nucleicos diana a partir de los cebadores usando los ácidos nucleicos diana originales como un molde. Después de generarse una cantidad umbral de ácidos nucleicos diana mediante la polimerasa, pueden detectarse mediante procedimientos más convencionales, tales como transferencias Southern. Cuando se usa el procedimiento de transferencia Southern, la sonda marcada hibridará con la secuencia de *Neisseria* (o su complemento).

Además, pueden detectarse ARNm o ADNc mediante técnicas de transferencia tradicionales descritas en Sambrook y col. [*supra*]. El ARNm, o el ADNc generado a partir de ARNm usando una enzima polimerasa, puede purificarse y separarse usando electroforesis en gel. Los ácidos nucleicos en el gel se transfieren después a un soporte sólido, tal como nitrocelulosa. El soporte sólido se expone a una sonda marcada y a continuación se lava para eliminar la sonda sin hibridar. A continuación, se detectan los dúplex que contienen la sonda marcada. Típicamente, la sonda se marca con un resto radioactivo.

### 30 Breve descripción de los dibujos

La figura 1 muestra datos bioquímicos y análisis de secuencias que pertenecen al Ejemplo 1 con ORF 40. M1 y M2 son marcadores de peso molecular. Las flechas indican la posición del producto recombinante principal. En el ensayo bactericida los resultados: un diamante (♦) muestra datos preinmunes; un triángulo (▲) muestra datos de control de GST; un círculo (●) muestra datos con la proteína recombinante de *N. meningitidis*. Los análisis computerizados muestran un gráfico de hidrofilia (superior), un gráfico de índice antigénico (medio) y un análisis de regiones AMPHI (inferior). El programa AMPHI se ha usado para predecir epítomos de células T [Gao y col. (1989) *J. Immunol.* 143: 3007; Roberts y col. (1996) *AIDS Res Hum Retrovir* 12: 593; Quakyi y col. (1992) *Scand J Immunol* supl. 11: 9] y está disponible en el paquete de programas Protean de DNASTAR, Inc (1228 South Park Street, Madison, Wisconsin 53715 EEUU).

#### Ejemplo

Los ejemplos describen secuencias de ácidos nucleicos que se han identificado en *N. meningitidis* junto con sus productos de traducción supuestos. No todas las secuencias de ácidos nucleicos son completas es decir, codifican menos de la proteína de tipo salvaje en su longitud total. Se cree actualmente que ninguna de las secuencias de ADN descritas en el presente documento tienen homólogos significativos en *N. gonorrhoeae*.

Los ejemplos están generalmente en el siguiente formato:

- 50 • una secuencia de nucleótidos que se ha identificado en *N. meningitidis* (cepa B)
- el supuesto producto de traducción de dicha secuencia
- 55 • un análisis computerizado del producto de traducción basado en comparaciones con bases de datos
- una secuencia genética y proteica correspondiente identificada en *N. meningitidis* (cepa A)
- una descripción de las características de las proteínas que indican que puede ser adecuadamente antigénica
- 60 • resultados del análisis bioquímico (expresión, purificación, ELISA, FACS, etc.).

El ejemplo incluye detalles de homología de secuencias entre especies y cepas. Las proteínas que son similares en secuencia son por lo general similares en estructura y función, y la homología indica con frecuencia un origen evolutivo común. La comparación con las secuencias de proteínas de función conocida se usa ampliamente como una guía para asignar la función supuesta de la proteína a una nueva secuencia y se ha demostrado que es particularmente útil en los análisis del genoma completo.

## ES 2 333 071 T3

Las comparaciones de secuencia se realizaron en NCBI (<http://www.ncbi.nlm.nih.gov>) usando los algoritmos BLAST, BLAST2, BLASTn, BLASTp, tBLASTn, BLASTx y tBLASTx [véase por ejemplo Altschul y col. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25: 2289-3402]. Las búsquedas se realizaron frente a las siguientes bases de datos: secuencias Gen-Bank+EMBL+DDBJ+PDB no redundantes y secuencias GenBank CDS translations+PBD+SwissProt+SPupdate+PIR no redundantes.

Los puntos dentro de las secuencias de nucleótidos (ninguno presente en el ejemplo) representarían nucleótidos que se han introducido arbitrariamente para mantener un marco de lectura. Del mismo modo, se eliminaron los nucleótidos doblemente subrayados (por ejemplo el residuo 40 en SEC ID 1). Las letras minúsculas (por ejemplo la posición 346 en SEC ID 1) representan ambigüedades que surgen durante el alineamiento de reacciones de secuenciación independientes (algunas de las secuencias de nucleótidos en los ejemplos derivan de combinar los resultados de dos o más experimentos).

Las secuencias de nucleótidos se exploraron en los seis marcos de lectura para predecir la presencia de dominios hidrófobos usando un algoritmo basado en los estudios estadísticos de Esposti y col. [Critical evaluation of the hydrophobicity of membrane proteins (1990) *Eur J Biochem* 190: 207-219]. Estos dominios representan potenciales regiones transmembranarias o secuencias líderes hidrófobas.

Los marcos de lectura abiertos se predijeron a partir de secuencias de nucleótidos fragmentadas usando el programa ORFFINDER (NCBI).

Pueden usarse diversas pruebas para evaluar *in vivo* la inmunogenicidad de las proteínas identificadas en los ejemplos. Por ejemplo, las proteínas pueden expresarse de manera recombinante y usarse para seleccionar los sueros de pacientes por inmunotransferencia. Una reacción positiva entre la proteína y el suero del paciente indica que el paciente previamente ha generado una respuesta inmune para la proteína en cuestión, es decir la proteína es un inmunógeno. Este procedimiento también puede usarse para identificar proteínas inmunodominantes.

La proteína recombinante puede también usarse de manera conveniente para preparar anticuerpos por ejemplo en un ratón. Estos pueden usarse para la confirmación directa de que una proteína está situada en la superficie celular. El anticuerpo marcado (por ejemplo marcación fluorescente para FACS) puede incubarse con la bacteria intacta y la presencia del marcador en la superficie bacteriana confirma la ubicación de la proteína.

En particular, se usaron los siguientes procedimientos (A) a (P) para expresar, purificar y caracterizar bioquímicamente las proteínas de la invención.

### 35 A) Preparación de ADN Cromosómico

Se cultivó la cepa 2996 de *N. meningitidis* hasta una fase exponencial en 100 ml de medio GC, se recogió mediante centrifugado y se resuspendió en 5 ml de tampón (sacarosa al 20%, Tris-HCl 50 mM, EDTA 50 mM, pH 8,0). Tras 10 minutos de incubación en hielo, las bacterias se lisaron añadiendo 10 ml de disolución de lisis (NaCl 50 mM, Na-Sarkosyl al 1%, proteinasa K 50 µg/ml) y la suspensión se incubó a 37°C durante 2 horas. Se realizaron 2 extracciones de fenol (equilibrado a pH 8) y una extracción con CHCl<sub>3</sub>/alcohol isoamílico (24:1). El ADN se precipitó mediante la adición de acetato sódico 0,3 M y 2 volúmenes de etanol, y se recogió mediante centrifugación. El sedimento se lavó una vez con etanol al 70% y se redisolvió en 4 ml de tampón (Tris-HCl 10 mM, EDTA 1 mM, pH 8). La concentración de ADN se midió leyendo la DO a 260 nm.

### B) Diseño de oligonucleótidos

Se diseñaron cebadores de oligonucleótidos sintéticos en base a la secuencia codificadora de cada ORF, usando (a) la secuencia de meningococos B cuando estaba disponible o (b) la secuencia de gonococos/meningococos A, adaptada para el uso preferente del codón de los meningococos según sea necesario. Se omitieron los péptidos señal predichos, deduciendo los cebadores 5' para la secuencia inmediatamente secuencia abajo en la cadena a partir de la secuencia líder predicha.

Los cebadores 5' incluyeron dos sitios de reconocimiento de enzimas de restricción (BamHI-NdeI, BamHI-NheI o EcoRI-NheI, dependiendo del patrón de restricción del gen de interés); los cebadores 3' incluyeron un sitio de restricción XhoI. Se estableció este procedimiento para dirigir la clonación de cada producto de amplificación (correspondiente a cada ORF) en dos sistemas de expresión diferentes: pGEX-KG (usando BamHI-XhoI o EcoRI-XhoI) y pET21 b+ (usando NdeI-XhoI o NheI-XhoI).

60 Cola del cebador del extremo 5': CGCGGATCCCATATG (BmHI-NdeI)

CGCGGATCCCGCTAGC (BmHI-NheI)

65 CCGGAATTCTAGCTAGC (EcoRI-NheI)

Cola del cebador del extremo 3': CCCGCTCGAG (XhoI)

## ES 2 333 071 T3

Así como contenían las secuencias de reconocimientos de enzimas de restricción, los cebadores incluyeron nucleótidos que hibridaron con la secuencia a amplificar. El número de nucleótidos que hibridaron dependió de la temperatura de fusión del cebador completo, y se determinó para cada cebador usando las fórmulas:

$$Tf = 4(G + C) + 2(A + T) \quad (\text{excluyendo la cola})$$

$$Tf = 64,9 + 0,41(\% \text{ de GC}) - 600/N \quad (\text{cebador completo})$$

La temperatura de fusión promedio de los oligos seleccionados fue de 65-70°C para el oligo completo y de 50-55°C para la región de hibridación sola.

La Tabla 1 muestra los cebadores en el sentido directo y en el sentido inverso usados para la amplificación. Los oligonucleótidos se sintetizaron usando un sintetizador Perkin Elmer 394 DNA/RNA, se eluyeron de las columnas en 2 ml de NH<sub>4</sub>OH y se desprotegeron por medio de 5 horas de incubación a 56°C. Los oligos precipitaron por medio de la adición de acetato de sodio 0,3 M y 2 volúmenes de etanol. Se centrifugaron las muestras y los sedimentos se resuspendieron en 100 µl o en 1,0 ml de agua. Se determinó la DO<sub>260</sub> usando un espectrofotómetro Perkin Elmer Lambda Bio y la concentración se ajustó a 2-10 pmol/µl.

### C) Amplificación

El protocolo de PCR convencional era el siguiente: se usaron 50-200 ng de ADN genómico como un molde en presencia de 20-40 µM de cada oligo, disolución de dNTP 400-800 µM, tampón de PCR 1x (incluyendo MgCl<sub>2</sub> 1,5 mM), 2,5 unidades de DNA polimerasa TaqI (usando Perkin Elmer AmpliTaq, GIBCO Platinum, DNA polimerasa Pwo, o Taq polimerasa Tahara Shuzo).

En algunos casos, la PCR se optimizaba mediante la adición de 10 µl de DMSO o 50 µl de betaína 2 M.

Después de un inicio con calor (añadiendo la polimerasa durante una incubación preliminar de 3 minutos de la mezcla completa a 95°C), cada muestra sufrió una amplificación de dos etapas. Los 5 primeros ciclos se realizaron usando la temperatura de hibridación que excluyó la cola de las enzimas de restricción del cebador, seguido por 30 ciclos realizados según la temperatura de hibridación de los oligos de longitud completa. Los ciclos se completaron con una etapa final de extensión de 10 minutos a 72°C.

Los ciclos convencionales fueron los siguientes:

	<b>Desnaturalización</b>	<b>Hibridación</b>	<b>Elongación</b>
<b>Primeros 5 ciclos</b>	30 segundos 95 °C	30 segundos 50-55 °C	30-60 segundos 72 °C
<b>Últimos 30 ciclos</b>	30 segundos 95 °C	30 segundos 65-70 °C	30-60 segundos 72 °C

Los tiempos de elongación variaron según la longitud del ORF a amplificar.

Las amplificaciones se realizaron usando un sistema Perkin Elmer GeneAmp PCR 9600 o un sistema Perkin Elmer GeneAmp PCR 2400. Para comprobar los resultados, se cargó 1/10 del volumen de amplificación en un gel de agarosa al 1-1,5% (p/v) y se comparó el tamaño de cada fragmento amplificado con un marcador de peso molecular de ADN.

El ADN amplificado se cargó directamente en un gel de agarosa al 1% o se precipitó en primer lugar con etanol y se resuspendió en un volumen adecuado para cargarse en un gel de agarosa al 1,0%. El fragmento de ADN correspondiente a la banda de tamaño correcto se eluyó a continuación y se purificó a partir del gel, usando el kit de Extracción de Gel de Qiagen, siguiendo el protocolo del fabricante. El volumen del fragmento de ADN era de 30 µl o 50 µl de agua o Tris 10 mM, pH 8,5.

## ES 2 333 071 T3

### D) Digestión de los fragmentos de PCR

El ADN purificado correspondiente al fragmento amplificado se separó en 2 alícuotas y se digirió dos veces con:

- 5 - *NdeI/XhoI* o *NheI/XhoI* para clonar en pET-21 b+ y posterior expresión de la proteína como una fusión His-tag C terminal.
- *BamHI/XhoI* o *EcoRI/XhoI* para clonar en pGEX-KG y posterior expresión de la proteína como una fusión GST N terminal.
- 10 - *EcoRI/PstI*, *EcoRI/SalI*, *SalI/PstI* para clonar en pGex-His y posterior expresión de la proteína como una fusión Histag N terminal.

15 Cada fragmento de ADN purificado se incubó (37°C durante 3 horas toda una noche) con 20 unidades de cada enzima de restricción (New England Biolabs) en un volumen final de 30 ó 40  $\mu$ l en presencia del tampón adecuado. A continuación los fragmentos digeridos se purificaron usando el kit de purificación QIAquick PCR siguiendo las instrucciones del fabricante y se eluyó en un volumen final de 30  $\mu$ l ó 50  $\mu$ l de agua o de Tris 10 mM, pH 8,5. La concentración final de ADN se determinó mediante electroforesis en gel de agarosa al 1,0% en presencia de un marcador de peso molecular valorado.

20

### E) Digestión de los vectores de clonación (*pET22B*, *pGEX-KG*, *pTRC-His A* y *pGEX-His*)

25 Se digirieron dos veces 10  $\mu$ g de plásmido con 50 unidades de cada enzima de restricción en 200  $\mu$ l de volumen de reacción en presencia del tampón apropiado mediante incubación durante la noche a 37°C. Después de cargar toda la digestión en un gel de agarosa al 1%, se purificó la banda correspondiente al vector digerido a partir del gel usando el kit de Extracción de Gel Qiaquick de Qiagen y el ADN se eluyó en 50  $\mu$ l de Tris-HCl 10 mM, pH 8,5. La concentración de ADN se evaluó midiendo la  $DO_{260}$  de la muestra y se ajustó a 50  $\mu$ g/ $\mu$ l. Se usó 1  $\mu$ l de plásmido para cada procedimiento de clonación.

30

El vector pGEX-His es un vector pGEX-2T modificado que lleva una región que codifica seis residuos de histidina secuencia arriba del sitio de escisión de trombina y que contiene el sitio de clonación múltiple del vector pTRC99 (Pharmacia).

35

### F) Clonación

40 Los fragmentos correspondientes a cada ORF, digeridos previamente y purificados, se unieron en pET22b y pGEX-KG. En un volumen final de 20  $\mu$ l, se unió una proporción molar de fragmento/vector 3:1 usando 0,5  $\mu$ l de T4 DNA ligasa NEB (400 unidades/ $\mu$ l), en presencia del tampón suministrado por el fabricante. La reacción se incubó a temperatura ambiente durante 3 horas. En algunos experimentos, la unión se realizó usando el "Rapid Ligation Kit" de Boheringer, siguiendo las instrucciones del fabricante.

45 Para introducir el plásmido recombinante en una cepa adecuada, se incubaron 100  $\mu$ l de células competentes DH5 de *E. coli* con la disolución de reacción de la ligasa durante 40 minutos en hielo, a continuación a 37°C durante 3 minutos, posteriormente, tras añadir 800  $\mu$ l de caldo LB, de nuevo a 37°C durante 20 minutos. Después se centrifugaron las células a velocidad máxima en una microcentrífuga Eppendorf y se resuspendieron en aproximadamente 200  $\mu$ l del sobrenadante. A continuación se plaqueó la suspensión en ampicilina LB (100 mg/ml).

50 La selección de los clones recombinantes se realizó cultivando 5 colonias seleccionadas al azar durante la noche a 37°C en 2 ml (clones pGEX o pTC) o en 5 ml (clones pET) en caldo LB + 100  $\mu$ g/ml de ampicilina. A continuación las células se centrifugaron y el ADN se extrajo usando el kit Qiagen QIAprep Spin Miniprep, siguiendo las instrucciones del fabricante, hasta un volumen final de 30  $\mu$ l. Se digirieron 5  $\mu$ l de cada minipreparación individual (aproximadamente 1 g) con *NdeI/XhoI* o con *BamHI/XhoI* y la digestión completa se cargó en un gel de agarosa al 1-1,5% (dependiendo del tamaño de inserto esperado), en paralelo con el marcador de peso molecular (DNA Ladder de 1 kb, GIBCO). La selección de los clones positivos se realizó de acuerdo con el tamaño correcto del inserto.

### 60 G) Expresión

Cada ORF clonado en el vector de expresión se transformó en la cepa adecuada para la expresión del producto de proteína recombinante. Se usó 1  $\mu$ l de cada construcción para transformar 30  $\mu$ l de BL21 de *E. coli* (vector pGEX), TOP 10 de *E. coli* (vector pTRC) o BL21-DE3 (pET) de *E. coli* como se describió anteriormente. En el caso del vector pGEX-His, se usó la misma cepa de *E. coli* (W3110) para la clonación inicial y la expresión. Se inocularon colonias recombinantes únicas en 2 ml de LB+Amp (100  $\mu$ g/ml), se incubaron a 37°C durante la noche, a continuación se diluyeron 1:30 en 20 ml de LB+Amp (100  $\mu$ g/ml) en matraces de 100 ml, asegurando de que la  $DO_{600}$  variaba entre 0,1 y 0,15. Los matraces se incubaron a 30°C en agitadores de baños de agua giratorios hasta que la DO indicaba

## ES 2 333 071 T3

un crecimiento exponencial adecuado para la inducción de la expresión (DO de 0,4-0,8 para vectores pET y pTRC; DO 0,8-1 para vectores pGEX y pGEX-His). Para los vectores pET, pTRC y pGEXHis, la expresión de la proteína se indujo mediante adición de IPTG 1 mM, mientras que en el caso del sistema de pGEX la concentración final de IPTG era 0,2 mM. Tras 3 horas de incubación a 30°C, se comprobó la concentración final de la muestra por medio de la DO.

5 Para comprobar la expresión, se extrajo 1 ml de cada muestra, se centrifugó en una microcentrífuga, el sedimento se resuspendió en PBS y se analizó mediante SDS-PAGE al 12% con tinción de azul de Coomassie. Toda la muestra se centrifugó a 6.000 g y el sedimento se resuspendió en PBS para uso posterior.

### 10 H) Purificación a gran escala de proteínas de fusión GST

Se cultivó una única colonia durante la noche a 37°C en placa de agar LB+Amp. Las bacterias se inocularon en 20 ml de cultivo líquido LB+Amp en un baño de agua con agitación y se cultivaron durante toda la noche. Las bacterias se diluyeron 1:30 en 600 ml de medio recién preparado y se dejaron crecer a la temperatura óptima (20-37°C) hasta una DO<sub>550</sub> de 0,8-1. Se indujo la expresión de las proteínas con IPTG 0,2 mM seguido por 3 horas de incubación. El cultivo se centrifugó a 8000 rpm a 4°C. El sobrenadante se descartó y el sedimento bacteriano se resuspendió en 7,5 ml de PBS frío. Las células se rompieron mediante sonicación en hielo durante 30 segundos a 40 W usando un sonicador Branson B-15, se congelaron y se descongelaron dos veces y se centrifugaron otra vez. El sobrenadante se recogió y se mezcló con 150 µl de resina Glutation-Sefarosa 4B (Pharmacia) (previamente lavada con PBS) y se incubaron a temperatura ambiente durante 30 minutos. La muestra se centrifugó a 700 g durante 5 minutos a 4°C. La resina se lavó dos veces con 10 ml de PBS frío durante 10 minutos, se resuspendió en 1 ml de PBS frío, y se cargó en una columna desechable. La resina se lavó dos veces con 2 ml de PBS frío hasta que el flujo a su través alcanzó una DO<sub>280</sub> de 0,02-0,06. La proteína de fusión GST se eluyó mediante la adición de 700 µl de tampón de elución de glutatión frío 10 mM reducido con glutatión, Tris-HCl 50 mM) y se recogieron las fracciones hasta que la DO<sub>280</sub> fue de 0,1. Se cargaron 25 21 µl de cada fracción en un gel SDS al 12% usando el patrón de peso molecular de SDS-PAGE de Biorad de amplio espectro (M1) (200, 116,25, 97,4, 66,2, 45, 31, 21,5, 14,4, 6,5 kDa) o el marcador Amersham Rainbow Marker (M2) (200, 66, 46, 30, 21,5, 14,3 kDa) como patrones. Como el PM de GST es de 26 kDa, este valor debe añadirse al PM de cada proteína de fusión GST.

### 30 I) Análisis de solubilidad de proteínas de fusión His

Para analizar la solubilidad de los productos de expresión de fusión His, se resuspendieron los sedimentos obtenidos de cultivos de 3 ml en tampón M1 [PBS 500 µl, pH 7,2]. Se añadieron 25 µl de lisozima (10 mg/ml) y las bacterias se incubaron durante 15 minutos a 4°C. Los sedimentos se sometieron a sonicación durante 30 segundos a 400 W usando un sonicador Branson B-15, se congelaron y descongelaron dos veces y a continuación se volvieron a separar en sedimento y sobrenadante por medio de una etapa de centrifugación. El sobrenadante se recogió y el sedimento se resuspendió en tampón M2 [urea 8 M, NaCl 0,5 M, imidazol 20 mM y NaH<sub>2</sub>PO<sub>4</sub> 0,1 M] y se incubó durante 3 a 4 horas a 4°C. Tras la centrifugación, el sobrenadante se recogió y el sedimento se resuspendió en tampón M3 [guanidinio-HCl 6 M, NaCl 0,5 M, imidazol 20 mM y NaH<sub>2</sub>PO<sub>4</sub> 0,1 M] durante la noche a 4°C. Los sobrenadantes de todas las etapas se analizaron mediante SDS-PAGE.

### 45 J) Purificación a gran escala de proteínas de fusión His

Se cultivó una única colonia durante la noche a 37°C en una placa de agar LB + Amp. Las bacterias se inocularon en 20 ml de cultivo líquido LB + Amp y se incubaron durante la noche en un agitador de baño de agua. Las bacterias se diluyeron 1:30 en 600 ml de medio recién preparado y se dejaron crecer a la temperatura óptima (20-37°C) a una DO<sub>550</sub> de 0,6-0,8. La expresión de las proteínas se indujo mediante la adición de IPTG 1 mM y el cultivo se incubó además durante tres horas. El cultivo se centrifugó a 8000 rpm a 4°C, el sobrenadante se descartó y el sedimento bacteriano se resuspendió en 7,5 ml de (i) tampón A frío (NaCl 300 mM, tampón fosfato 50 mM, imidazol 10 mM, pH 8) para proteínas solubles o (ii) tampón B (urea 8M, Tris-HCl 10 mM, tampón fosfato 100 mM, pH 8,8) para proteínas insolubles.

55 Las células se rompieron mediante sonicación en hielo durante 30 segundos a 40 W usando un sonicador Branson B-15, se congelaron y se descongelaron dos veces y se centrifugaron otra vez.

Para proteínas insolubles, el sobrenadante se conservó a -20°C, suspendiendo los sedimentos en 2 ml de tampón C (clorhidrato de guanidina 6 M, tampón fosfato 100 mM, Tris-HCl 10 mM, pH 7,5) y se trataron en un homogeneizador durante diez ciclos. El producto se centrifugó a 13000 rpm durante 40 minutos.

65 El sobrenadante se recogió y se mezcló con 150 µl de resina Ni<sup>2+</sup> (Pharmacia) (previamente lavada con tampón A o tampón B, según fuera adecuado) y se incubó a temperatura ambiente con agitación suave durante 30 minutos. La muestra se centrifugó a 700 g durante 5 minutos a 4°C. La resina se lavó dos veces con 10 ml de tampón A ó B durante 10 minutos, se resuspendió en 1 ml de tampón A ó B y se cargó en una columna desechable. La resina se lavó (i) a 4°C con 2 ml de tampón A frío o (ii) a temperatura ambiente con 2 ml de tampón B, hasta que el flujo a su través alcanzó una DO<sub>280</sub> de 0,02-0,06.

## ES 2 333 071 T3

La resina se lavó con (i) 2 ml de tampón de imidazol 20 mM frío (NaCl 300 mM, tampón fosfato 50 mM, imidazol 20 mM, pH 8) o (ii) tampón D (urea 8 M, Tris-HCl 10 mM, tampón fosfato 100 mM, pH 6,3) hasta que el flujo a su través alcanzó la  $DO_{280}$  de 0,02-0,06. La proteína de fusión His se eluyó mediante la adición de 700  $\mu$ l de (i) tampón de elución A frío (NaCl 300 mM, tampón fosfato 50 mM, imidazol 250 mM, pH 8) o (ii) tampón de elución B (urea 8 M, Tris-HCl 10 mM, tampón fosfato 100 mM, pH 4,5) y se recogieron las fracciones hasta que la  $DO_{280}$  fue de 0,1. Se cargaron 21  $\mu$ l de cada fracción en un gel SDS al 12%.

### K) Renaturalización de proteínas de fusión His

Se añadió glicerol al 10% a las proteínas desnaturalizadas. A continuación se diluyeron las proteínas a 20  $\mu$ g/ml usando tampón de diálisis I (glicerol al 10%, arginina 0,5 M, tampón fosfato 50 mM, glutatión reducido 5 mM, glutatión oxidado 0,5 mM, urea 2 M, pH 8,8) y se dializaron nuevamente frente al mismo tampón durante 12-14 horas a 4°C. La proteína se dializó otra vez frente a tampón de diálisis II (glicerol al 10%, arginina 0,5 M, tampón fosfato 50 mM, glutatión reducido 5 mM, glutatión oxidado 0,5 mM, pH 8,8) durante 12-14 horas a 4°C. La concentración de proteína se evaluó usando la siguiente fórmula:

$$\text{Proteína (mg/ml)} = (1,55 \times DO_{280}) - (0,76 \times DO_{260})$$

### L) Purificación a gran escala de proteínas de fusión His

Se indujeron 500 ml de cultivos bacterianos y se obtuvieron las proteínas de fusión solubles en tampón M1, M2 o M3 usando el procedimiento descrito anteriormente. Los extractos brutos se cargaron sobre una columna de superflujo de Ni-NTA (Quiagen) equilibrada con tampón M1, M2 o M3 según el tampón de solubilización de las proteínas de fusión. Se eluyó el material no unido lavando la columna con el mismo tampón. La proteína específica se eluyó con el tampón correspondiente que contenía imidazol 500 mM y se dializó frente al tampón correspondiente sin imidazol. Tras cada proceso se sanearon las columnas lavando con al menos dos volúmenes de columna de hidróxido de sodio 0,5 M y equilibrando nuevamente antes del siguiente uso.

### M) Inmunizaciones de ratones

Se usaron 20  $\mu$ g de cada proteína purificada para inmunizar ratones por vía intraperitoneal. Se inmunizaron ratones CD1 con adyuvante de Freund en los días 1, 21 y 42 y se controló la respuesta inmune en las muestras tomadas en el día 42.

### N) Ensayo ELISA (análisis de los sueros)

La cepa M7 de MenB no encapsulado se colocó en placas de agar chocolate y se incubó durante la noche a 37°C. Se recogieron colonias bacterianas a partir de las placas de agar usando un hisopo de dacrón estéril y se inocularon en 7 ml de caldo Mueller-Hinton (Difco) que contenía glucosa al 0,25%. El crecimiento bacteriano se controló cada 30 minutos siguiendo la  $DO_{620}$ . Se dejaron crecer las bacterias hasta que la DO alcanzó el valor de 0,3-0,4. El cultivo se centrifugó durante 10 minutos a 10.000 rpm. El sobrenadante se descartó y las bacterias se lavaron una vez con PBS, se resuspendieron en PBS que contenía formaldehído al 0,025% y se incubaron durante 2 horas a temperatura ambiente y después durante la noche a 4°C con agitación. Se añadieron 100  $\mu$ l de células bacterianas a cada pocillo de una placa Greiner de 96 pocillos y se incubaron durante toda la noche a 4°C. A continuación se lavaron los pocillos tres veces con tampón de lavado PBT (Tween 20 al 0,1% en PBS). Se añadieron 200  $\mu$ l de tampón de saturación (polivinilpirrolidona 10 al 2,7% en agua) a cada pocillo y las placas se incubaron durante 2 horas a 37°C. Se lavaron los pocillos tres veces con PBT. Se añadieron 200  $\mu$ l de suero diluido (tampón de dilución: BSA al 1%, Tween 20 al 0,1%,  $NaN_3$  al 0,1% en PBS) a cada pocillo y las placas se incubaron durante 90 minutos a 37°C. Los pocillos se lavaron tres veces con PBT. Se añadieron 100  $\mu$ l de suero anti-ratón de conejo conjugado con HRP (Dako) diluido 1:2000 en tampón de dilución a cada pocillo y las placas se incubaron durante 90 minutos a 37°C. Los pocillos se lavaron tres veces con tampón PBT. Se añadieron 100  $\mu$ l de tampón sustrato para HRP (25 ml de tampón citrato pH 5, 10 mg de O-fenildiamina y 10  $\mu$ l de  $H_2O$ ) a cada pocillo y las placas se dejaron a temperatura ambiente durante 20 minutos. Se añadieron 100  $\mu$ l de  $H_2SO_4$  a cada pocillo y se siguió la  $DO_{490}$ . El ensayo ELISA se consideró positivo cuando la  $DO_{490}$  era 2,5 veces la de los sueros pre-inmunes respectivos.

### O) Procedimiento de ensayo de Unión de Bacterias FACSscan

La cepa M7 de MenB encapsulados se plaqueó en placas de agar chocolate y se incubó durante la noche a 37°C. Se recogieron colonias bacterianas a partir de las placas de agar usando un hisopo de dracón estéril y se inocularon en 4 tubos que contenían 8 ml cada uno de caldo Mueller-Hinton (Difco) que contenía glucosa a 0,25%. El crecimiento bacteriano se controló cada 30 minutos siguiendo la  $DO_{620}$ . Se dejaron crecer las bacterias hasta que la DO alcanzó el valor de 0,35-0,5. El cultivo se centrifugó durante 10 minutos a 4.000 rpm. El sobrenadante se descartó y el sedimento

## ES 2 333 071 T3

se resuspendió en tampón de bloqueo (BSA al 1%, NaN<sub>3</sub> al 0,4%) y se centrifugó durante 5 minutos a 4.000 rpm. Las células se resuspendieron en tampón de bloqueo hasta alcanzar DO<sub>620</sub> de 0,07. Se añadieron 100 µl de células bacterianas a cada pocillo de una placa Costar de 96 pocillos. Se añadieron 100 µl de suero diluido (1:200) (en tampón de bloqueo) a cada pocillo y las placas se incubaron durante 2 horas a 4°C. Las células se centrifugaron durante 5 minutos a 4.000 rpm, se aspiró el sobrenadante y las células se lavaron mediante la adición de 200 µl/pocillo de tampón de bloqueo en cada pocillo. Se añadieron 100 µl de F(ab)<sub>2</sub> conjugado con R-ficoeritrina de cabra anti-ratón, diluido a 1:100, a cada pocillo y las placas se incubaron durante una hora a 4°C. Las células se sedimentaron mediante centrifugado a 4.000 rpm durante 5 minutos y se lavaron mediante adición de 200 µl/pocillo de tampón de bloqueo. El sobrenadante se aspiró y las células se resuspendieron en 200 µl/pocillo de PBS, formaldehído al 0,25%. Las muestras se transfirieron a tubos FACScan y se leyeron. Las condiciones para el ajuste del FACScan eran: FL1 activado, FL2 y FL3 desactivado, umbral de FSC-H: 92; voltaje de FSC PMT: E 02; SSC PMT: 474; aumentos de amplitud 7,1; FL-2 PMT: 539; valores de compensación: 0.

### 15 P) Ensayo bactericida

La cepa MC58 se cultivó durante la noche a 37°C en placas de agar chocolate. Se recogieron 5-7 colonias y se usaron para inocular 7 ml de caldo Mueller-Hinton. La suspensión se incubó a 37°C en un mezclador nutator y se dejó crecer hasta que la DO<sub>620</sub> era de 0,5-0,8. El cultivo se dividió en alícuotas en tubos Eppendorf estériles de 1,5 ml y se centrifugó durante 20 minutos a velocidad máxima en una microcentrífuga. El sedimento se lavó una vez en tampón de Gey (Gibco) y se resuspendió en el mismo tampón a una DO<sub>620</sub> de 0,5, se diluyó a 1:20000 en tampón de Gey y se almacenó a 25°C.

Se añadieron 50 µl de tampón Gey/BSA al 1% a cada pocillo de una placa de cultivo tisular de 96 pocillos. Se añadieron 25 µl de sueros de ratón diluidos (1:100 en tampón de Gey/BSA al 0,2%) a cada pocillo y la placa se incubó a 4°C. Se añadieron 25 µl de la suspensión bacteriana descrita anteriormente a cada pocillo. Se añadieron 25 µl de complemento de cría de conejo inactivado con calor (baño de agua de 56°C durante 30 minutos) o normal a cada pocillo. Inmediatamente después de la adición del complemento de cría de conejo, se colocaron 22 µl de cada muestra/pocillo en placas de agar Mueller-Hinton (tiempo 0). La placa de 96 pocillos se incubó durante 1 hora a 37°C con rotación y a continuación se plaquearon 22 µl de cada muestra/pocillo en placas de agar Mueller-Hinton (tiempo 1). Tras una incubación durante toda la noche se contaron las colonias correspondientes al tiempo 0 y al tiempo 1.

La Tabla II da un resumen de los resultados de clonación, expresión y purificación.

### 35 Ejemplo 1

La siguiente secuencia de ADN parcial se identificó en <SEC ID 1> de *N. meningitidis*:

```

40      1  . .ACACTGTTGT TTGCAACGGT TCAGGCAAGT GCTAACCAAT GAAGAGCAAG
      51  AAGAAGATTT ATATTTAGAC CCCGTACAAC GCCTGTTGC CGTGTGATA
     101  GTCAATTCCG ATAAAGAAGG CACGGGAGAA AAAGAAAAAG TAGAAGAAA
     151  TTCAGATTGG GCAGTATATT TCAACGAGAA AGGAGTACTA ACAGCCAGAG
     201  AAATCACCyT CAAAGCCGGC GACAACCTGA AAATCAAACA AAACGGCACA
45     251  AACTTCACCT ACTCGCTGAA AAAAGACCTC AcAGATCTGA CCAGTGTGG
     301  AACTGAAAAA TTATCGTTTA GCGCAAACGG CAATAAAGTC AACATcACAA
     351  GCGACACCAA AGGCTTGAAT TTTGCGAAAG AAACGGCTGG sACGAACGgC
     401  GACACCACGG TTCATCTGAA CGGTATTGGT TCGACTTTGA CCGATACGCT
     451  GCTGAATACC GGAGCGACCA CAAACGTAAC CAACGACAAC GTTACCGATG
50     501  ACGAGAAAAA ACGTGC GGCA AGCGTTAAAG ACGTATTAAA CGCTGGCTGG
     551  AACATTAAAG GCGTTAAACC CGGTACAACA GCTTCCGATA ACGTTGATTT
     601  CGTCCGCACT TACGACACAG TCGAGTTCTT GAGCGCAGAT ACGAAAACAA
     651  CGACTGTTAA TGTGGAAAGC AAAGACAACG GCAAGAAAAC CGAAGTTAAA
55     701  ATCGGTGCGA AGACTTCTGT TATTAAAGAA AAAGAC...

```

Esto corresponde a la secuencia de aminoácidos <SEC ID 2; ORF40>:

```

60      1  . .TLLFATVQAS ANQEEQEEDL YLDPVQRTVA VLVNSDKEG TGEKEKVEEN
      51  SDWAVYFNEK GVLTAAREITX KAGDNLKIKQ NGTNFTYSLK KDLTDLTSVG
     101  TEKLSFSANG NKVNITSDTK GLNFAKETAG TNGDTTVHLN GIGSTLTDTL
     151  LNTGATTNVT NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTASDNVDF
65     201  VRTYDTVEFL SADTKTTTVN VESKDNKKT EVKIGAKTSV IKEKD...

```

## ES 2 333 071 T3

Más trabajo reveló la secuencia completa de ADN <SEC ID 3>:

```

1   ATGAACAAAA TATACCGCAT CATTGGAAT AGTGCCCTCA ATGCCTGGGT
5   51  CGTCGTATCC GAGCTCACAC GCAACCACAC CAAACGCGCC TCCGCAACCG
    101  TGAAGACCGC CGTATTGGCG AACTGTGTGT TTGCAACGGT TCAGGCAAGT
    151  GCTAACAAATG AAGAGCAAGA AGAAGATTTA TATTTAGACC CCGTACAACG
    201  CACTGTTGCC GTGTTGATAG TCAATTCCGA TAAAGAAGGC ACGGGAGAAA
    251  AAGAAAAAGT  AGAAGAAAAT TCAGATTGGG CAGTATATTT CAACGAGAAA
10  301  GGAGTACTAA CAGCCAGAGA AATCACCTC  AAAGCCGGCG ACAACCTGAA
    351  AATCAAACAA AACGGCACAA ACTTACCTA  CTCGCTGAAA AAAGACCTCA
    401  CAGATCTGAC CAGTGTGGGA ACTGAAAAAT TATCGTTTAG CGCAAACGGC
    451  AATAAAGTCA ACATCACAAG CGACACCAA  GGCTTGAATT TTGCGAAAGA
    501  AACGGCTGGG ACGAACGGCG ACACCACGGT TCATCTGAAC GGTATTGGTT
15  551  CGACTTTGAC CGATACGCTG CTGAATACCG GAGCGACCAC AAACGTAACC
    601  AACGACAACG TTACCGATGA CGAGAAAAAA CGTGCGGCAA GCGTTAAAGA

    651  CGTATTA AAC GCTGGCTGGA ACATTAAGG CGTTAAACCC GGTACAACAG
20  701  CTTCCGATAA CGTTGATTTT GTCCGCACTT ACGACACAGT CGAGTTCTTG
    751  AGCGCAGATA CGAAAACAAC GACTGTTAAT GTGGAAAGCA AAGACAACGG
    801  CAAGAAAACC GAAGTTAAAA TCGGTGCGAA GACTTCTGTT ATTAAGAAA
    851  AAGACGGTAA GTTGGTTACT GGTAAAGACA AAGGCGAGAA TGTTCTTCT
    901  ACAGACGAAG GCGAAGGCTT AGTGACTGCA AAAGAAGTGA TTGATGCAGT
25  951  AAACAAGGCT GGTGGAGAA  TGAACAAC  AACCGCTAAT GGTCAAACAG
1001 GTCAAGCTGA CAAGTTTGAA ACCGTTACAT CAGGCACAAA TGTAACCTTT
1051 GCTAGTGGTA AAGGTACAAC TGCGACTGTA AGTAAAGATG ATCAAGGCAA
1101 CATCACTGTT ATGTATGATG TAAATGTCGG CGATGCCCTA AACGTC AATC
1151 AGCTGCAAAA CAGCGTTGG AATTTGGATT CCAAAGCGGT TGCAGGTTCT
30 1201 TCGGGCAAAG TCATCAGCGG CAATGTTTCG CCGAGCAAGG GAAAGATGGA
    1251 TGAAACCGTC AACATTAATG CCGGCAACA CATCGAGATT ACCCGCAACG
    1301 GTA AAAATAT CGACATCGCC ACTTCGATGA CCCCAGATT TTCAGCGTT
    1351 TCGCTCGGCG CGGGGCGGA TGCGCCCACT TTGAGCGTGG ATGGGGACGC
    1401 ATTGAATGTC GGCAGCAAGA AGGACAACA ACCCGTCCGC ATTACCAATG
35 1451 TCGCCCCGGG CGTTAAAGAG GGGGATGTTA CAAACGTCGC ACAACTTAAA
    1501 GGCGTGGCGC AAAACTTGAA CAACCGCATC GACAAATGTG ACGGCAACGC
    1551 GCGTGGCGGC ATCGCCCAAG CGATTGCAAC CGCAGGTCTG GTTCAGGCGT
    1601 ATTTGCCCGG CAAGAGTATG ATGGCGATCG GCGGCGGCAC TTATCGCGGC
    1651 GAAGCCGGTT ACGCCATCGG CTACTCCAGT ATTTCCGACG CCGGAAATTTG
40 1701 GATTATCAA  GGCACGGCTT CCGGCAATTC GCGCGGCCAT TTCGGTGCTT
    1751 CCGCATCTGT CGTTATCAG  TGGTAA

```

45 Esto corresponde a la secuencia de aminoácidos <SEC ID 4; ORF40-1>:

```

1   MNKIYRIWN SALNAWVVVS ELTRNHTKRA SATVKTAFLA TLLFATVQAS
50  51  ANNEEQEEDL YLDPVQRTVA VLIVNSDKEG TGEKEKVEEN SDWAVYFNEK
    101  GVLTAAREITL KAGDNLKIKQ NGTNFTYSLK KDLTDLTSVG TEKLSFSANG
    151  NKVNITSDTK GLNFAKETAG TNGDITVHLN GIGSTLDTL LNTGATTNVT
    201  NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTASDNVDF VRTYDTVEFL
    251  SADTKTTVN  VESKDNGKKT EVKIGAKTSV IKEKDGKLV T GKDKGENGSS
55  301  TDEGEGLVTA KEVIDAVNKA GWRMKTITAN GQTQADKFE TVTSGTNVTF
    351  ASGKGTATV  SKDDQGNITV MYDVNVGDAL NVNQLQNSGW NLDSKAVAGS
    401  SGKVISGNVS PSKGKMDTV  NINAGNNIEI TRNGKNIDIA TSMTPOFSSV
    451  SLGAGADAPT LSVGDALNV  GSKKDNKPVR ITNVAPGVKE GDVTNVAQLK
    501  GVAQNLNNRI DNVDGNARAG IAQAIATAGL VQAYLPGKSM MAIGGGTYRG
60  551  EAGYAIGYSS ISDGGNWI  IK GTASGNSRGH FGASASVGYQ W*

```

65

## ES 2 333 071 T3

Más trabajo identificó el gen correspondiente en la cepa A de *N. meningitidis* <SEC ID 5>:

```

5      1  ATGAACAAA  TATACCGCAT  CATTGGAAT  AGTGCCCTCA  ATGCCTGNGT
      51  CGCCGTATCC  GAGCTCACAC  GCAACCACAC  CAAACGCGCC  TCCGCAACCG
     101  TGAAGACCGC  CGTATTGGCG  AACTGTGTGT  TTGCAACGGT  TCAGGCGAAT
     151  GCTACCGATG  AAGATGAAGA  AGAAGAGTTA  GAATCCGTAC  AACGCTCTGT
     201  CGTAGGGAGC  ATTCAAGCCA  GTATGGAAGG  CAGCGGCGAA  TTGGAACGA
     251  TATCATTATC  AATGACTAAC  GACAGCAAGG  AATTGTAGA   CCCATACATA
     301  GTAGTTACCC  TCAAAGCCGG  CGACAACCTG  AAAATCAAAC  AAAACACCAA
     351  TGA AACACC  AATGCCAGTA  GCTTCACCTA  CTCGCTGAAA  AAAGACCTCA
     401  CAGCCTGAT  CAATGTTGAN  ACTGAAAAAT  TATCGTTTGG  CGCAAACGGC
     451  AAGAAAGTCA  ACATCATAAG  CGACACCAA   GGCTGAATT  TCGCGAAAGA
     501  AACGGCTGGG  ACGAACGGCG  ACACCACGGT  TCATCTGAAC  GGTATCGGTT
     551  CGACTTTGAC  CGATACGCTT  GCGGTTCTT  CTGCTTCTCA  CGTTGATGCG
     601  GGTAACCNAA  GTACACATTA  CACTCGTGCA  GCAAGTATTA  AGGATGTGTT
     651  GAATGCGGGT  TGAATATTA  AGGGTGTTAA  ANNNGGCTCA  ACAACTGGTC
     701  AATCAGAAAA  TGTCGATTTC  GTCCGCACTT  ACGACACAGT  CGAGTCTTG
     751  AGCGCAGATA  CGNAAACAAC  GACNGTTAAT  GTGGAAAGCA  AAGACAACGG
     801  CAAGAGAACC  GAAGTAAAA  TCGGTGCGAA  GACTTCTGTT  ATTAAGAAA
     851  AAGACGGTAA  GTTGGTTACT  GGTAAAGGCA  AAGGCGAGAA  TGGTCTTCT
     901  ACAGACGAAG  GCGAAGGCTT  AGTGACTGCA  AAAGAAGTGA  TTGATGCAGT
     951  AAACAAGGCT  GGTGGAGAA  TGAAACAAC  AACCCTAAT  GGTCAAACAG
    1001  GTC AAGCTGA  CAAGTTTGAA  ACCGTTACAT  CAGGCACAAA  TGTAACCTTT

    1451  ATGTCGCCCC  GGGCGTTAAA  GANGGGGATG  TTACAAACGT  CNCACAACCT
    1501  AAAGGCGTGG  CGCAAAACTT  GAACAACCGC  ATCGACAATG  TGGACGGCAA
    1551  CGCGCGTGCN  GGCATCGCCC  AAGCGATTGC  AACCGCAGGT  CTGGTTCAGG
    1601  CGTATCTGCC  CGGCAAGAGT  ATGATGGCGA  TCGGCGGCGG  CACTTATCGC
    1651  GGCGAAGCCG  GTTACGCCAT  CGCTACTCC  AGTATTTCCG  ACGGCGGAAA
    1701  TTGGATTATC  AAAGGCACGG  CTCCGGCAA  TTCGCGCGGC  CATTTCGGTG
    1751  CTTCCGCATC  TGTCGGTTAT  CAGTGGTAA

```

Esto codifica una proteína que tiene la secuencia de aminoácidos <SEC ID 6; ORF40a>:

```

      1  MNKIYRIWN  SALNAXVAVS  ELTRNHTKRA  SATVKTAVLA  TLLFATVQAN
     51  ATDEDEEEEL  ESVQRSVVGS  IQASMEGSGE  LETISLSMTN  DSKEFVDPYI
    101  VVTLKAGDNL  KIKQNTNENT  NASSFTYSLK  KDLTGLINX  TEKLSFGANG
    151  KKVNIISDTK  GLNFAKETAG  TNGD'TTVHLN  GIGSTLDTL  AGSSASHVDA
    201  GNXSTHYTRA  ASIKDVLNAG  WNIKGVKXGS  TTGQSENVDF  VRTYD'TVEFL
    251  SADTXTT'TVN  VESKDNGKRT  EVKIGAKTSV  IKEKDGKLV  T GKGKGENGSS
    301  TDEGEGLVTA  KEVIDAVNKA  GWRMKTTTAN  GQTGQADKFE  TVTSGTNVTF
    351  ASGKGTTATV  SKDDQGNITV  MYDVNVGDAL  NVNQLQNSGW  NLDSKAVAGS
    401  SGKVISGNVS  PSKGKMD'ETV  NINAGNNIEI  SRNGKNIDIA  TSMAPQFSSV
    451  SLGAGADAPT  LSV'DDEGALN  VGSKDANKPV  RITNVAPGVK  XGDVTNVXQL
    501  KGVAQN'LNR  IDNVDGNARA  GIAQAIATAG  LVQAYLPGKS  MMA'IGGGTYR
    551  GEAGYAIGYS  SISDGGNWII  KGTASGNSRG  HFGASASVGY  QW*

```

## ES 2 333 071 T3

La secuencia parcial de la cepa B identificada originalmente (ORF40) mostró una identidad del 65,7% sobre un solapamiento de 254 aa con ORF40a:

```

5
  orf40.pep                      10      20      30
                                TLLFATVQASANQEEQEEDLYLDPVQRTVA
  orf40a                          SALNAXVAVSELTRNHTKRASATVKTAVLATLLFATVQANATDEDEEEEL--ESVQRSV-
10                                20      30      40      50      60
                                40      50      60      70      80
  orf40.pep                      VLI VNSDKEGTGEKEKVEEN-SDWAVYFNEKGVLTAREITXKAGDNLKIKQN-----GT
  orf40a                          VGSIQASMEGSGELETISLSMTNDSKEFVDPYIV----VTLKAGDNLKIKQNTNENTNAS
15                                70      80      90      100     110     120
                                90      100     110     120     130     140
  orf40.pep                      NFTYSLKKDLTDLT SVGTEKLSF SANGNKVNITSDTKGLNFAKETAGTNGDTTVHNLGIG
  orf40a                          SFTYSLKKDLTGLIN VXT EKLSFGANGKKNII SDTKGLNFAKETAGTNGDTTVHNLGIG
20                                130     140     150     160     170     180
                                150     160     170     180     190     200
  orf40.pep                      STLTD TLLNTGATTNVTNDNVT DDEKKRAASVKDVLNAGWN IKGVKPGTTA--SDNVDFV
  orf40a                          STLTD TLAGSSAS-HVDAGNXST-HYTRAASIKDVLNAGWN IKGVKXGSTTGQSENVDFV
25                                190     200     210     220     230     240
                                210     220     230     240
  orf40.pep                      RTYDTVEFLSADTKTTTVNVESKDNGKKEVKIGAKTSVIKEKD
  orf40a                          RTYDTVEFLSADTXTTTVNVESKDNGKKEVKIGAKTSVIKEKDGKLV TGKKGGENGSST
30                                250     260     270     280     290     300
35

```

La secuencia completa de la cepa B (ORF40-1) y ORF40a muestran una identidad el 83,7% en un solapamiento de 601 aa:

```

40
  orf40-1.pep                    10      20      30      40      50      60
                                MNKIYRIIWN SALNAXVAVSELTRNHTKRASATVKTAVLATLLFATVQASANNEEQEEDL
  orf40a                          MNKIYRIIWN SALNAXVAVSELTRNHTKRASATVKTAVLATLLFATVQANATDEDEEEEL
45                                10      20      30      40      50      60
                                70      80      90      100     110     119
  orf40-1.pep                    YLDPVQRTVAVLI VNSDKEGTGEKEKVEEN-SDWAVYFNEKGVLTAREITLKAGDNLKIK
  orf40a                          : |||:| | :::: ||:| | : : : : | : : : : | ||| |||||
50
55
60
65

```

## ES 2 333 071 T3

```

orf40a      --ESVQRSV-VGSIQASMEGSGELETISLSMTNDSKEFVDPYIV----VTLKAGDNLKIK
              70      80      90      100      110

5  orf40-1.pep 120      130      140      150      160      170
    QN-----GTNFTYSLKKDLTDLTSSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGTNG
    ||      ::||| ||| ||| | :| ||| ||| :||| ||| ||| ||| ||| ||| ||| ||| |||
orf40a      QNTNENTNASSFTYSLKKDLTGLINVXTEKLSFGANGKKVNIISDTKGLNFAKETAGTNG
              120      130      140      150      160      170

10 orf40-1.pep 180      190      200      210      220      230
    DTTVHLNGIGSTLTDTLLNTGATTNVTNDNVDDEKKRAASVKDVLNAGWNKGVKPGTT
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
orf40a      DTTVHLNGIGSTLTDTLAGSSAS-HVDAGNXST-HYTRAASIKDVLNAGWNKGVKXGST
              180      190      200      210      220      230

15 orf40-1.pep 240      250      260      270      280      290
    A--SDNVDFVRTYDTVEFLSADTKTTTVNVESKDNKGKTEVKIGAKTSVIKEKDGKLVTG
    :  | : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
orf40a      TGQSENVDFVRTYDTVEFLSADTKTTTVNVESKDNKRTEVKIGAKTSVIKEKDGKLVTG
              240      250      260      270      280      290

20 orf40-1.pep 300      310      320      330      340      350
    KDKGENSSTDEGEGLVTAKEVIDAVNKAGWRMKT TTANGQTGQADKFETVTSNTVTF
    |  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
orf40a      KDKGENSSTDEGEGLVTAKEVIDAVNKAGWRMKT TTANGQTGQADKFETVTSNTVTF
              300      310      320      330      340      350

25 orf40-1.pep 360      370      380      390      400      410
    SCKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
orf40a      SCKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP
              360      370      380      390      400      410

30 orf40-1.pep 420      430      440      450      460      470
    SKGKMDETVNIAGNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSDVDG-ALNV
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
orf40a      SKGKMDETVNIAGNIEISRNGKNIDIATSMAPQFSSVSLGAGADAPTLSDVDEGALNV
              420      430      440      450      460      470

35 orf40-1.pep 480      490      500      510      520      530
    GSCKDNKPVRIITNVAPGVKEGDVTNVAQLKGVAQNLNRRIDNV DGNARAGIAQAIATAGL
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
orf40a      GSCKDANKPVRIITNVAPGVKXGDVTNVXQLKGVAQNLNRRIDNV DGNARAGIAQAIATAGL
              480      490      500      510      520      530

40 orf40-1.pep 540      550      560      570      580      590
    VQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWI IKG TASGNSRGHFGASASVGYQ
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
orf40a      VQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWI IKG TASGNSRGHFGASASVGYQ
              540      550      560      570      580      590

orf40-1.pep  WX
              ||
orf40a      WX

```

55 El análisis computerizado de estas secuencias de aminoácidos dio los siguientes resultados:

*Homología con la proteína Hsf codificada por el locus de fibrillas superficiales de H. influenzae tipo b (número de acceso U41852)*

60

65

## ES 2 333 071 T3

ORF40 y la proteína Hsf muestran una identidad de 54% en un solapamiento de 251 aa:

```

5      Orf40  1  TLLFATVQASANQEEQEEDLYLDPVQRTVAVLIVNSDXXXXXXXXXXXXXXXXNSDWAYVFNEK  60
      TLLFATVQA+A  E++E  LDPV RT VL +SD  NS+W +YF+ K
      Hsf    41  TLLFATVQANATDEDEE----LDPVVRTAPVLSFHSOKEGTGEKEVTE-NSNWGIYFDNK  95

      Orf40  61  GVL TAREITXKAGDNLKIKQN-----GTNFTYSLKKDLTDLTSVGTTEKLSFSANGNKVN 114
      GVL A IT KAGDNLKIKQN  ++FTYSLKKDLTDLTSV TEKLSF ANG+KV+
10     Hsf    96  GVLKAGAITLKAGDNLKIKQNTDESTNASSFTYSLKKDLTDLTSVATEKLSFGANGDKVD 155

      Orf40 115  ITS DTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGAXXXXXXXXXXXXXEKKRAAS 174
      ITSD GL AK G+ VHLNG+ STL D + NTG EK RAA+
15     Hsf   156  ITS DANGLKLAK-----TGNGNVHLNGLDSTLPDAVTNTGVLSSSFTPNVD-EKTRAAT 209

      Orf40 175  VKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVEKONGKKEVKI 234
      VKDVLNAGWNIK G ++VD V Y+ VEF++ D T V + +K+NGK TEVK
20     Hsf   210  VKDVLNAGWNIKAGKTAGGNVESVDLVSAYNNVEFITGDKNTLDVVLTAKENGKTTEVKF 269

      Orf40 235  GAKTSVIKEKD 245
      KTSVIKEKD
      Hsf   270  TPKTSVIKEKD 280
  
```

25

ORF40a también muestra homología con Hsf:

30

gi | 1666683 (U41852) producto del gen hsf [Haemophilus influenzae] Longitud = 2353

Puntuación = 153 (67,7 bits), Esperado = 1,5 e-116, Suma P(11) = 1,5 e-116

Identities = 33/36 (91%), Positivos = 34/36 (94%)

35

Incógnita: 16 VAVSELTRNHTKRASATVKTAVLATLLFATVQANAT 51  
 V VSELTR HTKRASATV+TAVLATLLFATVQANAT

Sujeto: 17 VVSELTRTHTKRASATVETAVLATLLFATVQANAT 52

40

Puntuación = 161 (71,2 bits), Esperado = 1,5 e-116, Suma P(11) = 1,5 e-116

Identities = 32/38 (84%), Positivos = 36/38 (94%)

45

Incógnita: 101 VTLKAGDNLKIKQNTNENTNASSFTYSLKKDLTGLINV 138  
 +TLKAGDNLKIKQNT+E+TNASSFTYSLKKDLT L +V

Sujeto: 103 ITLKAGDNLKIKQNTDESTNASSFTYSLKKDLTDLTSV 140

50

Puntuación = 110 (48,7 bits), Esperado = 1,5 e-116, Suma P(11) = 1,5 e-116

Identities = 21/29 (72%), Positivos = 25/29 (86%)

55

Incógnita: 138 VTEKLSFGANGKVNIIISDTKGLNFAKET 166  
 V++KLS G NG KVNI SDTKGLNFAK++

Sujeto: 1439 VSDKLSLGTNGNKVNITSDTKGLNFAKDS 1467

60

Puntuación = 85 (37,6 bits), Esperado = 1,5 e-116, Suma P(11) = 1,5 e-116

Identities = 18/32 (56%), Positivos = 20/32 (62%)

65

Incógnita: 169 TNGDTTVHLNGIGSTLTDTLAGSSASHVDAGN 200  
 T D +HLNGI STLTDTL S A+ GN

Sujeto: 1469 TGDDANIHLNGIASTLTDTLLNSGATTNLGGN 1500

## ES 2 333 071 T3

Puntuación = 92 (40,7 bits), Esperado = 1,5 e-116, Suma P(11) = 1,5 e-116

Identities = 16/19 (84%), Positivos = 19/19 (100%)

Incógnita: 206 RAASIKDVLNAGWNIKGVK 224  
RAAS+KDVLNAGWN++GVK

Sujeto: 1509 RAASVKDVLNAGWNVKGVK 1527

Puntuación = 90 (39,8 bits), Esperado = 1,5 e-116, Suma P(11) = 1,5 e-116

Identities = 17/28 (60%), Positivos = 20/28 (71%)

Incógnita: 226 STTGQSENVDFVRTYDTVEFLSADTTTT 253  
S Q EN+DFV TYDTV+F+S D TT

Sujeto: 1530 SANNQVENIDFVATYDTVDFVSGDKDTT 1557

Basándose en la homología con Hsf, se predijo que esta proteína de *N. meningitidis*, y sus epítomos, podrían ser útiles para vacunas o para diagnóstico.

Se clonó el ORF40-1 (61 kDa) en los vectores pET y pGex y se expresó en *E. coli*, como se describió anteriormente. Los productos de la expresión y purificación de proteínas se analizaron por SDS-PAGE. La Figura 1A muestra los resultados de purificación por afinidad de la proteína de fusión His, y la Figura 1B muestra los resultados de la expresión de la proteína de fusión GST en *E. coli*. La proteína de fusión His se usó para inmunizar ratones, cuyos sueros se usaron para el análisis FACS (Figura 1C), un ensayo bactericida (Figura 1D), y ELISA (resultado positivo). Estos experimentos confirman que ORF40-1 es una proteína expuesta de superficie, y que es un inmunógeno útil.

La Figura 1E muestra representaciones gráficas de hidrofilia, índice antigénico y regiones AMPHI para ORF40-1.

TABLA I

*Cebadores de PCR*

ORF	Cebador	Secuencia	Sitios de restricción
ORF40	Directo Inverso	CGCGGATCCCATATG-ACCGTGAAGACCGCC CCC <u>GCTCGAG</u> -CCACTGATAACCGACAGA	BamHI-NdeI XhoI

TABLA II

*Clonación, expresión y purificación*

ORF	PCR/clonación	Expresión de fusión His	Expresión de fusión GST	Purificación
orf 40	+	+	+	Fusión His

**REIVINDICACIONES**

- 5 1. Una proteína que comprende una secuencia de aminoácidos antigénica con homología de secuencia del 80% o superior con SEC ID N° 2 ó SEC ID N° 4, con la condición de que dicha proteína no sea (i) una proteína que tiene la secuencia de aminoácidos SEC ID N° 7, ó (ii) un fragmento de SEC ID N° 7.
- 10 2. La proteína de la reivindicación 1, que comprende una secuencia de aminoácidos con homología de secuencia del 80% o superior con SEC ID N° 2.
- 10 3. La proteína de la reivindicación 1, que comprende una secuencia de aminoácidos con homología de secuencia del 80% o superior con SEC ID N° 4.
- 15 4. La proteína de la reivindicación 2, que comprende una secuencia de aminoácidos con homología de secuencia del 90% o superior con SEC ID N° 2.
- 15 5. La proteína de la reivindicación 3, que comprende una secuencia de aminoácidos con homología de secuencia del 90% o superior con SEC ID N° 4.
- 20 6. La proteína de la reivindicación 4, que comprende una secuencia de aminoácidos con homología de secuencia del 95% o superior con SEC ID N° 2.
- 25 7. La proteína de la reivindicación 5, que comprende una secuencia de aminoácidos con homología de secuencia del 95% o superior con SEC ID N° 4.
- 25 8. La proteína de la reivindicación 6, que comprende una secuencia de aminoácidos con homología de secuencia del 99% o superior con SEC ID N° 2.
- 30 9. La proteína de la reivindicación 7, que comprende una secuencia de aminoácidos con homología de secuencia del 99% o superior con SEC ID N° 4.
- 30 10. La proteína de la reivindicación 1, que comprende la secuencia de aminoácidos SEC ID N° 2.
- 35 11. La proteína de la reivindicación 1, que comprende la secuencia de aminoácidos SEC ID N° 4.
- 35 12. La proteína de la reivindicación 1, que comprende la secuencia de aminoácidos SEC ID N° 6.
- 40 13. Una proteína que comprende un fragmento de al menos 20 aminoácidos consecutivos de la secuencia de aminoácidos SEC ID N° 2 ó SEC ID N° 4, en la que dicho fragmento comprende un epítipo de SEC ID N° 2 ó SEC ID N° 4, con la condición de que dicha proteína no sea (i) una proteína que tiene la secuencia de aminoácidos SEC ID N° 7, ó (ii) un fragmento de SEC ID N° 7.
- 40 14. La proteína de la reivindicación 13, en la que dicho fragmento comprende un epítipo de SEC ID N° 2.
- 45 15. La proteína de la reivindicación 13, que comprende un fragmento de al menos 20 aminoácidos consecutivos de la secuencia de aminoácidos SEC ID N° 2.
- 45 16. La proteína de la reivindicación 13, en la que dicho fragmento comprende un epítipo de SEC ID N° 4.
- 50 17. La proteína de la reivindicación 13, que comprende un fragmento de al menos 20 aminoácidos consecutivos de la secuencia de aminoácidos SEC ID N° 4.
- 50 18. La proteína de cualquiera de las reivindicaciones anteriores, en la que la proteína es una proteína de fusión.
- 55 19. La proteína de la reivindicación 18, que es una fusión con GST.
- 55 20. La proteína de la reivindicación 18, que es una fusión con seis residuos histidina.
- 60 21. Una proteína que comprende (i) una secuencia antigénica de aminoácidos con homología de secuencia del 80% o superior con SEC ID N° 2 o (ii) un fragmento de al menos 20 aminoácidos consecutivos de la secuencia de aminoácidos SEC ID N° 2 en la que dicho fragmento comprende un epítipo de SEC ID N° 2, en la que la proteína es una proteína de fusión que incluye seis residuos histidina.
- 60 22. La proteína de cualquiera de las reivindicaciones anteriores, que se prepara por medio de expresión recombinante.
- 65 23. La proteína de una cualquiera de las reivindicaciones 1 a 21, que se prepara por purificación a partir del cultivo celular.

## ES 2 333 071 T3

24. Un anticuerpo que se une a una proteína según una cualquiera de las reivindicaciones 1 a 17.
25. El anticuerpo de la reivindicación 24, que es monoclonal.
- 5 26. El anticuerpo de la reivindicación 24, que es policlonal.
27. Una molécula de ácido nucleico que codifica una proteína según una cualquiera de las reivindicaciones 1 a 21.
- 10 28. La molécula de ácido nucleico de la reivindicación 27, que comprende un fragmento de al menos 10 nucleótidos consecutivos de una secuencia de nucleótidos seleccionada del grupo constituido por SEC ID N°: 1, 3 y 5.
29. La molécula de ácido nucleico de la reivindicación 28, que comprende un fragmento de al menos 12 nucleótidos consecutivos de una secuencia de nucleótidos en dicho grupo.
- 15 30. La molécula de ácido nucleico de la reivindicación 28, que comprende un fragmento de al menos 14 nucleótidos consecutivos de una secuencia de nucleótidos en dicho grupo.
31. La molécula de ácido nucleico de la reivindicación 28, que comprende un fragmento de al menos 15 nucleótidos consecutivos de una secuencia de nucleótidos en dicho grupo.
- 20 32. La molécula de ácido nucleico de la reivindicación 28, que comprende un fragmento de al menos 18 nucleótidos consecutivos de una secuencia de nucleótidos en dicho grupo.
33. La molécula de ácido nucleico de la reivindicación 28, que comprende un fragmento de al menos 20 nucleótidos consecutivos de una secuencia de nucleótidos en dicho grupo.
- 25 34. La molécula de ácido nucleico de la reivindicación 28, que comprende un fragmento de al menos 25 nucleótidos consecutivos de una secuencia de nucleótidos en dicho grupo.
- 30 35. La molécula de ácido nucleico de la reivindicación 28, que comprende un fragmento de al menos 30 nucleótidos consecutivos de una secuencia de nucleótidos en dicho grupo.
36. La molécula de ácido nucleico de la reivindicación 28, que comprende un fragmento de al menos 35 nucleótidos consecutivos de una secuencia de nucleótidos en dicho grupo.
- 35 37. La molécula de ácido nucleico de la reivindicación 28, que comprende un fragmento de al menos 40 nucleótidos consecutivos de una secuencia de nucleótidos en dicho grupo.
- 40 38. La molécula de ácido nucleico de la reivindicación 28, que comprende la secuencia de nucleótidos SEC ID N° 1.
39. La molécula de ácido nucleico de la reivindicación 28, que comprende la secuencia de nucleótidos SEC ID N° 3.
- 45 40. La molécula de ácido nucleico de una cualquiera de las reivindicaciones 27 a 39, que es ADN.
41. Un vector que comprende el ácido nucleico de una cualquiera de las reivindicaciones 27 a 40.
- 50 42. El vector de la reivindicación 41, que es un vector de expresión.
43. Un vector de expresión bacteriano que codifica (i) una proteína que comprende una secuencia antigénica de aminoácidos con homología de secuencia del 80% o superior con SEC ID N° 2 ó (ii) una proteína que comprende un fragmento de al menos 20 aminoácidos consecutivos de la secuencia de aminoácidos SEC ID N° 2 en la que dicho fragmento comprende un epítipo de SEC ID N° 2; en el que el vector incluye una secuencia Shine-Dalgarno.
- 55 44. El vector de la reivindicación 42 ó reivindicación 43, en el que el vector puede estar integrado en un genoma bacteriano.
- 60 45. Un vector de expresión bacteriano que codifica (i) una proteína que comprende una secuencia antigénica de aminoácidos con homología de secuencia del 80% o superior con SEC ID N° 2 ó (ii) una proteína que comprende un fragmento de al menos 20 aminoácidos consecutivos de la secuencia de aminoácidos SEC ID N° 2 en la que dicho fragmento comprende un epítipo de SEC ID N° 2; en el que el vector es un vector de integración que puede estar integrado en un genoma bacteriano.
- 65 46. El vector de la reivindicación 44 ó reivindicación 45, en el que el vector incluye un marcador seleccionable para permitir la selección de una cepa bacteriana que se ha transformado.

## ES 2 333 071 T3

47. El vector de la reivindicación 46, en el que el marcador seleccionable es un gen que hace a la bacteria resistente a un fármaco seleccionado del grupo constituido por: ampicilina; cloramfenicol; eritromicina; kanamicina; neomicina; y tetraciclina.

5 48. Una célula huésped transformada con un vector de una cualquiera de las reivindicaciones 41 a 47.

49. La célula huésped de la reivindicación 48, que es una bacteria, una levadura o una célula de mamífero.

50. La célula huésped de la reivindicación 49, que es una bacteria.

10

51. La célula huésped de la reivindicación 50, en la que la bacteria es una *E. coli*.

52. Una célula huésped bacteriana transformada con un vector que codifica (i) una proteína que comprende una secuencia antigénica de aminoácidos con homología de secuencia del 80% o superior con SEC ID N° 2 ó (ii) una  
15 proteína que comprende un fragmento de al menos 20 aminoácidos consecutivos de la secuencia de aminoácidos SEC ID N° 2 en la que dicho fragmento comprende un epítipo de SEC ID N° 2; con la condición de que la célula huésped no sea una *Salmonella* o una *E. coli*.

20 53. La célula huésped de una cualquiera de las reivindicaciones 50 a 52, en la que el vector incluye un promotor bacteriano.

54. La célula huésped de la reivindicación 53, en la que el promotor bacteriano es de una enzima de ruta metabólica.

25 55. La célula huésped de una cualquiera de las reivindicaciones 50 a 54, en la que el vector es un vector de integración que puede estar integrado en el genoma bacteriano.

56. Una composición que comprende una proteína, una molécula de ácido nucleico, o un anticuerpo según una cualquiera de las reivindicaciones 1 a 40.

30 57. Una composición según la reivindicación 56 que es una composición de vacuna o una composición de diagnóstico.

58. La composición de la reivindicación 57, que es una composición de vacuna.

35 59. Una composición según la reivindicación 57, para uso como un medicamento.

60. La composición de una cualquiera de las reivindicaciones 56 a 59, que incluye un vehículo farmacéuticamente aceptable.

40 61. La composición de la reivindicación 60, en la que el vehículo comprende un polisacárido.

62. La composición de la reivindicación 60, en la que el vehículo comprende disolución salina.

45 63. La composición de la reivindicación 60, en la que el vehículo comprende una sustancia tamponadora del pH.

64. Una composición que comprende: (i) una proteína que comprende una secuencia antigénica de aminoácidos con homología de secuencia del 80% o superior con SEC ID N° 2 ó una proteína que comprende un fragmento de  
50 al menos 20 aminoácidos consecutivos de la secuencia de aminoácidos SEC ID N° 2 en la que dicho fragmento comprende un epítipo de SEC ID N° 2; y (ii) un vehículo farmacéuticamente aceptable que comprende disolución salina, un polisacárido, un ácido poliláctico, un ácido poliglicólico o una sustancia tamponadora del pH.

65. La composición de una cualquiera de las reivindicaciones 56 a 64, preparada como un inyectable.

55 66. La composición de una cualquiera de las reivindicaciones 56 a 64, preparada como una sólida adecuada para disolución en, o suspensión en, un vehículo líquido antes de la inyección.

67. Una composición que comprende: (i) una proteína que comprende una secuencia antigénica de aminoácidos con homología de secuencia del 80% o superior con SEC ID N° 2 ó una proteína que comprende un fragmento de  
60 al menos 20 aminoácidos consecutivos de la secuencia de aminoácidos SEC ID N° 2 en la que dicho fragmento comprende un epítipo de SEC ID N° 2; y (ii) un vehículo farmacéuticamente aceptable, en la que la composición es un inyectable.

68. La composición de una cualquiera de las reivindicaciones 56 a 67, que incluye un adyuvante.

65

69. La composición de la reivindicación 68, en la que el adyuvante se selecciona de: una sal de aluminio; una emulsión de aceite en agua; una saponina; o una citocina.

## ES 2 333 071 T3

70. La composición de la reivindicación 69, en la que el adyuvante es un hidróxido de aluminio o una sal de fosfato de aluminio.

5 71. Una composición que comprende: (i) una proteína que comprende una secuencia antigénica de aminoácidos con homología de secuencia del 80% o superior con SEC ID N° 2 ó una proteína que comprende un fragmento de al menos 20 aminoácidos consecutivos de la secuencia de aminoácidos SEC ID N° 2 en la que dicho fragmento comprende un epítipo de SEC ID N° 2; y (ii) un adyuvante seleccionado de hidróxido de aluminio, fosfato de aluminio o sulfato de aluminio.

10 72. La composición de una cualquiera de las reivindicaciones 56 a 59, que incluye quitosano.

73. La composición de una cualquiera de las reivindicaciones 56 a 59, que incluye poli(lactida-co-glicólido).

15 74. El uso de la proteína de una cualquiera de las reivindicaciones 1 a 21, en la fabricación de un medicamento para el tratamiento o la prevención de la infección por bacterias *Neisseria*, en particular *Neisseria meningitidis*.

20 75. El uso de (i) una proteína que comprende una secuencia antigénica de aminoácidos con homología de secuencia del 80% o superior con SEC ID N° 2, o (ii) una proteína que comprende un fragmento de al menos 20 aminoácidos consecutivos de la secuencia de aminoácidos SEC ID N° 2 en la que dicho fragmento comprende un epítipo de SEC ID N° 2, en la fabricación de un medicamento para el tratamiento o la prevención de la infección por *Neisseria meningitidis* serogrupo B.

25 76. Un procedimiento para producir la proteína de una cualquiera de las reivindicaciones 1 a 21, que comprende la etapa de cultivar una célula huésped de una cualquiera de las reivindicaciones 48 a 55 bajo condiciones que inducen la expresión de proteínas.

30

35

40

45

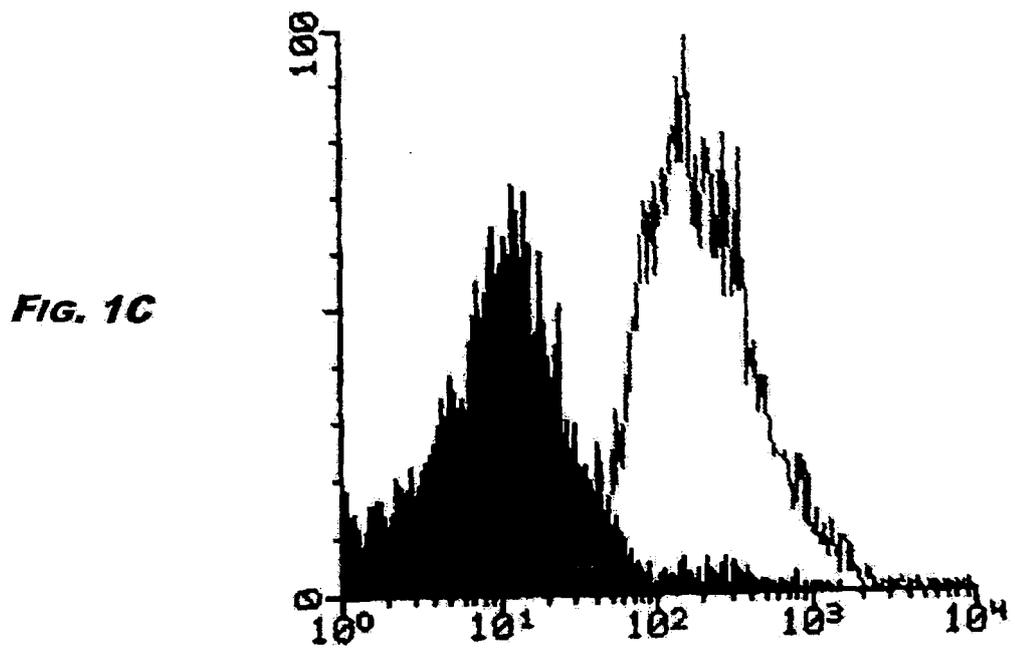
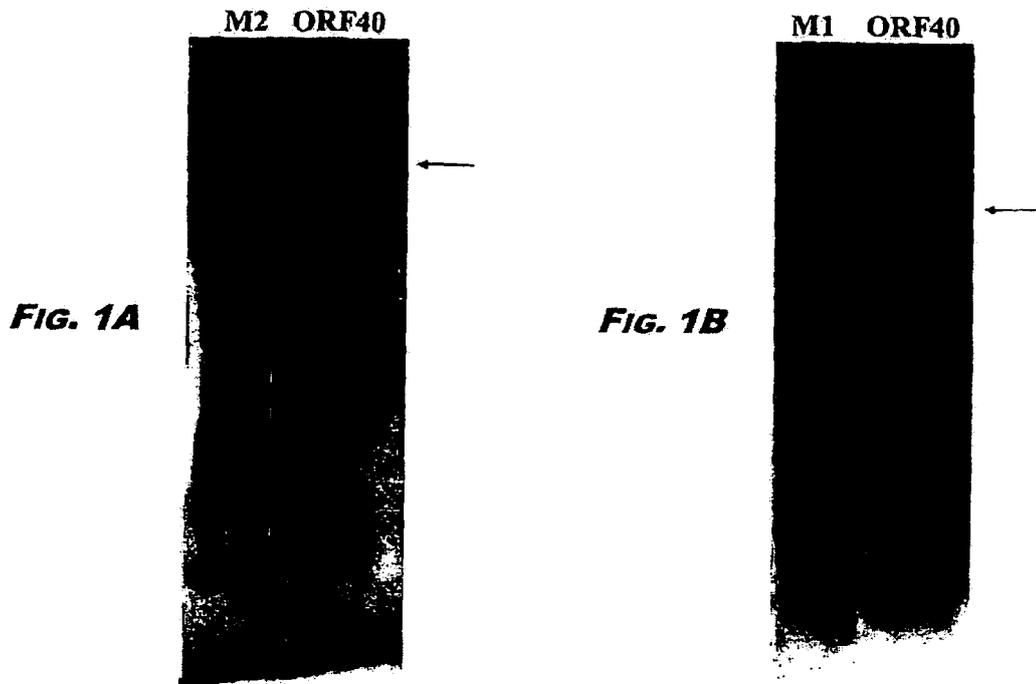
50

55

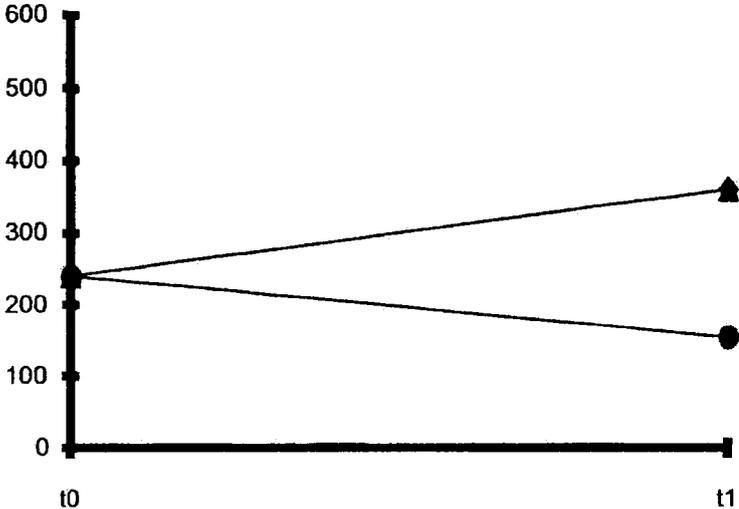
60

65

FIGURA 1



**FIG. 1D**



**FIG 1E**

