



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 278 381**

51 Int. Cl.:

A61K 39/00 (2006.01)

A61K 39/02 (2006.01)

A61K 39/40 (2006.01)

A61K 39/102 (2006.01)

A61K 39/395 (2006.01)

C07H 19/00 (2006.01)

C07K 16/00 (2006.01)

C12P 21/00 (2006.01)

C12P 21/08 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **95930245 .6**

86 Fecha de presentación : **16.08.1995**

87 Número de publicación de la solicitud: **0771213**

87 Fecha de publicación de la solicitud: **07.05.1997**

54 Título: **Proteínas de adherencia y penetración de *Haemophilus influenzae*.**

30 Prioridad: **25.08.1994 US 296791**

45 Fecha de publicación de la mención BOPI:
01.08.2007

45 Fecha de la publicación del folleto de la patente:
01.08.2007

73 Titular/es: **Washington University
1 Brookings Drive
St. Louis, Missouri 63130, US
The Board of Trustees of the Leland Stanford
Junior University**

72 Inventor/es: **St. Geme, Joseph W., III y
Falkow, Stanley**

74 Agente: **Ungría López, Javier**

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

Proteínas de adherencia y penetración de *Haemophilus influenzae*.

- 5 El gobierno tiene ciertos derechos en esta invención de acuerdo con los números de cesión AI23945 concedidos por el National Institute of Health.

Campo de la invención

- 10 La invención se refiere a las proteínas de adherencia y penetración de *Haemophilus*, a los ácidos nucleicos, y a las vacunas

Antecedentes de la invención

- 15 La mayoría de las enfermedades bacterianas comienzan con la colonización de una superficie mucosal concreta (Beachey *et al.*, 1981, J. Infect. Dis. 143:325-345). La colonización exitosa requiere que un organismo supere la limpieza mecánica de la superficie de la mucosa y evada la respuesta inmunitaria local. El procedimiento de colonización depende de factores microbianos especializados que promueven la unión a las células anfitrionas (Hultgren *et al.*, 1993 Cell, 73:887-901). En algunos casos el organismo colonizador entrará con posterioridad (invadirá) estas células y sobrevivirá intracelularmente (Falkow, 1991, Cell 65:1099-1102).

- Haemophilus influenzae* es un organismo comensal común del tracto respiratorio humano (Kuklinska and Kilian, 1984, Eur. J. Clin. Microbiol. 3:249-252). Es un organismo específico de humanos que normalmente reside en la nasofaringe humana y debe colonizar este sitio con el fin de evitar la extinción. Este microbio tiene numerosas estructuras superficiales capaces de promover el anclaje a las células anfitrionas (Guerina *et al.*, 1982, J. Infect. Dis. 146:564; Pichichero *et al.*, 1982, Lancet ii:960-962; St. Geme *et al.*, 1993, Proc. Natl. Acad. Sci. U.S.A. 90:2875-2879). Además, *H. influenzae* ha adquirido la capacidad de entrar y sobrevivir en estas células (Forsgren *et al.*, 1994, Infect. Immun. 62:673-679; St. Geme and Falkow, 1990, Infect. Immun. 58:4036-4044; St. Geme and Falkow, 1991, Infect. Immun. 59:1325-1333, Infect. Immun. 59:3366-3371). Como resultado, esta bacteria es una importante causa de enfermedad del tracto respiratorio tanto localizada como sistémica (Turk, 1984, J. Med. Microbiol. 18:1-16). Las cepas no tipificables, no encapsuladas representan la mayor parte de las enfermedades locales (Turk, 1984, *supra*); en contraste, las cepas de serotipo b, que expresan una cápsula compuesta por un polímero de ribosa y ribitol-5-fosfato (PRP), son responsables de más del 95% de los casos de las enfermedades sistémicas por *H. influenzae* (Turk, 1982, Clinical importance of *Haemophilus influenzae*, p. 3-9. En S.H. Sell and P.F. Wright (ed.), *Haemophilus influenzae* epidemiology, immunology, and prevention of disease. Elsevier/North-Holland Publishing Co., New York).

- La etapa inicial en la patogénesis de la enfermedad debida a *H. influenzae* implica la colonización de la mucosa respiratoria superior (Murphy *et al.*, 1987, J. Infect. Dis. 5:723-731). La colonización con una cepa concreta puede persistir durante semanas a meses, y la mayoría de los individuos permanecen asintomáticos durante todo este período (Spinola *et al.*, 1986, I. Infect. Dis. 154:100-109). No obstante, en ciertas circunstancias la colonización estará seguida de una continua diseminación en el tracto respiratorio, produciendo una enfermedad local en el oído medio, los senos, la conjuntiva, o los pulmones. Alternativamente, una bacteria ocasional penetrará en la barrera epitelial nasofaríngea y entrará en la corriente sanguínea.

- Las observaciones *in vitro* y los estudios en animales sugieren que los apéndices de la superficie bacteriana denominados pili (o fimbrias) juegan un importante papel en la colonización de *H. influenzae*. En 1982 dos grupos informaron sobre la correlación entre la piliación y el aumento de anclaje a células epiteliales orofaríngeas humanas y eritrocitos (Guerina *et al.*, *supra*; Pichichero *et al.*, *supra*). Otros investigadores han demostrado que los anticuerpos anti-pilus bloquean el anclaje *in vitro* por *H. influenzae* piliado (Forney *et al.*, 1992, J. Infect. Dis. 165:464-470; van Alphen *et al.*, 1988, Infect. Immun. 56:1800-1806). Recientemente Weber *et al.* inactivaron insersionalmente el gen estructural del pilus en una cepa de *H. influenzae* de tipo b y de ese modo eliminaron la expresión de los pili; el mutante resultante mostraba una capacidad de colonización reducida de monos adultos (Weber *et al.*, 1991, Infect. Immun. 59:4724-4728).

- Numerosos informes sugieren que los factores diferentes del pilus también facilitan la colonización por *Haemophilus*. Utilizando el modelo de cultivo en órgano nasofaríngeo humano, Farley *et al.* (1986, J. Infect. Dis. 161:274-280) y Loeb *et al.* (1988, Infect. Immun. 49:484-489) observaron que las cepas de tipo b no piliadas eran susceptibles de anclaje mucosal. Read y colaboradores realizaron observaciones similares al examinar cepas no tipificables en un modelo que emplea tejido de cornete nasal en cultivo de órganos (1991, J. Infect. Dis. 163:549-558). En el estudio de colonización de monos de Weber *et al.* (1991, *supra*), los organismos no piliados conservaban la capacidad de colonización, aunque a densidades reducidas; por otra parte, entre los monos infectados originalmente con la cepa piliada, virtualmente todos los organismos recuperados de la nasofaringe eran no piliados. Todas estas observaciones coinciden con el descubrimiento de que los productos aislados nasofaríngeos de niños colonizados con *H. influenzae* son frecuentemente no piliados (Mason *et al.*, 1985, Infect. Immun. 49:98-103; Brinton *et al.*, 1989, Pediatr. Infect. Dis. J. 8:554-561).

Los estudios previos han demostrado que *H. influenzae* son capaces de entrar (invadir) células epiteliales humanas cultivadas vía un mecanismo independiente de los pili (St. Geme and Falkow, 1990, *supra*; St. Geme and Falkow,

1991, *supra*). Aunque *H. influenzae* no es considerado generalmente un parásito intracelular, un informe reciente sugiere que estos descubrimientos *in vitro* pueden tener una correlación *in vivo* (Forsgren *et al.*, 1994, *supra*). Forsgren y colaboradores examinaron vegetaciones adenoides de 10 niños a los que se habían extirpado las vegetaciones adenoides debido a una otitis media secretora de larga duración o hipertrofia adenoidal. En los 10 casos había *H. influenzae* intracelular viable. La microscopía electrónica demostró que estos organismos estaban concentrados en el epitelio de la cripta reticular y en las células de tipo macrófago de la capa subepitelial del tejido. Una posibilidad es que la entrada de la bacteria en las células proporcione un mecanismo para la evasión de la respuesta inmune local, permitiendo de ese modo la persistencia en el tracto respiratorio.

De este modo, es deseable una vacuna para el tratamiento terapéutico y profiláctico de la infección por *Haemophilus*. Por consiguiente, un objeto de la presente invención es proporcionar proteínas de Adherencia y Penetración de *Haemophilus influenzae* recombinantes (HAP) y variantes de las mismas, y producir cantidades útiles de estas proteínas HAP utilizando técnicas de ADN recombinante.

Un objeto adicional de la invención es proporcionar ácidos nucleicos recombinantes que codifican proteínas HAP, y vectores de expresión y células anfitrionas que contienen el ácido nucleico que codifica la proteína HAP.

Un objeto adicional de la invención es proporcionar anticuerpos monoclonales para la diagnosis de la infección por *Haemophilus influenzae*.

Un objeto adicional de la invención es proporcionar métodos para producir las proteínas HAP, y una vacuna que comprende las proteínas HAP de la presente invención. También se proporcionan los métodos para el tratamiento terapéutico y profiláctico de la infección por *Haemophilus influenzae*.

Compendio de la invención

Según los objetos anteriores, la presente invención proporciona proteínas HAP recombinantes, concretamente una proteína de adherencia y penetración de *Haemophilus influenzae* recombinante, cuya proteína tiene una secuencia de aminoácidos con una homología de más del 80% con la secuencia de aminoácidos mostrada en la Figura 6, y ácidos nucleicos aislados o recombinantes que codifican las proteínas HAP de la presente invención. También se proporcionan vectores de expresión que comprenden ADN que codifica una proteína HAP de la invención conectada operablemente a ADN regulador transcripcional y traduccional, y células anfitrionas que contienen los vectores de expresión.

La invención también proporciona métodos para producir proteínas HAP de la invención que comprenden cultivar una célula anfitriona transformada con un vector de expresión y ocasionar la expresión del ácido nucleico que codifica la proteína HAP para producir una proteína HAP recombinante.

La invención también incluye vacunas para la infección por *Haemophilus influenzae* que comprenden una proteína HAP de la invención para uso profiláctico o terapéutico en la generación de una respuesta inmunitaria en el paciente. Los métodos para tratar o prevenir la infección por *Haemophilus influenzae* comprenden administrar una vacuna.

Breve descripción de los dibujos

Las Figuras 1A y 1B representan micrografías ópticas de cepas de *H. influenzae* DB117(pGJB103) y DB117 (pN187) incubadas con células epiteliales de Chang. Las bacterias fueron incubadas con una monocapa epitelial durante 30 minutos antes de aclarar y teñir con colorante Giemsa. Figura 1A: *H. influenzae* cepa DB117 que porta vector de clonación solo (pGJB103); Figura 1B: *H. influenzae* cepa DB117 que alberga el plásmido recombinante pN187. La barra representa 3,5 μ m.

Las Figuras 2A, 2B, 2C y 2D representan micrografías electrónicas de transmisión de sección delgada que demuestran la interacción entre las cepas de *H. influenzae* N187 y DB117(pN187) con células epiteliales de Chang. Las bacterias fueron incubadas con monocapas epiteliales durante cuatro horas antes de aclarar y transformar para el examen mediante microscopía electrónica de transmisión. Figura 2A: cepa N187 asociada con la superficie de la célula epitelial y presente en una localización intracelular; Figura 2B: *H. influenzae* DB117 (pN187) en íntimo contacto con la superficie de la célula epitelial; Figura 2C: cepa DB117 (pN187) en el proceso de entrada en la célula epitelial; Figura 2D: cepa DB117 (pN187) presente en una localización intracelular. La barra representa 1 μ m.

La Figura 3 representa perfiles de proteína de la membrana externa de diversas cepas. Las proteínas de la membrana externa fueron aisladas basándose en la insolubilidad del sarcosilo y resueltas sobre un gel de poliacrilamida-SDS al 10%. Las proteínas fueron visualizadas mediante tinción con Azul de Coomassie. Calle 1, *H. influenzae* cepa DB117 (pGJB103); calle 2, cepa DB117 (pN187); calle 3, cepa DB117 (pJS106); calle 4, *E. coli* HB101 (pGJB103); calle 5, HB101 (pN187). Obsérvense las proteínas novedosas en -160 kD y 45 kD marcadas por asteriscos en las calles 2 y 3.

La Figura 4 representa un mapa de restricción de pN187 y derivados y localizaciones de inserciones mini-Tn10 *kan*. pN187 es un derivado de pGJB103 que contiene un fragmento *Sau3AI* de 8,5 kb de ADN cromosómico de *H. influenzae* cepa N187. Las secuencias del vector se representan mediante cajas sombreadas. Las letras encima de la línea horizontal superior indican los sitios de las enzimas de restricción: Bg, *Bgl*II; C, *Cl*aI; E, *Eco*RI; P, *Pst*I. Los números y las piruletas encima de la línea horizontal superior muestran las posiciones de las inserciones mini-Tn10

kan; las piruletas vacías representan inserciones que no tienen efecto sobre la adherencia y la invasión, mientras las piruletas rellenas indican inserciones que eliminan la capacidad de pN187 para promover la asociación con monocapas epiteliales. La línea horizontal densa con flecha representa la localización del locus *hap* en pN187 y la dirección de la transcripción. (+): plásmidos recombinantes que promueven la adherencia y la invasión; (-): plásmidos recombinantes que fracasan al promover la adherencia y la invasión.

La Figura 5 representa la identificación de proteínas codificadas por plásmidos utilizando el sistema de expresión del bacteriófago T7. Las bacterias fueron radiomarcadas con metionina [S^{35}], y los productos lisados celulares completos fueron resueltos en un gel de poliacrilamida-SDS al 10%. Las proteínas fueron visualizadas mediante autoradiografía. Calle 1, *E. coli* XL-1 Blue(pT7-7) no inducido; calle 2, XL-1 Blue(pT7-7) inducido con IPTG; calle 3, XL-1 Blue(pJS103) no inducido; calle 4, XL-1 Blue(pJS103) inducido con IPTG; calle 5, XL-1 Blue(pJS104) no inducido; calle 6, XL-1 Blue(pJS104) inducido con IPTG. Los plásmidos pJS103 y pJS104 son derivados de pT7-7 que contienen el fragmento *PstI* de 6,5 kb de pN187 en orientaciones opuestas. El asterisco indica la proteína sobre-expresada en XL-1 Blue(pJS104).

Las Figuras 6A, 6B, y 6C representan la secuencia de nucleótidos y la secuencia de aminoácidos pronosticada del gen *hap*. Las supuestas secuencias 5' -10 y -35 con respecto a la secuencia codificadora de *hap* están subrayadas; un supuesto terminador independiente de *rho* 3' con respecto al codón de terminación de *hap* está indicado con flechas invertidas. Los primeros 25 aminoácidos de la proteína, que están en cajas, representan la secuencia señal.

Las Figuras 7A, 7B, 7C, 7D, 7E, 7F, 7G, y 7H representan una comparación de la secuencia del producto *hap* y las proteasas IgA1 de *H. influenzae* clonadas. Se muestran las homologías de aminoácidos entre el producto del gen *hap* deducido y los productos del gen *iga* de *H. influenzae* HK368, HK61, HK393, y HK793. Los guiones indican espacios introducidos en las secuencias con el fin de obtener la máxima homología. Se muestra una secuencia consenso para las cinco proteínas sobre la línea inferior. El dominio catalítico de la proteasa de tipo serina conservado está subrayado, y la serina del sitio activo común está indicada por un asterisco. Las cisteínas conservadas también están indicadas por asteriscos.

La Figura 8 representa el análisis de actividad proteasa IgA1. Los sobrenadantes de cultivo fueron analizados en cuanto a la capacidad para escindir IgA1. Las mezclas de reacción fueron resueltas en un gel de poliacrilamida-SDS al 10% y después transferidas a una membrana de nitrocelulosa. La membrana fue sondeada con anticuerpo para la cadena pesada de IgA1 humana. Calle 1, *H. influenzae* cepa N187; calle 2, cepa DB117(pGJB103); calle 3, cepa DB117(pN187). Los patrones del producto de escisión sugieren que la cepa N187 contiene una proteasa IgA1 de tipo 2 mientras las cepas DB117(pGJB103) y DB117(pN187) contienen una enzima de tipo 1. La banda superior ~70-kD observada con los derivados de DB117 representa una cadena pesada de IgA1 intacta.

Las Figuras 9A y 9B representan el análisis Southern del ADN cromosómico de la cepa de *H. influenzae* N187, sondeando con *hap* versus *iga*. Los fragmentos de ADN fueron separados en un gel de agarosa al 0,7% y transferidos bidireccionalmente a membranas de nitrocelulosa antes de sondear o bien con *hap* o bien con *iga*. Calle 1, ADN cromosómico de N187 digerido con *EcoRI*; calle 2, ADN cromosómico de N187 digerido con *BglII*; calle 3, ADN cromosómico de N187 digerido con *BamHI*; calle 4, el fragmento *ClaI-PstI* de 4,8 kb de pN167 que contiene el gen *hap* intacto. Figura 9A: Hibridación con el fragmento *ClaI-PstI* de 4,8 kb que contiene el gen *hap*; Figura 9B: hibridación con el gen *iga* de *H. influenzae* cepa Rd, llevado como un fragmento *ClaI-EcoRI* de 4,8 kb en pVD116.

La Figura 10 representa un gel de poliacrilamida-SDS de proteínas secretadas. Las bacterias se hicieron crecer hasta la fase log tardía, y los sobrenadantes de cultivo se hicieron precipitar con ácido tricloroacético y después se resolvió en un gel de poliacrilamida-SDS al 10%. Las proteínas fueron visualizadas mediante tinción con Azul de Coomassie. Calle 1, *H. influenzae* cepa DB117(pGJB103); calle 2, DB117(pN187); calle 3, DB117(pJS106); calle 4, DB117(pJS102); calle 5, DB117(pJS105); calle 6, DB117(Tn10-18); calle 7, DB117(Tn10-4'); calle 8, DB117(Tn10-30); calle 9, DB117(Tn10-16); calle 10, DB117(Tn10-10); calle 11, DB117(Tn10-8); calle 12, N187. El asterisco indica la proteína secretada de 110-kD codificada por *hap*.

Descripción detallada de la invención

La presente invención proporciona proteínas de Adherencia y Penetración (HAP) de *Haemophilus influenzae*.

Una proteína HAP puede ser identificada de diversas maneras. Un ácido nucleico HAP o una proteína HAP son identificados inicialmente por la homología de la secuencia de ácido nucleico y/o aminoácidos sustancial con las secuencias mostradas en la Figura 6. Semejante homología puede estar basada en las secuencias mostradas en la Figura 6. Semejante homología puede estar basada en la secuencia de ácido nucleico o aminoácidos total.

Las proteínas HAP de la presente invención tienen una homología limitada con las proteasas IgA1 de tipo serina de *Haemophilus influenzae* y *N. gonorrhoeae*. Esta homología, mostrada en la Figura 7, es de aproximadamente el 30-35% a nivel de aminoácidos, mostrando varios tramos una identidad del 55-60%, incluyendo los aminoácidos 457-549, 399-466, 572-622, y 233-261. No obstante, la homología entre la proteína HAP y la proteasa de IgA1 es considerablemente menor que la similitud entre las propias proteasas IgA1.

Además, la proteína HAP completa tiene homología con Tsh, una hemaglutinina expresada por una cepa de *E. coli* aviar (Provence and Curtiss 1994, Infect. Immun. 62:1369-1380). La mayor homología está en la mitad N-terminal de las proteínas, y la homología global es del 30,5%. La proteína HAP completa también tiene homología con la pertactina, una proteína de la membrana externa de 69 kD expresada por *B. pertussis*, mostrando la región media de las proteínas una homología del 39%. Finalmente, HAP tiene una homología del 34 - 52% con seis regiones de HpmA, una hemolisina independiente de calcio expresada por *Proteus mirabilis* (Uphoff and Welch, 1990, J. Bacteriol. 172:1206-1216).

Según se utiliza en la presente memoria, una proteína es una "proteína HAP" si la homología global de la secuencia de la proteína con la secuencia de aminoácidos mostrada en la Figura 6 es mayor del 80%. En algunas realizaciones la homología será tan alta como aproximadamente un 95 o 98%. Esta homología será determinada utilizando mecanismos normalizados en la técnica, tales como el programa de secuencias Best Fit descrito por Devereux *et al.*, Nucl. Acid Res. 12:387-395 (1984). La alineación puede incluir la introducción de espacios en las secuencias que se van a alinear. Además, para las secuencias que contienen más o menos aminoácidos que la proteína mostrada en la Figura 6, se entiende que el porcentaje de homología será determinado basándose en el número de aminoácidos homólogos en relación con el número total de aminoácidos. De este modo, por ejemplo, la homología de secuencias más cortas que la mostrada en la Figura 6, como se estudia más abajo, será determinada utilizando el número de aminoácidos de la secuencia más corta.

Las proteínas HAP de la presente invención pueden ser más cortas que la secuencia de aminoácidos mostrada en la Figura 6. Como se muestra en los Ejemplos, la proteína HAP puede experimentar una transformación post-traducciona similar a la observada para las proteasas IgA1 de tipo serina expresadas por *Haemophilus influenzae* y *N. gonorrhoeae*. Estas proteasas son sintetizadas como preproteínas con tres dominios funcionales: el péptido señal N-terminal, la proteasa, y un dominio coadyuvante C-terminal. Tras el movimiento de estas proteínas al espacio periplásmico, el dominio β carboxi terminal de la proenzima es insertado en la membrana externa, formando posiblemente un poro (Poulsen *et al.*, 1989, Infect. Immun. 57:3097-3105; Pohlner *et al.*, 1987, Nature (London). 325:458-462; Klauser *et al.*, 1992, EMBO J. 11:2327-2335; Klauser *et al.*, 1993, J. Mol. Biol. 234:579-593). Con posterioridad el extremo amino de la proteína es exportado a través de la membrana externa, y se produce una escisión autoproteolítica para dar lugar a la secreción de la proteasa de 100 a 106 kD madura. El dominio β C-terminal de 45 a 56 kD permanece asociado con la membrana externa después del evento de escisión. Como se muestra en los Ejemplos, el ácido nucleico de HAP está asociado con la expresión de una proteína de membrana externa de 106 kD. El producto génico secretado es una proteína de aproximadamente 110 kD, con la aparición simultánea de una proteína de membrana externa de 45 kD. La proteína de 45 kD parece corresponder a los aminoácidos desde aproximadamente el 960 a aproximadamente el 1394 de la Figura 6. Cualquiera de estas proteínas es considerada una proteína HAP para los fines de esta invención.

De este modo, en una realización preferida, se incluyen dentro de la definición de las proteínas HAP porciones o fragmentos de la secuencia mostrada en la Figura 6. Los fragmentos pueden ser fragmentos de la secuencia completa, la secuencia de 110 kD, o la secuencia de 45 kD. Generalmente, los fragmentos de la proteína HAP pueden oscilar de tamaño entre aproximadamente 10 aminoácidos y aproximadamente 1900 aminoácidos, prefiriéndose de aproximadamente 50 a aproximadamente 1000 aminoácidos, y también se prefiere de aproximadamente 100 a aproximadamente 500 aminoácidos. Los fragmentos particularmente preferidos son las secuencias únicas para HAP; estas secuencias tienen un uso concreto en la clonación de proteínas HAP de otros organismos o para generar anticuerpos específicos para las proteínas HAP. Las secuencias únicas son fácilmente identificadas por los expertos en la técnica después del examen de la secuencia de la proteína HAP y la comparación con otras proteínas; por ejemplo, mediante el examen de la alineación de la secuencia mostrado en la Figura 7. Por ejemplo, en comparación con las proteasas de IgA, las secuencias únicas incluyen, pero no están limitadas a, los aminoácidos 11-14, 16-22, 108-120, 155-164, 257-265, 281-288, 318-336, 345-353, 398-416, 684-693, 712-718, 753-761, 871-913, 935-953, 985-1008, 1023-1034, 1067-1076, 1440-1048, 1585-1592, 1631-1639, 1637-1648, 1735-1743, 1863-1871, 1882-1891, 1929-1941, y 1958-1966 (utilizando la numeración de la Figura 7). Los fragmentos de la proteína HAP que están incluidos en la definición de una proteína HAP incluyen truncamientos N- o C-terminales y delecciones que todavía permiten que la proteína sea biológicamente activa; por ejemplo, que todavía muestran actividad proteolítica en el caso de la secuencia de la supuesta proteasa de 110 kD. Además, cuando la proteína HAP va a ser utilizada para generar anticuerpos, por ejemplo como vacuna, la proteína HAP debe compartir al menos un epítipo o determinante con la proteína completa, la proteína de 110 kD o la proteína de 45 kD, mostradas en la Figura 6. En una realización preferida, el epítipo es único para la proteína HAP; esto es, los anticuerpos generados para un epítipo único muestran poca o ninguna reactividad cruzada con otras proteínas. Por "epítipo" o "determinante" se quiere significar en la presente memoria una porción de una proteína que generará y/o se unirá a un anticuerpo. Esto es, en la mayoría de los casos, los anticuerpos elaborados para una proteína HAP más pequeña serán capaces de unirse a la proteína completa.

En algunas realizaciones, el fragmento de la proteína HAP utilizado para generar anticuerpos es pequeño; de este modo, pueden ser utilizados como haptenos y acoplados a portadores proteicos para generar anticuerpos, como es sabido en la técnica.

Preferiblemente, los anticuerpos son generados para una porción de la proteína HAP que permanece anclada al organismo de *Haemophilus influenzae*. Por ejemplo, la proteína HAP puede ser utilizada para vacunar a un paciente para que produzca anticuerpos que tras la exposición al organismo de *Haemophilus influenzae* (por ejemplo durante una posterior infección) se unen al organismo y permiten una respuesta inmunitaria. De este modo, en una realización,

los anticuerpos son generados groseramente para el fragmento de 45 kD de la proteína HAP completa. Preferiblemente, los anticuerpos son generados para la porción del fragmento de 45 kD que está expuesta a la membrana externa.

En una realización alternativa, los anticuerpos se unen al fragmento de 110 kD secretado maduro. Por ejemplo, como se explica con detalle más abajo, las proteínas HAP de la presente invención pueden ser administradas terapéuticamente para generar anticuerpos neutralizadores para la supuesta proteasa de 110 kD, para disminuir los efectos indeseables del fragmento de 100 kD.

En el caso del ácido nucleico, la homología total de la secuencia de ácido nucleico guarda una proporción con la homología de los aminoácidos pero tiene en cuenta la degeneración del código genético y la preferencia codónica de diferentes organismos. Por consiguiente, la homología de la secuencia de ácido nucleico puede ser menor o mayor que la de la secuencia de proteínas. De este modo la homología de la secuencia de ácido nucleico en comparación con la secuencia de ácido nucleico de la Figura 6 es preferiblemente mayor del 40%, más preferiblemente mayor de aproximadamente el 60% y muy preferiblemente mayor del 80%. En algunas realizaciones la homología será tan alta como aproximadamente el 95 o 98%.

En una realización, la homología del ácido nucleico se determina por medio de estudios de hibridación. De este modo, por ejemplo, los ácidos nucleicos que hibridan en condiciones muy restrictivas con toda o parte de la secuencia de ácido nucleico mostrada en la Figura 6 son considerados genes de la proteína HAP. Las condiciones muy restrictivas incluyen lavados con 0,1XSSC a 65°C durante 2 horas.

Las proteínas HAP y los ácidos nucleicos de la presente invención son preferiblemente recombinantes. Según se utiliza en la presente memoria, “ácido nucleico” puede hacer referencia a ADN o ARN, o a moléculas que contienen tanto desoxirribonucleótidos como ribonucleótidos. Los ácidos nucleicos incluyen ADN genómico, ADNc y oligonucleótidos que incluyen ácidos nucleicos efectores y antisentido. Están específicamente incluidos en la definición de ácido nucleico los ácidos nucleicos antisentido. Un ácido nucleico antisentido hibridará con la correspondiente hebra no codificadora de la secuencia de ácido nucleico mostrada en la Figura 6, pero puede contener ribonucleótidos así como desoxirribonucleótidos. Generalmente, los ácidos nucleicos anti-sentido funcionan evitando la expresión del ARNm, de manera que la proteína HAP no es elaborada, o es elaborada a niveles reducidos. El ácido nucleico puede ser de doble hebra, de hebra sencilla, o contener porciones de secuencias tanto de doble hebra como de hebra sencilla. Por el término “ácido nucleico recombinante” en la presente memoria se quiere significar ácido nucleico, formado originalmente *in vitro* mediante la manipulación del ácido nucleico por endonucleasas, en una forma normalmente no encontrada en la naturaleza. De este modo un gen de la proteína HAP aislada, en forma lineal, o un vector de expresión formado *in vitro* ligando moléculas de ADN que no están normalmente unidas, son considerados ambos recombinantes para los fines de esta invención. Se entiende que una vez que el ácido nucleico recombinante es elaborado y reintroducido en una célula u organismo anfitrión, éste replicará no recombinantemente, esto es, utilizando la maquinaria celular *in vivo* de la célula anfitriona en lugar de las manipulaciones *in vitro*; no obstante, semejantes ácidos nucleicos, una vez producidos recombinantemente, aunque repliquen con posterioridad no recombinantemente, todavía son considerados recombinantes para los propósitos de la invención.

De un modo similar, una “proteína recombinante” es una proteína elaborada utilizando técnicas recombinantes, esto es por medio de la expresión de un ácido nucleico recombinante como se ha representado antes. Una proteína recombinante se distingue de una proteína de origen natural al menos en una o más características. Por ejemplo, la proteína puede ser aislada de algunas o de todas las proteínas y compuestos con los cuales está normalmente asociada en su anfitrión de tipo salvaje, o puede encontrarse en ausencia de las propias células anfitrionas. De este modo, la proteína puede ser parcialmente o sustancialmente purificada. La definición incluye la producción de una proteína HAP de un organismo en un organismo o célula anfitriona diferente. Alternativamente, la proteína puede ser elaborada a una concentración significativamente superior que la observada normalmente, por medio del uso de un promotor inducible o un promotor de alta expresión, de manera que la proteína se elabore a niveles de concentración incrementados. Alternativamente, la proteína puede estar en una forma no encontrada normalmente en la naturaleza, como en la adición de una etiqueta epitópica o en las sustituciones, inserciones y delecciones de aminoácidos.

También están incluidas en la definición de proteína HAP las proteínas HAP de otros organismos, que son clonadas y expresadas como se esboza más abajo.

En el caso de los ácidos nucleicos anti-sentido, un ácido nucleico anti-sentido se define como aquel que hibridará con toda o con parte de la secuencia no codificadora correspondiente mostrada en la Figura 6. Generalmente, las condiciones de hibridación utilizadas para la determinación de la hibridación anti-sentido serán las condiciones muy restrictivas, tales como 0,1 X SSC a 65°C.

Una vez que el ácido nucleico de la proteína HAP se ha identificado, puede ser clonado y, si fuera necesario, sus partes constituyentes recombinadas para formar el ácido nucleico de la proteína HAP completa. Una vez que se ha aislado de su fuente natural, p. ej., contenido en un plásmido u otro vector o separado por corte de allí en forma de un segmento de ácido nucleico lineal, el ácido nucleico de la proteína HAP recombinante puede ser utilizado adicionalmente como sonda para identificar y aislar otros ácidos nucleicos de proteínas HAP. También puede ser utilizado como ácido nucleico “precursor” para elaborar ácidos nucleicos y proteínas de la proteína HAP modificados o variantes.

Utilizando los ácidos nucleicos de la presente invención que codifican la proteína HAP, se elaboran una variedad de vectores de expresión. Los vectores de expresión pueden ser o bien vectores extracromosómicos auto-replicantes o bien vectores que se integran en el genoma del anfitrión. Generalmente, estos vectores de expresión incluyen un ácido nucleico regulador de la transcripción y de la traducción conectado operablemente al ácido nucleico que codifica la proteína HAP. “Conectado operablemente” en este contexto significa que el ADN regulador de la transcripción y la traducción está situado con respecto a la secuencia codificadora de la proteína HAP de tal manera que se inicie la transcripción. Generalmente, esto significará que las secuencias promotoras y de inicio o comienzo de la transcripción están situadas 5’ con respecto a la región codificadora de la proteína HAP. El ácido nucleico regulador de la transcripción y la traducción será generalmente apropiado para la célula anfitriona utilizada para expresar la proteína HAP; por ejemplo, las secuencias de ácido nucleico reguladoras de la transcripción y la traducción de *Bacillus* se utilizarán para expresar la proteína HAP en *Bacillus*. Se conocen numerosos tipos de vectores de expresión apropiados, y secuencias reguladoras adecuadas en la técnica para una variedad de células anfitrionas.

En general, las secuencias reguladoras de la transcripción y la traducción pueden incluir, pero no están limitadas a, secuencias promotoras, secuencias líder o señal, sitios de unión al ribosoma, secuencias de inicio y terminación de la traducción, y secuencias intensificadoras o activadoras. En una realización preferida, las secuencias reguladoras incluyen un promotor y secuencias de inicio y terminación de la transcripción.

Las secuencias promotoras codifican promotores constitutivos o inducibles. Los promotores pueden ser promotores de origen natural o promotores híbridos. Los promotores híbridos, que combinan elementos de más de un promotor, también son conocidos en la técnica, y son útiles en la presente invención.

Además, el vector de expresión puede comprender elementos adicionales, por ejemplo, el vector de expresión puede tener dos sistemas de replicación, permitiendo de este modo que sea mantenido en dos organismos, por ejemplo en células de mamífero o insecto para la expresión y en un anfitrión procariótico para la clonación y amplificación. Además, para los vectores de expresión integrantes, el vector de expresión contiene al menos una secuencia homóloga al genoma de la célula anfitriona, y preferiblemente dos secuencias homólogas que flanquean el constructo de expresión. El vector integrante puede estar dirigido a un locus específico en la célula anfitriona seleccionando la secuencia homóloga apropiada para la inclusión en el vector. Los constructos para los vectores integrantes son bien conocidos en la técnica.

Además, en una realización preferida, el vector de expresión contiene un gen marcador seleccionable para permitir la selección de las células anfitrionas transformadas. Los genes de selección son bien conocidos en la técnica y variarán con la célula anfitriona utilizada.

Las proteínas HAP de la presente invención son producidas cultivando una célula anfitriona transformada con un vector de expresión que contiene un ácido nucleico que codifica una proteína HAP, en las condiciones apropiadas para inducir o causar la expresión de la proteína HAP. Las condiciones apropiadas para la expresión de la proteína HAP variarán con la elección del vector de expresión y de la célula anfitriona, y serán fácilmente determinadas por un experto en la técnica por medio de la experimentación rutinaria. Por ejemplo, el uso de promotores constitutivos en el vector de expresión requerirá optimizar el crecimiento y la proliferación de la célula anfitriona, mientras el uso de un promotor inducible requiere las condiciones de crecimiento apropiadas para la inducción. Además, en algunas realizaciones, el momento de la cosecha es importante. Por ejemplo, los sistemas baculovirales utilizados en la expresión en células de insectos son virus líticos, y de este modo la selección del momento de la cosecha puede ser crucial para el rendimiento del producto.

Las células anfitrionas apropiadas incluyen células de levadura, de bacterias, de arqueobacterias, de hongos, y de insectos y animales. Son de particular interés las células de *Drosophila melanogaster*, de *Saccharomyces cerevisiae* y de otras levaduras, de *E. coli*, de *Bacillus subtilis*, las células SF9, las células C129, las células 293, las de *Neurospora*, BHK, CHO, COS, y HeLa, y las líneas celulares mieloides y linfoides de mamífero inmortalizadas.

En una realización preferida, las proteínas HAP son expresadas en sistemas bacterianos. Los sistemas de expresión bacterianos son bien conocidos en la técnica.

Un promotor bacteriano adecuado es cualquier secuencia de ácido nucleico capaz de unirse a la ARN polimerasa bacteriana e iniciar la transcripción aguas arriba (3’) de la secuencia codificadora de la proteína HAP en ARNm. Un promotor bacteriano tiene una región de inicio de la transcripción que se sitúa normalmente próxima al extremo 5’ de la secuencia codificadora. Esta región de inicio de la transcripción incluye típicamente un sitio de unión a la ARN polimerasa y un sitio de inicio de la transcripción. Las secuencias que codifican las enzimas de la ruta metabólica proporcionan secuencias promotoras particularmente útiles. Los ejemplos incluyen secuencias promotoras derivadas de enzimas metabolizadoras de azúcares, tales como galactosa, lactosa y maltosa, y secuencias derivadas de enzimas biosintéticas tales como el triptófano. Los promotores de bacteriófagos también pueden ser utilizados y son conocidos en la técnica. Además, también son útiles los promotores sintéticos; por ejemplo, el promotor tac es un híbrido de las secuencias *trp* y *lac*. Además, un promotor bacteriano puede incluir promotores de origen natural de origen no bacteriano que tienen la capacidad de unirse a la ARN polimerasa bacteriana e iniciar la transcripción.

Además de una secuencia promotora en funcionamiento, es deseable un sitio de unión al ribosoma eficaz. En *E. coli*, el sitio de unión al ribosoma se denomina secuencia de Shine-Delgarno (SD) e incluye un codón de inicia-

ción y una secuencia de 3-9 nucleótidos de longitud localizada 3-11 nucleótidos aguas arriba del codón de iniciación.

El vector de expresión también puede incluir una secuencia peptídica señal que proporciona la secreción de la proteína HAP a la bacteria. La secuencia señal codifica típicamente un péptido señal comprendido por aminoácidos hidrófobos que dirigen la secreción de la proteína desde la célula, como es bien sabido en la técnica. La proteína es secretada al medio de crecimiento (bacteria gram-positiva) o al espacio periplásmico, localizado entre la membrana interna y la externa de la célula (bacteria gram-negativa).

El vector de expresión bacteriano también puede incluir un gen marcador seleccionable para permitir la selección de cepas bacterianas que han sido transformadas. Los genes de selección adecuados incluyen genes que hacen a la bacteria resistente a los fármacos tales como ampicilina, cloranfenicol, eritromicina, kanamicina, neomicina y tetraciclina. Los marcadores seleccionables también incluyen genes biosintéticos, tales como los de las rutas biosintéticas de la histidina, el triptófano y la leucina.

Estos componentes son ensamblados en vectores de expresión. Los vectores de expresión para bacterias son bien conocidos en la técnica, e incluyen vectores para *Bacillus subtilis*, *E. coli*, *Streptococcus cremoris*, y *Streptococcus lividans*, entre otros.

Los vectores de expresión bacterianos son transformados en células anfitrionas bacterianas utilizando mecanismos bien conocidos en la técnica, tales como el tratamiento con cloruro de calcio, la electroporación, y otros.

En una realización, las proteínas HAP son producidas en células de insectos. Los vectores de expresión para la transformación de células de insectos, y en particular los vectores de expresión basados en baculovirus, son bien conocidos en la técnica. Brevemente, el baculovirus es un virus con ADN muy grande que produce su proteína de la envuelta a niveles muy elevados. Debido al tamaño del genoma baculoviral, los genes exógenos deben ser colocados en el genoma viral mediante recombinación. Por consiguiente, los componentes del sistema de expresión incluyen: un vector de transferencia, normalmente un plásmido bacteriano, que contiene tanto un fragmento del genoma de baculovirus, como un sitio de restricción conveniente para la inserción de la proteína HAP; 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 de baculovirus); y células anfitrionas de insecto y medios de crecimiento apropiados.

Los sistemas de expresión de mamífero también son conocidos en la técnica y se utilizan en una realización. Un promotor de mamífero es cualquier secuencia de ADN capaz de unirse a la ARN polimerasa de mamífero e iniciar la transcripción aguas abajo (3') de una secuencia codificadora para la proteína HAP en ARNm. Un promotor tendrá una región de inicio de la transcripción, que normalmente se sitúa próxima al extremo 5' de la secuencia codificadora, y una caja TATA, utilizando 25-30 pares de bases localizados aguas arriba del sitio de iniciación de la transcripción. Se piensa que la caja TATA dirige la ARN polimerasa II a comenzar la síntesis de ARN en el sitio correcto. Un promotor de mamífero también contendrá un elemento promotor aguas arriba, localizado típicamente de 100 a 200 pares de bases aguas arriba de la caja TATA. Un elemento promotor aguas arriba determina la velocidad a la cual se inicia la transcripción y puede actuar en cualquier orientación. En cuanto a los promotores de mamífero tienen un uso concreto los promotores de genes virales de mamífero, puesto que los genes virales a menudo son altamente expresados y tienen una amplia gama de anfitriones. Los ejemplos incluyen el promotor temprano de SV40, el promotor LTR del virus del tumor mamario de ratón, el promotor tardío principal de adenovirus, y el promotor del virus herpes simplex.

Típicamente, las secuencias de terminación de la transcripción y de poliadenilación reconocidas por las células de mamífero son regiones reguladoras localizadas 3' con respecto al codón de terminación de la traducción y de este modo, junto con los elementos promotores, flanquean la secuencia codificadora. El extremo 3' del ARN maduro está formado por la escisión y la poliadenilación post-traduccionales de sitio específico. Los ejemplos de las señales terminadora y de poliadenilación incluyen aquellas derivadas de SV40.

Los métodos de introducción de ácido nucleico exógeno en los anfitriones mamíferos, así como otros anfitriones, son bien conocidos en la técnica, y variarán con la célula anfitriona utilizada. Los mecanismos incluyen la transfección mediada por dextrano, la precipitación con fosfato de calcio, la transfección mediada por polibreno, la fusión de protoplastos, la electroporación, la encapsulación del polinucleótido o los polinucleótidos en liposomas, y la microinyección directa del ADN en los núcleos.

En una realización preferida, la proteína HAP es producida en células de levadura. Los sistemas de expresión de levadura son bien conocidos en la técnica, e incluyen vectores de expresión para *Saccharomyces cerevisiae*, *Candida albicans* y *C. maltosa*, *Hansenula polymorpha*, *Kluyveromyces fragilis* y *K. lactis*, *Pichia guilliermondii* y *P. pastoris*, *Schizosaccharomyces pombe*, y *Yarrowia lipolytica*. Las secuencias promotoras preferidas para la expresión en levadura incluyen el promotor GAL1 inducible, el promotor 10, los promotores de la alcohol deshidrogenasa, la enolasa, la glucocinasa, la glucosa-6-fosfato isomerasa, la gliceraldehído-3-fosfato-deshidrogenasa, la hexocinasa, la fosfofructocinasa, la 3-fosfoglicerato mutasa, la piruvato cinasa, y el gen de la fosfatasa ácida. Los marcadores seleccionables de levadura incluyen ADE2, HIS4, LEU2, TRP1, y ALG7, que confiere resistencia a la tunicamicina; el gen de resistencia G418, que confiere resistencia a G418; y el gen CUP1, que permite a la levadura crecer en presencia de iones cobre.

Una proteína HAP recombinante puede ser expresada intracelularmente o secretada. La proteína HAP también puede ser elaborada en forma de una proteína de fusión utilizando mecanismos bien conocidos en la técnica. De este modo, por ejemplo, si el epítipo deseado es pequeño, la proteína HAP puede ser fusionada a una proteína portadora para formar un inmunógeno. Alternativamente, la proteína HAP puede ser elaborada en forma de una proteína de fusión para incrementar la expresión.

También están incluidas dentro de la definición de proteínas HAP de la presente invención las variantes de la secuencia de aminoácidos. Estas variantes caen dentro de una o más de tres clases: variantes por sustitución, inserción o delección. Normalmente estas variantes se preparan mediante mutagénesis de sitio específico de los nucleótidos del ADN que codifica la proteína HAP, utilizando la mutagénesis por casete u otros mecanismos bien conocidos en la técnica, para producir ADN que codifica la variante, y después de eso expresar el ADN en un cultivo de células recombinantes como se ha esbozado antes. No obstante, los fragmentos de la proteína HAP variante que tienen hasta aproximadamente 100-150 restos pueden ser preparados mediante síntesis *in vitro* utilizando los mecanismos establecidos. Las variantes de la secuencia de aminoácidos se caracterizan por la naturaleza predeterminada de la variación, un rasgo que las aparta de la variación alélica o interespecie de origen natural de la secuencia de aminoácidos de la proteína HAP. Las variantes muestran típicamente la misma actividad biológica cualitativa que el análogo de origen natural, aunque también se pueden seleccionar variantes que tienen características modificadas como se esbozará más completamente más abajo.

Si bien el sitio o la región para introducir una variación en la secuencia de aminoácidos es predeterminado, la mutación *per se* no necesita ser predeterminada. Por ejemplo, con el fin de optimizar el funcionamiento de una mutación en un sitio dado, la mutagénesis al azar se puede llevar a cabo en el codón o la región diana y las variantes de la proteína HAP expresadas se pueden rastrear en cuanto a la combinación óptima de la actividad deseada. Las técnicas para elaborar mutaciones por sustitución en sitios predeterminados en un ADN que tiene una secuencia conocida son bien conocidas, por ejemplo, la mutagénesis del cebador M13. El rastreo de los mutantes se realiza utilizando el análisis de las actividades de la proteína HAP; por ejemplo, los genes de HAP mutados se colocan en cepas de delección de HAP y se someten a ensayo en cuanto a la actividad de HAP, como se describe en la presente memoria. La creación de cepas de delección, dada una secuencia génica, es conocida en la técnica. Por ejemplo, el ácido nucleico que codifica las variantes puede ser expresado en una cepa de *Haemophilus influenzae* deficiente en la proteína HAP, y la infectividad de la *Haemophilus influenzae* variante evaluada. Alternativamente, la proteína HAP variante puede ser expresada y sus características biológicas evaluadas, por ejemplo su actividad proteolítica.

Las sustituciones de aminoácidos son típicamente restos individuales; las inserciones serán normalmente del orden de aproximadamente 1 a 20 aminoácidos, aunque se pueden tolerar inserciones considerablemente más grandes. Las delecciones oscilan entre aproximadamente 1 y 30 restos, aunque en algunos casos las delecciones pueden ser mucho mayores, como por ejemplo cuando uno de los dominios de la proteína HAP es suprimido.

Las sustituciones, delecciones, inserciones o cualquier combinación de las mismas pueden ser utilizadas para llegar a un derivado final. Generalmente estos cambios se realizan en pocos aminoácidos para minimizar la alteración de la molécula. No obstante, se pueden tolerar cambios mayores en ciertas circunstancias.

Cuando se desean alteraciones pequeñas en las características de la proteína HAP, las sustituciones se realizan generalmente según el siguiente diagrama:

(Tabla pasa a página siguiente)

ES 2 278 381 T3

Diagrama I

<u>Resto Original</u>	<u>Sustituciones Ejemplares</u>
Ala	Ser
Arg	Lys
Asn	Gln, His
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Pro
His	Asn, Gln
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	Met, Leu, Tyr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp, Phe
Val	Ile, Leu

Los cambios sustanciales en la función o la identidad inmunológica se realizan seleccionando las sustituciones que son menos conservativas que las mostradas en el Diagrama I. Por ejemplo, se pueden realizar sustituciones que afectan más significativamente a: la estructura del esqueleto polipeptídico en el área de la alteración, por ejemplo la estructura alfa-helicoidal o en lámina beta; la carga o carácter hidrofóbico de la molécula en el sitio diana; o el volumen de la cadena lateral. Las sustituciones que se espera en general que produzcan los mayores cambios en las propiedades del polipéptido son aquellas en las que (a) un resto hidrófilo, v.g., serilo o treonilo, es sustituido con (o por) un resto hidrófobo, p. ej. leucilo, isoleucilo, fenilalanilo, valilo o alanilo; (b) una cisteína o prolina es sustituida con (o por) cualquier otro resto; un resto que tiene una cadena lateral electropositiva, p. ej. lisilo, arginilo, o histidilo, está sustituido con (o por) un resto electronegativo, p. ej. glutamilo o aspartilo; o (d) u resto que tiene una cadena lateral voluminosa, p. ej. fenilalanina, es sustituido con (o por) uno que no tiene una cadena lateral, p. ej. glicina.

Las variantes muestran típicamente la misma actividad biológica cualitativa y lograrán la misma respuesta inmunitaria que el análogo de origen natural, aunque las variantes también se seleccionan para modificar las características del polipéptido según se necesite. Alternativamente, la variante puede ser diseñada de manera que la actividad biológica de la proteína HAP resulte alterada. Por ejemplo, la actividad proteolítica del dominio de 110 kD más grande de la proteína HAP puede ser alterado, por medio de la sustitución de los aminoácidos del sitio activo. El supuesto dominio

catalítico de esta proteína es GDSGSPMF, correspondiendo la primera serina a la serina del sitio activo característica de las proteasas de tipo serina. Los restos del sitio activo pueden ser alterados individualmente o simultáneamente para disminuir o eliminar la actividad proteolítica. Esto se puede realizar para disminuir la toxicidad o los efectos secundarios de la vacuna. De un modo similar, el sitio de escisión entre el dominio de 45 kD y el dominio de 100 kD puede ser alterado, por ejemplo para eliminar el tratamiento proteolítico para formar los dos dominios, Supuestamente este sitio está en el resto 960.

En una realización preferida, la proteína HAP es purificada o aislada después de la expresión. Las proteínas HAP pueden ser aisladas o purificadas en una variedad de formas conocidas por los expertos en la técnica dependiendo de qué otros componentes están presentes en la muestra. Los métodos de purificación normalizados incluyen técnicas electroforéticas, moleculares, inmunológicas y cromatográficas, incluyendo la cromatografía HPLC de intercambio iónico, hidrófoba, de afinidad, y en fase reversa, y el cromatofoco. Por ejemplo, la proteína HAP puede ser purificada utilizando una columna de anticuerpo anti-HAP normalizada. También son útiles las técnicas de ultrafiltración y diafiltración, junto con la concentración de la proteína. Para una pauta general de las técnicas de purificación adecuadas, ver Scopes, R., Protein Purification, Springer-Verlag, NY (1982). El grado de purificación necesario variará dependiendo del uso de la proteína HAP. En algunos casos no será necesaria purificación.

Una vez expresadas y purificadas si es necesario, las proteínas HAP son útiles en numerosas aplicaciones.

Por ejemplo, las proteínas HAP pueden ser acopladas, utilizando tecnología normalizada, a columnas de cromatografía de afinidad. Estas columnas se pueden utilizar después para purificar anticuerpos de las muestras obtenidas de animales o pacientes expuestos a un organismo de *Haemophilus influenzae*. Los anticuerpos purificados se pueden utilizar después como se esboza más abajo.

Adicionalmente, las proteínas HAP son útiles para elaborar anticuerpos para las proteínas HAP. Estos anticuerpos encuentran uso en numerosas aplicaciones. En una realización preferida, los anticuerpos se utilizan para diagnosticar la presencia de una infección por *Haemophilus influenzae* en una muestra o un paciente. Esto se realizará utilizando mecanismos bien conocidos en la técnica; por ejemplo, se pueden obtener muestras tales como muestras de sangre o tejido de un paciente y someterlas a ensayo en cuanto a la reactividad con los anticuerpos, por ejemplo utilizando mecanismos normalizados tales como ELISA. En una realización preferida, se generan anticuerpos monoclonales para la proteína HAP, utilizando mecanismos bien conocidos en la técnica. Como se ha esbozado antes, los anticuerpos pueden ser generados para la proteína HAP completa, o para una porción de la proteína HAP.

Los anticuerpos generados para las proteínas HAP también pueden ser utilizados en tratamientos de inmunización pasiva, como es sabido en la técnica.

Los anticuerpos generados para secuencias únicas de proteínas HAP también pueden ser utilizados para rastrear genotecas de expresión de otros organismos para encontrar, y posteriormente clonar, ácidos nucleicos de HAP de otros organismos.

En una realización, los anticuerpos pueden estar marcados directamente o indirectamente. Por “marcados” se quiere significar en la presente memoria un compuesto que tiene al menos un elemento, isótopo o compuesto químico anclado para permitir la detección del compuesto. En general, las marcas se dividen en tres clases: a) marcas isotópicas, que pueden ser radiactivas o isótopos pesados; b) marcas inmunitarias, que pueden ser anticuerpos o antígenos; y c) tintes coloreados o fluorescentes. Las marcas se pueden incorporar al compuesto en cualquier posición. De este modo, por ejemplo, el anticuerpo de la proteína HAP puede ser marcado para la detección, o se puede crear y marcar un anticuerpo secundario para el anticuerpo de la proteína HAP.

En una realización, los anticuerpos generados para las proteínas HAP de la presente invención se utilizan para purificar o separar proteínas HAP o el organismo de *Haemophilus influenzae* de una muestra. De este modo, los anticuerpos generados para las proteínas HAP que se unirán al organismo de *Haemophilus influenzae* pueden ser acoplados, utilizando la tecnología normalizada, a columnas de cromatografía de afinidad. Estas columnas pueden ser utilizadas para retirar el organismo de *Haemophilus* de muestras medioambientales o de tejidos. Alternativamente, los anticuerpos generados para la porción de 110 kD soluble de la porción completa de la proteína mostrada en la Figura 7 pueden ser utilizados para purificar la proteína de 110 kD de las muestras.

En una realización preferida, las proteínas HAP de la presente invención se utilizan como vacunas para el tratamiento profiláctico o terapéutico de una infección por *Haemophilus influenzae* en un paciente. Por “vacuna” se quiere significar en la presente memoria un antígeno o compuesto que logra una respuesta inmunitaria en un animal o paciente. La vacuna puede ser administrada profilácticamente, por ejemplo a un paciente no expuesto nunca anteriormente al antígeno, de manera que se evite la posterior infección por el organismo *Haemophilus influenzae*. Alternativamente, la vacuna puede ser administrada terapéuticamente a un paciente expuesto o infectado previamente por el organismo *Haemophilus influenzae*. Si bien la infección no puede ser evitada, en este caso se genera una respuesta inmunitaria que permite al sistema inmunitario del paciente combatir más eficazmente la infección. De este modo, por ejemplo, puede haber una disminución o reducción de los síntomas asociados con la infección.

Un “paciente” para los fines de la presente invención incluye tanto seres humanos como otros animales y organismos. De este modo los métodos son aplicables tanto a la terapia en seres humanos como a aplicaciones veterinarias.

La administración de la proteína HAP como vacuna se realiza de diversas maneras. Generalmente, las proteínas HAP pueden ser formuladas según los métodos conocidos para preparar composiciones farmacéuticamente útiles, por medio de los cuales se combinan cantidades terapéuticamente eficaces de la proteína HAP mezcladas con un vehículo portador farmacéuticamente aceptable. Los vehículos adecuados y su formulación son bien conocidos en la técnica. Tales composiciones contendrán una cantidad eficaz de la proteína HAP junto con una cantidad adecuada de vehículo con el fin de preparar composiciones farmacéuticamente aceptables para la administración eficaz al anfitrión. La composición puede incluir sales, tampones, proteínas portadoras tales como seralbúmina, moléculas de redireccionamiento para localizar la proteína HAP en el sitio o tejido apropiado dentro del organismo, y otras moléculas. La composición puede incluir también coadyuvantes.

En una realización, la vacuna se administra en forma de una única dosis; esto es, una dosis es adecuada para inducir una respuesta inmunitaria suficiente para tratar profilácticamente o terapéuticamente una infección por *Haemophilus influenzae*. En realizaciones alternativas, la vacuna se administra a varias dosis a lo largo de un período de tiempo, en forma de una vacunación primaria y vacunaciones de “refuerzo”.

Por “cantidades terapéuticamente eficaces” se quiere significar en la presente memoria una cantidad de proteína HAP que es suficiente para inducir una respuesta inmunitaria. Esta cantidad puede ser diferente dependiendo de si se desea un tratamiento profiláctico o terapéutico. Generalmente, esta oscila de aproximadamente 0,001 mg a aproximadamente 1 gm, con un intervalo preferido de aproximadamente 0,05 a aproximadamente. Estas cantidades pueden ser ajustadas si se utilizan coadyuvantes.

Los siguientes ejemplos sirven para describir más completamente la manera de utilizar la invención descrita anteriormente, así como para mostrar los mejores modos contemplados para llevar a cabo los diversos aspectos de la invención. Se entiende que estos ejemplos no sirven en modo alguno para limitar el verdadero alcance de esta invención, sino que en vez de eso se presentan con fines ilustrativos.

Ejemplos

Ejemplo 1

Clonación de la proteína HAP

Cepas Bacterianas, plásmidos, y fagos. La cepa N187 de *H. influenzae* es un producto aislado que fue cultivado originalmente a partir del fluido del oído medio de un niño con otitis media aguda. Esta cepa fue clasificada como no tipificable basándose en la ausencia de aglutinación con antisuero de tipificación para los tipos a-f de *H. influenzae* (Burroughs Wellcome) y el fracaso al hibridar con pU038, un plásmido que contiene el locus *cap b* completo (Kroll and Moxon, 1988, J. Bacteriol. 170:859-864).

La cepa DB117 de *H. influenzae* es un mutante *recl* de Rd, una cepa de serotipo d carente de cápsula que ha estado en el laboratorio durante 40 años (Alexander and Leidy, 1951, J. Exp. Med. 83:345-359); DB117 fue obtenida de G. Barcak (University of Maryland, Baltimore, MD) (Sellow *et al.*, 1968). DB117 es deficiente para la adherencia y la invasión *in vitro*, como se analiza más abajo. La cepa 12 de *H. influenzae* es la cepa no tipificable de la cual se clonaron los genes que codifican las proteínas HMW1 y HMW2 (Barenkamp and Leininger, 1992, Infect. Immun. 60:1302-1313); HMW1 y HMW2 son los miembros prototipo de una familia de proteínas adherentes de elevado peso molecular antigénicamente relacionadas con *Haemophilus* no tipificables (St. Geme *et al.*, 1993).

E. coli HB101, que es no adherente y no invasiva, ha sido descrita previamente (Sambrook *et al.*, 1989, Molecular cloning: A laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.). *E. coli* DH5 α fue obtenido de Bethesda Research Laboratories. *E. coli* MC1061 fue obtenido de H. Kimsey (Tufts University, Boston, MA). *E. coli* XL-1 Blue y el plásmido pBluescript KS- fueron obtenidos de Stratagene. El plásmido pT7-7 y el fago mGP1-2 fueron proporcionados por S. Tabor (Harvard Medical School, Boston, MA) (Tabor and Richardson, 1985, Proc. Natl. Acad. Sci. USA. 82:1074-1078). El vector lanzadera de *E. coli*-*Haemophilus* pGJB103 (Tomb *et al.*, 1989, Rd. J. Bacteriol. 171:3796-3802) y el fago λ 1105 (Way *et al.*, 1984, Gene. 32:3 69-379) fueron proporcionados por G. Barcak (University of Maryland, Baltimore, MD). El plásmido pVD116 alberga el gen de la proteasa IgA1 de la cepa Rd de *H. influenzae* (Koohey and Falkow, 1984, Infect. Immun. 43:101-107) y fue obtenido de M. Koohey (University of Michigan, Ann Arbor, MI).

Condiciones de crecimiento. Las cepas de *H. influenzae* se hicieron crecer como se ha descrito (Anderson *et al.*, 1972, J. Clin. Invest. 51:31-38). Se almacenaron a -80°C en caldo de infusión de cerebro-corazón con 25% de glicerol. Las cepas de *E. coli* se hicieron crecer sobre agar LB o en caldo LB. Se almacenaron a -80°C en caldo LB con 50% de glicerol.

Para *H. influenzae*, se utilizó tetraciclina a una concentración de 5 μ g/ml y se utilizó kanamicina a una concentración de 25 μ g/ml. Para *E. coli*, se utilizaron antibióticos a las siguientes concentraciones: tetraciclina, 12,5 μ g/ml; kanamicina, 50 μ g/ml; ampicilina, 100 μ g/ml.

Métodos de ADN Recombinante. La ligación de ADN, las digestiones con endonucleasas de restricción, y las electroforesis en gel se realizaron según los mecanismos normalizados (Sambrook *et al.*, 1989, *supra*). Los plásmidos

fueron introducidos en cepas de *E. coli* mediante transformación química o electroporación, como describen (Sambrook *et al.*, 1989, *supra*; Dower *et al.*, 1988, Nucleic Acids Res. 16:617-6145). En *H. influenzae* la transformación se llevó a cabo utilizando el método de MIV de Herriott *et al.* (1970, J. Bacteriol. 101:517-524), y la electroporación se llevó a cabo utilizando el protocolo desarrollado para *E. coli* (Dower *et al.*, 1988, *supra*).

Construcción de genoteca genómica de la cepa N187 de H. influenzae. Se preparó ADN cromosómico de elevado peso molecular a partir de 3 ml de un cultivo en caldo durante la noche de *H. influenzae* N187 como se ha descrito previamente (Mekalanos, 1983, Cell. 35:253-263). Tras la digestión parcial con *Sau3AI*, se hicieron eluir los fragmentos de 8 a 12 kb en papel DEAE (Schleicher & Schuell, Keene, H.H.) y después se ligaron a pGJB103 tratado con fosfatasa de intestino de ternera digerido con *Bgl*III. La mezcla de ligación fue sometida a electroporación en *H. influenzae* DB117, y los transformantes fueron seleccionados sobre medio que contenía tetraciclina.

Mutagénesis de transposones

La mutagénesis del ADN plasmídico se realizó utilizando el elemento mini-Tn10 *kan* descrito por Way *et al.* (1984, *supra*). Inicialmente, se introdujo el plásmido apropiado en *E. coli* MC1061. La cepa resultante se infectó con λ 1105, que porta el transposón mini-Tn10 *kan*. Los transductantes se hicieron crecer durante la noche en presencia de kanamicina y un antibiótico para seleccionar el plásmido, y el ADN plasmídico se aisló utilizando el método de lisis alcalina. Con el fin de recuperar los plásmidos que contienen una inserción del transposón, se sometió a electroporación el ADN plasmídico en *E. coli* DH5 α , cultivando en placa sobre medio que contenía kanamicina y el segundo antibiótico apropiado.

Con el fin de establecer más precisamente la región de pN187 implicada en la promoción de la interacción con las células anfitrionas, inicialmente este plásmido se sometió análisis con endonucleasas de restricción. Con posterioridad, se construyeron numerosos subclones en el vector pGJB103 y se reintrodujeron en la cepa DB117 de *H. influenzae*. Las cepas resultantes se examinaron después en cuanto a la adherencia y la invasión. Como se resume en la Figura 4, los subclones que contienen o bien el fragmento *Pst*I-*Bgl*III de 3,9 kb (pJS105) o bien el fragmento *Bgl*III de 4,2 kb colindante (pJS102) fracasaba al conferir la capacidad para asociarse con las células Chang. En contraste, un subclon que contenía un inserto que incluía porciones de ambos de estos fragmentos (pJS106) no promovía la interacción con las monocapas epiteliales. La mutagénesis del transposón realizada sobre pN187 confirmó que las porciones limítrofes del inserto en este plásmido no requerían el fenotipo adherente/invasivo. Por otra parte, la inserción de transposón localizado adyacente con respecto al sitio *Bgl*III en pJS106 eliminaba la adherencia y la invasión. Una inserción entre los segundos sitios *Eco*RI y *Pst*I en este plásmido tenía un efecto similar (Figura 4).

Examen de las proteínas codificadas por el plásmido

Con el fin de examinar las proteínas codificadas por el plásmido, se ligó el ADN relevante en el vector de expresión del bacteriófago T7 pT7-7, y el constructo resultante se transformó en *E. coli* XL-1 Blue. El plásmido pT7-7 contiene el promotor ϕ 10 del fago T7 y el sitio de unión al ribosoma aguas arriba de un sitio de clonación múltiple (Tabor and Richardson, 1985, *supra*). El promotor de T7 fue inducido mediante infección con el fago de M13 recombinante mGP12 y adición de isopropil- β -D-tiogalactopiranosido (concentración final, 1 mM). El fago mGP1-2 contiene el gen que codifica la ARN polimerasa de T7, que activa el promotor ϕ 10 en pT7-7 (Tabor and Richardson, 1985, *supra*).

Como DB117(pN187), la cepa DB117 que porta pJS106 expresaba proteínas de la membrana externa nuevas de 160-kD y 45-kD de tamaño (Figura 3, calle 3). Con el fin de examinar si el inserto de 6,5 kb en pJS106 realmente codifica estas proteínas, este fragmento de ADN fue ligado en el vector de expresión del bacteriófago T7 pT7-7. El plásmido resultante que contenía el inserto en la misma orientación que en pN187 fue denominado pJS104, y el plásmido con el inserto en orientación opuesta fue denominado pJS103. Tanto pJS104, como pJS103 fueron introducidos en *E. coli* XL-1 Blue, produciendo XL-1 Blue (pJS104) y XL-1 Blue (pJS103), respectivamente. Como control negativo, también se transformó pT7-7 en XL-1 Blue. El promotor de T7 fue inducido en estas tres cepas por infección con el fago M13 recombinante mGP1-2 y la adición de isopropil- β -D-tiogalactopiranosido (concentración final, 1 mM), y las proteínas inducidas fueron detectadas utilizando metionina [S³⁵]. Como se muestra en la Figura 5, la inducción de XL-1 Blue (pJS104) daba como resultado la expresión de una proteína de 160 kD y numerosas proteínas más pequeñas que presumiblemente representan productos de degradación. En contraste, cuando se inducían XL-1 Blue(pJS103) y XL-1 Blue(pT7-7), no había expresión de estas proteínas. No había proteína de 45 kD indicada en ninguna de las tres cepas. Este experimento sugería que el inserto de 6,5 kb presente en pJS106 contiene el gen estructural para la proteína de la membrana externa de 160 kd identificada en DB117(pJS106). Por otra parte, este análisis fracasaba al establecer el origen de la proteína de membrana de 45 kD expresada por DB117 (pJS106).

Análisis de adherencia e invasión

Los análisis de adherencia e invasión se realizaron con células epiteliales de Chang [derivado Wong-Kilbourne, clon 1-5c-4 (conjuntiva humana)], que fueron sembradas en pocillos de placas para el cultivo de tejidos de 14 pocillos como se ha descrito previamente (St. Geme and Falkow, 1990). La adherencia fue medida después de incubar las bacterias con monocapas epiteliales durante 30 minutos como se ha descrito (St. Geme *et al.*, 1993). Los análisis de invasión se llevaron a cabo según el protocolo original de los autores de la presente invención e implicaban incubar bacterias con células epiteliales durante cuatro horas seguido de tratamiento con gentamicina durante dos horas (100 μ g/ml) (St. Geme and Falkow, 1990).

Determinación de la secuencia de nucleótidos y análisis

La secuencia de nucleótidos fue determinada utilizando un estuche Sequenase y un molde plasmídico de doble hebra. Los fragmentos de ADN fueron subclonados en pBluescript KS⁻ y secuenciados junto con ambas hebras mediante paseo con cebadores. El análisis de la secuencia de ADN se realizó utilizando el paquete de soporte lógico Genetics Computer Group (GCG) de la University of Wisconsin (Devereux *et al.*, 1984). Las búsquedas de similitud de la secuencia se llevaron a cabo utilizando el programa BLAST del National Center for Biotechnology Information (Altschul *et al.*, 1990, J. Mol. Biol. 215:403-410). La secuencia de ADN descrita en la presente memoria será depositada en EMBL/GenBank/DDBJ Nucleotide Sequence Data Libraries.

Basándose en los resultados de la subclonación de los autores de la presente invención, los autores de la presente invención razonaron que el sitio *Bgl*II central de pH187 estaba situado en un marco de lectura abierto. El examen de una serie de mutantes mini-Tn10 *kan* apoyaba esta conclusión (Figura 4). Por consiguiente, los autores de la presente invención secuenciaron DHA en ambos lados de este sitio *Bgl*II e identificaron un gen de 4182 pb, que ha sido denominado hap por *Haemophilus* adherence and penetration (adherencia y penetración de *Haemophilus*) (Figura 6). Este gen codifica un polipéptido de 1394 aminoácidos, que los autores de la presente invención han denominado Hap, con una masa molecular calculada de 155,4 kD, con forme a la masa molecular de las más grande de las dos proteína de la membrana externa novedosas expresadas por DB117(pN187) y la proteína expresada después de la inducción de XL-1 Blue/pJS104. El gen hap tiene un contenido en G+C de 39,1%, similar a la estimación publicada de 38,7% del genoma completo (Kilian, 1976, J. Gen. Microbiol. 93:9-62). Las supuestas secuencias promotoras - 10 y -35 están presentes aguas arriba del codón de iniciación. Falta un sitio de unión al ribosoma consenso. Una secuencia similar a un terminador de la transcripción independiente de rho está presente empezando 39 nucleótidos más allá del codón de terminación y contiene repeticiones invertidas interrumpidas con el potencial de formar una estructura en horquilla que contiene un bucle de tres bases y un tallo de ocho bases. Similar a la situación con los terminadores de *E. coli* típicos, esta estructura está seguida de un tramo rico en restos T. El análisis de la secuencia de aminoácidos pronosticada sugería la presencia de un péptido señal de 25 aminoácidos en el extremo amino. Esta región tiene características típicas de péptidos señal procarióticos, con tres cargas terminales H-positivas, una región hidrófoba central, y restos alanina en las posiciones 23 y 25 (-3 y -1 con respecto al supuesto sitio de escisión) (von Heijne, 1984, J. Mol. Biol. 173:243-251).

Comparación de la secuencia de aminoácidos deducida de Hap con otras proteínas. Se realizó una búsqueda de similitud de la secuencia de la proteína con la secuencia de aminoácidos pronosticada utilizando el servicio de la red BLAST del National Center for Biotechnology Information (Altschul *et al.*, 1990, *supra*). Esta búsqueda reveló homología con las proteasas IgA1 de *H. influenzae* y *Neisseria gonorrhoeae*. La alineación de las secuencias de aminoácidos derivadas para el producto génico hap y las proteasas IgA1 de cuatro cepas diferentes de *H. influenzae* reveló homología de un lado a otro del extremo de las proteínas (Figura 7), mostrando diversos tramos con una identidad del 55-60% y una similitud del 70-80%. Se observaban niveles similares de homología entre el producto hap y la proteasa IgA1 de la cepa MS11 de *N. gonorrhoeae*. Esta homología incluye la región identificada como sitio catalítico de las proteasas IgA1, que está comprendido por la secuencia GDSGSPLF, donde 2 es la serina del sitio activo característica de las serina proteasas (Brenner, 1988, Nature (London). 334:528-530; Poulsen *et al.*, 1992, J. Bacteriol. 174:2913-2921). En el caso de Hap, la secuencia correspondiente es GDSGSPMF. El producto hap también contiene dos cisteínas correspondientes a las cisteínas que se ha propuesto que son importantes en la formación del dominio catalítico de las proteasas IgA (Pohlner *et al.*, 1987, *supra*). En conjunto hay una identidad del 30-35% y una similitud del 51-55% entre el producto del gen hap y las proteasas IgA de *H. influenzae* y *N. gonorrhoeae*.

También se encontró que la secuencia de aminoácidos deducida codificada por hap contenía un homología significativa con Tsh, una hemaglutinina expresada por una cepa de *E. coli* aviar (Provence and Curtiss, 1994, *supra*). Esta homología se extiende a ambas proteínas pero es mayor en la mitad H-terminal de cada una. En conjunto las dos proteínas son idénticas en un 30,5% y similares en un 51,6%. Tsh también es sintetizada como una preproteína y es secretada en una forma más pequeña; como las proteasas IgA1 y quizás Hap, un péptido carboxi terminal permanece asociado con la membrana externa (D. Provence, comunicación personal). Si bien se presumía que esta proteína tenía actividad proteolítica, su sustrato todavía no ha sido determinado. Interesantemente, Tsh fue identificada primero basándose en su capacidad para promover la aglutinación de eritrocitos. De este modo Hap y Tsh son posiblemente los primeros miembros de una clase de proteínas adherentes novedosa que son transformadas análogamente a las proteasas IgA1.

Asimismo se ha observado homología con la pertactina, una proteína de la membrana externa de 69 kD expresada por *B. pertussis* (Charles *et al.*, 1989, Proc. Natl. Acad. Sci. USA. 86:3554-3558). Las porciones medias de estas dos moléculas son idénticas en un 39% y similares casi en un 60%. Esta proteína contiene el triplete de aminoácidos arginina-glicina-ácido aspártico (RGD) y se ha demostrado que promueve el anclaje a células de mamífero cultivadas por medio de su secuencia (Leininger *et al.*, 1991, Proc. Natl. Acad. Sci. USA. 88:345-349). Aunque las especies de *Bordetella* no se consideran generalmente parásitos intracelulares, el trabajo de Ewanowich y colaboradores indica que estos patógenos respiratorios son capaces de entrar en células epiteliales *in vitro* (Ewanowich *et al.*, 1989, Infect. Immun. 57:2698-2704; Ewanowich *et al.*, 1989, Infect. Immun. 57:1240-1247). Recientemente Leininger *et al.* informaron que la preincubación de las monocapas epiteliales con un péptido que contiene RGD derivado de la secuencia de pertactina inhibía específicamente la entrada de *B. pertussis* (Leininger *et al.*, 1992, Infect. Immun. 60:2380-2385). Además, estos investigadores descubrieron que el recubrimiento de *Staphylococcus aureus* con pertactina purificada producía una entrada de *S. aureus* más eficaz; el péptido que contenía RGD de pertactina inhibía esta entrada inten-

sificada por pertactina en un 75%. Aunque el producto *hap* carece de motivo RGD, es posible que Hap y pertactina cumplan funciones biológicas similares para *H. influenzae* y *Bordetella* sp., respectivamente.

- El análisis adicional reveló una homología significativa (identidad del 34 al 52%, similitud del 42 al 70%) con seis regiones de HpMA, una hemolisina independiente de calcio expresada por *Proteus mirabilis* (Uphoff and Welch, 1990, *supra*).

El locus hap es distinto del de la proteasa IgA1 de H. influenzae

- Dado el grado de similitud entre el producto génico de *hap* y la proteasa IgA1 de *H. influenzae*, los autores de la presente invención se preguntaron si habían aislado el gen de la proteasa IgA1 de la cepa N187. Para examinar esta posibilidad, los autores de la presente invención realizaron análisis de actividad de la proteasa IgA1. Entre las cepas de *H. influenzae*, se han encontrado dos tipos enzimáticamente distintos de proteasa IgA1 (Mulks *et al.*, 1982, J. Infect. Dis. 146:266-274). Las enzimas de tipo 1 escinden el enlace peptídico Pro-Ser entre los restos 231 y 232 en la región bisagra de la cadena pesada de IgA1 humana y general fragmentos de aproximadamente 28 kD y 31 kD; las enzimas de tipo 2 escinden el enlace Pro-Thr entre los restos 235 y 236 en la región bisagra y generan fragmentos de 26,5 kD y 32,5 kD. Los estudios previos de la cepa parental de la cual deriva DB117 han demostrado que esta cepa produce una proteasa IgA1 de tipo 1 (Koomey and Falkow, 1984, *supra*). Como se muestra en la Figura 8, la comparación de las actividades proteolíticas de la cepa DB117 y la cepa N187 sugieren que N187 produce una proteasa IgA1 de tipo 2. Los autores de la presente invención razonaron que DB117(pN187) podría generar un total de cuatro fragmentos a partir de la proteasa IgA1, coincidentes con dos especificidades de escisión distintas. El examen de DB117(pH187) reveló en lugar de eso que este transformante produce los mismos dos fragmentos de la cadena pesada de IgA1 que DB117, indicando que esta cepa produce solamente una enzima de tipo 1.

- En un esfuerzo para obtener una evidencia adicional frente a la posibilidad de que el plásmido pH187 contenga el gen de la proteasa IgA1 de N187, los autores de la presente invención realizaron una serie de transferencias Southern. Como se muestra en la Figura 9, cuando el ADN genómico de la cepa N187 era digerido con *EcoRI*, *BglII*, o *BamHI* y después sondeado con el gen *hap*, se detectaba un grupo de fragmentos de hibridación. El sondeo del mismo ADN con el gen *iga* de la cepa Rd de *H. influenzae* daba como resultado un grupo diferente de bandas de hibridación. Por otra parte, el *iga* fracasaba al hibridar con un fragmento de 4,8 kb purificado que contenía el gen *hap* intacto.

El plásmido recombinante asociado con la adherencia y la invasión codifica una proteína secretada

- La sorprendente homología entre el producto génico de *hap* y las proteasas IgA1 de *Haemophilus* y *Neisseria* sugería la posibilidad de que estas proteínas pudieran ser transformadas de una manera similar. Las proteasas IgA1 son sintetizadas como preproteínas con tres dominios funcionales: el péptido señal N-terminal, la proteasa, y un dominio coadyuvante C-terminal, que se postula que forma un poro en la membrana externa para la secreción de la proteasa (Poulsen *et al.*, 1989, *supra*; Pohlner *et al.*, 1987, *supra*). El péptido C-terminal permanece asociado con la membrana externa después de un evento de escisión autoproteolítico que produce la liberación de la enzima madura.

- De acuerdo con la posibilidad de que el producto del gen *hap* siga un destino similar, los autores de la presente invención encontraron que DB117(pN187) producía una proteína secretada de un tamaño de aproximadamente 110 kD que estaba ausente de DB117(pGJB103) (Figura 10). Esta proteína también era producida por DB117(pJS106), pero no por DB117(pJ5102) o DB117(pJS105). Además, los dos mutantes con inserciones de transposones en la región codificadora de *hap* eran deficientes en esta proteína. Con el fin de determinar la relación entre *hap* y la proteína secretada, esta proteína fue transferida a una membrana de PVDF y se realizó una secuenciación de aminoácidos N-terminales. Un fondo excesivo en el primer ciclo imposibilitaba la identificación del primer resto aminoácido del extremo amino libre. Se encontró que la secuencia de los siete restos siguientes era HTYFGID, que corresponde a los aminoácidos 27 a 33 del producto de *hap*.

- La introducción de *hap* en cepas de laboratorio de *E. coli* era incapaz de dotar a estos organismos de la capacidad de adherencia o invasión. Teniendo en cuenta estos resultados, es notable que los transformantes de *E. coli* fracasaran al expresar la proteína de membrana externa de 160 kD o de 45 kD. Por consiguiente, también fracasaban al expresar la proteína secretada de 110 kD. La explicación de esta carencia de expresión no está clara. Una posibilidad es que el promotor de *H. influenzae* o el sitio de unión al ribosoma eran escasamente reconocidos en *E. coli*. En efecto la supuesta secuencia -35 aguas arriba del codón de iniciación de *hap* es bastante divergente de la secuencia consenso $\sigma 70$, y el sitio de unión al ribosoma es irreconocible. Alternativamente, se puede requerir un gen accesorio para la exportación apropiada de la proteína Hap, aunque la sorprendente homología con las proteasas IgA, que son normalmente expresadas y secretadas en *E. coli*, indica argumenta en contra de esta hipótesis.

- Considerando la posibilidad de que el producto del gen *hap* promueva la adherencia y la invasión uniéndose directamente a una estructura de la superficie de la célula anfitriona, parece curioso que la proteína madura sea secretada del organismo. Sin embargo, existen ejemplos de otros factores de adherencia que también son secretados. La hemaglutinina filamentosa es una proteína de 220 kD expresada por *B. pertussis* que media la adherencia *in vitro* y facilita la colonización natural (Relman *et al.*, 1989, Proc. Natl. Acad. Sci. U.S.A. 86:2637-2641; Kimura *et al.*, 1990, Infect. Immun. 58:7-16). Esta proteína permanece asociada a la superficie hasta cierto punto pero también es liberada de la célula. El procedimiento de secreción de la hemaglutinina filamentosa implica una proteína accesorio denominada FhaC, que parece estar localizada en la membrana externa (Willems *et al.*, 1994, Molec. Microbiol. 11:337-347). De

un modo similar, las proteínas Ipa implicadas en la invasión de *Shigella* también son secretadas. La secreción de estas proteínas requiere los productos de múltiples genes en los loci *mxi* y *spa* loci (Allaoui *et al.*, 1993, Molec. Microbiol. 7:59-68; Andrews *et al.*, 1991, Infect. Immun. 59:1997-2005; Venkatsan *et al.*, 1992, J. Bacteriol. 174:1990-2001).

Es concebible que la secreción sea simplemente consecuencia del mecanismo para exportar el producto del gen *hap* a la superficie del organismo. No obstante, es notable que la proteína secretada contenga un dominio catalítico de proteasa de tipo serina y muestre homología con la hemolisina de *P. mirabilis*. Estos descubrimientos sugieren que la proteína Hap madura puede poseer actividad proteolítica y originar la posibilidad de que Hap promueva la interacción con la célula anfitriona a una distancia modificando la superficie de la célula anfitriona. Alternativamente, Hap puede modificar la superficie bacteriana con el fin de facilitar la interacción con un receptor de la célula anfitriona. Es posible que *hap* codifique una molécula con funciones duales, funcionando como adhesina y como proteasa.

Análisis de proteínas de la membrana externa y secretadas

Se aislaron proteínas de la membrana externa basándose en la insolubilidad del sarcosilo según el método de Carlone *et al.* (1986, J. Clin. Microbiol. 29:33C-332). Las proteínas secretadas se aislaron centrifugando cultivos bacterianos a 16.000 g durante 10 minutos, recuperando el sobrenadante, y precipitando con ácido trifluoroacético a una concentración final del 10%. Se realizó una electroforesis en gel de SDS-poliacrilamida como se ha descrito previamente (Laemmli, 1970, Nature (London). 227:680-685).

Para identificar las proteínas que podrían estar implicadas en la interacción con la superficie de la célula anfitriona, se compararon los perfiles de la proteína de la membrana externa para DB117 (pN187) y DB117 (pGJB103). Como se muestra en la Figura 3, DB117(pN187) expresaba dos nuevas proteínas de la membrana externa: una proteína de elevado peso molecular de aproximadamente 160 kD y una proteína de 45 kD. *E. coli* HB101 que albergaba pN187 fracasaba al expresar estas proteínas, sugiriendo una explicación para la observación de que HB101(pN187) es incapaz de adherirse o de invadir.

Los estudios previos han demostrado que una familia de proteínas de elevado peso molecular antigénicamente relacionadas con similitud a la hemaglutinina filamentosa de *Bordetella pertussis* media el anclaje por *H. influenzae* no tipificable a células epiteliales cultivadas (St. Geme *et al.*, 1993). Para explorar la posibilidad de que el gen que codifica el miembro de la cepa H187 de esta familia fuera clonado, se examinaron productos lisados de células completas de N187, DB117(pN187), y DB117(pGJB103) mediante inmunotransferencia Western. La cepa de control de los autores de la presente invención para este experimento era la cepa 12 de *H. influenzae*. Utilizando un antisuero policlonal dirigido contra HMW1 y HMW2, las proteínas prototipo de esta familia, los autores de la presente invención identificaron una proteína de 140 kD en la cepa H187 (no mostrada). En contraste, este antisuero fracasaba al reaccionar con DB117 (pN187) o DB117(pGJB103) (no mostrado), indicando que pN187 no tiene relación con la expresión de la proteína HMW.

Determinación de la secuencia amino terminal. Las proteínas secretadas fueron precipitadas con ácido tricloroacético, separadas en un gel de poliácridamida-SDS al 10%, y electrotransferidas a una membrana de difluoruro de polivinilideno (PVDF) (Matsudaira, 1987, J. Biol. Chem. 262:10035-10038). Después de la tinción con Azul Brillante de Coomassie R-250, la proteína de 110 kD fue cortada de la membrana de PVDF y sometida a determinación de la secuencia amino terminal en el Protein Chemistry Laboratory en el Washington University School of Medicine. El análisis de la secuencia se realizó mediante degradación de Edman automatizada utilizando un secuenciador de proteínas Modelo 470A de Applied Biosystems.

Examen de la actividad proteasa IgA1. Con el fin de evaluar la actividad proteasa IgA1, las bacterias fueron inoculadas en caldo y se hicieron crecer aerobiamente durante la noche. Las muestras se centrifugaron después en un microfago durante dos minutos, y se recogieron los sobrenadantes. Se mezcló un volumen de 10 µl con 16 µl de 0,5 µg/ml de IgA1 humana (Calbiochem), y se añadió cloramfenicol a una concentración final de 2 µg/ml. Después de incubar durante la noche a 37°C, las mezclas de reacción se sometieron a electroforesis en un gel de poliácridamida SDS al 10%, se transfirieron a una membrana de nitrocelulosa, y se sondearon con cadena pesada de IgA1 anti-humana de cabra conjugada con fosfatasa alcalina (Kirkegaard & Perry). La membrana se reveló mediante inmersión en solución sustrato de fosfatasa (sistema de sustrato 5-bromo-4-cloro-3-indolilfosfato toluidinio-nitroazul de tetrazolio; Kirkegaard & Perry).

Análisis de inmunotransferencia. El análisis de inmunotransferencia de los productos lisados celulares completos se llevó a cabo como se ha descrito (St. Geme *et al.*, 1991).

Hibridación Southern. La transferencia Southern se realizó utilizando condiciones restrictivas como se ha descrito previamente (St. Geme and Falkow, 1991).

Microscopía

i. Microscopía óptica. Las muestras de células epiteliales con bacterias asociadas se tiñeron con colorante Giemsa y se examinaron mediante microscopía óptica como se ha descrito (St. Geme and Falkow, 1990).

ii. *Microscopía de Transmisión de electrones.* Para la microscopía de transmisión de electrones, las bacterias fueron incubadas con monocapas de células epiteliales durante cuatro horas y después fueron aclaradas cuatro veces con PBS, fijadas con 2% de glutaraldehído/1% de tetróxido de osmio en tampón fosfato de sodio 0,1 M pH 6,4 durante dos horas sobre hielo, y teñidas con acetato de uranilo acuoso durante la noche. Después las muestras se deshidrataron en soluciones de etanol graduadas y se embebieron en Polybed. Se examinaron secciones ultra finas (0,4 μ m) en un microscopio electrónico Phillips 201c.

Como se muestra en la Figura 2, DB117(pN187) incubado con monocapas durante cuatro horas demostró una íntima interacción con la superficie de la célula epitelial y ocasionalmente se encontró que era intracelular. En una sección fina dada, las células invadidas contenían generalmente uno o más organismo intracelulares. Como se puede notar, las bacterias intracelulares eran más comunes en las secciones preparadas con la cepa N187, una observación coherente con los resultados utilizando el análisis con gentamicina. En contraste, el examen de las muestras preparadas con la cepa DB117 que portaba el vector de clonación solo (pGJB103) fracasaba al revelar la bacteria internalizada (no mostrado).

Habiendo descrito las realizaciones preferidas de la presente invención a los expertos normales en la técnica les parecerá que se pueden realizar diversas modificaciones de las realizaciones descritas, y que tales modificaciones están destinadas a estar dentro del alcance de la presente invención.

REIVINDICACIONES

- 5 1. Una proteína de adherencia y penetración *Haemophilus influenzae* recombinante, cuya proteína tiene una secuencia de aminoácidos con una homología de más del 80% con la secuencia de aminoácidos mostrada en la Figura 6.
- 10 2. Una proteína de adherencia y penetración de *Haemophilus influenzae* recombinante como se reivindica en la reivindicación 1 codificada por un ácido nucleico, cuyo ácido nucleico hibrida en condiciones muy restrictivas con el ácido nucleico que tiene la secuencia mostrada en la Figura 6.
- 15 3. Una proteína de adherencia y penetración de *Haemophilus influenzae* recombinante como se reivindica en la reivindicación 1 que tiene una secuencia de aminoácidos mostrada en la Figura 6.
- 20 4. Una composición que comprende un portador farmacéuticamente aceptable y una proteína de adherencia y penetración de *Haemophilus influenzae* según cualquiera de las reivindicaciones 1-3.
- 25 5. Un ácido nucleico recombinante que codifica una proteína de adherencia y penetración de *Haemophilus influenzae*, cuya proteína tiene una secuencia de aminoácidos con una homología de más del 80% con la secuencia de aminoácidos mostrada en la Figura 6.
- 30 6. Un ácido nucleico recombinante que codifica una proteína de adherencia y penetración de *Haemophilus influenzae* como se reivindica en la reivindicación 5, cuyo ácido nucleico hibrida en condiciones muy restrictivas con el ácido nucleico que tiene la secuencia mostrada en la Figura 6.
- 35 7. Un vector de expresión que comprende ácido nucleico regulador de la transcripción y la traducción conectado operablemente al ácido nucleico recombinante de las reivindicaciones 5 o 6.
- 40 8. Una célula anfitriona transformada con el vector de expresión de la reivindicación 7.
9. Un método para producir una proteína de adherencia y penetración de *Haemophilus influenzae* que comprende:
 - a) cultivar una célula anfitriona transformada con un vector de expresión que comprende un ácido nucleico recombinante que codifica una proteína de adherencia y penetración de *Haemophilus influenzae*, cuya proteína tiene una secuencia de aminoácidos con una homología de más del 80% con la secuencia de aminoácidos mostrada en la Figura 6; y
 - b) expresar dicho ácido nucleico para producir una proteína de adherencia y penetración de *Haemophilus influenzae*.
10. Un método reivindicado en la reivindicación 9, donde el ácido nucleico recombinante hibrida en condiciones muy restrictivas con el ácido nucleico que tiene la secuencia mostrada en la Figura 6.
11. Una vacuna que comprende un portador farmacéuticamente aceptable y una proteína de adherencia y penetración de *Haemophilus influenzae* para el uso profiláctico o terapéutico en la generación de una respuesta inmunitaria, cuya proteína tiene una secuencia de aminoácidos con una homología de más del 80% con la secuencia de aminoácidos mostrada en la Figura 6.
12. Una vacuna como se ha reivindicado en la reivindicación 11, donde el ácido nucleico que codifica dicha proteína de adherencia y penetración de *Haemophilus influenzae* hibrida en condiciones muy restrictivas con el ácido nucleico que tiene la secuencia mostrada en la Figura 6.
13. Una vacuna como se ha reivindicado en las reivindicaciones 11 o 12, donde dicha proteína de adherencia y penetración de *Haemophilus influenzae* tiene la secuencia de aminoácidos mostrada en la Figura 6.
14. El uso de una proteína de adherencia y penetración de *Haemophilus influenzae* de cualquiera de las reivindicaciones 1-3 en la fabricación de una vacuna para tratar o prevenir la infección por *Haemophilus influenzae*.
15. Un anticuerpo monoclonal capaz de unirse a la proteína de adherencia y penetración de *Haemophilus influenzae* de cualquiera de las reivindicaciones 1-3.



FIG._1B



FIG._1A



FIG._2A



FIG._2B

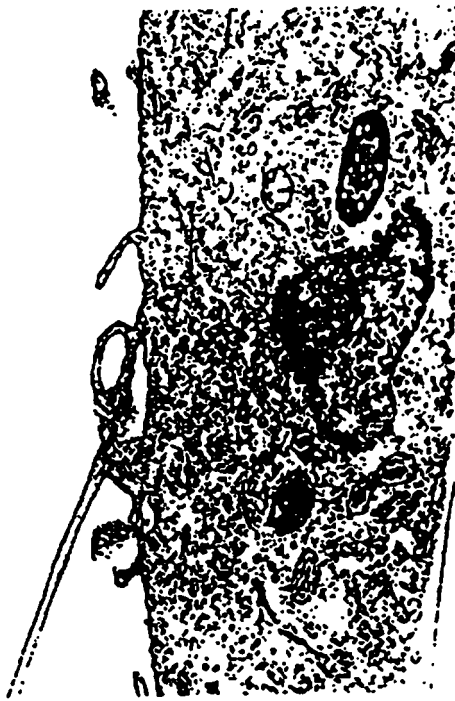


FIG..2D

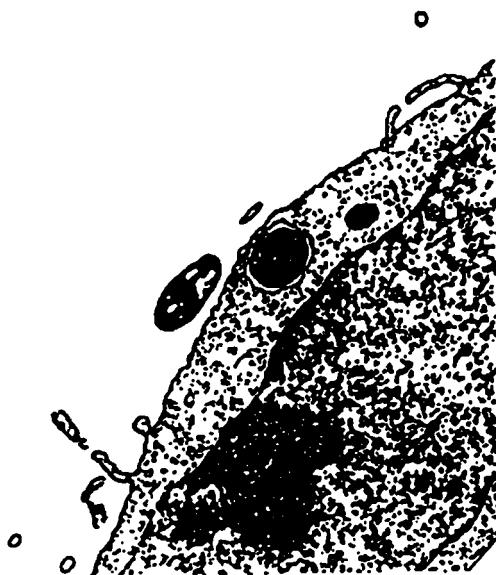
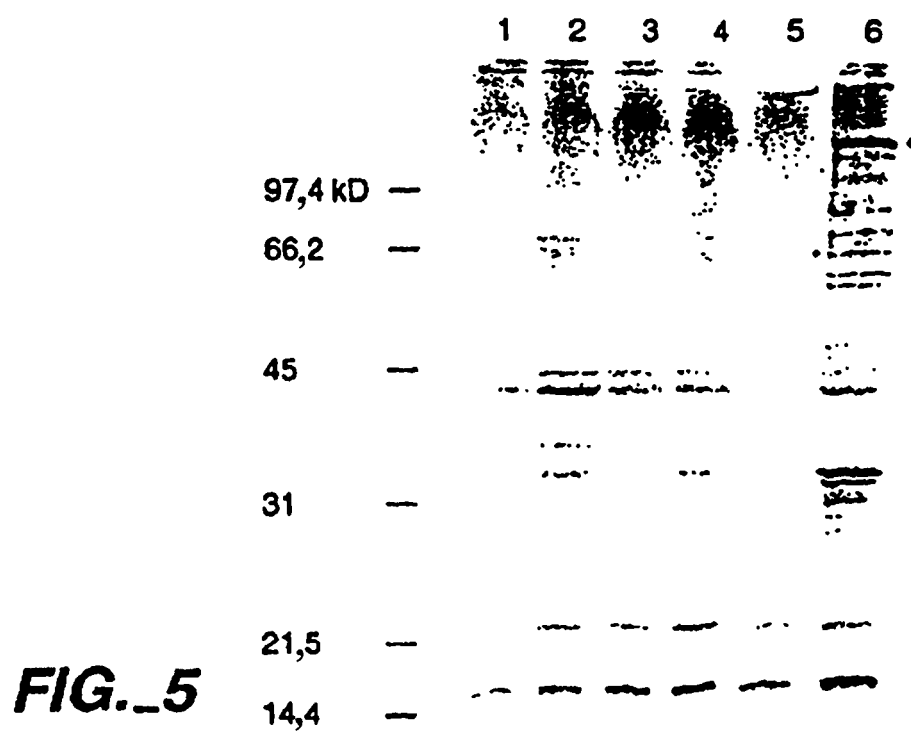
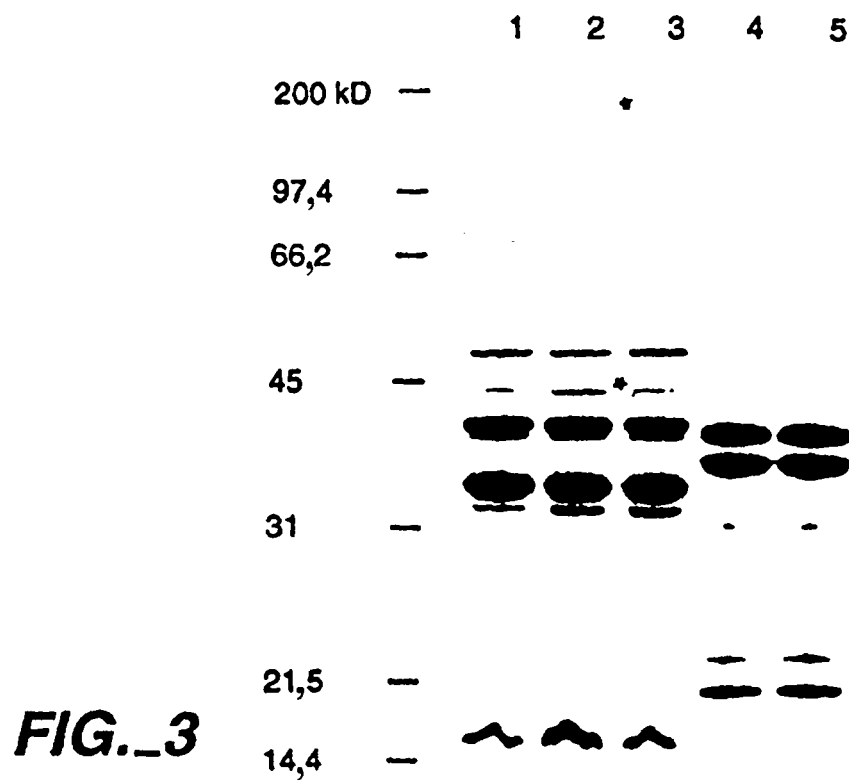
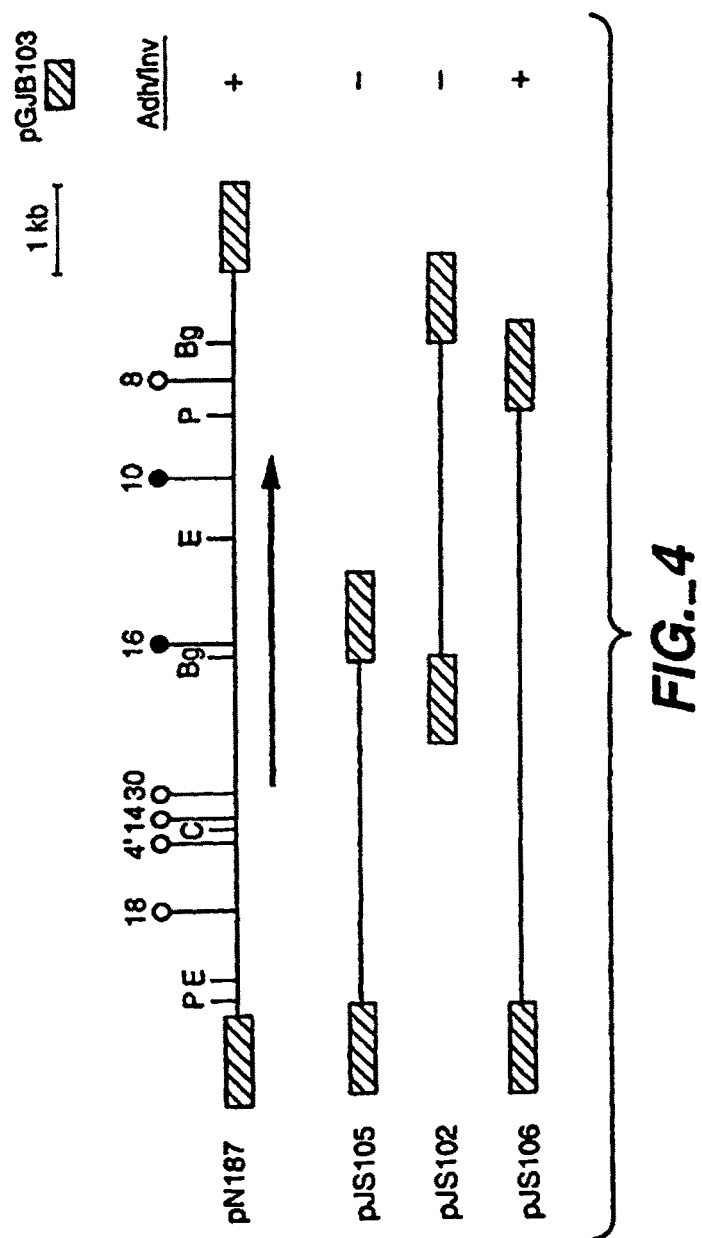


FIG..2C





10 30 50 70 90
 TCAATAGTCGTTTAACTAGTATTTTAAATACGAAATACTAAATTAATAAACAATTATGAAAAAAGCTGTATTTCTGCTCTTAATTTT
 -35 -10 M K K T V F R L N F
 110 130 150 170
 TTAACCGCTTGCAATTCATTAGGGATAGTATCGCAAGCGTGGCTGTCACACTTATTTGGGATTGATTACCAATATATTATCGTGATTTT
 L T A C I S L G I V S Q A W A G H T Y F G I D Y Q Y Y R D F
 190 210 230 250 270
 GCCGAGAAATAAAGGGAAGTTCCACAGTTGGGGCTCAAAATATTAAAGTTTATAACAACAAGGGCAATTAAGTTGGCACATCAATGACAAAA
 A E N K G K F T V G A Q N I K V Y N K Q G Q L V G T S M T K
 290 310 330 350
 GCCCGATGATTGATTTTCTGTAGTGTACGTACGTAACGCGTGGCAGCCCTGGTTGAAAATCAATATATTGTGAGCGGTGGCACATAACGTA
 A P M I D F S V V S R N G V A A L V E N Q Y I V S V A H N V
 370 390 410 430 450
 GGATATACAGATGTTGATTTTGGTGCAGAGGGAACAACCCCGATCAACATCGTTTACTTATAAGATTGTAAACGAAATAACTACAA
 G Y T D V D P G A E G N N P D Q H R F T Y K I V K R N N Y K
 470 490 510 530
 AAAGATAATTACATCCTTATGAGGACGATTACCATTAATCCACGATTACATAAATTCGTTACAGAGCGGCTCCAATTGATATGACTTCG
 K D N L H P Y E D D Y H N P R L H K F V T E A A P I D M T S
 550 570 590 610 630
 AATATGAAATGGCAGTACTTATTACAGATAGAACAAATAATCCAGAACGTGTTCGTTATCGGCTCTGGACGGCAGTTTGGCGAAATGATCAA
 N M N G S T Y S D R T K Y P E R V R I G S G R Q F W R N D Q
 650 670 690 710
 GACAAAGGCGACCAAGTTGCCGGTGCATATCATTTATCTGACAGCTGGCAATACACACAATCAGCGTGGAGCAGGTAAATGGATATTCGTAT
 D K G D Q V A G A Y H Y L T A G N T H N Q R G A G N G Y S Y

FIG.-6A

```

730      750      770      790      810
TTGGAGCGCATGTTTCGTAAGCGGGAGAAATATGGTCCATTACCGATTSCAGGCTCAAGGGGGACAGTGGTTCTCCGATGTTTATTTAT
L G G D V R K A G E Y G P L P I A G S K G D S G S P M P I Y

830      850      870      890
GATGCTGAAAACAAAATGGTTAATTAATGGGATATTACGGGAAGCAACCCCTTTGAAAGGCCAAAGAAAATGGGTTTCAATGGTTCGC
D A E K Q K W L I N G I L R E G N P F E G K E N G P Q L V R

910      930      950      970      990
AAATCTTATTTTGATGAAATTTTCGAAAGAGATTACATACATCAGTCTTACACCCGAGCTGGTAATGGAGTGTACACAATTAGTGGAAAT
K S Y F D E I F E R D L H T S L Y T R A G N G V Y T I S G N

1010     1030     1050     1070
GATAATGGTCAGGGGTCTATAACTCAGAAATCAGGAATACCATCAGAAAATTAAATTAACGTTAGCTTAGCAAAATAGAGTTTACCTTTGAAAAGAG
D N G Q G S I T Q K S G I P S E I K I T L A N M S L P L K E

1090     1110     1130     1150     1170
AAGGATAAAGTTCATAATCCTAGATATGACGGACCTAAATATTATCTCCACGTTTAAACAATGGAGAAAACGCTATATTTTATGGATCAA
K D K V H N P R Y D G P N I Y S P R L N N G E T L Y F M D Q

1190     1210     1230     1250
AAACAAGGATCATTAATCTTCGCATCTGACATTAACCAAGGGCGGGTGGTCTTTATTTTGAGGGTAATTTTACAGTATCTCCAAATCT
K Q G S L I F A S D I N Q G A G G L Y F E G N F T V S P N S

1270     1290     1310     1330     1350
AACCAAACTTGGCAAGGAGCTGGCATACATGTAAAGTGAATAAGCAGCCGTTACTTGGAAAGTAATAATGSCGTGGAACATGATCGACTTCT
N Q T W Q G A G I H V S E N S T V T W K V N G V E H D R L S

1370     1390     1410     1430
AAAATTGGTAAAGGAACATTGCACGTTCAAGCCAAAAGGGGAAAATAAAGGTTTCGATCAGCGTAGCCGATGCTAAAGTCATTTTGGAGCAG
K I G K G T L H V Q A K G E N K G S I S V G D G K V I L E Q

```

FIG._6B

1450 1470 1510 1530
 CAGGCAGCGATCAAGGCAACAACAGCCCTTTAGTGAATTGGCTTGTAGCGGCAGAGGGACTGTTCATTAATAACGATGATAAACAA
 Q A D D Q G N K Q A F S E I G L V S G R G T V Q L N D D K Q

1550 1570 1590 1610
 TTTGATACCGATAAATTTATTTTCGGCTTTCGTGGTGGTCGCTTAGATCTTAACGGGCATTCATTAACCTTTAAACGTAATCCAAAATACG
 F D T D K F Y F G F R G G R L D L N G H S L T F K R I Q N T

1630 1650 1670 1690 1710
 GACGAGGGGCAATGATTGTGAACCATATACAACTCAAGCCGCTAATGTCACTATTACTGGGAACGAAAGCATTTGTTCTACCTAATGGA
 D E G A M I V N H N T T Q A A N V T I T G N E S I V L P N G

1730 1750 1770 1790
 AATAATATTAACTTGATTACAGAAAGAAATTCCTACACGGTTGGTTGGCGAAACAGATATAAATAACACACAATGGGCGATTA
 N N I N K L D Y R K E I A Y N G W F G E T D K N K H N G R L

1810 1830 1850 1870 1890
 AACCTTATTATTAACCAACCAAGATCGTACTTTGCTACTTTTCAGGTGGTGACAAAATTAAGGCGGATATTACCCAAACAAAGGT
 N L I Y K P T T E D R T L L L S G G T N L K G D I T Q T K G

1910 1930 1950 1970
 AAACATATTTTTCAGCGGTAGACCGACACCGCACCGCTACAAATCAATTAATAAATGAAGTGAAGGATATACCAACAGGCGAA
 K L F F S G R P T P H A Y N H L N K R W S E M E G I P Q G E

1990 2010 2030 2050 2070
 ATTGTGTGGGATCACGATTGGATCAACCGTACATTTAAAGCTGAATACTCCAAATTAAGGCGGAAGTGCAGTGGTTCTCGCAATGTT
 I V W D H D W I N R T F K A E N F Q I K G G S A V V S R N V

2090 2110 2130 2150
 TCCTCAATTGAGGGAAATTGGACAGTCAGCAATAATGCAAAATGCCACATTTGGTGTGTGCGCAATCAACAAAATACCATTTGCACGCGT
 S S I E G N W T V S N N A N A T F G V V P N Q Q N T I C T R

FIG.-6C

2170 2190 2210 2230 2250
 TCAGATTGGACAGGATTACGACCTTGTCTAAAAGTGGATTTAACCGGATACAAAAGTTATTATCTATACCAAAAACACAAATCAATGGC
 S D W T G L T T C Q K V D L T D T K V I N S I P K T Q I N G

 2270 2290 2310 2330
 TCTATTAAATTAACTGATAATGCAACGGCGAATGTTAAAGGTTTAGCAAAACTTAATGGCAATGTCACCTTTAAACAAATCACAGCCAAATTT
 S I N L T D N A T A N V K G L A K L N G N V T L T N H S Q P

 2350 2370 2390 2410 2430
 ACATTAAAGCAACAATGCCACCCCAATAGGCAATATTCGACTTTCGACAAATTCAACTGCAACGGGTGGATAATGCAAACTTGAACGGTAAT
 T L S N N A T Q I G N I R L S D N S T A T V D N A N L N G N

 2450 2470 2490 2510
 GTGCATTAAACGGATTCAGCTCAATTTCTTTAAAAAACAGCCATTTTCGCACCAAAATTCAGGGAGACAAAGGCACAAACAGTGACGTTG
 V H L T D S A Q F S L K N S H F S H Q I Q G D K G T T V T L

 2530 2550 2570 2590 2610
 GAAAATCGGACTTGGACAAATGCCCTAGCGATACATTCAGCAATTTAAACGCTAAATAACAGTACGATCAGCTTAATTCAGCTTATTCA
 E N A T W T M P S D T T L Q N L T L N N S T I T L N S A Y S

 2630 2650 2670 2690
 GCTAGCTCAAAACAATACGCCACGTCGCCGTTTCATTAGAGACGGAAACAACGCCCAACATCGGCGAGAACATCGTTTCAACACATTGACAGTA
 A S S N N T P R R R S L E T E T P T S A E H R F N T L T V

 2710 2730 2750 2770 2790
 AATGGTAAATTGAGTGGGCAAGGCACATTCCCAATTACTTCATCTTTATTGGCTATAAAAGCGATAAAATTAAATTAATCCAAATGACGCT
 N G K L S G Q G T F Q F T S S L F G Y K S D K L K L S N D A

 2810 2830 2850 2870
 GAGGGCGATTACATATTATCTGTTCGCAACACAGGCAAGAAACCGGAACCCCTTGAGCAATTAACTTTGGTTGAAAGCAAGATTAATCAA
 E G D Y I L S V R N T G K E P E T L E Q L T L V E S K D N Q

FIG..6D

2890 2910 2930 2950 2970
 CCGTATCAGATAAGCTCAAATTTACTTTAGAAAATGACCACCGTTGATGCAGGTGCAATTACGTTATATAAATTAGTGAAGATGATGGCGAA
 P L S D K L K F T L E N D H V D A G A L R Y K L V K N D G E

 2990 3010 3030 3050
 TTCCGCTTGCAATAACCCAAATAAAGAGCAGGAATTCACAAATGATTTAGTAAGAGCAGAGCAAGCAGAACGAACATTAGAAGCCAAACAA
 F R L H N P I K E Q E L H N D L V R A E Q A E R T L E A K Q

 3070 3090 3110 3130 3150
 GTTGAACCGACTGCTAAACACAAACAGGTGAGCCAAAAGTGCGGTCAAGAAGAGCAGCGAGAGCAGCGTTTCCTGATACCCCTGCCCTGAT
 V E P T A K T Q T G E P K V R S R R A A R A A F P D T L P D

 3170 3190 3210 3230
 CAAAGCCTGTAAACGCATTAGAGCCAAACAGCTGAACCTGACTGCTGAAACACAAAAGTAAGGCCAAAACAAAAGTGCGGTCA
 Q S L L N A L E A K Q A E L T A E T Q K S K A K T K V R S

 3250 3270 3290 3310 3330
 AAAAGAGCAGTGTTTCTGATCCCTGCTTGATCMAAGCCTGTTCGCATTAGAAAGCCGCACTTGAGGTTATTGATGCCCCACAGCAATCG
 K R A V F S D P L L D Q S L F A L E A A L E V I D A P Q Q S

 3350 3370 3390 3410
 GAAAAGATCGTCTAGCTCMAAGAGCGGAAAACMACGCAACAAAAGACTTGATCAGCCGTTATTCAAATAGTGCGTTATCAGAA
 E K D R L A Q E E A E K Q R K Q K D L I S R Y S N S A L S E

 3430 3450 3470 3490 3510
 TTATCTGCAACAGTAAATAGTATGCTTCTGTTCMAAGATGAATTAGATCGTCTTTTGTAGATCAAGCACAATCTGCCGTGTGGACAAAT
 L S A T V N S M L S V Q D E L D R L F V D Q A Q S A V W T N

 3530 3550 3570 3590
 ATCGCACAGGATAAAAGACGCTATGATTCTGATGCGTTCGCTTATCAGCAGCAGAAAACGAACCTTACGTCAAATTGGGGTGCAAAAA
 I A Q D K R R Y D S D A F R A Y Q Q Q K T N L R Q I G V Q K

FIG._6E

3610 3630 3650 3670 3690
 GCCTTAGCTAATGGACGAATTGGGGCAGTTTCTCGCATAGCCGTTTCAGATTAATACCTTTTGATGAAACAGGTTAAAAATCACGCGACATTA
 A L A N G R I G A V F S H S R S D N T F D E Q V K N H A T L

 3710 3730 3750 3770
 ACGATGATGTCGGGTTTGGCCCAATATCAATGGGCGGATTACAAATTGGTGTAACGTTGGGAACGGGAATCAGTCCGAGTAAAAATGGCT
 T M M S G F A Q Y Q W G D L Q F G V N V G T G I S A S K M A

 3790 3810 3830 3850 3870
 GAAAGAACAAAGCCGAAAAATTCATCGAAAAAGCGATAAATTATGGCGTGAATGCAAGTTATCAGTTCCTCGTTTAGGGCAATTGGGCGCATTCAG
 E E Q S R K I H R K A I N Y G V N A S Y Q F R L G Q L G I Q

 3890 3910 3930 3950
 CCTTATTTTGGAGTTAATCGCTATTTTATTGAACGTGAAATTTATCAATCTGAGGAAGTGAGAGTGAAACGCCCTTAGCCCTTGCAATTAAAT
 P Y F G V N R Y F I E R E N Y Q S E E V R V K T P S L A F N

 3970 3990 4010 4030 4050
 CGCTATAATGCTGGCATTCGAGTTGATTTATACATTTTACTCCGACAGATAATATCAGCGTTAAGCCTTATTTCTTTCGTCAATTATGTGAT
 R Y N A G I R V D Y T F T P T D N I S V K P Y F F V N Y V D

 4070 4090 4110 4130
 GTTTCAAAACGCTAACGTACAAAACCCACGGTAAATCTCACGGTGTGCAACAACCACTTTGGACGTTATTGGCAAAAAGAGTGGGATTAAG
 V S N A N V Q T T V N L T V L Q Q P F G R Y W Q K E V G L K

 4150 4170 4190 4210 4230
 GCAGAAATTTTACATTTCCAAATTTCCGCTTTTATCTCAAAATCTCAAGGTTTCACAACTCGGCAACAGCAAAATGTGGGCGGTGAAATTG
 A E I L H F Q I S A F I S K S Q G S Q L G K Q Q N V G V K L

 4250 4270 4290 4310
 GGCTATCGTTGGTAAAAATCAACATATTTTATTCGTTTATTTGATAAACAAGGTGGTCCAGATCAGATCCCACTTTTATTTATCCAAATAAT
 G Y R W *

FIG.-6F

Hap	1	MKKTVERLNF	LTACISLGIV	SQAWAGHTYF	GIDYQYRRDF	AENKGRFTVG	50
HK368IGA		MLNKKFKLNF	IALTVAALV	PYTEAALVRD	DVDYQIFRDF	AENKGRFTVG	
HK393IGA		MLNKKFKLNF	IALTVAALV	PYTEAALVRD	DVDYQIFRDF	AENKGRFTVG	
HK715IGA		MLNKKFKLNF	IALTVAALV	PYTEAALVRD	DVDYQIFRDF	AENKGRFTVG	
HK61IGA		MLNKKFKLNF	IALTVAALV	PYTEAALVRD	DVDYQIFRDF	AENKGRFTVG	
Consenso		M-----F-LNF	-----A-----	-----A-----	--DYQ--RDF	AENKGF-P-VG	
Hap	51	AQNIKVYNKQ	GQLVGTSMK	A.PMIDFSV	SRNG.VAALV	ENQYIVSVAH	100
HK368IGA		ATNVLVKDKN	NKDLGTALPN	GIPMIDFSV	DVDKRIATLI	NPQYVVGVKH	
HK393IGA		ATNVEVRDKN	NRPLGNVLPN	GIPMIDFSV	DVDKRIATLV	NPQYVVGVKH	
HK715IGA		ATNVEVRDKN	NHSLGNVLPN	GIPMIDFSV	DVDKRIATLI	NPQYVVGVKH	
HK61IGA		ATNVEVRDKN	NHSLGNVLPN	GIPMIDFSV	DVDKRIATLV	NPQYVVGVKH	
Consenso		A-N--V--K-	-----G-----	--PMIDFSV	-----A-L-	--QY-V-V-H	
Hap	101	NVGY	TDVDFGAEGN	NPQHR....	..FTYKIVKR	NNY.....	150
HK368IGA		VSNQVSELHF	GNLNGNMNNG	NAKAHRDVSS	EENRYFSVEK	NEYPTKLNGK	
HK393IGA		VSNQVSELHF	GNLNGNMNNG	NAKAHRDVSS	EENRYYTVEK	NEYPTKLNGK	
HK715IGA		VSNQVSELHF	GNLNGNMNNG	NDKSHRDVSS	EENRYFSVEK	NEYPTKLNGK	
HK61IGA		VSNQVSELHF	GNLNGNMNNG	NAKSHRDVSS	EENRYYTVEK	NNPPTENVT	
Consenso		-----	-----	N---HR---	-----Y--V--	N-----	
Hap	151	KKDNLH	PYEDDYHNPR	LHKFVTEAAP	IDM.TSNMNG	STYSDRTKYP	200
HK368IGA		TVTTEDQ.TQ	KRREDYIMPR	LDKFVTEVAP	IEASTASSDA	GTYNQNKYP	
HK393IGA		AVTTEDQ.AQ	KRREDYIMPR	LDKFVTEVAP	IEASTSSDA	GTYNKDKYP	
HK715IGA		AVTTEDQ.TQ	KRREDYIMPR	LDKFVTEVAP	IEASTASSDA	GTYNQNKYP	
HK61IGA		PTTKEEQDAQ	KRREDYIMPR	LDKFVTEVAP	IEASTANNK	GEYNNSDKYP	
Consenso		-----	-----DY--PR	I-KFVTE-AP	I---T-----	--Y----KYP	

FIG..7A

Hap	201	ERVRIGSRQ F.....	WRNDQ	250	DKGDQVAGAY
HK368IGA		AFVRLGSCSQ FIYKKGDNYS LIL.....N NH.....	EVGG		NNLKLVGDAY
HK393IGA		YFVRLGSGTQ FIYENGTRYE LWL.....G KEGQKSDAGG			YNLKLVGDAY
HK715IGA		AFVRLGSCSQ FIYKKGDNYS LIL.....N NH.....	EVGG		NNLKLVGDAY
HK61IGA		AFVRLGSCSQ FIYKKGSRQY LILTERDKQG NLLRNWDVGG			DNLELVGNAY
Consenso		--VR-GSG-Q F-----	-----		-----V--AY
Hap	251	HYLTAGNTHN ORGAGNGYSY LGG.....D VRRAGEYGPL		300	PIAGSKGDSG
HK368IGA		TYGIAGTPYK VNHENGLIG FGNSKEEHS D			PKGILSQDPL TNYAVLGDSG
HK393IGA		TYGIAGTPYE VNHENDGLIG FGNSNNEYIN			PKEILSKKPL TNYAVLGDSG
HK715IGA		TYGIAGTPYK VNHENGLIG FGNSKEEHS D			PKGILSQDPL TNYAVLGDSG
HK61IGA		TYGIAGTPYK VNHENGLIG FGNSKEEHS D			PKGILSQDPL TNYAVLGDSG
Consenso		-Y--AG-----	-G-----		-----PL -----GDSC
Hap	301	SPMFIYDAEK QKWLINGILR EGNPFEGKEN	GFQLVRKSYF	350	D.EIFERDLH
HK368IGA		SPLFVYDREK GKWLFLGSDYD FWAGYN....	KKSQ		EWNIYKSQFT
HK393IGA		SPLFVYDREK GKWLFLGSDYD YWAGYN....	KKSQ		EWNIYKPEFA
HK715IGA		SPLFVYDREK GKWLFLGSDYD FWAGYN....	KKSQ		EWNIYKPEFA
HK61IGA		SPLFVYDREK GKWLFLGSDYD FWAGYN....	KKSQ		EWNIYKHEFA
Consenso		SP-F-YD-EK -KWL--G---	-----KS--		---I-----
Hap	351	TSLYTRAGNG VYTISGNDNG QGSITQKSGI	PSEIKITLAN	400	MSLPLKEKDK
HK368IGA		KDVLNPKDSAG SLIGSKTDYS WSSNGKTSTI	TGGEK....S		LNVDLAD...
HK393IGA		EKIYEYSAG SLIGSKTDYS WSSNGKTSTI	TGGEK....S		LNVDLAD...
HK715IGA		KTVLDKDTAG SLTGSNTQYN WNPTGKTSTI	SNGSE....S		LNVDLFD...
HK61IGA		EKIYQYSAG SLTGSNTQYT WQATGSTSTI	TGGEK....P		LSVDLTD...
Consenso		-----G -----S-----	-----S-I		-----L-----

FIG..7B

Hap	401	VHNPRYDGP	IYSPRLNNGE	TLYFMDQKQG	SLIFASDINQ	GAGGLYFEGN	450
HK368IGA		GKD.	KPNHGK	SVTFEG..SG	TLTLNNDIDQ	GAGGLPFEGD	
HK393IGA		GKD.	KPNHGK	SVTFEG..SG	TLTLNNDIDQ	GAGGLPFEGD	
HK715IGA		SSQD	TDSKKNHGK	SVTLRG..SG	TLTLNNDIDQ	GAGGLPFEGD	
HK61IGA		GKD.	KPNHGK	SITLKG..SG	TLTLNNDIDQ	GAGGLPFEGD	
Consenso		-----	-----N-G-	-----G-	-L-----I-Q	GAGGL-FEG-	
Hap	451	FTVSPNSNQ.	TWQGAGIHVS	ENSTVTWKVN	GVEHDLRSKI	GKGTLVQAK	500
HK368IGA		YEVKGTSDNT	TWKGAGVSA	EGKTVTWKVH	NPQYDRLAKI	GKGTLLIVEGT	
HK393IGA		YEVKGTSDNT	TWKGAGVSA	EGKTVTWKVH	NPQYDRLAKI	GKGTLLIVEGT	
HK715IGA		YEVKGTSDST	TWKGAGVSA	DGKTVTWKVH	NPKSDRLAKI	GKGTLLIVEGK	
HK61IGA		YEVKGTSDST	TWKGAGVSA	DGKTVTWKVH	NPQYDRLAKI	GKGTLLIVEGK	
Consenso		--V---S---	TW-GAG--V-	---TVTWKV-	----DRL-KI	GKGTLL-V---	
Hap	501	GENKGSISVG	DGKVILEQQA	DDQGNKQAPS	EIGLVSGRGT	VQLNDDKQPD	550
HK368IGA		GDNKGSLLKVG	DGTVILKQQT	NGSGQ.HAFA	SVGIVSGRST	LVLNDDKQVD	
HK393IGA		GDNKGSLLKVG	DGTVILKQQT	NGSGQ.HAFA	SVGIVSGRST	LVLNDDKQVD	
HK715IGA		GENKGSLLKVG	DGTVILKQQA	DANNKVAFS	QVGIVSGRST	VVLNDDKQVD	
HK61IGA		GKNEGLLKVG	DGTVILKQKA	DANNKVQAPS	QVGIVSGRST	LVLNDDKQVD	
Consenso		G-N-G---VG	DG-VIL-Q--	-----AF-	--G-VSGR-T	--LNDDKQ-D	
Hap	551	TDKFYFGFRG	GRDLNGHSL	TFKRIQNTDE	GAMIVNHNNT	QAANVTITGN	600
HK368IGA		PNSIYFGFRG	GRDLNGNSL	TFDHIRNIDD	GARLVNHNMT	NASNITITGE	
HK393IGA		PNSIYFGFRG	GRDLNGNSL	TFDHIRNIDE	GARLVNHNMT	NASNITITGE	
HK715IGA		PNSIYFGFRG	GRDLANGNNL	TFEHIRNIDD	GARLVNHNMT	KTSTVTITGE	
HK61IGA		PNSIYFGFRG	GRDLNGNSL	TFDHIRNIDD	GARLVNHNMT	NTSNITITGE	
Consenso		----YFGFRG	GRLD-NG--L	TF--I-N-D-	GA--VNH---	-----TITG-	

FIG..7C

Hap	601	ESIVLPNG...					650
HK368IGA		SLITDPNTIT	PYNIDAPDED	NPYAFRRIKD	GGQLYLNLEN	YTYVALRKGA	
HK393IGA		NLITDPNNVS	IYYVKPLEDD	NPYAIRQIKY	GYQLYFNEEN	RTYYVALKKDA	
HK715IGA		SLITDPNTIT	PYNIDAPDED	NPYAFRRIKD	GGQLYLNLEN	YTYVALRKGA	
HK61IGA		SLITDPNTIT	SYNIEAQDDD	HPLRIRSIPY	R.QLYFNQDN	RSYYTLKKGA	
Consenso		--I--PN--	-----	-----	-----	-----	
Hap	651						700
HK368IGA		STRSELPKNS	GESNENWLYM	GKTSDEAKRN	VNMHINNRM	NGFNGYFGEE	
HK393IGA		SIRSEFPQNR	GESNNSWLYM	GTEKADAQKN	AMNHINNRM	NGFNGYFGEE	
HK715IGA		STRSELPKNS	GESNENWLYM	GKTSDEAKRN	VNMHINNRM	NGFNGYFGEE	
HK61IGA		STRSELPQNS	GESNENWLYM	GRTSDEAKRN	VNMHINNRM	NGFNGYFGEE	
Consenso		-----	-----	-----	-----	-----	
Hap	701	D.KNKHNGRL	NLIYKPTTED	RTLILLGGTN	LKGDITQTKG	KLFFSGRPTP	750
HK368IGA		EGK..NNGNL	NVTFKGKSEQ	NRFLLTGGTN	LNGDLTVEKG	TLFLSGRPTP	
HK393IGA		EGK..NNGNL	NVTFKGKSEQ	NRFLLTGGTN	LNGDLNVQQG	TLFLSGRPTP	
HK715IGA		EGK..NNGNL	NVTFKGKSEQ	NRFLLTGGTN	LNGDLKVEKG	TLFLSGRPTP	
HK61IGA		ETKATQNGKL	NVTFNGKSDQ	NRFLLTGGTN	LNGDLNVEKG	TLFLSGRPTP	
Consenso		--K--NG-L	N-----	---LL-GGTN	L-GD-----	-LP-SGRPTP	
Hap	751	HAYNHLNKRW	SEMEG..IPQ	GEIVWDHDWI	NRTFKAENFQ	IKGGSVVVS.	800
HK368IGA		HARDIAGISS	TKKDPHPAEN	NEVVVEDDWI	NRNFKATTMN	VTGNASLYSG	
HK393IGA		HARDIAGISS	TKKDSHPFSEN	NEVVVEDDWI	NRNFKATNIN	VTNNATLYSG	
HK715IGA		HARDIAGISS	TKKDDQHPAEN	NEVVVEDDWI	NFNERATNIN	VTNNATLYSG	
HK61IGA		HARDIAGISS	TKKDDPHFTEN	NEVVVEDDWI	NRNFKATTMN	VTGNASLYSG	
Consenso		HA-----	-----	-E-V---DWI	NR-FKA----	-----S-	

FIG._7D

Hap	801	RNVSSIEGNW	TVSNNAATP	GVVPNQNTI	CTRSDWGTGLT	TCQKVDLTD	850
HK368IGA		RNVANITSNI	TASNKAQVHI	GY..KTGDTV	CVRSDYTGIV	TCCTDKLSD.	
HK393IGA		RNVESITSNI	TASNNAKVHI	GY..KAGDTV	CVRSDYTGIV	TCCTDKLSD.	
HK715IGA		RNVANITSNI	TASDNAKVHI	GY..KAGDTV	CVRSDYTGIV	TCCTDKLSD.	
HK61IGA		RNVANITSNI	TASNNAQVHI	GY..KTGDTV	CVRSDYTGIV	TCHNSNLSE.	
Consenso		RNV--I--N-	T-S--A----	G-----T-	C-RSD-TG--	TC-----L----	
					*	*	
Hap	851	KVINSIPKTQ	INGSINLTDN	ATANVKGLAK	LNGNVTLTNH	SQFTLSNNAT	900
HK368IGA		KALNSFNPTN	LRGNVNLTES	A.....			
HK393IGA		KALNSFNPTN	LRGNVNLTES	A.....			
HK715IGA		KALNSFNATN	VSGNVNLSCN	A.....			
HK61IGA		KALNSFNPTN	LRGNVNLTEN	A.....			
Consenso		K--NS---T-	--G--NL---	A-----	-----	-----	
Hap	901	QIGNIRLSDN	STATVDNANL	NGNVHLTDSA	QFSLKNSHPS	HQIQGDKGTT	950
HK368IGA			.NEVLGKANL	FGTIQSRGNS	QVRLT.....		
HK393IGA			.NEVLGKANL	FGTIQSRGNS	QVRLT.....		
HK715IGA			.NEVLGKANL	FGTIQSGTNS	QVRLT.....		
HK61IGA			.SFTLGKANL	FGTIQSIGTS	QVNLK.....		
Consenso		-----	-----ANL	-G-----	Q--L-----	-----	
Hap	951	VTLENATWTM	PSDATTLQNL	LNNSTITLNS	AYSASSNNTTP	RRRSLETETT	1000
HK368IGA		...ENSHWHL	TGNSDVHQLD	LANGHIHLNS	ADNSNNVTK.		
HK393IGA		...ENSHWHL	TGNSDVHQLD	LANGHIHLNS	ADNSNNVTK.		
HK715IGA		...ENSHMHL	TGDSNVNQLN	LDRGHIHLNA	QNDANKVTT.		
HK61IGA		...ENSHWHL	TGNSNVNQLN	LTNGHIHLNA	QNDANKVTT.		
Consenso		---EN--W--	-----L-	L-----I-LN-	-----	-----	

FIG..7E

Hap	1001	PTSAEHRFNT	LTVNGKLSGQ	GTFQFTSSLF	GKSDKCLKS	1050
HK368IGA		YNT	LTVNS.LSGN	GSFYLLTDLS	NKQGDKVVVT	NDAEGDYILS
HK393IGA		YNT	LTVNS.LSGN	GSFYLLTDLS	NKQGDKVVVT	KSATGNFTLQ
HK715IGA		YNT	LTVNS.LSGN	GSFYLLTDLS	NKQGDKVVVT	KSATGNFTLQ
HK61IGA		YNT	LTVNS.LSGN	GSFYLLTDLS	NKQGDKVVVT	KSATGNFTLQ
Consenso		-----NT	LTVN---LSG-	G-F-----	-----K----	---A-G---L-
Hap	1051	VRNTGKEPET	LEQLTLVESK	DNQPLSDCLK	FTLENDHVDA	1100
HK368IGA		VADKTGEPNH	.NELTLFDAS	KAQR..DHLN	VSLVGNTVDL	GALRYKLVKN
HK393IGA		VADKTGEPNH	.NELTLFDAS	KAQR..DHLN	VSLVGNTVDL	GAWKYKLRNV
HK715IGA		VADKTGEPTK	.NELTLFDAS	NATR..NNLN	VSLVGNTVDL	GAWKYKLRNV
HK61IGA		VADKTGEPNH	.NELTLFDAS	NATR..NNLE	VSLVGNTVDL	GAWKYKLRNV
Consenso		V-----EP--	---LTL----	-----L-	--L-----VD-	GA--YKL----
Hap	1101	DGEFRLHNPI	KEQELHNDLV			1150
HK368IGA		NGRYDLNYP.	.EVEKRNTQTV	DTTNITTPNN	IQADVPSVPS	NNEEIARVDE
HK393IGA		NGRYDLNYP.	.EVEKRNTQTV	DTTNITTPNN	IQADVPSVPS	NNEEIARVDE
HK715IGA		NGRYDLNYP.	.EVEKRNTQTV	DTTNITTPNN	IQADVPSVPS	NNEEIARV.E
HK61IGA		NGRYDLNYP.	.EVEKRNTQTV	DTTNITTPND	IQADAPSAQS	NNEEIARV.E
Consenso		-G---L-NP-	-E-E--N--V	-----	-----	-----
Hap	1151					1200
HK368IGA		APVPPPPAPAT				
HK393IGA		APVPPPPAPAT				
HK715IGA		TPVPPPPAPAT				
HK61IGA		TPVPPPPAPAT	ESAIASEQPE	TRPAETAQPA	MEETNTANST	ETAPKSDTAT
Consenso		-----	-----	-----	-----	-----

FIG..7F

Hap	1201			RAEQAERTLE	AKQVEPT		1250
HK368IGA			PSETTETVAE	NSKQESKTVE	KNEQDATETT	AQNREVAKEA	
HK393IGA			PSETTETVAE	NSKQESKTVE	KNEQDATETT	AQNREVAKEA	
HK715IGA			PSETTETVAE	NSKQESKTVE	KNEQDATETT	AQNREVAKEA	
HK61IGA		QTEPNSESV	PSETTEKVAE	NPPQENETVA	KNEQEATEPT	PQNGEVAKEE	
Consenso		-----	---Q---T---	-----T---	-----	-----	
Hap	1251		AKTQT GE				1300
HK368IGA		KSNVKANTQT	NEVAQSGSET	KETQTTETK.		ETATVE	
HK393IGA		KSNVKANTQT	NEVAQSGSET	KETQTTETK.		ETATVE	
HK715IGA		KPNVKANTQT	NEVAQSGSET	EETQTTETK.		ETAKVE	
HK61IGA		QPTVEANTQT	NEATQSEGKT	EETQTAETKS	EPTESVTVSE	NQPEKTVSQS	
Consenso		-----A-TQT -E-----	-----	-----	-----	-----	
Hap	1301						1350
HK368IGA		KEEK.....					
HK393IGA		KEEK.....					
HK715IGA		KEEKAKVEKE	EKAKVEKDEI	QEAPQMASET	SPKQAKPAPK	EVSTDTKVEE	
HK61IGA		TEDKVVVEKE	EKAKVETEET	QKAPQVTSKE	PPKQAEPAPE	EVPTDTNAEE	
Consenso		-----	-----	-----	-----	-----	
Hap	1351						1400
HK368IGA							
HK393IGA							
HK715IGA		TQVQAQPQTQ	STTVAAAEAT	SPNSKPAEET	.QPSEKTNAE	PVTPVVSKNQ	
HK61IGA		A..QALQQTQ	PTTVAAAEET	SPNSKPAEET	QQPSEKTNAE	PVTPVVS...	
Consenso		-----	-----	-----	-----	-----	

FIG._7G

1401			PKVRS	RRAARAAPPD	TLP.....	1450
Hap						
HK368IGA						
HK393IGA						
HK715IGA						
HK611IGA						
Consenso						
1451						1500
Hap						
HK368IGA						
HK393IGA						
HK715IGA						
HK611IGA						
Consenso						
1501	D	QSLNLEA.	KQAEI	TAETQKSKAK	TKK.....	1550
Hap						
HK368IGA						
HK393IGA						
HK715IGA						
HK611IGA						
Consenso						
1551						1600
Hap						
HK368IGA						
HK393IGA						
HK715IGA						
HK611IGA						
Consenso						

FIG..7H

1601	F	ALBALEVID	APQSEKDR	1650
Hap		PENTPATQ	PTVNSESSN.	..KPK.NRHR
HK368IGA		PENTPATQ	PTVNSESSN.	..KPK.NRHR
HK393IGA		RSISQPKETS	ABETTAASD	ETIADNSKR
HK715IGA	ESSEPKSRRR	RSVSQPKETS	ABETTVASTQ	ETTVDNSVST
HK61IGA	ESSEKSRRR	RSVSQPKETS	ABETTVASTQ	ETTVDNSVST
Consenso	-----	-----	-----	-----R-
1651			QKDLI	1700
Hap		PATSSND..		SRYNSALS
HK368IGA	SVRSVPHNVE	PATSSND..		RSTVALCDLT
HK393IGA	SVRSVPHNVE	PATSSND..		RSTVALCDLT
HK715IGA	SVRS.....E	PTVTNGSD		RSTVALCDLT
HK61IGA	SVQTSNYPEV	ELPTENAENA	ENVQSGNNVA	NSQPALRLNT
Consenso	-----	-----	-----L-	S---N---S-
1701	LSA.....TV	NSMLSVQDEL	DRL.FVDQAQ	1750
Hap		NVGKAVSQHI	SOLEMNNEGQ	YVWVVSNTSM
HK368IGA	ARAKAQFVAL	NVGKAVSQHI	SOLEMNNEGQ	YVWVVSNTSM
HK393IGA	ARAKAQFVAL	NVGKAVSQHI	SOLEMNNEGQ	YVWVVSNTSM
HK715IGA	AMAKAQFVAL	NVGKAVSQHI	SOLEMNNEGQ	YVWVVSNTSM
HK61IGA	AMAKAQFVAL	NVGKAVSQHI	SOLEMNNEGQ	YVWVVSNTSM
Consenso	----A-----	N-----V----	--L-----Q	---Y-S---R
1751	AYQQQKTNLR	QIGVQKALAN	GRIGAVFSHS	1800
Hap	RFSSKSTQTQ	LGWDQTISNN	VQLGGVFTYV	RNSNFDKAT
HK368IGA	RFSSKSTQTQ	LGWDQTISNN	VQLGGVFTYV	RNSNFDKAT
HK393IGA	RFSSKSTQTQ	LGWDQTISNN	VQLGGVFTYV	RNSNFDKAT
HK715IGA	RFSSKSTQTQ	LGWDQTISNN	VQLGGVFTYV	RNSNFDKAS
HK61IGA	RFSSKSTQTQ	LGWDQTISNN	VQLGGVFTYV	RNSNFDKAS
Consenso	-----T----	-----Q-----N	---G-VF---	R--N-FD---

FIG..71

Hap	1801	GFΛQYQWGDΛ	QF..GVNVGT	GISASKMAEE	QSRKIHRKAI	NYGVNASYQF	1850
HK368IGA		FYSKY.YADN	HWYLGIDLGY	GKFQSKLQTN	HNAKPARHTA	QFGLTAGKAF	
HK393IGA		FYSKY.YADN	HWYLGIDLGY	GKFQSKLQTN	HNAKPARHTA	QFGLTAGKAF	
HK715IGA		FYSKY.YADN	HWYLGIDLGY	GKFQSNLKTN	HNAKPARHTA	QFGLTAGKAF	
HK61IGA		FYSKY.YADN	HWYLGIDLGY	GKFQSNLQTN	NNAKPARHTA	QIGLTAGKAF	
Consenso		----Y---D-	----G---G-	G---S-----	---K--R---	--G--A---P	
Hap	1851	RLGQLGIGQPY	FGVNRYPFIER	ENYQSEEV RV	KTPSLAFNRY	NAGIRVDYTF	1900
HK368IGA		NLGNFGITPI	VGVRYSYLSN	ADFALDQARI	KVNPISVKTA	FAQVDLSYTY	
HK393IGA		NLGNFGITPI	VGVRYSYLSN	ADFALDQARI	KVNPISVKTA	FAQVDLSYTY	
HK715IGA		NLGNFGITPI	VGVRYSYLSN	ANFALAKDRI	KVNPISVKTA	FAQVDLSYTY	
HK61IGA		NLGNFAVKPT	VGVRYSYLSN	ADFALAQDRI	KVNPISVKTA	FAQVDLSYTY	
Consenso		-LG-----P-	-GV-----	-----R-	K-----	-A-----YT-	
Hap	1901	TPTDNISVKP	YFFVNVVDVS	NANVQTTVNL	TVLQPPFGRY	WQKEVGLKAE	1950
HK368IGA		.HLGEFSVTP	ILSARY.DAN	QSGSKINVNG	YDFAYNVENQ	QQYNAGLRLK	
HK393IGA		.HLGEFSVTP	ILSARY.DAN	QSGSKINVNG	YDFAYNVENQ	QQYNAGLRLK	
HK715IGA		.HLGEFSVTP	ILSARY.DTN	QSGSKINVNQ	YDFAYNVENQ	QQYNAGLRLK	
HK61IGA		.HLGEFSITP	ILSARY.DAN	QNGKINVSV	YDFAYNVENQ	QQYNAGLRLK	
Consenso		-----S--P	-----Y-D--	-----V--	-----	-Q---GLK--	
Hap	1951	ILHFQISAFI	SKSQGSQLGK	QONVGVKLG Y	RW		1982
HK368IGA		YHNVKLSLIG	GLTKAKQAEK	QKTAELKLSF	SF		
HK393IGA		YHNVKLSLIG	GLTKAKQAEK	QKTAELKLSF	SF		
HK715IGA		YHNVKLSLIG	GLTKAKQAEK	QKTAELKLSF	SF		
HK61IGA		YHNVKLSLIG	GLTKAKQAEK	QKTAELKLSF	SF		
Consenso		-----S---	-----Q--K	Q-----KL--	--		

FIG._7J

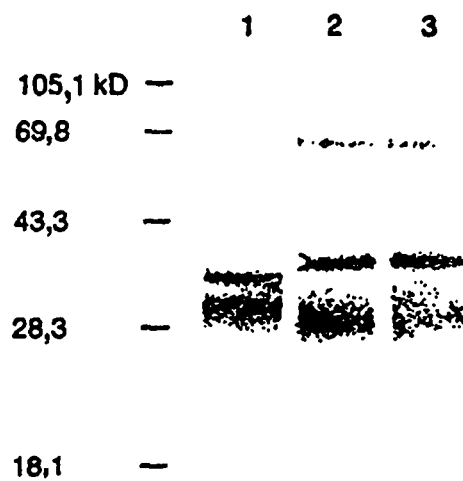


FIG._8

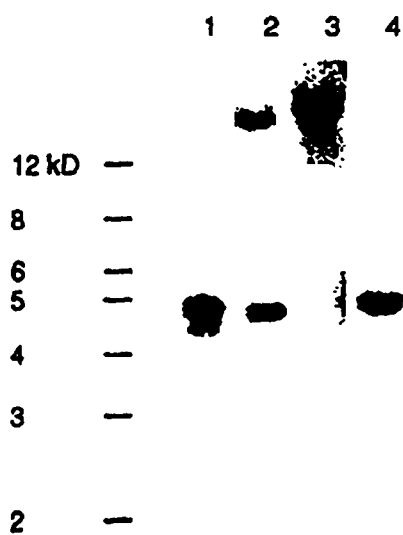


FIG._9A

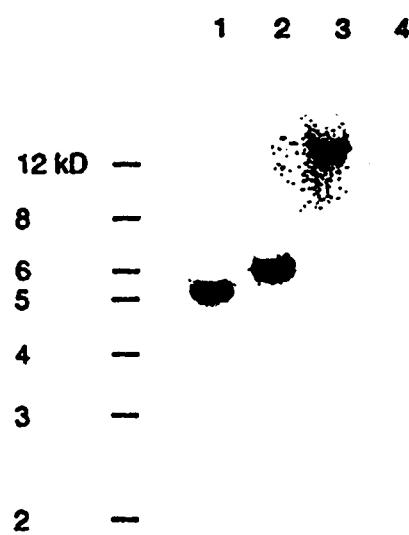


FIG._9B

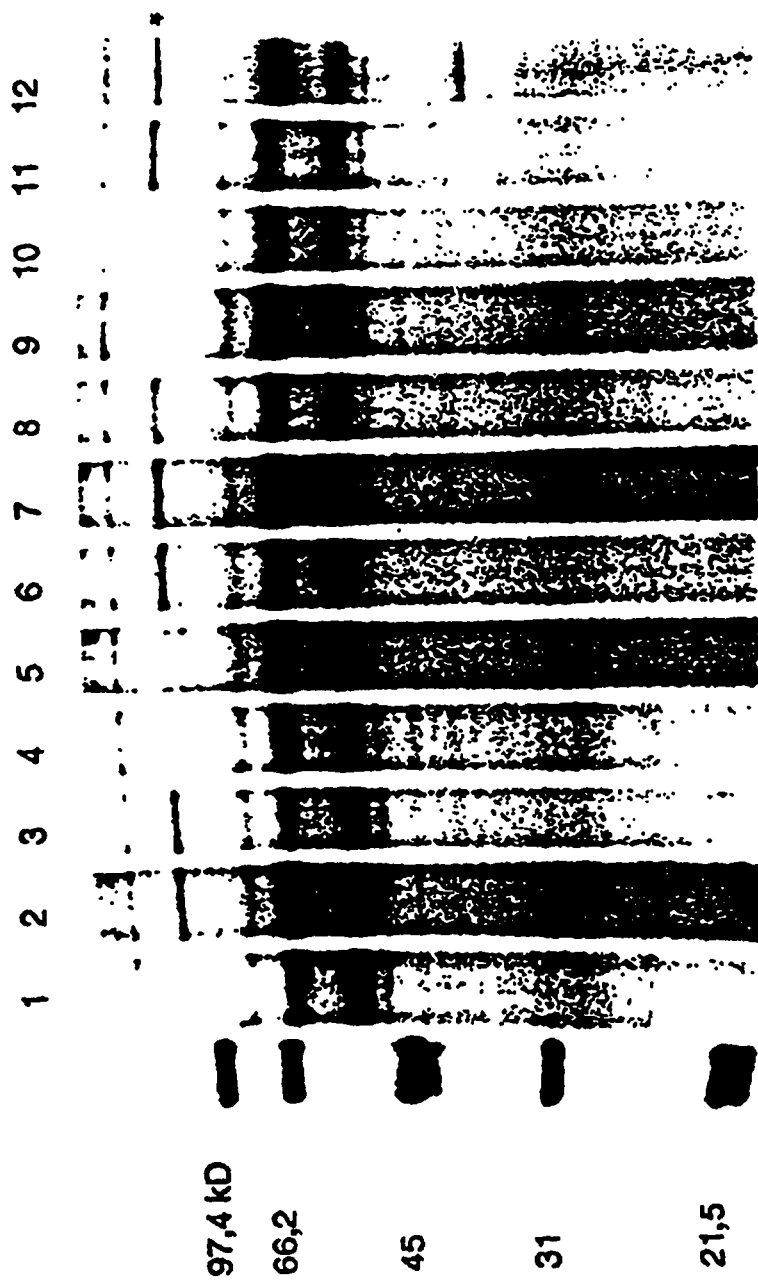


FIG. 10

LISTA DE SECUENCIAS

(1) INFORMACIÓN GENERAL:

- 5 (i) SOLICITANTE: Washington University *et al.*
- (ii) TÍTULO DE LA INVENCION: *Haemophilus Adherence* and Penetration Protein
- (iii) NÚMERO DE SECUENCIAS: 9
- 10 (iv) DIRECCIÓN DE CORRESPONDENCIA:
- (A) DESTINATARIO: Flehr, Hohbach, Test, Albritton & Herbert
- (B) CALLE: 4 Embarcadero Center, Suite 3400
- 15 (C) CIUDAD: San Francisco
- (D) ESTADO: California
- (E) PAÍS: Estados Unidos
- 20 (F) CÓDIGO POSTAL: 94111-4187
- (v) FORMA LEGIBLE CON ORDENADOR:
- (A) TIPO MEDIO: Disco Flexible
- (B) ORDENADOR: PC compatible con IBM
- 25 (C) SISTEMA OPERATIVO: PC-DOS/MS-DOS
- (D) SOPORTE LÓGICO: PatentIn Release #1.0, Versión #1.25
- (vi) DATOS DE LA SOLICITUD ACTUAL:
- 30 (A) NÚMERO DE SOLICITUD: PCT/US 95/10661
- (B) FECHA DE PRESENTACIÓN: 16-AUG-1995
- (C) CLASIFICACIÓN:
- 35 (vii) DATOS DE LA SOLICITUD ANTERIOR:
- (A) NÚMERO DE SOLICITUD: US 08/296,791
- (B) FECHA DE PRESENTACIÓN: 25-AUG-1994
- 40 (C) CLASIFICACIÓN:
- (viii) INFORMACIÓN DEL APODERADO/AGENTE:
- (A) NOMBRE: Trecartin, Richard F.
- 45 (B) NÚMERO DE REGISTRO: 31,801
- (C) NÚMERO DE REFERENCIA/EXPEDIENTE: FE-59941/RFT/RMS
- (ix) INFORMACIÓN DE TELECOMUNICACIONES:
- 50 (A) TELÉFONO: (415) 781-1989
- (B) TELEFAX: (415) 398-3249
- (C) TELEX: 910 277299

(2) INFORMACIÓN PARA EL SEQ ID NO: 1:

- 55 (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 4319 pares de bases
- 60 (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: doble
- (D) TOPOLOGÍA: both
- (ix) FEATURE:
- 65 (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 60..4241

ES 2 278 381 T3

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 1:

	TC AATAGTCG TT TAACTAGT AT TTTT TTAAT ACGAAAAATT ACTTAATTAA ATAAACATT	59
5		
	ATG AAA AAA ACT GTA TTT CGT CTT AAT TTT TTA ACC GCT TGC ATT TCA Met Lys Lys Thr Val Phe Arg Leu Asn Phe Leu Thr Ala Cys Ile Ser 1 5 10 15	107
10		
	TTA GGG ATA GTA TCG CAA GCG TGG GCT GGT CAC ACT TAT TTT GGG ATT Leu Gly Ile Val Ser Gln Ala Trp Ala Gly His Thr Tyr Phe Gly Ile 20 25 30	155
15		
	GAT TAC CAA TAT TAT CGT GAT TTT GCC GAG AAT AAA GGG AAG TTC ACA Asp Tyr Gln Tyr Tyr Arg Asp Phe Ala Glu Asn Lys Gly Lys Phe Thr 35 40 45	203
20		
	GTT GGG GCT CAA AAT ATT AAG GTT TAT AAC AAA CAA GGG CAA TTA GTT Val Gly Ala Gln Asn Ile Lys Val Tyr Asn Lys Gln Gly Gln Leu Val 50 55 60	251
25		
	GGC ACA TCA ATG ACA AAA GCC CCG ATG ATT GAT TTT TCT GTA GTG TCA Gly Thr Ser Met Thr Lys Ala Pro Met Ile Asp Phe Ser Val Val Ser 65 70 75 80	299
30		
	CGT AAC GGC GTG GCA GCC TTG GTT GAA AAT CAA TAT ATT GTG AGC GTG Arg Asn Gly Val Ala Ala Leu Val Glu Asn Gln Tyr Ile Val Ser Val 85 90 95	347
35		
	GCA CAT AAC GTA GGA TAT ACA GAT GTT GAT TTT GGT GCA GAG GGA AAC Ala His Asn Val Gly Tyr Thr Asp Val Asp Phe Gly Ala Glu Gly Asn 100 105 110	395
40		
	AAC CCC GAT CAA CAT CGT TTT ACT TAT AAG ATT GTA AAA CGA AAT AAC Asn Pro Asp Gln His Arg Phe Thr Tyr Lys Ile Val Lys Arg Asn Asn 115 120 125	443
45		
	TAC AAA AAA GAT AAT TTA CAT CCT TAT GAG GAC GAT TAC CAT AAT CCA Tyr Lys Lys Asp Asn Leu His Pro Tyr Glu Asp Asp Tyr His Asn Pro 130 135 140	491
50		
	CGA TTA CAT AAA TTC GTT ACA GAA GCG GCT CCA ATT GAT ATG ACT TCG Arg Leu His Lys Phe Val Thr Glu Ala Ala Pro Ile Asp Met Thr Ser 145 150 155 160	539
55		
	AAT ATG AAT GGC AGT ACT TAT TCA GAT AGA ACA AAA TAT CCA GAA CGT Asn Met Asn Gly Ser Thr Tyr Ser Asp Arg Thr Lys Tyr Pro Glu Arg 165 170 175	587
60		
	GTT CGT ATC GGC TCT GGA CGG CAG TTT TGG CGA AAT GAT CAA GAC AAA Val Arg Ile Gly Ser Gly Arg Gln Phe Trp Arg Asn Asp Gln Asp Lys 180 185 190	635
65		
	GGC GAC CAA GTT GCC GGT GCA TAT CAT TAT CTG ACA GCT GGC AAT ACA Gly Asp Gln Val Ala Gly Ala Tyr His Tyr Leu Thr Ala Gly Asn Thr 195 200 205	683
70		
	CAC AAT CAG CGT GGA GCA GGT AAT GGA TAT TCG TAT TTG GGA GGC GAT His Asn Gln Arg Gly Ala Gly Asn Gly Tyr Ser Tyr Leu Gly Gly Asp 210 215 220	731
75		
	GTT CGT AAA GCG GGA GAA TAT GGT CCA TTA CCG ATT GCA GGC TCA AAG Val Arg Lys Ala Gly Glu Tyr Gly Pro Leu Pro Ile Ala Gly Ser Lys 225 230 235 240	779
80		
	GGG GAC AGT GGT TCT CCG ATG TTT ATT TAT GAT GCT GAA AAA CAA AAA Gly Asp Ser Gly Ser Pro Met Phe Ile Tyr Asp Ala Glu Lys Gln Lys 245 250 255	827

ES 2 278 381 T3

5
10
15
20
25
30
35
40
45
50
55
60
65

TGG TTA ATT AAT GGG ATA TTA CGG GAA GGC AAC CCT TTT GAA GGC AAA 875
Trp Leu Ile Asn Gly Ile Leu Arg Glu Gly Asn Pro Phe Glu Gly Lys
260 265 270

GAA AAT GGG TTT CAA TTG GTT CGC AAA TCT TAT TTT GAT GAA ATT TTC 923
Glu Asn Gly Phe Gln Leu Val Arg Lys Ser Tyr Phe Asp Glu Ile Phe
275 280 285

GAA AGA GAT TTA CAT ACA TCA CTT TAC ACC CGA GCT GGT AAT GGA GTG 971
Glu Arg Asp Leu His Thr Ser Leu Tyr Thr Arg Ala Gly Asn Gly Val
290 295 300

TAC ACA ATT AGT GGA AAT GAT AAT GGT CAG GGG TCT ATA ACT CAG AAA 1019
Tyr Thr Ile Ser Gly Asn Asp Asn Gly Gln Gly Ser Ile Thr Gln Lys
305 310 315 320

TCA GGA ATA CCA TCA GAA ATT AAA ATT ACG TTA GCA AAT ATG AGT TTA 1067
Ser Gly Ile Pro Ser Glu Ile Lys Ile Thr Leu Ala Asn Met Ser Leu
325 330 335

CCT TTG AAA GAG AAG GAT AAA GTT CAT AAT CCT AGA TAT GAC GGA CCT 1115
Pro Leu Lys Glu Lys Asp Lys Val His Asn Pro Arg Tyr Asp Gly Pro
340 345 350

AAT ATT TAT TCT CCA CGT TTA AAC AAT GGA GAA ACG CTA TAT TTT ATG 1163
Asn Ile Tyr Ser Pro Arg Leu Asn Asn Gly Glu Thr Leu Tyr Phe Met
355 360 365

GAT CAA AAA CAA GGA TCA TTA ATC TTC GCA TCT GAC ATT AAC CAA GGG 1211
Asp Gln Lys Gln Gly Ser Leu Ile Phe Ala Ser Asp Ile Asn Gln Gly
370 375 380

GCG GGT GGT CTT TAT TTT GAG GGT AAT TTT ACA GTA TCT CCA AAT TCT 1259
Ala Gly Gly Leu Tyr Phe Glu Gly Asn Phe Thr Val Ser Pro Asn Ser
385 390 395 400

AAC CAA ACT TGG CAA GGA GCT GGC ATA CAT GTA AGT GAA AAT AGC ACC 1307
Asn Gln Thr Trp Gln Gly Ala Gly Ile His Val Ser Glu Asn Ser Thr
405 410 415

GTT ACT TGG AAA GTA AAT GGC GTG GAA CAT GAT CGA CTT TCT AAA ATT 1355
Val Thr Trp Lys Val Asn Gly Val Glu His Asp Arg Leu Ser Lys Ile
420 425 430

GGT AAA GGA ACA TTG CAC GTT CAA GCC AAA GGG GAA AAT AAA GGT TCG 1403
Gly Lys Gly Thr Leu His Val Gln Ala Lys Gly Glu Asn Lys Gly Ser
435 440 445

ATC AGC GTA GGC GAT GGT AAA GTC ATT TTG GAG CAG CAG GCA GAC GAT 1451
Ile Ser Val Gly Asp Gly Lys Val Ile Leu Glu Gln Gln Ala Asp Asp
450 455 460

CAA GGC AAC AAA CAA GCC TTT AGT GAA ATT GGC TTG GTT AGC GGC AGA 1499
Gln Gly Asn Lys Gln Ala Phe Ser Glu Ile Gly Leu Val Ser Gly Arg
465 470 475 480

GGG ACT GTT CAA TTA AAC GAT GAT AAA CAA TTT GAT ACC GAT AAA TTT 1547
Gly Thr Val Gln Leu Asn Asp Asp Lys Gln Phe Asp Thr Asp Lys Phe
485 490 495

TAT TTC GGC TTT CGT GGT GGT CGC TTA GAT CTT AAC GGG CAT TCA TTA 1595
Tyr Phe Gly Phe Arg Gly Gly Arg Leu Asp Leu Asn Gly His Ser Leu

ES 2 278 381 T3

	500	505	510	
5	ACC TTT AAA CGT ATC CAA AAT	ACG GAC GAG GGG GCA ATG ATT GTG AAC	1643	
	Thr Phe Lys Arg Ile Gln Asn	Thr Asp Glu Gly Ala Met Ile Val Asn		
	515	520	525	
10	CAT AAT ACA ACT CAA GCC GCT AAT GTC ACT ATT ACT GGG AAC GAA AGC	1691		
	His Asn Thr Thr Gln Ala Ala Asn Val Thr Ile Thr Gly Asn Glu Ser			
	530	535	540	
15	ATT GTT CTA CCT AAT GGA AAT AAT ATT AAT AAA CTT GAT TAC AGA AAA	1739		
	Ile Val Leu Pro Asn Gly Asn Asn Ile Asn Lys Leu Asp Tyr Arg Lys			
	545	550	555	560
20	GAA ATT GCC TAC AAC GGT TGG TTT GGC GAA ACA GAT AAA AAT AAA CAC	1787		
	Glu Ile Ala Tyr Asn Gly Trp Phe Gly Glu Thr Asp Lys Asn Arg His			
	565	570	575	
25	AAT GGG CGA TTA AAC CTT ATT TAT AAA CCA ACC ACA GAA GAT CGT ACT	1835		
	Asn Gly Arg Leu Asn Leu Ile Tyr Lys Pro Thr Thr Glu Asp Arg Thr			
	580	585	590	
30	TTG CTA CTT TCA GGT GGT ACA AAT TTA AAA GGC GAT ATT ACC CAA ACA	1883		
	Leu Leu Leu Ser Gly Gly Thr Asn Leu Lys Gly Asp Ile Thr Gln Thr			
	595	600	605	
35	AAA GGT AAA CTA TTT TTC AGC GGT AGA CCG ACA CCG CAC GCC TAC AAT	1931		
	Lys Gly Lys Leu Phe Phe Ser Gly Arg Pro Thr Pro His Ala Tyr Asn			
	610	615	620	
40	CAT TTA AAT AAA CGT TGG TCA GAA ATG GAA GGT ATA CCA CAA GGC GAA	1979		
	His Leu Asn Lys Arg Trp Ser Glu Met Glu Gly Ile Pro Gln Gly Glu			
	625	630	635	640
45	ATT GTG TGG GAT CAC GAT TGG ATC AAC CGT ACA TTT AAA GCT GAA AAC	2027		
	Ile Val Trp Asp His Asp Trp Ile Asn Arg Thr Phe Lys Ala Glu Asn			
	645	650	655	
50	TTC CAA ATT AAA GGC GGA AGT GCG GTG GTT TCT CGC AAT GTT TCT TCA	2075		
	Phe Gln Ile Lys Gly Gly Ser Ala Val Val Ser Arg Asn Val Ser Ser			
	660	665	670	
55	ATT GAG GGA AAT TGG ACA GTC AGC AAT AAT GCA AAT GCC ACA TTT GGT	2123		
	Ile Glu Gly Asn Trp Thr Val Ser Asn Asn Ala Asn Ala Thr Phe Gly			
	675	680	685	
60	GTT GTG CCA AAT CAA CAA AAT ACC ATT TGC ACG CGT TCA GAT TGG ACA	2171		
	Val Val Pro Asn Gln Gln Asn Thr Ile Cys Thr Arg Ser Asp Trp Thr			
	690	695	700	
65	GGA TTA ACG ACT TGT CAA AAA GTG GAT TTA ACC GAT ACA AAA GTT ATT	2219		
	Gly Leu Thr Thr Cys Gln Lys Val Asp Leu Thr Asp Thr Lys Val Ile			
	705	710	715	720
70	AAT TCT ATA CCA AAA ACA CAA ATC AAT GGC TCT ATT AAT TTA ACT GAT	2267		
	Asn Ser Ile Pro Lys Thr Gln Ile Asn Gly Ser Ile Asn Leu Thr Asp			
	725	730	735	
75	AAT GCA ACG GCG AAT GTT AAA GGT TTA GCA AAA CTT AAT GGC AAT GTC	2315		
	Asn Ala Thr Ala Asn Val Lys Gly Leu Ala Lys Leu Asn Gly Asn Val			
	740	745	750	

ES 2 278 381 T3

ACT TTA ACA AAT CAC AGC CAA TTT ACA TTA AGC AAC AAT GCC ACC CAA 2363
 Thr Leu Thr Asn His Ser Gln Phe Thr Leu Ser Asn Asn Ala Thr Gln
 755 760 765

5 ATA GGC AAT ATT CGA CTT TCC GAC AAT TCA ACT GCA ACG GTG GAT AAT 2411
 Ile Gly Asn Ile Arg Leu Ser Asp Asn Ser Thr Ala Thr Val Asp Asn
 770 775 780

10 GCA AAC TTG AAC GGT AAT GTG CAT TTA ACG GAT TCA GCT CAA TTT TCT 2459
 Ala Asn Leu Asn Gly Asn Val His Leu Thr Asp Ser Ala Gln Phe Ser
 785 790 795 800

15 TTA AAA AAC AGC CAT TTT TCG CAC CAA ATT CAG GGA GAC AAA GGC ACA 2507
 Leu Lys Asn Ser His Phe Ser His Gln Ile Gln Gly Asp Lys Gly Thr
 805 810 815

20 ACA GTG ACG TTG GAA AAT GCG ACT TGG ACA ATG CCT AGC GAT ACT ACA 2555
 Thr Val Thr Leu Glu Asn Ala Thr Trp Thr Met Pro Ser Asp Thr Thr
 820 825 830

25 TTG CAG AAT TTA ACG CTA AAT AAC AGT ACG ATC ACG TTA AAT TCA GCT 2603
 Leu Gln Asn Leu Thr Leu Asn Asn Ser Thr Ile Thr Leu Asn Ser Ala
 835 840 845

TAT TCA GCT AGC TCA AAC AAT ACG CCA CGT CGC CGT TCA TTA GAG ACG 2651
 Tyr Ser Ala Ser Ser Asn Asn Thr Pro Arg Arg Arg Ser Leu Glu Thr
 850 855 860

GAA ACA ACG CCA ACA TCG GCA GAA CAT CGT TTC AAC ACA TTG ACA GTA 2699
 Glu Thr Thr Pro Thr Ser Ala Glu His Arg Phe Asn Thr Leu Thr Val
 865 870 875 880

30 AAT GGT AAA TTG AGT GGG CAA GGC ACA TTC CAA TTT ACT TCA TCT TTA 2747
 Asn Gly Lys Leu Ser Gly Gln Gly Thr Phe Gln Phe Thr Ser Ser Leu
 885 890 895

35 TTT GGC TAT AAA AGC GAT AAA TTA AAA TTA TCC AAT GAC GCT GAG GGC 2795
 Phe Gly Tyr Lys Ser Asp Lys Leu Lys Leu Ser Asn Asp Ala Glu Gly
 900 905 910

GAT TAC ATA TTA TCT GTT CGC AAC ACA GGC AAA GAA CCC GAA ACC CTT 2843
 Asp Tyr Ile Leu Ser Val Arg Asn Thr Gly Lys Glu Pro Glu Thr Leu
 915 920 925

40 GAG CAA TTA ACT TTG GTT GAA AGC AAA GAT AAT CAA CCG TTA TCA GAT 2891
 Glu Gln Leu Thr Leu Val Glu Ser Lys Asp Asn Gln Pro Leu Ser Asp
 930 935 940

45 AAG CTC AAA TTT ACT TTA GAA AAT GAC CAC GTT GAT GCA GGT GCA TTA 2939
 Lys Leu Lys Phe Thr Leu Glu Asn Asp His Val Asp Ala Gly Ala Leu
 945 950 955 960

CGT TAT AAA TTA GTG AAG AAT GAT GGC GAA TTC CGC TTG CAT AAC CCA 2987
 Arg Tyr Lys Leu Val Lys Asn Asp Gly Glu Phe Arg Leu His Asn Pro
 965 970 975

50 ATA AAA GAG CAG GAA TTG CAC AAT GAT TTA GTA AGA GCA GAG CAA GCA 3035
 Ile Lys Glu Gln Glu Leu His Asn Asp Leu Val Arg Ala Glu Gln Ala
 980 985 990

55 GAA CGA ACA TTA GAA GCC AAA CAA GTT GAA CCG ACT GCT AAA ACA CAA 3083
 Glu Arg Thr Leu Glu Ala Lys Gln Val Glu Pro Thr Ala Lys Thr Gln

ES 2 278 381 T3

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

995	1000	1005	
ACA GGT GAG CCA AAA GTG CGG TCA AGA AGA GCA GCG AGA GCA GCG TTT			3131
Thr Gly Glu Pro Lys Val Arg Ser Arg Arg Ala Ala Arg Ala Ala Phe			
1010	1015	1020	
CCT GAT ACC CTG CCT GAT CAA AGC CTG TTA AAC GCA TTA GAA GCC AAA			3179
Pro Asp Thr Leu Pro Asp Gln Ser Leu Leu Asn Ala Leu Glu Ala Lys			
1025	1030	1035	1040
CAA GCT GAA CTG ACT GCT GAA ACA CAA AAA AGT AAG GCA AAA ACA AAA			3227
Gln Ala Glu Leu Thr Ala Glu Thr Gln Lys Ser Lys Ala Lys Thr Lys			
1045	1050	1055	
AAA GTG CGG TCA AAA AGA GCA GTG TTT TCT GAT CCC CTG CTT GAT CAA			3275
Lys Val Arg Ser Lys Arg Ala Val Phe Ser Asp Pro Leu Leu Asp Gln			
1060	1065	1070	
AGC CTG TTC GCA TTA GAA GCC GCA CTT GAG GTT ATT GAT GCC CCA CAG			3323
Ser Leu Phe Ala Leu Glu Ala Ala Leu Glu Val Ile Asp Ala Pro Gln			
1075	1080	1085	
CAA TCG GAA AAA GAT CGT CTA GCT CAA GAA GAA GCG GAA AAA CAA CGC			3371
Gln Ser Glu Lys Asp Arg Leu Ala Gln Glu Glu Ala Glu Lys Gln Arg			
1090	1095	1100	
AAA CAA AAA GAC TTG ATC AGC CGT TAT TCA AAT AGT GCG TTA TCA GAA			3419
Lys Gln Lys Asp Leu Ile Ser Arg Tyr Ser Asn Ser Ala Leu Ser Glu			
1105	1110	1115	1120
TTA TCT GCA ACA GTA AAT AGT ATG CTT TCT GTT CAA GAT GAA TTA GAT			3467
Leu Ser Ala Thr Val Asn Ser Met Leu Ser Val Gln Asp Glu Leu Asp			
1125	1130	1135	
CGT CTT TTT GTA GAT CAA GCA CAA TCT GCC GTG TGG ACA AAT ATC GCA			3515
Arg Leu Phe Val Asp Gln Ala Gln Ser Ala Val Trp Thr Asn Ile Ala			
1140	1145	1150	
CAG GAT AAA AGA CGC TAT GAT TCT GAT GCG TTC CGT GCT TAT CAG CAG			3563
Gln Asp Lys Arg Arg Tyr Asp Ser Asp Ala Phe Arg Ala Tyr Gln Gln			
1155	1160	1165	
CAG AAA ACG AAC TTA CGT CAA ATT GGG GTG CAA AAA GCC TTA GCT AAT			3611
Gln Lys Thr Asn Leu Arg Gln Ile Gly Val Gln Lys Ala Leu Ala Asn			
1170	1175	1180	
GGA CGA ATT GGG GCA GTT TTC TCG CAT AGC CGT TCA GAT AAT ACC TTT			3659
Gly Arg Ile Gly Ala Val Phe Ser His Ser Arg Ser Asp Asn Thr Phe			
1185	1190	1195	1200
GAT GAA CAG GTT AAA AAT CAC GCG ACA TTA ACG ATG ATG TCG GGT TTT			3707
Asp Glu Gln Val Lys Asn His Ala Thr Leu Thr Met Met Ser Gly Phe			
1205	1210	1215	
GCC CAA TAT CAA TGG GGC GAT TTA CAA TTT GGT GTA AAC GTG GGA ACG			3755
Ala Gln Tyr Gln Trp Gly Asp Leu Gln Phe Gly Val Asn Val Gly Thr			
1220	1225	1230	
GGA ATC AGT GCG AGT AAA ATG GCT GAA GAA CAA AGC CGA AAA ATT CAT			3803
Gly Ile Ser Ala Ser Lys Met Ala Glu Glu Gln Ser Arg Lys Ile His			
1235	1240	1245	

ES 2 278 381 T3

CGA AAA GCG ATA AAT TAT GGC GTG AAT GCA AGT TAT CAG TTC CGT TTA 3851
Arg Lys Ala Ile Asn Tyr Gly Val Asn Ala Ser Tyr Gln Phe Arg Leu
1250 1255 1260

GGG CAA TTG GGC ATT CAG CCT TAT TTT GGA GTT AAT CGC TAT TTT ATT 3899
Gly Gln Leu Gly Ile Gln Pro Tyr Phe Gly Val Asn Arg Tyr Phe Ile
1265 1270 1275 1280

GAA CGT GAA AAT TAT CAA TCT GAG GAA GTG AGA GTG AAA ACG CCT AGC 3947
Glu Arg Glu Asn Tyr Gln Ser Glu Glu Val Arg Val Lys Thr Pro Ser
1285 1290 1295

CTT GCA TTT AAT CGC TAT AAT GCT GGC ATT CGA GTT GAT TAT ACA TTT 3995
Leu Ala Phe Asn Arg Tyr Asn Ala Gly Ile Arg Val Asp Tyr Thr Phe
1300 1305 1310

ACT CCG ACA GAT AAT ATC AGC GTT AAG CCT TAT TTC TTC GTC AAT TAT 4043
Thr Pro Thr Asp Asn Ile Ser Val Lys Pro Tyr Phe Phe Val Asn Tyr
1315 1320 1325

GTT GAT GTT TCA AAC GCT AAC GTA CAA ACC ACG GTA AAT CTC ACG GTG 4091
Val Asp Val Ser Asn Ala Asn Val Gln Thr Thr Val Asn Leu Thr Val
1330 1335 1340

TTG CAA CAA CCA TTT GGA CGT TAT TGG CAA AAA GAA GTG GGA TTA AAG 4139
Leu Gln Gln Pro Phe Gly Arg Tyr Trp Gln Lys Glu Val Gly Leu Lys
1345 1350 1355 1360

GCA GAA ATT TTA CAT TTC CAA ATT TCC GCT TTT ATC TCA AAA TCT CAA 4187
Ala Glu Ile Leu His Phe Gln Ile Ser Ala Phe Ile Ser Lys Ser Gln
1365 1370 1375

GGT TCA CAA CTC GGC AAA CAG CAA AAT GTG GGC GTG AAA TTG GGC TAT 4235
Gly Ser Gln Leu Gly Lys Gln Gln Asn Val Gly Val Lys Leu Gly Tyr
1380 1385 1390

CGT TGG TAAAAATCA ACATAATTTT ATCGTTTATT GATAAACAAG GTGGGTCAGA 4290
Arg Trp

TCAGATCCCA CCTTTTTTAT TCCAATAAT 4319

(2) INFORMACIÓN PARA EL SEQ ID NO: 2:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1394 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLECULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 2:

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

Leu Gly Ile Val Ser Gln Ala Trp Ala Gly His Thr Tyr Phe Gly Ile
20 25 30

Asp Tyr Gln Tyr Tyr Arg Asp Phe Ala Glu Asn Lys Gly Lys Phe Thr

ES 2 278 381 T3

	35	40	45
5	Val Gly Ala Gln Asn Ile Lys	Val Tyr Asn Lys	Gln Gly Gln Leu Val
	50	55	60
	Gly Thr Ser Met Thr Lys Ala Pro Met Ile Asp Phe Ser Val Val Ser		
	65	70	75
10	Arg Asn Gly Val Ala Ala Leu Val Glu Asn Gln Tyr Ile Val Ser Val		
	85	90	95
	Ala His Asn Val Gly Tyr Thr Asp Val Asp Phe Gly Ala Glu Gly Asn		
	100	105	110
15	Asn Pro Asp Gln His Arg Phe Thr Tyr Lys Ile Val Lys Arg Asn Asn		
	115	120	125
	Tyr Lys Lys Asp Asn Leu His Pro Tyr Glu Asp Asp Tyr His Asn Pro		
	130	135	140
20	Arg Leu His Lys Phe Val Thr Glu Ala Ala Pro Ile Asp Met Thr Ser		
	145	150	155
	Asn Met Asn Gly Ser Thr Tyr Ser Asp Arg Thr Lys Tyr Pro Glu Arg		
	165	170	175
25	Val Arg Ile Gly Ser Gly Arg Gln Phe Trp Arg Asn Asp Gln Asp Lys		
	180	185	190
	Gly Asp Gln Val Ala Gly Ala Tyr His Tyr Leu Thr Ala Gly Asn Thr		
	195	200	205
30	His Asn Gln Arg Gly Ala Gly Asn Gly Tyr Ser Tyr Leu Gly Gly Asp		
	210	215	220
	Val Arg Lys Ala Gly Glu Tyr Gly Pro Leu Pro Ile Ala Gly Ser Lys		
	225	230	235
35	Gly Asp Ser Gly Ser Pro Met Phe Ile Tyr Asp Ala Glu Lys Gln Lys		
	245	250	255
	Trp Leu Ile Asn Gly Ile Leu Arg Glu Gly Asn Pro Phe Glu Gly Lys		
	260	265	270
40	Glu Asn Gly Phe Gln Leu Val Arg Lys Ser Tyr Phe Asp Glu Ile Phe		
	275	280	285
	Glu Arg Asp Leu His Thr Ser Leu Tyr Thr Arg Ala Gly Asn Gly Val		
	290	295	300
45	Tyr Thr Ile Ser Gly Asn Asp Asn Gly Gln Gly Ser Ile Thr Gln Lys		
	305	310	315
	Ser Gly Ile Pro Ser Glu Ile Lys Ile Thr Leu Ala Asn Met Ser Leu		
	325	330	335
50	Pro Leu Lys Glu Lys Asp Lys Val His Asn Pro Arg Tyr Asp Gly Pro		
	340	345	350
	Asn Ile Tyr Ser Pro Arg Leu Asn Asn Gly Glu Thr Leu Tyr Phe Met		
	355	360	365
55	Asp Gln Lys Gln Gly Ser Leu Ile Phe Ala Ser Asp Ile Asn Gln Gly		
	370	375	380

ES 2 278 381 T3

5 Ala Gly Gly Leu Tyr Phe Glu Gly Asn Phe Thr Val Ser Pro Asn Ser
 385 390 395 400
 Asn Gln Thr Trp Gln Gly Ala Gly Ile His Val Ser Glu Asn Ser Thr
 405 410 415
 Val Thr Trp Lys Val Asn Gly Val Glu His Asp Arg Leu Ser Lys Ile
 420 425 430
 10 Gly Lys Gly Thr Leu His Val Gln Ala Lys Gly Glu Asn Lys Gly Ser
 435 440 445
 Ile Ser Val Gly Asp Gly Lys Val Ile Leu Glu Gln Gln Ala Asp Asp
 450 455 460
 15 Gln Gly Asn Lys Gln Ala Phe Ser Glu Ile Gly Leu Val Ser Gly Arg
 465 470 475 480
 Gly Thr Val Gln Leu Asn Asp Asp Lys Gln Phe Asp Thr Asp Lys Phe
 485 490 495
 20 Tyr Phe Gly Phe Arg Gly Gly Arg Leu Asp Leu Asn Gly His Ser Leu
 500 505 510
 Thr Phe Lys Arg Ile Gln Asn Thr Asp Glu Gly Ala Met Ile Val Asn
 515 520 525
 25 His Asn Thr Thr Gln Ala Ala Asn Val Thr Ile Thr Gly Asn Glu Ser
 530 535 540
 Ile Val Leu Pro Asn Gly Asn Asn Ile Asn Lys Leu Asp Tyr Arg Lys
 545 550 555 560
 30 Glu Ile Ala Tyr Asn Gly Trp Phe Gly Glu Thr Asp Lys Asn Lys His
 565 570 575
 Asn Gly Arg Leu Asn Leu Ile Tyr Lys Pro Thr Thr Glu Asp Arg Thr
 580 585 590
 35 Leu Leu Leu Ser Gly Gly Thr Asn Leu Lys Gly Asp Ile Thr Gln Thr
 595 600 605
 Lys Gly Lys Leu Phe Phe Ser Gly Arg Pro Thr Pro His Ala Tyr Asn
 610 615 620
 40 His Leu Asn Lys Arg Trp Ser Glu Met Glu Gly Ile Pro Gln Gly Glu
 625 630 635 640
 Ile Val Trp Asp His Asp Trp Ile Asn Arg Thr Phe Lys Ala Glu Asn
 645 650 655
 45 Phe Gln Ile Lys Gly Gly Ser Ala Val Val Ser Arg Asn Val Ser Ser
 660 665 670
 Ile Glu Gly Asn Trp Thr Val Ser Asn Asn Ala Asn Ala Thr Phe Gly
 675 680 685
 50 Val Val Pro Asn Gln Gln Asn Thr Ile Cys Thr Arg Ser Asp Trp Thr
 690 695 700
 Gly Leu Thr Thr Cys Gln Lys Val Asp Leu Thr Asp Thr Lys Val Ile
 705 710 715 720
 55 Asn Ser Ile Pro Lys Thr Gln Ile Asn Gly Ser Ile Asn Leu Thr Asp

ES 2 278 381 T3

	725	730	735
5	Asn Ala Thr 740	Ala Asn Val Lys Gly Leu 745	Ala Lys Leu Asn Gly Asn Val 750
	Thr Leu Thr 755	Asn His Ser Gln Phe Thr 760	Leu Ser Asn Asn Ala Thr Gln 765
10	Ile Gly Asn 770	Ile Arg Leu Ser Asp 775	Asn Ser Thr Ala Thr Val Asp Asn 780
	Ala Asn Leu 785	Asn Gly Asn Val His Leu 790	Thr Asp Ser Ala Gln Phe Ser 795 800
15	Leu Lys Asn Ser 805	His Phe Ser His Gln 810	Ile Gln Gly Asp Lys Gly Thr 815
	Thr Val Thr 820	Leu Glu Asn Ala Thr 825	Trp Thr Met Pro Ser Asp Thr Thr 830
20	Leu Gln Asn 835	Leu Thr Leu Asn Asn Ser 840	Thr Ile Thr Leu Asn Ser Ala 845
	Tyr Ser Ala 850	Ser Ser Asn Asn Thr Pro 855	Arg Arg Arg Ser Leu Glu Thr 860
25	Glu Thr Thr 865	Pro Thr Ser Ala Glu His Arg 870 875	Phe Asn Thr Leu Thr Val 880
	Asn Gly Lys 885	Leu Ser Gly Gln Gly Thr 890	Phe Gln Phe Thr Ser Ser Leu 895
30	Phe Gly Tyr 900	Lys Ser Asp Lys Leu Lys 905	Leu Ser Asn Asp Ala Glu Gly 910
	Asp Tyr Ile 915	Leu Ser Val Arg Asn Thr 920	Gly Lys Glu Pro Glu Thr Leu 925
35	Glu Gln Leu 930	Thr Leu Val Glu Ser Lys Asp 935	Asn Gln Pro Leu Ser Asp 940
	Lys Leu Lys 945	Phe Thr Leu Glu Asn Asp 950	His Val Asp Ala Gly Ala Leu 955 960
40	Arg Tyr Lys 965	Leu Val Lys Asn Asp Gly 970	Glu Phe Arg Leu His Asn Pro 975
	Ile Lys Glu 980	Gln Glu Leu His Asn Asp 985	Leu Val Arg Ala Glu Gln Ala 990
45	Glu Arg Thr 995	Leu Glu Ala Lys Gln Val 1000	Glu Pro Thr Ala Lys Thr Gln 1005
	Thr Gly Glu 1010	Pro Lys Val Arg Ser Arg 1015	Arg Ala Ala Arg Ala Ala Phe 1020
50	Pro Asp Thr 1025	Leu Pro Asp Gln Ser Leu 1030	Leu Asn Ala Leu Glu Ala Lys 1035 1040
	Gln Ala Glu 1045	Leu Thr Ala Glu Thr Gln 1050	Lys Ser Lys Ala Lys Thr Lys 1055
55	Lys Val Arg 1060	Ser Lys Arg Ala Val Phe 1065	Ser Asp Pro Leu Leu Asp Gln 1070

ES 2 278 381 T3

Ser Leu Phe Ala Leu Glu Ala Ala Leu Glu Val Ile Asp Ala Pro Gln
1075 1080 1085

Gln Ser Glu Lys Asp Arg Leu Ala Gln Glu Glu Ala Glu Lys Gln Arg
1090 1095 1100

Lys Gln Lys Asp Leu Ile Ser Arg Tyr Ser Asn Ser Ala Leu Ser Glu
1105 1110 1115 1120

Leu Ser Ala Thr Val Asn Ser Met Leu Ser Val Gln Asp Glu Leu Asp
1125 1130 1135

Arg Leu Phe Val Asp Gln Ala Gln Ser Ala Val Trp Thr Asn Ile Ala
1140 1145 1150

Gln Asp Lys Arg Arg Tyr Asp Ser Asp Ala Phe Arg Ala Tyr Gln Gln
1155 1160 1165

Gln Lys Thr Asn Leu Arg Gln Ile Gly Val Gln Lys Ala Leu Ala Asn
1170 1175 1180

Gly Arg Ile Gly Ala Val Phe Ser His Ser Arg Ser Asp Asn Thr Phe
1185 1190 1195 1200

Asp Glu Gln Val Lys Asn His Ala Thr Leu Thr Met Met Ser Gly Phe
1205 1210 1215

Ala Gln Tyr Gln Trp Gly Asp Leu Gln Phe Gly Val Asn Val Gly Thr
1220 1225 1230

Gly Ile Ser Ala Ser Lys Met Ala Glu Glu Gln Ser Arg Lys Ile His
1235 1240 1245

Arg Lys Ala Ile Asn Tyr Gly Val Asn Ala Ser Tyr Gln Phe Arg Leu
1250 1255 1260

Gly Gln Leu Gly Ile Gln Pro Tyr Phe Gly Val Asn Arg Tyr Phe Ile
1265 1270 1275 1280

Glu Arg Glu Asn Tyr Gln Ser Glu Glu Val Arg Val Lys Thr Pro Ser
1285 1290 1295

Leu Ala Phe Asn Arg Tyr Asn Ala Gly Ile Arg Val Asp Tyr Thr Phe
1300 1305 1310

Thr Pro Thr Asp Asn Ile Ser Val Lys Pro Tyr Phe Phe Val Asn Tyr
1315 1320 1325

Val Asp Val Ser Asn Ala Asn Val Gln Thr Thr Val Asn Leu Thr Val
1330 1335 1340

Leu Gln Gln Pro Phe Gly Arg Tyr Trp Gln Lys Glu Val Gly Leu Lys
1345 1350 1355 1360

Ala Glu Ile Leu His Phe Gln Ile Ser Ala Phe Ile Ser Lys Ser Gln
1365 1370 1375

Gly Ser Gln Leu Gly Lys Gln Gln Asn Val Gly Val Lys Leu Gly Tyr
1380 1385 1390

Arg Trp

(2) INFORMACIÓN PARA EL SEQ ID NO: 3:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1541 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGÍA: desconocida

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 3:

ES 2 278 381 T3

Met Leu Asn Lys Lys Phe Lys Leu Asn Phe Ile Ala Leu Thr Val Ala
 1 5 10 15
 Tyr Ala Leu Thr Pro Tyr Thr Glu Ala Ala Leu Val Arg Asp Asp Val
 5 20 25 30
 Asp Tyr Gln Ile Phe Arg Asp Phe Ala Glu Asn Lys Gly Lys Phe Ser
 35 40 45
 Val Gly Ala Thr Asn Val Leu Val Lys Asp Lys Asn Asn Lys Asp Leu
 10 50 55 60
 Gly Thr Ala Leu Pro Asn Gly Ile Pro Met Ile Asp Phe Ser Val Val
 65 70 75 80
 Asp Val Asp Lys Arg Ile Ala Thr Leu Ile Asn Pro Gln Tyr Val Val
 15 85 90 95
 Gly Val Lys His Val Ser Asn Gly Val Ser Glu Leu His Phe Gly Asn
 100 105 110
 Leu Asn Gly Asn Met Asn Asn Gly Asn Ala Lys Ala His Arg Asp Val
 115 120 125
 Ser Ser Glu Glu Asn Arg Tyr Phe Ser Val Glu Lys Asn Glu Tyr Pro
 130 135 140
 Thr Lys Leu Asn Gly Lys Thr Val Thr Thr Glu Asp Gln Thr Gln Lys
 145 150 155 160
 Arg Arg Glu Asp Tyr Tyr Met Pro Arg Leu Asp Lys Phe Val Thr Glu
 165 170 175
 Val Ala Pro Ile Glu Ala Ser Thr Ala Ser Ser Asp Ala Gly Thr Tyr
 180 185 190
 Asn Asp Gln Asn Lys Tyr Pro Ala Phe Val Arg Leu Gly Ser Gly Ser
 195 200 205
 Gln Phe Ile Tyr Lys Lys Gly Asp Asn Tyr Ser Leu Ile Leu Asn Asn
 210 215 220
 His Glu Val Gly Gly Asn Asn Leu Lys Leu Val Gly Asp Ala Tyr Thr
 225 230 235 240
 Tyr Gly Ile Ala Gly Thr Pro Tyr Lys Val Asn His Glu Asn Asn Gly
 245 250 255
 Leu Ile Gly Phe Gly Asn Ser Lys Glu Glu His Ser Asp Pro Lys Gly
 260 265 270
 Ile Leu Ser Gln Asp Pro Leu Thr Asn Tyr Ala Val Leu Gly Asp Ser
 275 280 285

ES 2 278 381 T3

Gly Ser Pro Leu Phe Val Tyr Asp Arg Glu Lys Gly Lys Trp Leu Phe
 290 295 300
 5 Leu Gly Ser Tyr Asp Phe Trp Ala Gly Tyr Asn Lys Lys Ser Trp Gln
 305 310 315 320
 Glu Trp Asn Ile Tyr Lys Ser Gln Phe Thr Lys Asp Val Leu Asn Lys
 325 330 335
 10 Asp Ser Ala Gly Ser Leu Ile Gly Ser Lys Thr Asp Tyr Ser Trp Ser
 340 345 350
 Ser Asn Gly Lys Thr Ser Thr Ile Thr Gly Gly Glu Lys Ser Leu Asn
 355 360 365
 15 Val Asp Leu Ala Asp Gly Lys Asp Lys Pro Asn His Gly Lys Ser Val
 370 375 380
 Thr Phe Glu Gly Ser Gly Thr Leu Thr Leu Asn Asn Asn Ile Asp Gln
 385 390 395 400
 20 Gly Ala Gly Gly Leu Phe Phe Glu Gly Asp Tyr Glu Val Lys Gly Thr
 405 410 415
 Ser Asp Asn Thr Thr Trp Lys Gly Ala Gly Val Ser Val Ala Glu Gly
 420 425 430
 25 Lys Thr Val Thr Trp Lys Val His Asn Pro Gln Tyr Asp Arg Leu Ala
 435 440 445
 Lys Ile Gly Lys Gly Thr Leu Ile Val Glu Gly Thr Gly Asp Asn Lys
 450 455 460
 30 Gly Ser Leu Lys Val Gly Asp Gly Thr Val Ile Leu Lys Gln Gln Thr
 465 470 475 480
 Asn Gly Ser Gly Gln His Ala Phe Ala Ser Val Gly Ile Val Ser Gly
 485 490 495
 35 Arg Ser Thr Leu Val Leu Asn Asp Asp Lys Gln Val Asp Pro Asn Ser
 500 505 510
 Ile Tyr Phe Gly Phe Arg Gly Gly Arg Leu Asp Leu Asn Gly Asn Ser
 515 520 525
 40 Leu Thr Phe Asp His Ile Arg Asn Ile Asp Asp Gly Ala Arg Leu Val
 530 535 540
 Asn His Asn Met Thr Asn Ala Ser Asn Ile Thr Ile Thr Gly Glu Ser
 545 550 555 560
 45 Leu Ile Thr Asp Pro Asn Thr Ile Thr Pro Tyr Asn Ile Asp Ala Pro
 565 570 575
 Asp Glu Asp Asn Pro Tyr Ala Phe Arg Arg Ile Lys Asp Gly Gly Gln
 580 585 590
 50 Leu Tyr Leu Asn Leu Glu Asn Tyr Thr Tyr Tyr Ala Leu Arg Lys Gly
 595 600 605
 Ala Ser Thr Arg Ser Glu Leu Pro Lys Asn Ser Gly Glu Ser Asn Glu
 610 615 620
 55 Asn Trp Leu Tyr Met Gly Lys Thr Ser Asp Glu Ala Lys Arg Asn Val

ES 2 278 381 T3

	625				630					635				640
5	Met	Asn	His	Ile	Asn	Asn	Glu	Arg	Met	Asn	Gly	Phe	Asn	Gly
					645					650			Tyr	Phe
	Gly	Glu	Glu	Glu	Gly	Lys	Asn	Asn	Gly	Asn	Leu	Asn	Val	Thr
					660				665				670	Lys
10	Gly	Lys	Ser	Glu	Gln	Asn	Arg	Phe	Leu	Leu	Thr	Gly	Gly	Thr
			675					680					685	Asn
	Asn	Gly	Asp	Leu	Thr	Val	Glu	Lys	Gly	Thr	Leu	Phe	Leu	Ser
		690					695					700	Gly	Arg
15	Pro	Thr	Pro	His	Ala	Arg	Asp	Ile	Ala	Gly	Ile	Ser	Ser	Thr
	705					710					715			Lys
	Asp	Pro	His	Phe	Ala	Glu	Asn	Asn	Glu	Val	Val	Val	Glu	Asp
					725					730				735
20	Ile	Asn	Arg	Asn	Phe	Lys	Ala	Thr	Thr	Met	Asn	Val	Thr	Gly
					740				745					750
	Ser	Leu	Tyr	Ser	Gly	Arg	Asn	Val	Ala	Asn	Ile	Thr	Ser	Asn
			755				760						765	Ile
25	Ala	Ser	Asn	Lys	Ala	Gln	Val	His	Ile	Gly	Tyr	Lys	Thr	Gly
			770				775					780		Asp
	Val	Cys	Val	Arg	Ser	Asp	Tyr	Thr	Gly	Tyr	Val	Thr	Cys	Thr
						790					795			800
30	Lys	Leu	Ser	Asp	Lys	Ala	Leu	Asn	Ser	Phe	Asn	Pro	Thr	Asn
					805					810				815
	Gly	Asn	Val	Asn	Leu	Thr	Glu	Ser	Ala	Asn	Phe	Val	Leu	Gly
					820				825					830
35	Asn	Leu	Phe	Gly	Thr	Ile	Gln	Ser	Arg	Gly	Asn	Ser	Gln	Val
			835					840					845	Arg
	Thr	Glu	Asn	Ser	His	Trp	His	Leu	Thr	Gly	Asn	Ser	Asp	Val
		850				855					860			His
40	Leu	Asp	Leu	Ala	Asn	Gly	His	Ile	His	Leu	Asn	Ser	Ala	Asp
		865				870					875			880
	Asn	Asn	Val	Thr	Lys	Tyr	Asn	Thr	Leu	Thr	Val	Asn	Ser	Leu
					885					890				895
45	Asn	Gly	Ser	Phe	Tyr	Tyr	Leu	Thr	Asp	Leu	Ser	Asn	Lys	Gln
				900					905					910
	Lys	Val	Val	Val	Thr	Lys	Ser	Ala	Thr	Gly	Asn	Phe	Thr	Leu
					915			920					925	Gln
50	Ala	Asp	Lys	Thr	Gly	Glu	Pro	Asn	His	Asn	Glu	Leu	Thr	Leu
		930					935					940		Phe
	Ala	Ser	Lys	Ala	Gln	Arg	Asp	His	Leu	Asn	Val	Ser	Leu	Val
					945		950			955				Gly
55	Thr	Val	Asp	Leu	Gly	Ala	Trp	Lys	Tyr	Lys	Leu	Arg	Asn	Val
					965					970				975

ES 2 278 381 T3

Arg Tyr Asp Leu Tyr Asn Pro Glu Val Glu Lys Arg Asn Gln Thr Val
 980 985 990
 5 Asp Thr Thr Asn Ile Thr Thr Pro Asn Asn Ile Gln Ala Asp Val Pro
 995 1000 1005
 Ser Val Pro Ser Asn Asn Glu Glu Ile Ala Arg Val Asp Glu Ala Pro
 1010 1015 1020
 10 Val Pro Pro Pro Ala Pro Ala Thr Pro Ser Glu Thr Thr Glu Thr Val
 1025 1030 1035 1040
 Ala Glu Asn Ser Lys Gln Glu Ser Lys Thr Val Glu Lys Asn Glu Gln
 1045 1050 1055
 15 Asp Ala Thr Glu Thr Thr Ala Gln Asn Arg Glu Val Ala Lys Glu Ala
 1060 1065 1070
 Lys Ser Asn Val Lys Ala Asn Thr Gln Thr Asn Glu Val Ala Gln Ser
 1075 1080 1085
 20 Gly Ser Glu Thr Lys Glu Thr Gln Thr Thr Glu Thr Lys Glu Thr Ala
 1090 1095 1100
 Thr Val Glu Lys Glu Glu Lys Ala Lys Val Glu Thr Glu Lys Thr Gln
 1105 1110 1115 1120
 25 Glu Val Pro Lys Val Thr Ser Gln Val Ser Pro Lys Gln Glu Gln Ser
 1125 1130 1135
 Glu Thr Val Gln Pro Gln Ala Glu Pro Ala Arg Glu Asn Asp Pro Thr
 1140 1145 1150
 30 Val Asn Ile Lys Glu Pro Gln Ser Gln Thr Asn Thr Thr Ala Asp Thr
 1155 1160 1165
 Glu Gln Pro Ala Lys Glu Thr Ser Ser Asn Val Glu Gln Pro Val Thr
 1170 1175 1180
 35 Glu Ser Thr Thr Val Asn Thr Gly Asn Ser Val Val Glu Asn Pro Glu
 1185 1190 1195 1200
 Asn Thr Thr Pro Ala Thr Thr Gln Pro Thr Val Asn Ser Glu Ser Ser
 1205 1210 1215
 40 Asn Lys Pro Lys Asn Arg His Arg Arg Ser Val Arg Ser Val Pro His
 1220 1225 1230
 Asn Val Glu Pro Ala Thr Thr Ser Ser Asn Asp Arg Ser Thr Val Ala
 1235 1240 1245
 45 Leu Cys Asp Leu Thr Ser Thr Asn Thr Asn Ala Val Leu Ser Asp Ala
 1250 1255 1260
 Arg Ala Lys Ala Gln Phe Val Ala Leu Asn Val Gly Lys Ala Val Ser
 1265 1270 1275 1280
 50 Gln His Ile Ser Gln Leu Glu Met Asn Asn Glu Gly Gln Tyr Asn Val
 1285 1290 1295
 Trp Val Ser Asn Thr Ser Met Asn Lys Asn Tyr Ser Ser Ser Gln Tyr
 1300 1305 1310
 55 Arg Arg Phe Ser Ser Lys Ser Thr Gln Thr Gln Leu Gly Trp Asp Gln

ES 2 278 381 T3

1315 1320 1325
 Thr Ile Ser Asn Asn Val Gln Leu Gly Gly Val Phe Thr Tyr Val Arg
 1330 1335 1340
 Asn Ser Asn Asn Phe Asp Lys Ala Thr Ser Lys Asn Thr Leu Ala Gln
 1345 1350 1355 1360
 Val Asn Phe Tyr Ser Lys Tyr Tyr Ala Asp Asn His Trp Tyr Leu Gly
 1365 1370 1375
 Ile Asp Leu Gly Tyr Gly Lys Phe Gln Ser Lys Leu Gln Thr Asn His
 1380 1385 1390
 Asn Ala Lys Phe Ala Arg His Thr Ala Gln Phe Gly Leu Thr Ala Gly
 1395 1400 1405
 Lys Ala Phe Asn Leu Gly Asn Phe Gly Ile Thr Pro Ile Val Gly Val
 1410 1415 1420
 Arg Tyr Ser Tyr Leu Ser Asn Ala Asp Phe Ala Leu Asp Gln Ala Arg
 1425 1430 1435 1440
 Ile Lys Val Asn Pro Ile Ser Val Lys Thr Ala Phe Ala Gln Val Asp
 1445 1450 1455
 Leu Ser Tyr Thr Tyr His Leu Gly Glu Phe Ser Val Thr Pro Ile Leu
 1460 1465 1470
 Ser Ala Arg Tyr Asp Ala Asn Gln Gly Ser Gly Lys Ile Asn Val Asn
 1475 1480 1485
 Gly Tyr Asp Phe Ala Tyr Asn Val Glu Asn Gln Gln Gln Tyr Asn Ala
 1490 1495 1500
 Gly Leu Lys Leu Lys Tyr His Asn Val Lys Leu Ser Leu Ile Gly Gly
 1505 1510 1515 1520
 Leu Thr Lys Ala Lys Gln Ala Glu Lys Gln Lys Thr Ala Glu Leu Lys
 1525 1530 1535
 Leu Ser Phe Ser Phe
 1540

(2) INFORMACIÓN PARA EL SEQ ID NO: 4:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1545 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGÍA: desconocida

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 4:

Met Leu Asn Lys Lys Phe Lys Leu Asn Phe Ile Ala Leu Thr Val Ala
 1 5 10 15
 Tyr Ala Leu Thr Pro Tyr Thr Glu Ala Ala Leu Val Arg Asp Asp Val
 20 25 30
 Asp Tyr Gln Ile Phe Arg Asp Phe Ala Glu Asn Lys Gly Lys Phe Ser
 35 40 45

ES 2 278 381 T3

Val Gly Ala Thr Asn Val Glu Val Arg Asp Lys Asn Asn Arg Pro Leu
 50 55 60
 5 Gly Asn Val Leu Pro Asn Gly Ile Pro Met Ile Asp Phe Ser Val Val
 65 70 75 80
 Asp Val Asp Lys Arg Ile Ala Thr Leu Val Asn Pro Gln Tyr Val Val
 85 90 95
 10 Gly Val Lys His Val Ser Asn Gly Val Ser Glu Leu His Phe Gly Asn
 100 105 110
 Leu Asn Gly Asn Met Asn Asn Gly Asn Ala Lys Ala His Arg Asp Val
 115 120 125
 15 Ser Ser Glu Glu Asn Arg Tyr Tyr Thr Val Glu Lys Asn Glu Tyr Pro
 130 135 140
 Thr Lys Leu Asn Gly Lys Ala Val Thr Thr Glu Asp Gln Ala Gln Lys
 145 150 155 160
 20 Arg Arg Glu Asp Tyr Tyr Met Pro Arg Leu Asp Lys Phe Val Thr Glu
 165 170 175
 Val Ala Pro Ile Glu Ala Ser Thr Asp Ser Ser Thr Ala Gly Thr Tyr
 180 185 190
 25 Asn Asn Lys Asp Lys Tyr Pro Tyr Phe Val Arg Leu Gly Ser Gly Thr
 195 200 205
 Gln Phe Ile Tyr Glu Asn Gly Thr Arg Tyr Glu Leu Trp Leu Gly Lys
 210 215 220
 30 Glu Gly Gln Lys Ser Asp Ala Gly Gly Tyr Asn Leu Lys Leu Val Gly
 225 230 235 240
 Asn Ala Tyr Thr Tyr Gly Ile Ala Gly Thr Pro Tyr Glu Val Asn His
 245 250 255
 35 Glu Asn Asp Gly Leu Ile Gly Phe Gly Asn Ser Asn Asn Glu Tyr Ile
 260 265 270
 Asn Pro Lys Glu Ile Leu Ser Lys Lys Pro Leu Thr Asn Tyr Ala Val
 275 280 285
 40 Leu Gly Asp Ser Gly Ser Pro Leu Phe Val Tyr Asp Arg Glu Lys Gly
 290 295 300
 Lys Trp Leu Phe Leu Gly Ser Tyr Asp Tyr Trp Ala Gly Tyr Asn Lys
 305 310 315 320
 45 Lys Ser Trp Gln Glu Trp Asn Ile Tyr Lys Pro Glu Phe Ala Glu Lys
 325 330 335
 Ile Tyr Glu Gln Tyr Ser Ala Gly Ser Leu Ile Gly Ser Lys Thr Asp
 340 345 350
 50 Tyr Ser Trp Ser Ser Asn Gly Lys Thr Ser Thr Ile Thr Gly Gly Glu
 355 360 365
 Lys Ser Leu Asn Val Asp Leu Ala Asp Gly Lys Asp Lys Pro Asn His
 370 375 380

ES 2 278 381 T3

Gly Lys Ser Val Thr Phe Glu Gly Ser Gly Thr Leu Thr Leu Asn Asn
 385 390 395 400
 5 Asn Ile Asp Gln Gly Ala Gly Gly Leu Phe Phe Glu Gly Asp Tyr Glu
 405 410 415
 Val Lys Gly Thr Ser Asp Asn Thr Thr Trp Lys Gly Ala Gly Val Ser
 420 425 430
 10 Val Ala Glu Gly Lys Thr Val Thr Trp Lys Val His Asn Pro Gln Tyr
 435 440 445
 Asp Arg Leu Ala Lys Ile Gly Lys Gly Thr Leu Ile Val Glu Gly Thr
 450 455 460
 15 Gly Asp Asn Lys Gly Ser Leu Lys Val Gly Asp Gly Thr Val Ile Leu
 465 470 475 480
 Lys Gln Gln Thr Asn Gly Ser Gly Gln His Ala Phe Ala Ser Val Gly
 485 490 495
 20 Ile Val Ser Gly Arg Ser Thr Leu Val Leu Asn Asp Asp Lys Gln Val
 500 505 510
 Asp Pro Asn Ser Ile Tyr Phe Gly Phe Arg Gly Gly Arg Leu Asp Leu
 515 520 525
 25 Asn Gly Asn Ser Leu Thr Phe Asp His Ile Arg Asn Ile Asp Glu Gly
 530 535 540
 Ala Arg Leu Val Asn His Ser Thr Ser Lys His Ser Thr Val Thr Ile
 545 550 555 560
 30 Thr Gly Asp Asn Leu Ile Thr Asp Pro Asn Asn Val Ser Ile Tyr Tyr
 565 570 575
 Val Lys Pro Leu Glu Asp Asp Asn Pro Tyr Ala Ile Arg Gln Ile Lys
 580 585 590
 35 Tyr Gly Tyr Gln Leu Tyr Phe Asn Glu Glu Asn Arg Thr Tyr Tyr Ala
 595 600 605
 Leu Lys Lys Asp Ala Ser Ile Arg Ser Glu Phe Pro Gln Asn Arg Gly
 610 615 620
 40 Glu Ser Asn Asn Ser Trp Leu Tyr Met Gly Thr Glu Lys Ala Asp Ala
 625 630 635 640
 Gln Lys Asn Ala Met Asn His Ile Asn Asn Glu Arg Met Asn Gly Phe
 645 650 655
 45 Asn Gly Tyr Phe Gly Glu Glu Glu Gly Lys Asn Asn Gly Asn Leu Asn
 660 665 670
 Val Thr Phe Lys Gly Lys Ser Glu Gln Asn Arg Phe Leu Leu Thr Gly
 675 680 685
 50 Gly Thr Asn Leu Asn Gly Asp Leu Asn Val Gln Gln Gly Thr Leu Phe
 690 695 700
 Leu Ser Gly Arg Pro Thr Pro His Ala Arg Asp Ile Ala Gly Ile Ser
 705 710 715 720
 55 Ser Thr Lys Lys Asp Ser His Phe Ser Glu Asn Asn Glu Val Val Val

ES 2 278 381 T3

	725					730					735					
5	Glu	Asp	Asp	Trp	Ile	Asn	Arg	Asn	Phe	Lys	Ala	Thr	Asn	Ile	Asn	Val
				740					745					750		
	Thr	Asn	Asn	Ala	Thr	Leu	Tyr	Ser	Gly	Arg	Asn	Val	Glu	Ser	Ile	Thr
			755					760					765			
10	Ser	Asn	Ile	Thr	Ala	Ser	Asn	Asn	Ala	Lys	Val	His	Ile	Gly	Tyr	Lys
		770					775					780				
	Ala	Gly	Asp	Thr	Val	Cys	Val	Arg	Ser	Asp	Tyr	Thr	Gly	Tyr	Val	Thr
	785					790				795						800
15	Cys	Thr	Thr	Asp	Lys	Leu	Ser	Asp	Lys	Ala	Leu	Asn	Ser	Phe	Asn	Pro
				805						810					815	
	Thr	Asn	Leu	Arg	Gly	Asn	Val	Asn	Leu	Thr	Glu	Ser	Ala	Asn	Phe	Val
			820						825					830		
20	Leu	Gly	Lys	Ala	Asn	Leu	Phe	Gly	Thr	Ile	Gln	Ser	Arg	Gly	Asn	Ser
		835					840						845			
	Gln	Val	Arg	Leu	Thr	Glu	Asn	Ser	His	Trp	His	Leu	Thr	Gly	Asn	Ser
	850					855					860					
25	Asp	Val	His	Gln	Leu	Asp	Leu	Ala	Asn	Gly	His	Ile	His	Leu	Asn	Ser
	865				870					875					880	
	Ala	Asp	Asn	Ser	Asn	Asn	Val	Thr	Lys	Tyr	Asn	Thr	Leu	Thr	Val	Asn
				885					890						895	
30	Ser	Leu	Ser	Gly	Asn	Gly	Ser	Phe	Tyr	Tyr	Leu	Thr	Asp	Leu	Ser	Asn
			900					905						910		
	Lys	Gln	Gly	Asp	Lys	Val	Val	Val	Thr	Lys	Ser	Ala	Thr	Gly	Asn	Phe
		915						920						925		
35	Thr	Leu	Gln	Val	Ala	Asp	Lys	Thr	Gly	Glu	Pro	Asn	His	Asn	Glu	Leu
	930					935					940					
	Thr	Leu	Phe	Asp	Ala	Ser	Lys	Ala	Gln	Arg	Asp	His	Leu	Asn	Val	Ser
	945				950					955					960	
40	Leu	Val	Gly	Asn	Thr	Val	Asp	Leu	Gly	Ala	Trp	Lys	Tyr	Lys	Leu	Arg
				965					970						975	
	Asn	Val	Asn	Gly	Arg	Tyr	Asp	Leu	Tyr	Asn	Pro	Glu	Val	Glu	Lys	Arg
			980					985						990		
45	Asn	Gln	Thr	Val	Asp	Thr	Thr	Asn	Ile	Thr	Thr	Pro	Asn	Asn	Ile	Gln
		995						1000					1005			
	Ala	Asp	Val	Pro	Ser	Val	Pro	Ser	Asn	Asn	Glu	Glu	Ile	Ala	Arg	Val
	1010					1015					1020					
50	Asp	Glu	Ala	Pro	Val	Pro	Pro	Pro	Ala	Pro	Ala	Thr	Pro	Ser	Glu	Thr
	1025				1030					1035					1040	
	Thr	Glu	Thr	Val	Ala	Glu	Asn	Ser	Lys	Gln	Glu	Ser	Lys	Thr	Val	Glu
				1045					1050						1055	
55	Lys	Asn	Glu	Gln	Asp	Ala	Thr	Glu	Thr	Thr	Ala	Gln	Asn	Arg	Glu	Val
				1060				1065						1070		

ES 2 278 381 T3

	Ala	Lys	Glu	Ala	Lys	Ser	Asn	Val	Lys	Ala	Asn	Thr	Gln	Thr	Asn	Glu	
				1075					1080							1085	
5	Val	Ala	Gln	Ser	Gly	Ser	Glu	Thr	Lys	Glu	Thr	Gln	Thr	Thr	Glu	Thr	
				1090					1095								
	Lys	Glu	Thr	Ala	Thr	Val	Glu	Lys	Glu	Glu	Lys	Ala	Lys	Val	Glu	Thr	
						1110						1115				1120	
10	Glu	Lys	Thr	Gln	Glu	Val	Pro	Lys	Val	Thr	Ser	Gln	Val	Ser	Pro	Lys	
						1125						1130				1135	
	Gln	Glu	Gln	Ser	Glu	Thr	Val	Gln	Pro	Gln	Ala	Glu	Pro	Ala	Arg	Glu	
						1140						1145				1150	
15	Asn	Asp	Pro	Thr	Val	Asn	Ile	Lys	Glu	Pro	Gln	Ser	Gln	Thr	Asn	Thr	
						1155								1165			
	Thr	Ala	Asp	Thr	Glu	Gln	Pro	Ala	Lys	Glu	Thr	Ser	Ser	Asn	Val	Glu	
								1175						1180			
20	Gln	Pro	Val	Thr	Glu	Ser	Thr	Thr	Val	Asn	Thr	Gly	Asn	Ser	Val	Val	
							1190					1195				1200	
	Glu	Asn	Pro	Glu	Asn	Thr	Thr	Pro	Ala	Thr	Thr	Gln	Pro	Thr	Val	Asn	
						1205						1210				1215	
25	Ser	Glu	Ser	Ser	Asn	Lys	Pro	Lys	Asn	Arg	His	Arg	Arg	Ser	Val	Arg	
						1220				1225					1230		
	Ser	Val	Pro	His	Asn	Val	Glu	Pro	Ala	Thr	Thr	Ser	Ser	Asn	Asp	Arg	
						1235				1240				1245			
30	Ser	Thr	Val	Ala	Leu	Cys	Asp	Leu	Thr	Ser	Thr	Asn	Thr	Asn	Ala	Val	
							1255						1260				
	Leu	Ser	Asp	Ala	Arg	Ala	Lys	Ala	Gln	Phe	Val	Ala	Leu	Asn	Val	Gly	
						1270						1275				1280	
35	Lys	Ala	Val	Ser	Gln	His	Ile	Ser	Gln	Leu	Glu	Met	Asn	Asn	Glu	Gly	
						1285						1290				1295	
	Gln	Tyr	Asn	Val	Trp	Val	Ser	Asn	Thr	Ser	Met	Asn	Lys	Asn	Tyr	Ser	
						1300					1305				1310		
40	Ser	Ser	Gln	Tyr	Arg	Arg	Phe	Ser	Ser	Lys	Ser	Thr	Gln	Thr	Gln	Leu	
						1315					1320				1325		
	Gly	Trp	Asp	Gln	Thr	Ile	Ser	Asn	Asn	Val	Gln	Leu	Gly	Gly	Val	Phe	
							1335					1340					
45	Thr	Tyr	Val	Arg	Asn	Ser	Asn	Asn	Phe	Asp	Lys	Ala	Thr	Ser	Lys	Asn	
						1350					1355					1360	
	Thr	Leu	Ala	Gln	Val	Asn	Phe	Tyr	Ser	Lys	Tyr	Tyr	Ala	Asp	Asn	His	
						1365					1370				1375		
50	Trp	Tyr	Leu	Gly	Ile	Asp	Leu	Gly	Tyr	Gly	Lys	Phe	Gln	Ser	Lys	Leu	
						1380	</										

ES 2 278 381 T3

1410 1415 1420
 Ile Val Gly Val Arg Tyr Ser Tyr Leu Ser Asn Ala Asp Phe Ala Leu
 1425 1430 1435 1440
 5 Asp Gln Ala Arg Ile Lys Val Asn Pro Ile Ser Val Lys Thr Ala Phe
 1445 1450 1455
 Ala Gln Val Asp Leu Ser Tyr Thr Tyr His Leu Gly Glu Phe Ser Val
 1460 1465 1470
 10 Thr Pro Ile Leu Ser Ala Arg Tyr Asp Ala Asn Gln Gly Ser Gly Lys
 1475 1480 1485
 Ile Asn Val Asn Gly Tyr Asp Phe Ala Tyr Asn Val Glu Asn Gln Gln
 1490 1495 1500
 15 Gln Tyr Asn Ala Gly Leu Lys Leu Lys Tyr His Asn Val Lys Leu Ser
 1505 1510 1515 1520
 Leu Ile Gly Gly Leu Thr Lys Ala Lys Gln Ala Glu Lys Gln Lys Thr
 1525 1530 1535
 20 Ala Glu Leu Lys Leu Ser Phe Ser Phe
 1540 1545

25 (2) INFORMACIÓN PARA EL SEQ ID NO: 5:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1702 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: desconocida

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 5:

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

ES 2 278 381 T3

	145		150		155		160
5	Arg	Arg	Glu	Asp	Tyr	Tyr	Met
				165			Pro
						Arg	Leu
						170	Asp
						Lys	Phe
						Val	Thr
						175	Glu
	Val	Ala	Pro	Ile	Glu	Ala	Ser
				180			Thr
						185	Ala
						Ser	Ser
						Asp	Ala
						190	Gly
						Thr	Tyr
10	Asn	Asp	Gln	Asn	Lys	Tyr	Pro
				195			Ala
						200	Phe
						Val	Arg
						Leu	Gly
						205	Ser
						Gly	Thr
	Gln	Phe	Ile	Tyr	Lys	Lys	Gly
						215	Asn
						Tyr	Ser
						Leu	Ile
						220	Leu
						Asn	Asn
15	His	Glu	Val	Gly	Gly	Asn	Asn
						230	Leu
						Lys	Leu
						Val	Gly
						235	Asp
						Ala	Tyr
						Thr	240
	Tyr	Gly	Ile	Ala	Gly	Thr	Pro
						245	Tyr
						Lys	Val
						250	Asn
						His	Glu
						Asn	Asn
						255	Gly
20	Leu	Ile	Gly	Phe	Gly	Asn	Ser
						260	Lys
						Glu	Glu
						265	His
						Ser	Asp
						Pro	Lys
						270	Gly
						Asp	Ser
	Ile	Leu	Ser	Gln	Asp	Pro	Leu
						275	Thr
						280	Asn
						Tyr	Ala
						Val	Leu
						285	Gly
						Asp	Ser
25	Gly	Ser	Pro	Leu	Phe	Val	Tyr
						295	Asp
						Arg	Glu
						Lys	Gly
						300	Lys
						Trp	Leu
						Phe	320
	Leu	Gly	Ser	Tyr	Asp	Phe	Trp
						310	Lys
						315	Lys
						Ser	Trp
						Gln	320
30	Glu	Trp	Asn	Ile	Tyr	Lys	Pro
						325	Glu
						Phe	Ala
						330	Lys
						Thr	Val
						Leu	Asp
						335	Lys
	Asp	Thr	Ala	Gly	Ser	Leu	Ile
						340	Gly
						345	Asn
						Thr	Gln
						Tyr	Asn
						350	Trp
						Asn	350
35	Pro	Thr	Gly	Lys	Thr	Ser	Val
						355	Ile
						360	Ser
						Asn	Gly
						Ser	Glu
						365	Ser
						Leu	Asn
	Val	Asp	Leu	Phe	Asp	Ser	Ser
						370	Lys
						375	Lys
						380	Asn
						Asn	380
40	His	Gly	Lys	Ser	Val	Thr	Leu
						385	Arg
						390	Gly
						Ser	Gly
						395	Thr
						Leu	Thr
						400	Leu
						Asn	400
	Asn	Asn	Ile	Asp	Gln	Gly	Ala
						405	Gly
						410	Phe
						Phe	Glu
						Gly	Asp
						415	Tyr
45	Glu	Val	Lys	Gly	Thr	Ser	Asp
						420	Ser
						425	Thr
						Thr	Trp
						Lys	Gly
						430	Ala
						Gly	Val
	Ser	Val	Ala	Asp	Gly	Lys	Thr
						435	Val
						440	His
						445	Asn
						Pro	Lys
50	Ser	Asp	Arg	Leu	Ala	Lys	Ile
						450	Gly
						455	Lys
						Gly	Thr
						Leu	Ile
						460	Val
						Glu	Gly
	Lys	Gly	Glu	Asn	Lys	Gly	Ser
						465	Leu
						470	Val
						475	Gly
						480	Thr
						Val	Ile
55	Leu	Lys	Gln	Gln	Ala	Asp	Ala
						485	Asn
						490	Lys
						Val	Lys
						Ala	Phe
						Ser	Gln
						495	

ES 2 278 381 T3

Val Gly Ile Val Ser Gly Arg Ser Thr Val Val Leu Asn Asp Asp Lys
 500 505 510
 5 Gln Val Asp Pro Asn Ser Ile Tyr Phe Gly Phe Arg Gly Gly Arg Leu
 515 520 525
 Asp Ala Asn Gly Asn Asn Leu Thr Phe Glu His Ile Arg Asn Ile Asp
 530 535 540
 10 Asp Gly Ala Arg Leu Val Asn His Asn Thr Ser Lys Thr Ser Thr Val
 545 550 555 560
 Thr Ile Thr Gly Glu Ser Leu Ile Thr Asp Pro Asn Thr Ile Thr Pro
 565 570 575
 15 Tyr Asn Ile Asp Ala Pro Asp Glu Asp Asn Pro Tyr Ala Phe Arg Arg
 580 585 590
 Ile Lys Asp Gly Gly Gln Leu Tyr Leu Asn Leu Glu Asn Tyr Thr Tyr
 595 600 605
 20 Tyr Ala Leu Arg Lys Gly Ala Ser Thr Arg Ser Glu Leu Pro Lys Asn
 610 615 620
 Ser Gly Glu Ser Asn Glu Asn Trp Leu Tyr Met Gly Lys Thr Ser Asp
 625 630 635 640
 25 Ala Ala Lys Arg Asn Val Met Asn His Ile Asn Asn Glu Arg Met Asn
 645 650 655
 Gly Phe Asn Gly Tyr Phe Gly Glu Glu Glu Gly Lys Asn Asn Gly Asn
 660 665 670
 30 Leu Asn Val Thr Phe Lys Gly Lys Ser Glu Gln Asn Arg Phe Leu Leu
 675 680 685
 Thr Gly Gly Thr Asn Leu Asn Gly Asp Leu Lys Val Glu Lys Gly Thr
 690 695 700
 35 Leu Phe Leu Ser Gly Arg Pro Thr Pro His Ala Arg Asp Ile Ala Gly
 705 710 715 720
 Ile Ser Ser Thr Lys Lys Asp Gln His Phe Ala Glu Asn Asn Glu Val
 725 730 735
 40 Val Val Glu Asp Asp Trp Ile Asn Arg Asn Phe Lys Ala Thr Asn Ile
 740 745 750
 Asn Val Thr Asn Asn Ala Thr Leu Tyr Ser Gly Arg Asn Val Ala Asn
 755 760 765
 45 Ile Thr Ser Asn Ile Thr Ala Ser Asp Asn Ala Lys Val His Ile Gly
 770 775 780
 Tyr Lys Ala Gly Asp Thr Val Cys Val Arg Ser Asp Tyr Thr Gly Tyr
 785 790 795 800
 50 Val Thr Cys Thr Thr Asp Lys Leu Ser Asp Lys Ala Leu Asn Ser Phe
 805 810 815
 Asn Ala Thr Asn Val Ser Gly Asn Val Asn Leu Ser Gly Asn Ala Asn
 820 825 830
 55 Phe Val Leu Gly Lys Ala Asn Leu Phe Gly Thr Ile Ser Gly Thr Gly

ES 2 278 381 T3

	835	840	845
5	Asn Ser Gln Val Arg Leu Thr	Glu Asn Ser His Trp	His Leu Thr Gly
	850	855	860
	Asp Ser Asn Val Asn Gln Leu	Asn Leu Asp Lys Gly	His Ile His Leu
	865	870	875
10	Asn Ala Gln Asn Asp Ala Asn Lys Val Thr	Thr Tyr Asn Thr Leu Thr	
		885	890
	Val Asn Ser Leu Ser Gly Asn Gly Ser Phe Tyr Tyr	Leu Thr Asp Leu	
		900	905
15	Ser Asn Lys Gln Gly Asp Lys Val Val Val Thr Lys	Ser Ala Thr Gly	
		915	920
	Asn Phe Thr Leu Gln Val Ala Asp Lys Thr Gly	Glu Pro Thr Lys Asn	
		930	935
20	Glu Leu Thr Leu Phe Asp Ala Ser Asn Ala Thr Arg	Asn Asn Leu Asn	
		945	950
	Val Ser Leu Val Gly Asn Thr Val Asp Leu Gly Ala Trp Lys Tyr Lys		
		965	970
25	Leu Arg Asn Val Asn Gly Arg Tyr Asp Leu Tyr Asn Pro Glu Val Glu		
		980	985
	Lys Arg Asn Gln Thr Val Asp Thr Thr Asn Ile Thr Thr Pro Asn Asn		
		995	1000
30	Ile Gln Ala Asp Val Pro Ser Val Pro Ser Asn Asn Glu Glu Ile Ala		
		1010	1015
	Arg Val Glu Thr Pro Val Pro Pro Pro Ala Pro Ala Thr Pro Ser Glu		
		1025	1030
35	Thr Thr Glu Thr Val Ala Glu Asn Ser Lys Gln Glu Ser Lys Thr Val		
		1045	1050
	Glu Lys Asn Glu Gln Asp Ala Thr Glu Thr Thr Ala Gln Asn Gly Glu		
		1060	1065
40	Val Ala Glu Glu Ala Lys Pro Ser Val Lys Ala Asn Thr Gln Thr Asn		
		1075	1080
	Glu Val Ala Gln Ser Gly Ser Glu Thr Glu Glu Thr Gln Thr Thr Glu		
		1090	1095
45	Ile Lys Glu Thr Ala Lys Val Glu Lys Glu Glu Lys Ala Lys Val Glu		
		1105	1110
	Lys Glu Glu Lys Ala Lys Val Glu Lys Asp Glu Ile Gln Glu Ala Pro		
		1125	1130
50	Gln Met Ala Ser Glu Thr Ser Pro Lys Gln Ala Lys Pro Ala Pro Lys		
		1140	1145
	Glu Val Ser Thr Asp Thr Lys Val Glu Glu Thr Gln Val Gln Ala Gln		
		1155	1160
55	Pro Gln Thr Gln Ser Thr Thr Val Ala Ala Ala Glu Ala Thr Ser Pro		
		1170	1175

ES 2 278 381 T3

Asn Ser Lys Pro Ala Glu Glu Thr Gln Pro Ser Glu Lys Thr Asn Ala
 1185 1190 1195 1200
 5 Glu Pro Val Thr Pro Val Val Ser Lys Asn Gln Thr Glu Asn Thr Thr
 1205 1210 1215
 Asp Gln Pro Thr Glu Arg Glu Lys Thr Ala Lys Val Glu Thr Glu Lys
 1220 1225 1230
 10 Thr Gln Glu Pro Pro Gln Val Ala Ser Gln Ala Ser Pro Lys Gln Glu
 1235 1240 1245
 Gln Ser Glu Thr Val Gln Pro Gln Ala Val Leu Glu Ser Glu Asn Val
 1250 1255 1260
 15 Pro Thr Val Asn Asn Ala Glu Glu Val Gln Ala Gln Leu Gln Thr Gln
 1265 1270 1275 1280
 Thr Ser Ala Thr Val Ser Thr Lys Gln Pro Ala Pro Glu Asn Ser Ile
 1285 1290 1295
 20 Asn Thr Gly Ser Ala Thr Ala Ile Thr Glu Thr Ala Glu Lys Ser Asp.
 1300 1305 1310
 Lys Pro Gln Thr Glu Thr Ala Ala Ser Thr Glu Asp Ala Ser Gln His
 1315 1320 1325
 25 Lys Ala Asn Thr Val Ala Asp Asn Ser Val Ala Asn Asn Ser Glu Ser
 1330 1335 1340
 Ser Glu Pro Lys Ser Arg Arg Arg Arg Ser Ile Ser Gln Pro Gln Glu
 1345 1350 1355 1360
 30 Thr Ser Ala Glu Glu Thr Thr Ala Ala Ser Thr Asp Glu Thr Thr Ile
 1365 1370 1375
 Ala Asp Asn Ser Lys Arg Ser Lys Pro Asn Arg Arg Ser Arg Arg Ser
 1380 1385 1390
 35 Val Arg Ser Glu Pro Thr Val Thr Asn Gly Ser Asp Arg Ser Thr Val
 1395 1400 1405
 Ala Leu Arg Asp Leu Thr Ser Thr Asn Thr Asn Ala Val Ile Ser Asp
 1410 1415 1420
 40 Ala Met Ala Lys Ala Gln Phe Val Ala Leu Asn Val Gly Lys Ala Val
 1425 1430 1435 1440
 Ser Gln His Ile Ser Gln Leu Glu Met Asn Asn Glu Gly Gln Tyr Asn
 1445 1450 1455
 45 Val Trp Val Ser Asn Thr Ser Met Asn Glu Asn Tyr Ser Ser Ser Gln
 1460 1465 1470
 Tyr Arg Arg Phe Ser Ser Lys Ser Thr Gln Thr Gln Leu Gly Trp Asp
 1475 1480 1485
 50 Gln Thr Ile Ser Asn Asn Val Gln Leu Gly Gly Val Phe Thr Tyr Val
 1490 1495 1500
 Arg Asn Ser Asn Asn Phe Asp Lys Ala Ser Ser Lys Asn Thr Leu Ala
 1505 1510 1515 1520
 55 Gln Val Asn Phe Tyr Ser Lys Tyr Tyr Ala Asp Asn His Trp Tyr Leu

ES 2 278 381 T3

	1525	1530	1535
5	Gly Ile Asp Leu Gly Tyr Gly Lys Phe Gln Ser Asn Leu Lys Thr Asn 1540 1545 1550		
	His Asn Ala Lys Phe Ala Arg His Thr Ala Gln Phe Gly Leu Thr Ala 1555 1560 1565		
10	Gly Lys Ala Phe Asn Leu Gly Asn Phe Gly Ile Thr Pro Ile Val Gly 1570 1575 1580		
	Val Arg Tyr Ser Tyr Leu Ser Asn Ala Asn Phe Ala Leu Ala Lys Asp 1585 1590 1595 1600		
15	Arg Ile Lys Val Asn Pro Ile Ser Val Lys Thr Ala Phe Ala Gln Val 1605 1610 1615		
	Asp Leu Ser Tyr Thr Tyr His Leu Gly Glu Phe Ser Val Thr Pro Ile 1620 1625 1630		
20	Leu Ser Ala Arg Tyr Asp Thr Asn Gln Gly Ser Gly Lys Ile Asn Val 1635 1640 1645		
	Asn Gln Tyr Asp Phe Ala Tyr Asn Val Glu Asn Gln Gln Gln Tyr Asn 1650 1655 1660		
25	Ala Gly Leu Lys Leu Lys Tyr His Asn Val Lys Leu Ser Leu Ile Gly 1665 1670 1675 1680		
	Gly Leu Thr Lys Ala Lys Gln Ala Glu Lys Gln Lys Thr Ala Glu Leu 1685 1690 1695		
30	Lys Leu Ser Phe Ser Phe 1700		

(2) INFORMACIÓN PARA EL SEQ ID NO: 6:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1848 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGÍA: desconocida

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 6:

45	Met Leu Asn Lys Lys Phe Lys Leu Asn Phe Ile Ala Leu Thr Val Ala 1 5 10 15
	Tyr Ala Leu Thr Pro Tyr Thr Glu Ala Ala Leu Val Arg Asp Asp Val 20 25 30
50	Asp Tyr Gln Ile Phe Arg Asp Phe Ala Glu Asn Lys Gly Lys Phe Ser 35 40 45
	Val Gly Ala Thr Asn Val Glu Val Arg Asp Lys Lys Asn Gln Ser Leu 50 55 60
55	Gly Ser Ala Leu Pro Asn Gly Ile Pro Met Ile Asp Phe Ser Val Val 65 70 75 80
	Asp Val Asp Lys Arg Ile Ala Thr Leu Val Asn Pro Gln Tyr Val Val 85 90 95
60	Gly Val Lys His Val Ser Asn Gly Val Ser Glu Leu His Phe Gly Asn

65

ES 2 278 381 T3

	100	105	110
5	Leu Asn Gly 115	Asn Met Asn Asn Gly 120	Asn Ala Lys Ser His 125
	Ser Ser 130	Glu Glu Asn Arg Tyr 135	Tyr Thr Val Glu Lys 140
10	Thr 145	Glu Asn Val Thr Ser 150	Phe Thr Lys Glu Glu Gln Asp Ala Gln Lys 160
	Arg Arg Glu Asp Tyr 165	Tyr Tyr Met Pro Arg Leu 170	Asp Lys Phe Val Thr Glu 175
15	Val Ala Pro 180	Ile Glu Ala Ser Thr Ala 185	Asn Asn Asn Lys Gly Glu Tyr 190
	Asn Asn Ser 195	Asp Lys Tyr Pro Ala Phe Val Arg Leu 205	Gly Ser Gly Thr
20	Gln Phe Ile Tyr Lys Lys 210	Gly 215	Ser Arg Tyr Gln Leu Ile Leu Thr Glu 220
	Lys Asp Lys Gln Gly Asn Leu Leu Arg Asn Trp Asp Val Gly Gly Asp 225	230	235
25	Asn Leu Glu Leu Val Gly Asn Ala Tyr Thr Tyr Gly Ile Ala Gly Thr 245	250	255
	Pro Tyr Lys Val Asn His Glu Asn Asn Gly Leu Ile Gly Phe Gly Asn 260	265	270
30	Ser Lys Glu Glu His Ser Asp Pro Lys Gly Ile Leu Ser Gln Asp Pro 275	280	285
	Leu Thr Asn Tyr Ala Val Leu Gly Asp Ser Gly Ser Pro Leu Phe Val 290	295	300
35	Tyr Asp Arg Glu Lys Gly Lys Trp Leu Phe Leu Gly Ser Tyr Asp Phe 305	310	315
	Trp Ala Gly Tyr Asn Lys Lys Ser Trp Gln Glu Trp Asn Ile Tyr Lys 325	330	335
40	His Glu Phe Ala Glu Lys Ile Tyr Gln Gln Tyr Ser Ala Gly Ser Leu 340	345	350
	Ile Gly Ser Asn Thr Gln Tyr Thr Trp Gln Ala Thr Gly Ser Thr Ser 355	360	365
45	Thr Ile Thr Gly Gly Gly Glu Pro Leu Ser Val Asp Leu Thr Asp Gly 370	375	380
	Lys Asp Lys Pro Asn His Gly Lys Ser Ile Thr Leu Lys Gly Ser Gly 385	390	395
50	Thr Leu Thr Leu Asn Asn His Ile Asp Gln Gly Ala Gly Gly Leu Phe 405	410	415
	Phe Glu Gly Asp Tyr Glu Val Lys Gly Thr Ser Asp Ser Thr Thr Trp 420	425	430
55	Lys Gly Ala Gly Val Ser Val Ala Asp Gly Lys Thr Val Thr Trp Lys 435	440	445

ES 2 278 381 T3

Val His Asn Pro Lys Tyr Asp Arg Leu Ala Lys Ile Gly Lys Gly Thr
 450 455 460
 5 Leu Val Val Glu Gly Lys Gly Lys Asn Glu Gly Leu Leu Lys Val Gly
 465 470 475 480
 Asp Gly Thr Val Ile Leu Lys Gln Lys Ala Asp Ala Asn Asn Lys Val
 485 490 495
 10 Gln Ala Phe Ser Gln Val Gly Ile Val Ser Gly Arg Ser Thr Leu Val
 500 505 510
 Leu Asn Asp Asp Lys Gln Val Asp Pro Asn Ser Ile Tyr Phe Gly Phe
 515 520 525
 15 Arg Gly Gly Arg Leu Asp Leu Asn Gly Asn Ser Leu Thr Phe Asp His
 530 535 540
 Ile Arg Asn Ile Asp Asp Gly Ala Arg Val Val Asn His Asn Met Thr
 545 550 555 560
 20 Asn Thr Ser Asn Ile Thr Ile Thr Gly Glu Ser Leu Ile Thr Asn Pro
 565 570 575
 Asn Thr Ile Thr Ser Tyr Asn Ile Glu Ala Gln Asp Asp Asp His Pro
 580 585 590
 25 Leu Arg Ile Arg Ser Ile Pro Tyr Arg Gln Leu Tyr Phe Asn Gln Asp
 595 600 605
 Asn Arg Ser Tyr Tyr Thr Leu Lys Lys Gly Ala Ser Thr Arg Ser Glu
 610 615 620
 30 Leu Pro Gln Asn Ser Gly Glu Ser Asn Glu Asn Trp Leu Tyr Met Gly
 625 630 635 640
 Arg Thr Ser Asp Ala Ala Lys Arg Asn Val Met Asn His Ile Asn Asn
 645 650 655
 35 Glu Arg Met Asn Gly Phe Asn Gly Tyr Phe Gly Glu Glu Glu Thr Lys
 660 665 670
 Ala Thr Gln Asn Gly Lys Leu Asn Val Thr Phe Asn Gly Lys Ser Asp
 675 680 685
 40 Gln Asn Arg Phe Leu Leu Thr Gly Gly Thr Asn Leu Asn Gly Asp Leu
 690 695 700
 Asn Val Glu Lys Gly Thr Leu Phe Leu Ser Gly Arg Pro Thr Pro His
 705 710 715 720
 45 Ala Arg Asp Ile Ala Gly Ile Ser Ser Thr Lys Lys Asp Pro His Phe
 725 730 735
 Thr Glu Asn Asn Glu Val Val Val Glu Asp Asp Trp Ile Asn Arg Asn
 740 745 750
 50 Phe Lys Ala Thr Thr Met Asn Val Thr Gly Asn Ala Ser Leu Tyr Ser
 755 760 765
 Gly Arg Asn Val Ala Asn Ile Thr Ser Asn Ile Thr Ala Ser Asn Asn
 770 775 780
 55 Ala Gln Val His Ile Gly Tyr Lys Thr Gly Asp Thr Val Cys Val Arg

ES 2 278 381 T3

	785		790		795		800									
5	Ser	Asp	Tyr	Thr	Gly	Tyr	Val	Thr	Cys	His	Asn	Ser	Asn	Leu	Ser	Glu
					805					810					815	
	Lys	Ala	Leu	Asn	Ser	Phe	Asn	Pro	Thr	Asn	Leu	Arg	Gly	Asn	Val	Asn
			820					825					830			
10	Leu	Thr	Glu	Asn	Ala	Ser	Phe	Thr	Leu	Gly	Lys	Ala	Asn	Leu	Phe	Gly
		835						840					845			
	Thr	Ile	Gln	Ser	Ile	Gly	Thr	Ser	Gln	Val	Asn	Leu	Lys	Glu	Asn	Ser
		850				855						860				
15	His	Trp	His	Leu	Thr	Gly	Asn	Ser	Asn	Val	Asn	Gln	Leu	Asn	Leu	Thr
	865					870					875				880	
	Asn	Gly	His	Ile	His	Leu	Asn	Ala	Gln	Asn	Asp	Ala	Asn	Lys	Val	Thr
				885					890					895		
20	Thr	Tyr	Asn	Thr	Leu	Thr	Val	Asn	Ser	Leu	Ser	Gly	Asn	Gly	Ser	Phe
			900					905					910			
	Tyr	Tyr	Trp	Val	Asp	Phe	Thr	Asn	Asn	Lys	Ser	Asn	Lys	Val	Val	Val
			915					920					925			
25	Asn	Lys	Ser	Ala	Thr	Gly	Asn	Phe	Thr	Leu	Gln	Val	Ala	Asp	Lys	Thr
		930					935					940				
	Gly	Glu	Pro	Asn	His	Asn	Glu	Leu	Thr	Leu	Phe	Asp	Ala	Ser	Asn	Ala
	945					950					955				960	
30	Thr	Arg	Asn	Asn	Leu	Glu	Val	Thr	Leu	Ala	Asn	Gly	Ser	Val	Asp	Arg
				965					970						975	
	Gly	Ala	Trp	Lys	Tyr	Lys	Leu	Arg	Asn	Val	Asn	Gly	Arg	Tyr	Asp	Leu
			980					985						990		
35	Tyr	Asn	Pro	Glu	Val	Glu	Lys	Arg	Asn	Gln	Thr	Val	Asp	Thr	Thr	Asn
		995						1000					1005			
	Ile	Thr	Thr	Pro	Asn	Asp	Ile	Gln	Ala	Asp	Ala	Pro	Ser	Ala	Gln	Ser
	1010						1015					1020				
40	Asn	Asn	Glu	Glu	Ile	Ala	Arg	Val	Glu	Thr	Pro	Val	Pro	Pro	Pro	Ala
	1025				1030						1035					1040
	Pro	Ala	Thr	Glu	Ser	Ala	Ile	Ala	Ser	Glu	Gln	Pro	Glu	Thr	Arg	Pro
				1045					1050						1055	
45	Ala	Glu	Thr	Ala	Gln	Pro	Ala	Met	Glu	Glu	Thr	Asn	Thr	Ala	Asn	Ser
			1060						1065					1070		
	Thr	Glu	Thr	Ala	Pro	Lys	Ser	Asp	Thr	Ala	Thr	Gln	Thr	Glu	Asn	Pro
		1075						1080					1085			
50	Asn	Ser	Glu	Ser	Val	Pro	Ser	Glu	Thr	Thr	Glu	Lys	Val	Ala	Glu	Asn
	1090						1095					1100				
	Pro	Pro	Gln	Glu	Asn	Glu	Thr	Val	Ala	Lys	Asn	Glu	Gln	Glu	Ala	Thr
	1105				1110					1115					1120	
55	Glu	Pro	Thr	Pro	Gln	Asn	Gly	Glu	Val	Ala	Lys	Glu	Asp	Gln	Pro	Thr
				1125						1130					1135	

ES 2 278 381 T3

Val Glu Ala Asn Thr Gln Thr Asn Glu Ala Thr Gln Ser Glu Gly Lys
 1140 1145 1150
 5 Thr Glu Glu Thr Gln Thr Ala Glu Thr Lys Ser Glu Pro Thr Glu Ser
 1155 1160 1165
 Val Thr Val Ser Glu Asn Gln Pro Glu Lys Thr Val Ser Gln Ser Thr
 1170 1175 1180
 10 Glu Asp Lys Val Val Val Glu Lys Glu Glu Lys Ala Lys Val Glu Thr
 1185 1190 1195 1200
 Glu Glu Thr Gln Lys Ala Pro Gln Val Thr Ser Lys Glu Pro Pro Lys
 1205 1210 1215
 15 Gln Ala Glu Pro Ala Pro Glu Glu Val Pro Thr Asp Thr Asn Ala Glu
 1220 1225 1230
 Glu Ala Gln Ala Leu Gln Gln Thr Gln Pro Thr Thr Val Ala Ala Ala
 1235 1240 1245
 20 Glu Thr Thr Ser Pro Asn Ser Lys Pro Ala Glu Glu Thr Gln Gln Pro
 1250 1255 1260
 Ser Glu Lys Thr Asn Ala Glu Pro Val Thr Pro Val Val Ser Glu Asn
 1265 1270 1275 1280
 25 Thr Ala Thr Gln Pro Thr Glu Thr Glu Glu Thr Ala Lys Val Glu Lys
 1285 1290 1295
 Glu Lys Thr Gln Glu Val Pro Gln Val Ala Ser Gln Glu Ser Pro Lys
 1300 1305 1310
 30 Gln Glu Gln Pro Ala Ala Lys Pro Gln Ala Gln Thr Lys Pro Gln Ala
 1315 1320 1325
 Glu Pro Ala Arg Glu Asn Val Leu Thr Thr Lys Asn Val Gly Glu Pro
 1330 1335 1340
 35 Gln Pro Gln Ala Gln Pro Gln Thr Gln Ser Thr Ala Val Pro Thr Thr
 1345 1350 1355 1360
 Gly Glu Thr Ala Ala Asn Ser Lys Pro Ala Ala Lys Pro Gln Ala Gln
 1365 1370 1375
 40 Ala Lys Pro Gln Thr Glu Pro Ala Arg Glu Asn Val Ser Thr Val Asn
 1380 1385 1390
 Thr Lys Glu Pro Gln Ser Gln Thr Ser Ala Thr Val Ser Thr Glu Gln
 1395 1400 1405
 45 Pro Ala Lys Glu Thr Ser Ser Asn Val Glu Gln Pro Ala Pro Glu Asn
 1410 1415 1420
 Ser Ile Asn Thr Gly Ser Ala Thr Thr Met Thr Glu Thr Ala Glu Lys
 1425 1430 1435 1440
 50 Ser Asp Lys Pro Gln Met Glu Thr Val Thr Glu Asn Asp Arg Gln Pro
 1445 1450 1455
 Glu Ala Asn Thr Val Ala Asp Asn Ser Val Ala Asn Asn Ser Glu Ser
 1460 1465 1470
 55 Ser Glu Ser Lys Ser Arg Arg Arg Arg Ser Val Ser Gln Pro Lys Glu

ES 2 278 381 T3

	1475	1480	1485
5	Thr Ser Ala Glu Glu Thr 1490	Thr Val Ala Ser Thr 1495	Gln Glu Thr Thr Val 1500
	Asp Asn Ser Val Ser Thr 1505	Pro Lys Pro Arg Ser 1510	Arg Arg Thr Arg Arg 1515
10	Ser Val Gln Thr Asn Ser 1525	Tyr Glu Pro Val Glu 1530	Leu Pro Thr Glu Asn 1535
	Ala Glu Asn Ala Glu Asn 1540	Val Gln Ser Gly Asn 1545	Asn Val Ala Asn Ser 1550
15	Gln Pro Ala Leu Arg Asn 1555	Leu Thr Ser Lys Asn 1560	Thr Asn Ala Val Ile 1565
	Ser Asn Ala Met Ala Lys 1570	Ala Gln Phe Val Ala 1575	Leu Asn Val Gly Lys 1580
20	Ala Val Ser Gln His 1585	Ile Ser Gln Leu Glu 1590	Met Asn Asn Glu Gly Gln 1595
	Tyr Asn Val Trp Ile Ser 1605	Asn Thr Ser Met Asn 1610	Lys Asn Tyr Ser Ser 1615
25	Glu Gln Tyr Arg Arg Phe 1620	Ser Ser Lys Ser Thr 1625	Gln Thr Gln Leu Gly 1630
	Trp Asp Gln Thr Ile Ser 1635	Asn Asn Val Gln Leu 1640	Gly Gly Val Phe Thr 1645
30	Tyr Val Arg Asn Ser Asn 1650	Asn Phe Asp Lys Ala 1655	Ser Ser Lys Asn Thr 1660
	Leu Ala Gln Val Asn Phe 1665	Tyr Ser Lys Tyr Tyr 1670	Ala Asp Asn His Trp 1675
35	Tyr Leu Gly Ile Asp Leu 1685	Gly Tyr Gly Lys Phe 1690	Gln Ser Asn Leu Gln 1695
	Thr Asn Asn Asn Ala Lys 1700	Phe Ala Arg His Thr 1705	Ala Gln Ile Gly Leu 1710
40	Thr Ala Gly Lys Ala Phe 1715	Asn Leu Gly Asn Phe 1720	Ala Val Lys Pro Thr 1725
	Val Gly Val Arg Tyr Ser 1730	Tyr Leu Ser Asn Ala 1735	Asp Phe Ala Leu Ala 1740
45	Gln Asp Arg Ile Lys Val 1745	Asn Pro Ile Ser Val 1750	Lys Thr Ala Phe Ala 1755
	Gln Val Asp Leu Ser Tyr 1765	Thr Tyr His Leu Gly 1770	Glu Phe Ser Ile Thr 1775
50	Pro Ile Leu Ser Ala Arg 1780	Tyr Asp Ala Asn Gln 1785	Gly Asn Gly Lys Ile 1790
	Asn Val Ser Val Tyr Asp 1795	Phe Ala Tyr Asn Val 1800	Glu Asn Gln Gln Gln 1805
55	Tyr Asn Ala Gly Leu Lys 1810	Leu Lys Tyr His Asn 1815	Val Lys Leu Ser Leu 1820
	Ile Gly Gly Leu Thr Lys 1825	Ala Lys Gln Ala Glu 1830	Lys Gln Lys Thr Ala 1835
60	Glu Val Lys Leu Ser Phe 1845	Ser Phe	

(2) INFORMACIÓN PARA EL SEQ ID NO: 7:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 8 aminoácidos

ES 2 278 381 T3

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 7:

Gly Asp Ser Gly Ser Pro Met Phe
1 5

(2) INFORMACIÓN PARA EL SEQ ID NO: 8:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 8 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 8:

Gly Asp Ser Gly Ser Pro Leu Phe
1 5

(2) INFORMACIÓN PARA EL SEQ ID NO: 9:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 7 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 9:

His Thr Tyr Phe Gly Ile Asp
1 5