



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 355 952**

51 Int. Cl.:
C07K 14/71 (2006.01)
C12N 5/10 (2006.01)
C12N 15/85 (2006.01)
C12N 15/12 (2006.01)
C07K 19/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **95916208 .2**
96 Fecha de presentación : **04.04.1995**
97 Número de publicación de la solicitud: **0804572**
97 Fecha de publicación de la solicitud: **05.11.1997**

54 Título: **Proteína de fusión que comprende bpTK7.**

30 Prioridad: **04.04.1994 US 222616**

45 Fecha de publicación de la mención BOPI:
01.04.2011

45 Fecha de la publicación del folleto de la patente:
01.04.2011

73 Titular/es: **GENENTECH, Inc.**
460 Point San Bruno Boulevard
South San Francisco, California 94080, US

72 Inventor/es: **Bennett, Brian, D.;**
Goeddel, David;
Lee, James, M.;
Matthews, William;
Tsai, Siao, Ping y
Wood, William, I.

74 Agente: **Ponti Sales, Adelaida**

ES 2 355 952 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓNANTECEDENTES DE LA INVENCIONCAMPO DE LA INVENCION

[0001] La presente invención se refiere a quimeras pTK-inmunoglobulina.

5 DESCRIPCIÓN DE LA TÉCNICA ANTERIOR

[0002] La transducción de señales que regulan el crecimiento y diferenciación celular está regulada en parte por la fosforilación de varias proteínas celulares. Las proteínas tirosina quinasa son enzimas que catalizan este proceso. Además, muchas actúan como receptores de factores de crecimiento. El subgrupo c-kit de receptores de tirosina quinasa catalizan la fosforilación de sustratos exógenos, así como residuos de tirosina en sus propias cadenas de polipéptidos (Ullrich et al., Cell 61:203 [1990]). Los miembros del subgrupo c-kit incluyen FLT/FLK (Kinasa de hígado fetal), FGF (Receptor del Factor de Crecimiento de Fibroblastos) y NGF (Receptor del Factor de Crecimiento Nervioso)

[0003] La subfamilia *EPH* de tirosina quinasa, Eph, Elk, Eck, Eek, Hek, Hek2, Sek, Ekh-1, Ekh-2, Cek-4 to -10, Tyro 1, 4, 5 y 6, parece ser la subfamilia más grande de tirosinas quinasas transmembrana (Hirai et al., Science 238:1717-1720 [1987]; Letwin et al., Oncogene 3:621-678 [1988]; Lhotak et al., Mol. Cell. Biol. 13:7071-7079 [1993]; Lindberg et al., Mol. cell. Biol. 10:6316-6324 [1990]; Bohme et al., Oncogene 8:2857-2862 [1993]; y Wicks et al., Proc. Natl. Acad. Sci. USA. 89:1611-1615 [1992]; Pasquale et al. Cell Regulation 2:523-534 [1991]; Sajjadi et al., New Biol. 3:769-778 [1991]; Wicks et al., Proc. Natl. Acad. Sci. USA. 89:1611-1615 [1992]; Lhotak et al., Mol. Cell. Bio. 11:2496-2502 [1991]; Gilardi-Hebenstreit et al., Oncogene 7:2499-2506 [1992]; Lai et al., Neuron 6:691-704 [1991]; Sajjadi et al., Oncogene 8:1807-1813 [1993]; and Maisonpierre et al., Oncogene 8:3277-3288 [1993]).

[0004] Las pTK adicionales y anticuerpos agonistas a las mismas son necesarias para estudiar adicionalmente el crecimiento y diferenciación de las células, para su uso como agentes terapéuticos y para fines de diagnóstico.

[0005] Ashman et al. (Journal of Cellular Physiology 158: 545-554; 1994) describe tres anticuerpos monoclonales al receptor tirosina quinasa c-kit humano.

25 [0006] WO 93/15201 A (New England Deaconess Hospital) describe información de la secuencia para las proteínas tirosina quinasa HpTK5, SAL-S1 y bpTK7, y propone el uso de las proteínas en ensayos para ligandos, tales como factores de crecimiento o factores de diferenciación que inhiben o aumentan la actividad quinasa y como inmunógenos para producir anticuerpos monoclonales o policlonales.

30 [0007] Mark et al. (The Journal of Biological Chemistry 267: 26166-26171; 1992) describe que el receptor del factor de crecimiento de hepatocitos (HGF α) es un receptor pTK y describe una proteína de fusión HGF α -IgG.

[0008] La presente invención proporciona anticuerpos agonistas que se unen al dominio extracelular del dominio quinasa de un receptor de proteína tirosina quinasa (pTK) seleccionado del grupo que consiste en:

a) SAL-S1;

b) HpTK 5; y

35 c) bpTK 7.

y lo activan

La presente invención también proporciona métodos de utilización de los mismos, quimeras del receptor bpTK7 y métodos para producir las quimeras.

DESCRIPCIÓN RESUMIDA DE LA INVENCION

40 [0009] Los genes aislados tal como se describen en la presente invención son referidos, colectivamente, como "genes de proteína tirosina quinasa" o "genes de pTK". Las secuencias de ácidos nucleicos de algunos de estos genes, aislados tal como se describe aquí, muestran una homología significativa con proteínas tirosina quinasas identificadas previamente que contienen dominios extracelulares, que actúan como receptores de factores de crecimiento (por ejemplo, pTKs del subgrupo c-kit). Se ha observado que algunos de los genes de pTK están
45 presentes tanto en células megacariocíticas como células linfocíticas.

[0010] En particular, se han identificado catorce genes de pTK. Se identificaron dos genes de pTK, referidos como SAL-S1 y SAL-D4, en células megacariocíticas. SAL-D4 está relacionado con la familia CSK de pTK intracelulares y SAL-S1 está relacionado con la familia de receptores FGF de pTK. Se identificaron cinco genes de pTK, referidos como LpTK, en células linfocíticas y se ha observado que están presentes también en megacariocitos. Se identificó
50 un gen de pTK, referido como HpTK5, en células de hematoma humano. Se hallaron seis genes de pTK, referidos

como genes de bpTK, en tejido de cerebro humano.

- [0011]** Los genes de pTK se identificaron en general utilizando dos grupos de cebadores de oligonucleótidos degenerados: un primer grupo que amplifica todos los segmentos de ADN de pTK (SEC ID NOS: 1-2), y un segundo grupo que amplifica secuencias altamente conservadas presentes en el dominio catalítico del subgrupo c-kit de pTKs (SEC ID NOS: 3-4). Los genes de pTK identificados de esta manera se describen a continuación.
- [0012]** SAL-S1 se expresa en varias líneas de células megacariocíticas, pero no en líneas de células eritroides. Se obtuvo la secuencia de nucleótidos de parte de SAL-S1, revelando una secuencia que contenía 160 pares de bases (SEC ID NO: 5). Este fragmento de ADN aislado codificaba una secuencia de aminoácidos (SEC ID NO: 6) que mostraba una homología de secuencia significativa con proteínas tirosina quinasas conocidas de la familia FLT/FLK.
- [0013]** SAL-D4, también expresado en células megacariocíticas, es un fragmento de ADN que contiene la secuencia de nucleótidos de 147 pares de bases. (SEC ID NO: 7). Este fragmento de ADN aislado codificaba una secuencia de aminoácidos (SEC ID NO: 8) que mostraba una homología de secuencia significativa con proteínas tirosina quinasas conocidas de la familia de pTK intracelulares CSK.
- [0014]** Las LpTKs, incluyendo LpTK 2, LpTK 3, LpTK 4, LpTK 13 y LpTK 25, se expresan en células linfocíticas, así como células megacariocíticas. Se obtuvo la secuencia de nucleótidos (151 pares de bases) del gen de LpTK 3 (SEC ID NO: 11). Las secuencia de nucleótidos de los genes de LpTK 2, LpTK 4, y LpTK 13 contenían 149 pares de bases (SEC ID NO: 9), 137 pares de bases (SEC ID NO: 13), y 211 pares de bases (SEC ID NO: 15) respectivamente. La LpTK 25 tiene una secuencia de nucleótidos de 3120 b.p. (SEC ID NO: 22). Se ha obtenido una secuencia de genes de longitud completa para LpTK 2 (SEC ID NO: 19) que contiene 7607 b.p. La secuenciación adicional de LpTK 4 reveló una secuencia de 404 b.p. (SEC ID NO: 21).
- [0015]** El gen de HpTK5, expresado en células de hematoma humano, presenta una secuencia de nucleótidos de 3969 b.p. (SEC ID NO: 23).
- [0016]** Las secuencias de nucleótidos de las bpTKs, incluyendo bpTK 1, bpTK 2, bpTK 3, bpTK 4, bpTK 5 y bpTK 7, se expresan en tejido de cerebro humano y codifican proteínas que tienen las secuencias de aminoácidos de SEC ID NOS: 25-30 y 36, respectivamente.
- [0017]** De este modo, se puede aislar el ADN de una línea de células megacariocíticas humanas, que se hibrida al ADN que codifica una secuencia de aminoácidos que es altamente conservada en el dominio catalítico de proteínas tirosina quinasas del subgrupo c-kit.
- [0018]** Las proteínas codificadas por los genes de pTK identificadas tal como se describe en la presente invención muestran una homología de secuencia significativa con los miembros del subgrupo c-kit de pTKs, así como las proteínas codificadas por HpTK5 y las bpTKs. Son posibles homólogos o equivalentes de SAL-S1, SAL-D4, LpTK, HpTK5 y bpTK (es decir, proteínas que tienen secuencias de aminoácidos sustancialmente similares, pero no idénticas, a las de SAL-S1, SAL-D4, las LpTKs, HpTK5 y las bpTKs, que muestran actividad tirosina quinasa). Los péptidos (fragmentos de SAL-S1, SAL-D4, LpTK, HpTK5 y bpTK) que mantienen la actividad tirosina quinasa, aunque son menos que las secuencias completas de SAL-S1, SAL-D4, LpTK, HpTK5 y bpTK también son posibles; y se pueden prever usos para las secuencias de ácidos nucleicos de SAL-S1, SAL-D4, las LpTK, HpTK y las bpTK y equivalentes de SAL-S1, SAL-D4, LpTK, HpTK y bpTK.
- [0019]** Las secuencias de ácidos nucleicos se pueden hibridar con ADN o ARN que codifican las proteínas descritas en la presente invención y pueden mostrar una homología de secuencia significativa con FLT/FLK, el receptor de FGF o la familia de receptores NGF de proteínas tirosina quinasas contenidas en el subgrupo c-kit. Dichas secuencias de ácido nucleico son útiles como sondas para identificar genes de pTK en otros vertebrados, particularmente mamíferos, y en otros tipos de células. También se pueden utilizar como oligonucleótidos antisentido para inhibir la actividad de proteína tirosina quinasa, tanto in vitro como in vivo.
- [0020]** Las tirosina quinasas SAL-S1, SAL-D4, LpTK, HpTK y bpTK se pueden utilizar como proteínas diana conjuntamente con el desarrollo de fármacos y agentes terapéuticos para modular el crecimiento y diferenciación celular y otras funciones metabólicas. Las proteínas SAL-S1, SAL-D4, LpTK, HpTK o bpTK se pueden utilizar como agonistas o antagonistas a otras tirosina quinasas. Las pTKs también pueden jugar un papel decisivo en la modulación de interacciones de adhesión a megacariocitos y/o plaquetas.
- [0021]** Además, las tirosina quinasas SAL-S1, SAL-D4, LpTK, HpTK y bpTK se pueden utilizar en ensayos de screening para detectar factores de crecimiento y/o diferenciación celular. Utilizando técnicas de laboratorio estándar, se pueden identificar los ligandos de las pTK. En particular, la presente invención proporciona proteínas de fusión quiméricas bpTK7-inmunoglobulina que son útiles para aislar ligandos para las pTK descritas en la presente invención. Las proteínas quiméricas también son útiles para ensayos de diagnóstico diseñados para detectar estos ligandos presentes de forma endógena, en las células, así como de forma exógena, en fluidos extracelulares. También se pueden diseñar ensayos, utilizando proteínas quiméricas, como herramientas de diagnóstico para detectar estos ligandos en fluidos corporales, tales como sangre y orina.

[0022] Los anticuerpos específicos para SAL-S1, SAL-D4, las LpTK, HpTK5 y las bpTK pueden ser agonistas para su respectiva pTK (donde la pTK es un receptor). En la presente invención se definen una línea de células de hibridoma y un ácido nucleico aislado que codifica un anticuerpo monoclonal.

5 [0023] Un método para activar una pTK tal como se describe en la presente invención, puede comprender hacer reaccionar la pTK con un anticuerpo agonista para la misma. Un método para aumentar el crecimiento y/o diferenciación celular puede comprender administrar a un paciente humano con necesidad de dicho tratamiento una cantidad fisiológicamente eficaz de un anticuerpo agonista que activa una pTK tal como se describe en la presente invención.

10 [0024] Un método para detectar una pTK puede implicar poner en contacto una fuente sospechosa de contener la pTK con un anticuerpo monoclonal marcado de forma detectable que reacciona de forma inmunológica con la pTK y determinar si el anticuerpo se une a la fuente.

BREVE DESCRIPCIÓN DE LAS FIGURAS

[0025] Las figuras 1A y 1B representan la secuencia de nucleótidos de SAL-S1 (SEC ID NO: 5) y su secuencia de aminoácidos deducida (SEC ID NO: 6).

15 Las figuras 2A y 2B representan la secuencia de nucleótidos de SAL-D4 (SEC ID NO: 7) y su secuencia de aminoácidos deducida (SEC ID NO: 8).

La figura 3A representa la secuencia de nucleótidos de LpTK 2 (SEC ID NO: 9) y su secuencia de aminoácidos deducida (SEC ID NO: 10).

20 La figura 3B representa la secuencia de nucleótidos de LpTK 3 (SEC ID NO: 11) y su secuencia de aminoácidos deducida (SEC ID NO: 12).

La figura 3C representa la secuencia de nucleótidos de LpTK 4 (SEC ID NO: 13) y su secuencia de aminoácidos deducida (SEC ID NO: 14).

La figura 3D representa la secuencia de nucleótidos de LpTK 13 (SEC ID NO: 15) y su secuencia de aminoácidos deducida (SEC ID NO: 16).

25 Las figuras 4A-4I representan la secuencia de nucleótidos (SEC ID NO: 17) de SAL-S1 y su secuencia de aminoácidos deducida (SEC ID NO: 18).

Las figuras 5A-5K representan la secuencia de nucleótidos de longitud completa (SEC ID NO: 19) de LpTK2 y su secuencia de aminoácidos deducida (SEC ID NO: 20).

La figura 6 representa la secuencia de nucleótidos parcial (SEC ID NO: 21) para LpTK4.

30 Las figuras 7A-7C representan la secuencia de nucleótidos de longitud completa (SEC ID NO: 22) para LpTK25.

Las figuras 8A-8I representan la secuencia de nucleótidos de longitud completa (SEC ID NO: 23) y la secuencia de aminoácidos deducida de HpTK5 (SEC ID NO: 24).

La figura 9 representa la secuencia de aminoácidos (SEC ID NO: 25) de bpTK1.

La figura 10 representa la secuencia de aminoácidos (SEC ID NO: 26) de bpTK2.

35 La figura 11 representa la secuencia de aminoácidos (SEC ID NO: 27) de bpTK3.

La figura 12 representa la secuencia de aminoácidos (SEC ID NO: 28) de bpTK4.

La figura 13 representa la secuencia de aminoácidos (SEC ID NO: 29) de bpTK5.

La figura 14 representa la secuencia de aminoácidos (SEC ID NO: 30) de bpTK7.

40 Las figuras 15A-15F representan la secuencia de nucleótidos de longitud completa de SAL-S1 (SEC ID NO: 31) y su secuencia de aminoácidos deducida (SEC ID NO: 33).

Las figuras 16A-16H representan la secuencia de nucleótidos de longitud completa de bpTK7 (SEC ID NO: 34) y su secuencia de aminoácidos deducida (SEC ID NO: 36).

DESCRIPCIÓN DETALLADA DE LAS REALIZACIONES PREFERIDAS

45 [0026] Se han identificado nuevos genes de proteínas tirosina quinasa, se han determinado sus secuencias de ácido nucleicos y las secuencias de aminoácidos de las proteínas codificadas deducidas. Los genes aislados tal como se describen en la presente invención son referidos, colectivamente, como "genes de proteínas tirosina quinasa" o

“genes de pTK”.

- [0027]** Para facilitar el aislamiento e identificación de estas pTKs nuevas, se utilizaron dos grupos de sondas de ADN, tal como se describe en el ejemplo 1. El primer grupo consistía en general de dos secuencias de oligonucleótidos degeneradas, pTK 1 (SEC ID NO: 1) y pTK 2 (SEC ID NO: 2) (Matthews, Cell 65:1143 [1991]; y Wilks, Proc. Natl. Acad. Sci. USA 86:1603 [1989]). Estas secuencias se utilizaron como cebadores en una reacción en cadena de la polimerasa para amplificar los segmentos de ADN de tirosina quinasa (Mullis, et al., Cold Spring Harbor Symp. Advan. Biol. 51:263 [1986]).
- [0028]** El segundo grupo consistía en general de dos secuencias de oligonucleótidos, pTK 3 (SEC ID NO: 3) y pTKKW (SEC ID NO: 4) diseñadas para amplificar la secuencia de ácidos nucleicos que codifica las regiones altamente conservadas de los dominios catalíticos de la familia c-kit de proteínas tirosina quinasa. Estas secuencias se utilizaron como cebadores en la reacción en cadena de la polimerasa (PCR) en una segunda ronda de amplificación de ADN. Utilizando este procedimiento de amplificación de dos etapas, se identificaron fragmentos de ADN que se hibridaban a estos cebadores de pTK, se aislaron y posteriormente se secuenciaron.
- [0029]** En particular, se han identificado catorce genes de pTK. Se identificaron dos genes de pTK, referidos como SAL-S1 y SAL-D4, en varias líneas de células megacariocíticas, incluyendo CMK 11-5, DAMI, UT-7 y UT-7 desarrolladas en eritropoyetina, pero no en las líneas de células eritroides HEL, células HEL estimuladas con PMA, o K562. Se identificaron cinco genes de pTK, referidos como LpTKs, en células linfocíticas, así como en células megacariocíticas. Se identificó un gen de pTK, referido como HpTK5, en células de hematoma humano, y se identificaron seis genes, referidos como bpTKs, en tejido de cerebro humano.
- [0030]** La SAL-S1 (SEC ID NOS: 6, 18 y 33) codificada por la secuencia de ácidos nucleicos de SEC ID NOS: 5, 17 y 31 muestra una homología significativa con la familia FLT/FLK de pTKs. SAL-S1 tiene un péptido señal (es decir, residuos de aminoácidos 1 a 24 de la figura 15); dominios extracelulares (es decir, residuos de aminoácidos 25 a 775 de la figura 15); dominio transmembrana (es decir, residuos de aminoácidos 776 a 800 de la figura 15) y un dominio citoplasmático de tirosina quinasa (es decir, residuos de aminoácidos 801 a 1298 de la figura 15). La SAL-D4 (SEC ID NO: 8) codificada por SEC ID NO: 7 se refiere a la familia CSK de pTKs intracelulares. Las LpTKs, LpTK 2 (SEC ID NOS: 10 y 20) codificadas por SEC ID NOS: 9 y 19; LpTK 3 (SEC ID NO: 12) codificada por SEC ID NO: 11; LpTK4 (SEC ID NO: 14) codificada por SEC ID NOS: 13 y 21; LpTK13 (SEC ID NO: 16) codificada por SEC ID NO: 15; y LpTK25 codificada por SEC ID NO: 22, también muestran una homología de secuencia con proteínas tirosina quinasa conocidas.
- [0031]** La HpTK5 (SEC ID NO: 24) codificada por SEC ID NO: 23 y las bpTKs 1, 2, 3, 4, 5 y 7 (SEC ID NOS: 25-29 y 36 respectivamente), muestran de manera similar una homología de secuencia con proteínas tirosina quinasa conocidas. La BpTK7 codifica un receptor de pTK con un péptido señal (es decir, residuos de aminoácidos 1-19 de la figura 16); dominio extracelular (es decir, residuos de aminoácidos 20-547 de la figura 16); y dominio transmembrana (es decir, residuos de aminoácidos 548-570 de la figura 16). La secuencia restante comprende el dominio tirosina quinasa intracelular.
- [0032]** De este modo, tal como se ha descrito anteriormente, se han aislado y secuenciado moléculas de ADN que se hibridan con ADN que codifica las secuencias de aminoácidos presentes en el dominio catalítico de una proteína tirosina quinasa del subgrupo c-kit de proteínas quinasa. Se ha observado que estas secuencias de ADN aisladas, referidas colectivamente como “genes de pTK”, (y sus secuencias de aminoácidos deducidas) muestran una homología de secuencia significativa con miembros conocidos de familias pTK.
- [0033]** Una vez aislados, estos fragmentos de ADN se pueden amplificar utilizando técnicas estándar conocidas, tales como PCR. Estos fragmentos amplificados se pueden clonar a continuación en vectores de clonación apropiados y determinar sus secuencias de ADN.
- [0034]** Estas secuencias de ADN se pueden escindir de los vectores de clonación, marcarse con un nucleótido radiomarcado, tal como ³²P, y utilizarse para cribar las bibliotecas de ADNc apropiadas para obtener el clon de ADNc de longitud completa.
- [0035]** Los genes de pTK, tal como se han descrito anteriormente, se han aislado de la fuente en que están de forma natural, por ejemplo, células megacariocíticas y linfocíticas. La presente invención pretende incluir genes de pTK producidos utilizando técnicas de ingeniería genética, tales como tecnología recombinante, así como genes de pTK que se sintetizan químicamente.
- [0036]** Las secuencias de aminoácidos deducidas de los genes de pTK incluyen secuencias de aminoácidos que codifican péptidos que muestran una homología significativa con el dominio catalítico de proteínas tirosina quinasa del subgrupo c-kit de tirosina quinasa. Estas proteínas, codificadas por los genes de pTK, pueden incluir secuencias en las que residuos de aminoácidos funcionalmente equivalentes están sustituidos por residuos en la secuencia, que dan lugar a un cambio silencioso, que es un cambio no detectado fenotípicamente. Por ejemplo, se pueden sustituir uno o más residuos de aminoácidos en la secuencia por otro aminoácido de una polaridad similar que actúa como equivalente funcional, dando lugar a una sustitución silenciosa.

[0037] Además, la estructura de la proteína se puede modificar mediante deleciones, adiciones, inversión, inserciones o sustituciones de uno o más residuos de aminoácidos en la secuencia que no perjudican sustancialmente a las propiedades funcionales deseadas de tirosina quinasa del péptido.

5 **[0038]** Las pTK modificadas, con actividad tirosina quinasa, se pueden producir utilizando técnicas de ADN recombinante, tales como escisión de un vector que contiene un ADNc que codifica dicha proteína, o mediante la síntesis del ADN que codifica la proteína deseada mecánicamente y/o químicamente utilizando técnicas conocidas.

10 **[0039]** Una estrategia alternativa para producir pTK es utilizar la síntesis de péptidos para producir un péptido o polipéptido que tiene la secuencia de aminoácidos de dicha proteína, dependiendo de la longitud de pTK deseada. Los péptidos o equivalentes modificados de los mismos se pueden sintetizar directamente mediante química de fase sólida o líquida estándar para la síntesis de péptidos.

[0040] Preferiblemente, las pTK se producirán mediante la inserción de ADN que codifica las proteínas en un sistema vector/huésped apropiado donde se expresará. Las secuencias de ADN se pueden obtener de fuentes en las que tienen lugar de forma natural, se pueden sintetizar químicamente o se pueden producir utilizando tecnología recombinante estándar.

15 **[0041]** Un vector de expresión puede comprender un gen de pTK, que codifica una proteína que muestra actividad de receptor tirosina quinasa.

[0042] Los genes de pTK se pueden utilizar para un conjunto de fines de diagnóstico y terapéuticos. Por ejemplo, las secuencias de ácido nucleico de los genes de pTK se pueden utilizar como sondas para identificar otras proteínas tirosina quinasa presentes en otros tipos de células, incluyendo tipos de células eucariotas y procariotas.

20 **[0043]** Las secuencias de ácido nucleico también se pueden utilizar para diseñar fármacos que inhiben directamente la actividad quinasa de proteínas tirosina quinasa o para diseñar péptidos que se unen al dominio catalítico de tirosina quinasa, inhibiendo de este modo su actividad. Estas secuencias también se pueden utilizar para diseñar nucleótidos anti-sentido que también pueden inhibir o destruir actividad tirosina quinasa. Dicha inhibición de actividad tirosina quinasa sería deseable en estados patológicos donde la proliferación celular disminuida sería
25 beneficiosa, tales como leucemias u otros cánceres.

[0044] Las secuencias de ácidos nucleicos también se pueden utilizar para diseñar fármacos, péptidos o nucleótidos antisentido como antes, pero con efectos potenciadores, en lugar de inhibidores, sobre tirosina quinasa. Dicha actividad aumentada de tirosina quinasa daría lugar a un incremento de la fosforilación de sustratos (exógenos, así como residuos de tirosina endógenos). Serían deseables efectos potenciados en estados en los que una mayor
30 proliferación celular sería beneficiosa, tales como anemias, trastornos del sangrado y durante procedimientos quirúrgicos.

[0045] Los genes de pTK también se pueden utilizar para obtener fragmentos solubles de receptores tirosina quinasa, capaces de unirse a sus ligandos respectivos. Los genes de pTK que codifican fragmentos solubles de tirosina quinasa se pueden producir utilizando técnicas de ADN recombinante o sintéticamente. En cualquier caso, el
35 ADN obtenido codifica un fragmento de pTK soluble que carece de una parte sustancial de la región transmembrana hidrofóbica para permitir la solubilización del fragmento.

[0046] Estos fragmentos de proteínas pTK se pueden introducir exógenamente para actuar como competidores con las pTK endógenas unidas a membrana por sus respectivos ligandos, inhibiendo así la actividad de tirosina quinasa. Alternativamente, se puede introducir un fragmento de proteína pTK soluble modificado que se une al ligando pero
40 no activa la actividad quinasa.

[0047] Estos fragmentos solubles de proteínas pTK también se pueden utilizar en ensayos de unión para detectar ligandos, tales como factores de crecimiento y diferenciación. Una vez se han identificado estos ligandos, se pueden alterar o modificar para inhibir o potenciar la actividad quinasa. Por ejemplo, los ligandos se pueden modificar o unir a sustancias que son tóxicas para la célula, tal como ricina, destruyendo así la célula diana. La sustancia puede ser
45 una sustancia superactivante que, después de la unión a la pTK, puede incrementar sustancialmente la actividad quinasa, o activar otros factores de crecimiento.

[0048] Los genes de pTK también serían útiles para desarrollar herramientas de diagnóstico para ensayos de cribado in vitro para ligandos, tales como factores de crecimiento o factores de diferenciación que inhiben o potencian la actividad quinasa. Las proteínas codificadas por los genes de pTK también se pueden utilizar en dichos
50 ensayos, o como inmunógenos, para producir anticuerpos monoclonales o policlonales a utilizar en dichos ensayos.

[0049] En una realización de la invención, se puede construir una quimera que comprende una fusión del dominio extracelular de la bPTK7 y un dominio constante de inmunoglobulina que se puede utilizar para ensayar los ligandos para el receptor y se puede utilizar para la producción de anticuerpos contra el dominio extracelular del receptor.

55 **[0050]** La expresión "dominio extracelular" o "ECD" cuando se utiliza en la presente invención se refiere a cualquier secuencia de polipéptido que comparte una función de unión a ligando del dominio extracelular del receptor natural

de pTKs descrito aquí. La función de unión a ligando del dominio extracelular se refiere a la capacidad del polipéptido de unirse a por lo menos un ligando de pTK. Por consiguiente, no es necesario incluir el dominio extracelular completo, ya que se encuentra que normalmente segmentos más pequeños son adecuados para unirse al ligando. El dominio extracelular truncado es en general soluble. El término ECD comprende secuencias de polipéptido en las que la secuencia transmembrana hidrofóbica (y, opcionalmente los aminoácidos 1-20 C-terminales y/o N-terminales con respecto al dominio transmembrana) de la pTK madura ha sido eliminada. De este modo, el polipéptido que contiene el dominio extracelular soluble puede comprender el dominio extracelular y el dominio citoplasmático de la pTK. Alternativamente, en la realización preferida, el polipéptido comprende sólo el dominio extracelular de la pTK. Los dominio extracelular y transmembrana de la pTK se pueden determinar fácilmente por el experto en la materia mediante la alineación de la pTK de interés con secuencias de aminoácidos de pTK conocidas para los que se han trazado estos dominios. Alternativamente, el dominio transmembrana hidrofóbico se puede trazar fácilmente en base a la representación de la hidrofobicidad de la secuencia. El dominio extracelular es N-terminal con respecto al dominio transmembrana.

[0051] El término “inmunoglobulina” se refiere en general a polipéptidos que comprende una cadea aligera o pesada normalmente ambas unidas a disulfuro en la configuración “Y” nativa, aunque otras uniones entre ellos, incluyendo tetrameros o agregados de los mismos, se encuentra en el alcance de la invención.

[0052] Las inmunoglobulinas (Ig) y ciertas variantes de las mismas son conocidas y muchas se han preparado en cultivo celular recombinante. Por ejemplo, véase la Patente de Estados Unidos 4,745,055; EP 256,654; Faulkner et al., Nature 398:286 [1982]; EP 120,694; EP 125,023; Morrison, J. Immun. 123:793 [1979]; Köhler et al., Proc. Nat'l. Acad. Sci. USA 77:2197 [1980]; Raso et al., Cancer Res. 41:2073 [1981]; Morrison et al., Ann. Rev. Immunol. 2:239 [1984]; Morrison, Science 229:1202 [1985]; Morrison et al., Proc. Nat'l. Acad. Sci. USA 81:6851 [1984]; EP 255,694; EP 266,663; y WO 88/03559. También se conocen cadenas de inmunoglobulina recombinadas. Véase, por ejemplo, la patente de Estados Unidos 4,444,878; WO 88/03565; y EP 68,763 y las referencias citadas en las mismas. La parte de inmunoglobulina en la quimera de la presente invención se puede obtener de los subtipos IgG1, IgG2, IgG3, o IgG4, IgA, IgE, IgD o IgM, pero preferiblemente IgG1 o IgG3. Más preferiblemente, la parte inmunoglobulina es la parte Fc de IgG-γ.

[0053] Los términos “quimera que comprende una fusión de un dominio extracelular de una pTK con una secuencia de dominio constante de inmunoglobulina” o “quimera pTK-inmunoglobulina” se refieren a un polipéptido que comprende un dominio extracelular que codifica una secuencia de aminoácido de una pTK conjugada a una secuencia de dominio constante de inmunoglobulina. Esta definición incluye quimeras en formas monoméricas, homomultiméricas o heteromultiméricas, y particularmente homodiméricas o heterodiméricas o tetraméricas.

[0054] Una realización preferida es la fusión del extremo C-terminal del dominio extracelular de una pTK, al N-terminal de la parte C-terminal de un anticuerpo (en particular, el dominio Fc) que contiene las funciones efectoras de inmunoglobulina G1. En una realización preferida, la región constante de cadena pesada completa se fusiona al dominio extracelular. En otra realización preferida, una secuencia que empieza en la región bisagra justo en dirección 5' del sitio de división por papaína (que define la Fv de IgG químicamente; el residuo 216, tomando el primer residuo de la región constante de la cadena pesada el 114 (Kabat et al., Sequences of Immunological Interest, National Institutes of Health, Bethesda, MD, [1987]), o sitios análogos de otras inmunoglobulinas) se fusiona con la ECD de la pTK.

[0055] En una realización particularmente preferida, el dominio extracelular de pTK se fusiona a la región bisagra y la bisagra CH2 y CH3 o CH1, dominio CH2 y CH3 de una cadena pesada de IgG1, IgG2 o IgG3. El sitio preciso en que la fusión se realiza no es crítico, y el sitio óptimo se puede determinar mediante experimentación de rutina. Una ventaja principal de las quimeras es que se secretan en el medio de cultivo de huéspedes recombinantes, aunque el grado de secreción podría ser diferente para varios sistemas de expresión.

[0056] En general, las quimeras de la presente invención se construyen de una manera similar a los anticuerpos quiméricos en los que un dominio variable de un anticuerpo de una especie se sustituye por el dominio variable de otra especie. Véase, por ejemplo, EP 0 125 023; EP 173,494; Munro, Nature 312: [13 Diciembre 1984]; Neuberger et al., Nature 312: [13 Diciembre 1984]; Sharon et al., Nature 309: [24 May 1984]; Morrison et al., Proc. Nat'l. Acad. Sci. USA 81:6851-6855 [1984]; Morrison et al. Science 229:1202-1207 [1985]; Boulianne et al., Nature 312:643-646 [13 Diciembre 1984]; Capon et al., Nature 337, 525-531 [1989]; Trauneker et al., Nature 339, 68-70 [1989].

[0057] Para preparar los polipéptidos quiméricos pTK-IgG, se divide el ADN que incluye una región que codifica la secuencia de pTK deseada mediante una enzima de restricción en el extremo 3' o próximo al mismo del ADN que codifica el dominio o dominios de tipo inmunoglobulina en un punto o próximo al ADN que codifica el extremo N-terminal de la pTK madura (donde se contempla el uso de una secuencia líder diferente) o en la región codificante N-terminal para pTK o próxima a la misma (donde se utiliza la secuencia señal nativa). Este fragmento de ADN se inserta fácilmente a continuación de forma próxima al ADN que codifica una región constante

de cadena ligera o pesada de inmunoglobulina y, si es necesario, la construcción resultante se ajusta mediante mutagénesis por delección. Preferiblemente, la Ig es una inmunoglobulina humana cuando la variante está destinada a terapia in vivo para humanos. El ADN que codifica las regiones constantes de cadena ligera o pesada de

inmunoglobulina es conocido o fácilmente disponible a partir de bibliotecas de ADNc o se sintetiza. Véase, por ejemplo, Adams et al., *Biochemistry* 19:2711-2719 [1980]; Gough et al., *Biochemistry* 19:2702-2710 [1980]; Dolby et al., *P.N.A.S. USA*, 22:6027-6031 [1980]; Rice et al., *P.N.A.S. USA* 79:7862-7865 [1982]; Falkner et al., *Nature* 298:286-288 [1982]; y Morrison et al., *Ann. Rev. Immunol.* 2:239.256 [1984].

- 5 **[0058]** Las proteínas quiméricas descritas aquí son útiles como agentes de diagnóstico para aislar o cribar ligandos para la pTK de interés utilizando las técnicas de Lyman et al., *Cell* 75:1157-1167 [1993], por ejemplo. Además, las proteínas quiméricas son útiles para fines de diagnóstico para estudiar la interacción de varios ligandos con el dominio extracelular de las diversas pTK (véase, por ejemplo, Bennett et al., *J. Biol. Chem.* 266(34):23060-23067 [1991]). Las proteínas quiméricas son útiles además para la producción de anticuerpos contra el dominio extracelular de la pTK (véase los ejemplos 3 y 5 de la presente invención). Las proteínas quiméricas también presentan una utilidad terapéutica adicional siempre que proporcionen una forma soluble del dominio extracelular de la pTK que generalmente presenta una mayor vida media en plasma (en comparación con el dominio extracelular solo) y, por lo tanto, se puede formular en un portador farmacéuticamente aceptable y se administra a un paciente. Las proteínas quiméricas se cree que son útiles como agentes terapéuticos para la eliminación del exceso de ligando de pTK
- 10 sistémico o localizado en un tejido que se ha administrado a un paciente. La eliminación del ligando en exceso es particularmente de forma deseable cuando el ligando puede ser tóxico para el paciente. La proteína quimérica actúa uniéndose al ligando en competición con la pTK endógena en el paciente. De forma similar, se contempla que la proteína quimérica se puede administrar a un paciente de manera simultánea, o posteriormente a la administración del ligando en forma de una composición de liberación prolongada. La proteína quimérica actúa como proteína de
- 15 unión soluble para el ligando, extendiendo de este modo la vida media del ligando.

[0059] El término "anticuerpo" se utiliza en la presente invención en el sentido más amplio y cubre específicamente anticuerpos policlonales, anticuerpos monoclonales, cadenas de inmunoglobulinas o fragmentos de las mismas, que reaccionan inmunológicamente con una pTK.

- 25 **[0060]** En la realización preferida de la invención, los anticuerpos son anticuerpos monoclonales producidos utilizando técnicas que son conocidas en el sector. Por ejemplo, se puede utilizar la técnica de hibridoma descrita originalmente por Kohler y Milstein, *Eur. J. Immunol.*, 6:511 [1976], y también descrita por Hammerling et al., In: *Monoclonal Antibodies and T-Cell Hybridomas*, Elsevier, N.Y., pp. 563-681 [1981]. Las técnicas de Cote et al. y Boerner et al. también están disponibles para la preparación de anticuerpos monoclonales humanos [Cote et al., *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, p. 77 [1985] y Boerner et al., *J. Immunol.*, 147(1):86-95 [1991]).

- [0061]** El término "anticuerpo monoclonal" tal como se utiliza en la presente invención se refiere a un anticuerpo (tal como se ha definido anteriormente en la presente invención) obtenido de una población de anticuerpos sustancialmente homogéneos, es decir, los anticuerpos individuales que comprenden la población son idénticos a excepción de posibles mutaciones naturales que pueden estar presentes en cantidades menores. Los anticuerpos monoclonales son muy específicos, dirigiéndose contra un único sitio antigénico. Además, a diferencia de las preparaciones de anticuerpos convencionales (policlonales) que incluyen habitualmente diferentes anticuerpos dirigidos contra diferentes determinantes (epítomos), cada anticuerpo monoclonal está dirigido contra un único determinante en el antígeno. Además de su especificidad, los anticuerpos monoclonales son ventajosos en que se pueden sintetizar mediante un cultivo de hibridoma, no contaminado por otras inmunoglobulinas.

- 40 **[0062]** Las formas "humanizadas" de anticuerpos no humanos (por ejemplo, murinos) son inmunoglobulinas, cadenas de inmunoglobulinas o fragmentos de las mismas (tales como, Fv, Fab, Fab', F(ab')₂ u otras subsecuencias de anticuerpos de unión a antígeno) que contienen mínimos residuos de aminoácidos derivados de una inmunoglobulina no humana. Para la mayor parte, los anticuerpos humanizados son inmunoglobulinas humanas (anticuerpo receptor) en que los residuos de una región determinante de complementariedad (CDR) del receptor se
- 45 sustituyen por residuos de una CDR de una especie no humana (anticuerpo dador), tal como ratón, rata o conejo que tiene la especificidad, afinidad y capacidad deseadas. En algunos casos, la región armazón Fv (FR) de la inmunoglobulina humana se sustituye por los correspondientes residuos de FR no humana. Además, un anticuerpo humanizado puede comprender residuos que no se encuentran ni en el anticuerpo receptor ni en la CDR o secuencias del armazón importadas. Estas modificaciones se realizan para refinar adicionalmente y optimizar la
- 50 acción del anticuerpo.

- [0063]** Los anticuerpos monoclonales de la presente invención incluyen anticuerpos híbridos (quiméricos) y recombinantes producidos por el corte y empalme ("splicing") de un dominio variable (incluyendo hipervariable) de un anticuerpo anti-pTK con un dominio constante (por ejemplo, anticuerpos "humanizados"), sólo uno de los cuales dirigido contra una pTK, o cadena ligera con una cadena pesada, o una cadena de una especie con una cadena de
- 55 otra especie, o fusiones con proteínas heterólogas, independientemente de especies de origen o designación de clase o subclase de inmunoglobulina, siempre que sean capaces de unirse a la pTK de interés [Véase, por ejemplo, Cabilly, et al., *Patente de Estados Unidos No. 4,816,567*; y Mage & Lamoyi, in *Monoclonal Antibody Production Techniques and Applications*, pág. 79-97 (Marcel Dekker, Inc., New York [1987]).

- [0064]** Para anticuerpos "quiméricos" y "humanizados", véase, por ejemplo, la patente de Estados Unidos No. 60 4,816,567; WO 91/09968; EP452,508; y WO 91/16927.

[0065] De este modo, el modificador “monoclonal” indica el carácter del anticuerpo obtenido de una población de anticuerpos sustancialmente homogéneos y no debe interpretarse que se requiere la producción del anticuerpo mediante ningún método particular.

[0066] Los anticuerpos pueden ser anticuerpos agonistas. Por “anticuerpo agonista” se entiende un anticuerpo que es un anticuerpo agonista. Por “anticuerpo agonista” se entiende un anticuerpo que es capaz de unirse a una pTK particular y activarla. Por ejemplo, el agonista se puede unir al dominio extracelular de la pTK y causar de este modo la dimerización de la pTK, dando lugar a la transfosforilación y activación del dominio quinasa catalítico intracelular. Consecuentemente, esto puede dar lugar a la estimulación del crecimiento y/o diferenciación de células que expresan el receptor in vitro y/o in vivo. Los anticuerpos agonistas de la presente invención son preferiblemente contra epítomos en el dominio extracelular de la pTK, y preferiblemente, presentan las mismas características biológicas que el anticuerpo monoclonal producido por la línea de células de hibridoma depositada en la American TIPO Culture Collection No. de Acceso. ATCC HB 11,583. Por “características biológicas” se entiende las actividades in vitro y/o in vivo del anticuerpo monoclonal, por ejemplo, capacidad de activar el dominio quinasa de una pTK particular, capacidad de estimular el crecimiento y/o diferenciación celular de células que expresan la pTK, y características de unión del anticuerpo, etc. Por consiguiente, el anticuerpo se une preferiblemente a sustancialmente el mismo epítipo que el anticuerpo monoclonal anti-HpTK5 descrito específicamente en la presente invención. Más preferiblemente, el anticuerpo tendrá sustancialmente la misma afinidad de unión a antígeno o superior del anticuerpo monoclonal anti-HpTK5 descrito en la presente invención. Para determinar si un anticuerpo monoclonal presenta la misma especificidad que el anticuerpo anti-HpTK5 descrito específicamente (es decir, el anticuerpo que tiene el depósito ATCC No. HB 11,583), se puede utilizar, por ejemplo, un ensayo de unión ELISA competitivo.

[0067] El ADN que codifica anticuerpos monoclonales se aísla y secuencía fácilmente utilizando procedimientos convencionales (por ejemplo, mediante la utilización de sondas de oligonucleótidos que son capaces de unirse específicamente a genes que codifican las cadenas pesada y ligera de anticuerpos murinos). Las células de hibridoma sirven como una fuente preferida de dicho ADN. Una vez aislado, el ADN se puede colocar en vectores de expresión que, a continuación, se transfectan en células huésped, tales como células E. coli, células COS de simio, células de Ovario de Hámster Chino (CHO), o células de mieloma que no producen de otro modo proteína inmunoglobulina, para obtener la síntesis de anticuerpos monoclonales en las células huésped recombinantes.

[0068] Los anticuerpos agonistas descritos en la presente invención son útiles para ensayos de diagnóstico in vitro para activar el receptor de pTK de interés. Esto es útil a efectos de estudiar el papel del receptor en el crecimiento y/o diferenciación celular.

[0069] Los anticuerpos agonistas de pTK presentan una utilidad terapéutica adicional en un método para aumentar el crecimiento y/o diferenciación celular que comprende administrar a un paciente humano con necesidad de dicho tratamiento una cantidad fisiológicamente eficaz de un anticuerpo agonista de pTK exógeno. Los anticuerpos agonistas para la pTK SAL-S1 pueden ser útiles en el tratamiento de trastornos con el sangrado y anemias, ya que se observó que esta pTK se expresaba en células megacariocíticas. Los anticuerpos agonistas de bpTK se pueden utilizar de forma similar para aumentar la diferenciación y/o proliferación de células de cerebro en enfermedades neurodegenerativas (tales como la enfermedad de Alzheimer) en base a la expresión de estos receptores en tejido cerebral. Finalmente, los anticuerpos agonistas de HpTK5 se pueden utilizar para aumentar la proliferación de células hematopoyéticas primitivas en pacientes que han experimentado quimioterapia o terapia por radiación o trasplante de médula ósea.

[0070] Un compuesto terapéutico “exógeno” se define en la presente invención para significar un compuesto terapéutico que es exógeno para el paciente mamífero, u homólogo a un compuesto que se encuentra en el paciente mamífero, pero producido fuera del paciente mamífero.

[0071] Los anticuerpos también son adecuados para detectar una pTK mediante el contacto de una fuente sospechosa de contener la pTK con un anticuerpo monoclonal marcado de forma detectable y determinar si el anticuerpo se une a la fuente. Existen muchos marcadores diferentes y métodos de marcaje conocidos en la técnica. Entre los marcadores adecuados se incluyen, por ejemplo, enzimas, radioisótopos, compuestos fluorescentes, compuestos quimioluminiscentes y bioluminiscentes, isótopos paramagnéticos. La pTK puede estar presente en muestras biológicas, tales como fluidos biológicos o tejidos. Para fines analíticos o de diagnóstico, los anticuerpos se administran en una cantidad suficiente para permitir la detección de un sitio en una pTK para el que el anticuerpo monoclonal es específico. La concentración del anticuerpo monoclonal marcado de forma detectable debe ser suficiente para proporcionar una señal detectable por encima de la línea base cuando se une a un epítipo de pTK.

[0072] Los anticuerpos agonistas de pTL descritos en la presente invención se pueden administrar a un mamífero, preferiblemente un humano, en una forma de dosificación farmacéuticamente aceptable, incluyendo aquellas que se pueden administrar a un humano de forma intravenosa como un bolo o mediante infusión continua durante un periodo de tiempo, mediante rutas intramuscular, subcutánea, intraarticular, intrasinoval, intratecal, oral, tópica o por inhalación.

[0073] Dichas formas de dosificación comprenden portadores farmacéuticamente aceptables que son

intrínsecamente no tóxicos y no terapéuticos. Entre los ejemplos de dichos portadores se incluyen intercambiadores de iones, alúmina, estearato de aluminio, lecitina, proteínas del suero, tales como albúmina de suero humano, sustancias tamponadoras, tales como fosfatos, glicina, ácido sórbico, sorbato de potasio, mezclas de glicéridos parciales de aceites grasos vegetales, agua, sales, o electrolitos, tales como sulfato de protamina, hidrogeno fosfato de disodio, hidrogeno fosfato de potasio, cloruro de sodio, sales de zinc, sílice coloidal, trisilicato de magnesio, polivinil pirrolidona, sustancias basadas en celulosa y polietilenglicol. Los portadores para formas tópicas o de base gel de anticuerpo incluyen polisacáridos, tales como carboximetilcelulosa de sodio o metilcelulosa, polivinilpirrolidona, poliacrilatos, polímeros en bloque de polioxietileno-polioxipropileno, polietilenglicol, y alcoholes de cera de madera. Para todas las administraciones, se utilizan de forma adecuada formas de depósito. Dichas formas incluyen, por ejemplo, microcápsulas, nanocápsulas, liposomas, vendajes adhesivos, formas de inhalación, pulverizadores nasales y comprimidos sublinguales. El anticuerpo se formulará habitualmente en dichos vehículos a una concentración de aproximadamente 0,1 mg/ml a 100 mg/ml.

[0074] Las composiciones farmacéuticas se pueden preparar y formular en formas de dosificación mediante métodos conocidos en la técnica; por ejemplo, véase, Remington's Pharmaceutical Sciences, Mack Publishing Company, Easton, Pennsylvania, 15a Edición 1975.

[0075] Una cantidad eficaz del anticuerpo agonista de pTK a utilizar terapéuticamente dependerá, por ejemplo, de los objetivos terapéuticos, la ruta de administración, y la condición del paciente. Por consiguiente, será necesario para el terapeuta titular la dosis y modificar la ruta de administración según se desee para obtener el efecto terapéutico óptimo. Una dosis diaria habitual podría variar desde aproximadamente 1 µg/kg hasta 1000 mg/kg o más, dependiendo de los factores mencionados anteriormente. Habitualmente, el médico administrará la molécula hasta alcanzar una dosis que consigue el efecto deseado. El proceso de esta terapia se controla fácilmente mediante ensayos convencionales.

[0076] Dependiendo del tipo y gravedad de la enfermedad, de aproximadamente 0,001 mg/kg a aproximadamente 1000 mg/kg, más preferiblemente aproximadamente 0,01 mg a 100 mg/kg, más preferiblemente aproximadamente 0,010 a 20 mg/kg del anticuerpo agonista podría ser una dosificación candidata inicial para la administración al paciente, mediante, por ejemplo una o más administraciones separadas o mediante infusión continua. Para administraciones repetidas durante varios días o más, dependiendo de la enfermedad, el tratamiento se repite hasta que tenga lugar una supresión deseada de síntomas de la enfermedad o se consigue la mejora deseada en la enfermedad del paciente. Sin embargo, también pueden ser útiles otras pautas de dosificación.

[0077] La presente invención se ilustrará a continuación mediante los siguientes ejemplos, que no pretenden ser limitantes en ningún caso.

EJEMPLO 1

IDENTIFICACIÓN Y AISLAMIENTO DE GENES DE pTK

[0078] Para facilitar el aislamiento e identificación de genes de pTK, se utilizaron en general dos grupos de sondas de ADN (véase la Tabla 1).

[0079] El primer grupo consistía en dos secuencias de oligonucleótidos degenerados, pTK1 (SEC ID NO: 1) y pTK 2 (SEC ID NO: 2). Estas secuencias se utilizaron como cebadores de la reacción en cadena de la polimerasa (PCR) utilizando técnicas PCR estándar, para amplificar segmentos de ADN de tirosina quinasa.

[0080] El segundo grupo consistía en dos secuencias de oligonucleótidos, pTK 3 (SEC ID NO: 3) y pTKKW (SEC ID NO: 4) seleccionadas de las regiones altamente conservadas de los dominios catalíticos del subgrupo c-kit de proteínas tirosina quinasa. Estas secuencias también se utilizaron como cebadores de la reacción en cadena de la polimerasa en una segunda ronda de amplificación del ADN. Utilizando este procedimiento de amplificación de dos etapas, se identificaron, aislaron y posteriormente se secuenciaron fragmentos de ADN que se hibridaban a estos cebadores de pTK utilizando técnicas de laboratorio conocidas.

TABLA 1

Primera ronda de amplificación

<u>Nombre de la sonda</u>	<u>Secuencia</u>
pTK1	5'-CGGATCCACAGNGACCT-3'
pTK2	5'-GGAATTCCAAAGGACCAGACGTC-3'

Segunda ronda de amplificación

pTK3 (específica de la familia kit) 5'-CGGATCCATCCACAGAGATGT-3'
 pTKKW (específica de la familia kit) 5'-GGAATTCCTTCAGGAGCCATCCACTT-3'

EJEMPLO 2 AISLAMIENTO Y CARACTERIZACIÓN DE HpTK55 A. Amplificación de ADN y clonación de HpTK5

[0081] Las células mononucleares de médula ósea humana de densidad ligera, obtenidas de voluntarios normales utilizando protocolos aprobados por el Deaconess Hospital Institutional Review Board y con el consentimiento por escrito de los voluntarios, se separaron mediante anticuerpo anti-CD34 (AMAC, Westbrook, ME) y esferas inmunomagnéticas (Dyna, Oslo, Noruega). El análisis citométrico de flujo que utiliza anticuerpos anti-CD34 conjugados a FITC (AMAC) confirmó una positividad CD34 de células aisladas del ~95%. La línea celular de hematoma, Hep3B, se cultivó en medio alfa (Gibco, Grand Island, NY) suplementado con penicilina (100U/mL), estreptomycin (100 µg/mL) y suero bovino fetal al 10% (Gibco) a 37°C en un incubador de CO₂ al 5%. El ARN total extraído de células mononucleares de médula ósea CD34+ o células Hep3B se transcribió de forma inversa con cebadores aleatorios y la transcriptasa inversa (RT) del virus de leucemia murina de Moloney siguiendo las condiciones del fabricante (Gibco-BRL) en una reacción de 20 µl. La PCR se realizó en el producto de reacción de RT en una reacción de 100 µl que contenía KCl 50 mM, Tris- HCl 10 mM (pH 8,4), NaCl 1,5 mM, gelatina 20 µg/ml, dNTPs 0,2 mM, 2,5 unidades de Taq polimerasa (Perkin-Elmer/Cetus) y 50 pmol de cada uno de los cebadores degenerados específicos de pTK

[pTK1 5'TCGGATCCACA/CGNGAC/TC/TTGGC 3' (SEC ID NO. 37),

20 pTK1B 5'TCGGATCCAC/TC/AGNGAC/TC/TTNGCNGC 3' (SEC ID NO. 38),

pTK2 5'CTCGAATTCCA/GA/TAA/GC/GT/ACCAG/CACA/GTC 3' (SEC ID NO. 39),

pTK2B 5'CTCGAATTCCA/GA/TAT/CC/GT/ACCAT/AACA/GTC 3'(SEC ID NO. 40)]

derivados de regiones de consenso entre pTK conocidas tal como se ha descrito por otros (Hanks et al., Science, 241:42-52 [1988]; Wilks, Proc. Nat. Acad. Sci.. USA 86:1603-1607 [1989]; y Matthews et al., Cell 65:1143-1152 [1991]). El ciclo PCR fue de 1,5 min a 95°C, 2 min a 37°C y 3 min a 63°C repetido 35 veces. El producto de reacción se separó electroforéticamente en un gel de agarosa de punto de fusión al 2%, se purificó en una columna Elutip-D (Schleicher & Schuell) digerida con EcoR1 y BamH1, y se subclonó en pUC19.

[0082] Se secuenciaron los recombinantes mediante el método didesoxi de Sanger y se evaluaron mediante el programa de análisis de secuencia de ácidos nucleicos FASTA. Se marcó radioactivamente un clon denominado HpTK5 (214 pb) mediante cebado aleatorio y se utilizó para cribar una biblioteca de ADN de lambda gt10 de Hep3B cebado con oligo dT. El ADN aislado 17 placas de fagos positivos y se subclonaron insertos en el sitio EcoRI de pBluescript (Stratagene La Jolla, CA). El inserto mayor, un ADNc de 3969 pb, se sonico hasta un tamaño promedio de 800-2000 pb y se clonó en el sitio SmaI de M13. Se secuenciaron clones solapantes utilizando el Método de Ciclo de Cebador de Colorante Taq (CABI) en la Catalyst 800 Molecular Biology Lab Station (ABI). A continuación, las reacciones de secuenciación se analizaron en el Secuenciador de ADN Automatizado ABI 373A.

[0083] Se aisló y secuenció un ADNc único de 3969 pb de longitud completa (figuras 8A-8F). El clon de longitud completa, denominado quinasa transmembrana de hepatoma (HTK) o HpTK5, incluía un marco de lectura abierto que se extendía del nucleótido 90 al 350 previsto para codificar una proteína de 987 aminoácidos de 108.270 Dalton. El codón de iniciación potencial está precedido por un codón de parada en el marco que empieza en la base 78. Antes del marco de lectura abierto se encuentra una región 5' no traducida que es rica en GC como es característico para muchos factores de crecimiento o receptores de factores de crecimiento (Kozak, J. Cell Biol. 115:887-903 [1991]).

[0084] La secuencia de proteína prevista incluye una región transmembrana (aminoácidos 538-563) que divide la HpTK5 en dominios extracelulares (ECD) e intracelulares (ICD). El ECD de 538 aminoácidos incluye un péptido señal de 15 aminoácidos y una caja rica en cisteína que contiene 20 residuos de Cys. Además, existen dos repeticiones de fibronectina tipo III que abarcan los aminoácidos 321 a 425 y 435 a 526. La Asn en las posiciones 208, 340 y 341 son posibles sitios para la N-glicosilación.

[0085] El dominio intracelular potencial (ICD) contiene una región consenso quinasa de la posición 613 a 881. Esta región quinasa incluye un consenso potencial de unión a ATP (Gly-X-Gly-X-X-Gly) en el subdominio I en las

posiciones 622-627. Una Lys en la posición 647 (subdominio II) corresponde a una Lys invariante entre tirosina quinasa que se cree que son críticas para la reacción de fosfotransferencia. Las regiones distintivas indicativas de especificidad de sustrato sugieren que la HpTK5 es una tirosina en lugar de una serina/treonina quinasa. Éstas incluyen la secuencia en las posiciones 740-745 en el subdominio VI y la secuencia en las posiciones 783-790 en el subdominio VIII. Los residuos de tirosina en las posiciones 601, 619 y 741 son posibles sustratos para actividad tirosina quinasa.

[0086] La secuencia de aminoácidos prevista de HpTK5 que se parece más a la de la subfamilia definida originalmente por EPH. El patrón de expresión de la subfamilia de EPH sugiere un papel en la diferenciación y desarrollo. En particular, la aparición de elementos neurales corresponde con la expresión de ciertos genes relacionados con EPH. Los receptores de la familia de EPH, Hek2 y Elk, son las pTK más estrechamente relacionadas con HpTK5. Comparten una identidad del 79,3 y 76,5% en el ICD, respectivamente, y una identidad del 45 y 42% en el CD, respectivamente.

B. Mapeo cromosómico de HpTK5

[0087] Se utilizaron ADN híbridos de células somáticas de un panel de 25 líneas de células humanas-hámster (Bios, New Haven, CN) para la localización cromosómica por PCR. Se eligieron dos grupos de cebadores de la región 3' no traducida de HpTK5. Se realizó la PCR con 250 ng de ADN y 50 pmol de cada uno de los cebadores 5' y 3', KCl 50 mM, MgCl₂ 1,5 mM, gelatina 20 µg/ml, dNTP 0,2 mM y 2,5 unidades de Taq polimerasa en un volumen final de 100 µl. Los ciclos de 94°C durante 30 s, 60°C durante 30 s y 72°C durante 30 s se repitieron 30 veces. Se sometió a electroforesis una parte de cada muestra (15 µl) a través de un gel de agarosa al 1,5%, se transfirió a una membrana de nylon y se hibridó a una sonda de ADNc de HpTK5 de longitud completa marcada con ³²P antes de una autorradiografía de 5 horas. Se anotaron los positivos y se compararon con un resumen de la matriz de material cromosómico humano presente en cada uno de los ADN híbridos de células somáticas.

[0088] La región 3' no traducida contiene de forma característica algunas, si están presentes, secuencias intermedias y presenta un grado elevado de diversidad entre los miembros de familias de genes haciéndola preferida en este tipo de análisis. Ambos grupos de cebadores produjeron resultados que eran concordantes con sólo el cromosoma 7 humano. El cromosoma 7 humano también incluye los genes para el receptor de EGF, el receptor del factor de crecimiento de hepatocitos (HGF), HGF, el factor de crecimiento derivado de plaquetas (PDGF) y la interleuquina-6. Las anomalías cariotípicas que implican este cromosoma son habituales entre las leucemias humanas, particularmente en leucemias mieloides agresivas que aparecen después de la radiación, quimioterapia con agente alquilante o una enfermedad mielodisplásica existente (Baer et al., Curr. Opin. Oncol. 4,:24-32 [1992]).

C. Transferencia Northern de HpTK5

[0089] Se sometió a electroforesis un ARN seleccionado de Poli-A a través de una agarosa al 1,2%, gel de formaldehído 2,2 M y se transfirió a un filtro de nylon. Se hibridaron filtros preparados u obtenidos comercialmente en formamida al 50% a 42°C a HpTK5 marcada con ³²P, gliceraldehído-3-fosfato deshidrogenasa (GAPDH) o inserciones de ADNc de actina y se lavaron bajo condiciones astringentes (lavado final: 0,1 x SSC, SDS al 0,2% SDS a 65°C). SSC es NaCl 0,15 M/Na₃-citrato 0,015M, pH 7,6. Las transferencias Northern de ARN de tejido fetal o adulto humano se obtuvieron de Clontech (Palo Alto, CA) y contenían 2 µg/carril de ARN seleccionado de poly A.

[0090] El análisis por transferencia Northern de tejidos fetales humanos reveló un único transcrito de ~4Kb en el corazón, pulmón, hígado y riñón, con una señal menor detectable en el cerebro. En tejido humano adulto, no se detectó señal en cerebro, mientras que la placenta presentaba una señal particularmente intensa seguida del riñón, hígado, pulmón y páncreas. El músculo esquelético y el corazón presentaban una intensidad de señal inferior.

[0091] También se analizó la expresión de HpTK5 en líneas celulares de tumor humano mediante análisis de transferencia Northern realizado tal como se ha descrito anteriormente. Las líneas celulares derivadas de hígado, mama (MCF 7), (Colo 205), pulmón (NCI 69), melanocito (HM-1) o cuello del útero (HeLa) presentaban una señal detectable de tamaño apropiado. El mensaje estaba presente en líneas celulares seleccionadas de origen hematopoyético. K562 (una célula mielóide primitiva con multipotencial), THP-1 (una célula monocitoide), U937 (una línea celular mielomonocítica), Hep3B (una línea celular de hepatocarcinoma humano), y CMK (de origen megacariocítico) fueron todas positivas para el mensaje de HpTK5, pero linfóide (H9, Jurkat, JH-1, Raji, Ramos) u otras células mieloides seleccionadas (KG-1 o KMT2) no presentaban un transcrito detectable mediante análisis Northern.

[0092] La expresión diferencial del transcrito de HpTK5 en cerebro fetal frente a cerebro adulto sugiere que la HPTK5 puede compartir, con otros miembros de la subfamilia de EPH, un papel en sucesos relacionados con el desarrollo neural. Sin embargo, a diferencia de algunos miembros de la subfamilia de EPH que se expresan exclusivamente en neuronas (Maisonpierre et al., supra), la HpTK5 se expresa ampliamente en otros tejidos. En particular, la HpTK5 se expresa en células hematopoyéticas que incluyen células progenitoras hematopoyéticas CD34+. La presencia del mensaje de HpTK5 en células hematopoyéticas tempranas y líneas celulares de linaje mielóide, pero no en líneas celulares derivadas de células linfoides, sugiere que la HpTK5 puede presentar una

expresión limitada al linaje.

EJEMPLO 3

PRODUCCIÓN DE ANTICUERPOS POLICLONALES PARA HPTK5

- [0093] Se construyó un gen de fusión de Fc de IgG1 humana y el dominio extracelular (ECD) de HpTK5 y se produjo una proteína de fusión tal como se ha descrito previamente (Bennett et al., J. Biol. Chem. 266:23060-23067 [1991]). Se generaron anticuerpos policlonaes en conejos blancos de Nueva Zelanda contra la proteína de fusión; se emulsionaron 4 µg en 100 µl de PBS con 100 µl de adyuvante de Freund (adyuvante completo para la inyección primaria y adyuvante incompleto para todos los refuerzos). Para la inmunización primaria y el primer refuerzo, se inyectó la proteína directamente en los nódulos linfáticos poplíteos (Sigel et al., Methods Enzymol. 93:3-12 [1983]).
- 10 Para los refuerzos posteriores, la proteína se inyectó en sitios subcutáneos e intramusculares. Se inyectaron 1,3 µg proteína/kg de peso corporal cada 3 semanas con sangrados tomados 1 y 2 semanas después de cada refuerzo. La especificidad de HpTK5 del suero de conejo inmunizado se valoró mediante análisis citométrico de flujo de células NIH3T3 transfectadas con HpTK5 de longitud completa o vector solo utilizando una dilución 1:200 de suero preinmune o suero de Fc de IgG anti-HpTK5. Se observaron desplazamientos significativos en los picos en varios clones que expresaban HpTK5 en comparación con los controles transfectantes de suero preinmune o vector solo.

EJEMPLO 4

UTILIDAD Y ACTIVIDAD AGONISTA DE ANTICUERPOS POLICLONALES PARA HPTK5

A. Construcción de la fusión FLAG-HpTK5

- [0094] Se ligaron oligonucleótidos solapantes que codifican un péptido de 12 aminoácidos que tiene la secuencia MDYKDDDDKKLAM (SEC ID NO: 41) que incluye el sitio de reconocimiento de anticuerpo de 4 aminoácidos "FLAG" (IBI, New Haven, CT) un sitio de restricción 5'-EcoRV y un sitio de restricción 3'-NcoI (5'-CCGGATATCATGGACTACAAGGACGACGATGACAAGAAGCTTGCCATGGAGCTC; SEC ID NO: 42), en el sitio NcoI (base 88) de HpTK5 en el vector Bluescript digerido con EcoRV (Stratagene, La Jolla, CA).

B. Transcripción y traducción *in vitro*

- 25 [0095] La transcripción se realiza en 2 pmol de PTK5 linealizado o plásmido que contenía FLAG-HpTK5 a 37°C durante 1 hora en un volumen de 50 µl que contenía ditiotreitol 10 mM, 2,5 µg de albúmina de suero bovino, 0,25 mM de cada dNTP, recubrimiento m7GRNA 0,5 M (New England Biolabs, Beverly, MA), 2,5 unidades de RNasin (Promega, Madison, WI), 3 unidades T3 ARN polimerasa (Pharmacia, Piscataway, NJ). Se añadió 1 µg de DNAasa (New England Biolabs, Beverly MA) durante 15 min a 37°C antes de la extracción con fenol/cloroformo y la precipitación con etanol. La traducción se realizó utilizando el kit de lisado de reticulocitos de conejo Promega según las instrucciones de fabricante con o sin el marcaje con ³⁵S-metionina (350 µCi). Se añadió tampón de muestra que contenía SDS y beta-mercaptoetanol 2-ME) antes de la ebullición y SDS-PAGE al 10%.

C. Expresión de HpTK5 en células NIH3T3

- [0096] Se subclonaron un fragmento de ADNc ClaI - XbaI de 4038 pb que contenía 23 pb de secuencia enlazadora, 37 pb de polienlazador pBluescript (Stratagene La Jolla, CA) y el ADNc de HpTK5 de 3969 pb completo en el vector de pRIS (Genentech, Inc.) bajo el control del promotor LTR del virus de sarcoma de Rous. Las células NIH3T3 mantenidas en un Medio de Eagle Modificado por Dulbecco con glucosa elevada (DMEM) suplementado con FCS al 10% se cotransfectaron con pRIS-HpTK5 y pNeo (un vector basado en Sv40 que contenía el marcador de resistencia a neomicina) mediante el método de fosfato de calcio descrito por Gorman et al., en DNA Prot. Engineer. Tech. 2:3-10 [1990]. Las colonias de resistencia a neomicina se seleccionaron 48 horas después de la transfección con Geneticina (Gibco/BRL) a 400 µg/ml. Catorce días después se aislaron las colonias resistentes individuales, se expandieron y analizaron mediante citometría de flujo para la expresión de HpTK5 utilizando antisuero policlonal de conejo.

D. Inmunoprecipitación

- 45 [0097] Se utilizaron células (Hep3B, control NIH3T3 o NIH3T3 transfectadas con HpTK5) o proteína traducida *in vitro* (HpTK5 o FLAGHpTK5) para la inmunoprecipitación con suero (pre-inmune o anti-HpTK5-IgG Fc) o anticuerpo monoclonal (específico de FLAG, M2, o control de isotipo) (IBI, Rochester, NY). Las células subconfluentes se marcaron con 200 µCi/ml de ³⁵S-metionina durante 18 horas y se lisaron en tampón de lisis (NaCl 150 mM, Tris-HCl 50 mM, pH 8,0, EDTA 1 mM, azida sódica 0,025, N-40 al 1%, SDS al 0,1%, glicerol al 10%, desoxicolato sódico al 0,5%, fluoruro de fenilmetilsulfonilo (PMSF) 1 mM, apronitina 10 µg/ml, leupeptina 10 µg/ml y vanadato de sodio 50 µM) durante 30 min en hielo. El lisado celular se centrifugó (12,000 X g) durante 10 min a 4°C. El sobrenadante del lisado celular o mezcla de traducción *in vitro* se preaclará con 0,05 volumen de suero de conejo normal y se adsorbió con 0,05 volumen de proteína A Sefarosa CL4B de Staphylococcus aureus. Después de la centrifugación, el suero preinmune o inmune (dilución 1:100), o anticuerpo monoclonal, se añadieron y agitaron durante la noche a 4°C antes de añadir 100 µl de proteína A Sefarosa CL4B y la solución se agitó a 4°C durante 2 h adicionales. Se lavaron los

inmunoprecipitados, se suspendieron en un tampón de carga de SDS/PAGE (glicerol al 10%, 2-ME al 5%, SDS al 2,3% y Tris-HCl 62,5 mM pH 6,8), se calentaron hasta 95°C durante 5 min y se analizaron mediante SDS-PAGE al 7,5%.

E. Fraccionamiento celular

- 5 [0098] El fraccionamiento celular de células Hep3B se realizó para confirmar la localización en la membrana de HpTK5 prevista mediante su secuencia de aminoácidos. Las células Hep-3B (1×10^7) se marcaron durante toda la noche con 200 μ Ci/ml de 35 S-metionina en un medio alfa MEM que contenía FCS dializado al 10%. Las células se lavaron dos veces con PBS frío, se extrajeron en 1 ml de tampón frío (Tris-HCl 20 mM pH 7,5, EDTA 2 mM, EGTA 5 mM, sacarosa 0,25 M, leupeptina al 0,01%, PMSF 4 mM, 2-ME 10 mM) y se alteraron mediante sonicación durante 10 40 segundos. Se centrifugaron los homogenatos completos a 12.000 X g durante 15 min, se aislaron los residuos nucleares y el sobrenadante decantado se centrifugó a 140.000 X g durante 40 min a 4°C para obtener residuos de membranas. El sobrenadante resultante se utilizó como la fracción citosólica (C). Las fracciones nucleares (N) y de membrana (M) se lavaron y disolvieron en tampón que contenía NP-40 al 0,5% antes de la inmunoprecipitación. Las fracciones c, N o M se inmunoprecipitaron con un suero anti-HpTK5 o pre-inmune (control), se sometieron a SDS-PAGE al 12% y se autorradiografiaron. La HpTK5 se segregó predominantemente con la fracción de membrana, aunque el material inmunoprecipitado era evidente en un grado menor en el citosol.

F. Ensayo de proteína quinasa

- [0099] Se lavaron los inmunoprecipitados una vez con tampón quinasa (Hepes 25 mM pH 7,4, DTT 1 mM, $MgCl_2$ 10 mM, $MnCl$ 10 mM) y se resuspendió en 40 μ l de tampón quinasa que contenía ATP no marcado o 10 μ Ci de 32 P-ATP (3000 Ci/mM). Después de una incubación de 10 minutos a 30°C, la reacción se detuvo mediante la adición de 40 μ l de 2 X tampón de muestra y la ebullición de las muestras durante 3 min antes de la electroforesis en gel SDS-PAGE al 8,0%. El gel secado se cubrió con 4 láminas de papel de aluminio para bloquear la autorradiografía de proteína marcada con 35 S y el gel se colocó bajo una película durante 5 horas durante la noche.

G. Transferencia Western y ensayo de fosfotirosina

- 25 [0100] Las proteínas se transfirieron electroforéticamente a una membrana de nitrocelulosa de 0,2 μ m (Bio-Rad) o una membrana de difluoruro de polivinilideno de 0,45 μ m (Millipore) en un tampón que contenía Tris-HCl 25 mM (pH 7,5), glicina 192 mM y metanol al 20% a 100 mA durante 2 horas. Los filtros se lavaron en TBS (Tris-HCl 10 mM pH 8,0, NaCl 150 mM), se bloquearon durante la noche mediante la incubación en TBST (TBS con Tween-20 al 0,05%) más BSA al 5%. Los filtros se lavaron cuatro veces durante 5 minutos cada uno en TBST y se incubaron durante 2 30 horas con anticuerpo anti-fosfotirosina 4G10 de UBI (dilución 1:1000 en TBST). Los filtros se lavaron cuatro veces durante 5 minutos cada uno en TBST y se incubaron durante 1 hora con el anticuerpo secundario anti-ratón marcado con fosfatasa alcalina (Promega) a una dilución 1:7500 en TBST. Después de lavarse cuatro veces, se reveló la transferencia durante 30-60 min en tampón AP (Tris-HCl 100 mM, NaCl 100 mM, $MgCl_2$ 5 mM) más sustratos BCIP, NBT

35 H. Ensayo de fosforilación inducido por anticuerpo

- [0101] Se analizó el anticuerpo de conejo para HptK5-Fc de IgG por su capacidad de inducir la fosforilación de HpTK5 en células NIH3T3 transfectadas con HpTK5. Las células se emplacaron a una densidad de 5×10^5 células/pocillo en una placa de 6 pocillos y, después de 24 horas, se privaron de suero durante 1 hora antes de la adición de suero pre-inmune o inmune a una dilución de 1:50 durante 30 minutos. A continuación, las células se lavaron en PBS y se lisaron en 2X tampón de muestra o tampón de lisis NP-40 tal como se ha descrito anteriormente. A continuación, los lisados crudos o lisados de células inmunoprecipitadas se separaron a través de un gradiente de SDS-PAGE al 4-12% y se analizaron mediante inmunotransferencia anti-fosfotirosina tal como se ha descrito anteriormente. Las células que expresaban HpTK5 se expusieron a antisuero y se separaron mediante SDS-PAGE con sin inmunoprecipitación. El gel electrotransferido se inmunotransfirió con anticuerpo anti-fosfotirosina. Se observó una mayor fosforilación de tirosina de HpTK5 después de la exposición a antisuero policlonal mostrando un efecto de tipo agonista de la unión a anticuerpo. La interacción de HpTK5 con un anticuerpo dirigido contra su ECD induce la fosforilación. Esto apoya además que el HpTK5 pueda servir como receptor para un ligando que desencadena la activación de quinasa. Los detalles del mecanismo de señalización de HpTK5 se pueden explorar adicionalmente utilizando antisuero como ligando sustituto.

50 I. Conclusiones

- [0102] Se expresó una proteína de fusión HpTK5-Fc de IgG, se purificó y utilizó para generar anti-suero de conejo que inmunoprecipitó una proteína de 120 kD de células Hep3B. La especificidad del antisuero se confirmó mediante inmunoprecipitación de ARN de HpTK5 traducido *in vitro* y células NIH3T3 transfectadas con HpTK5. Para determinar la capacidad funcional de HpTK5, se inmunoprecipitó la HpTK5 traducida *in vitro*, se expuso a 55 condiciones de quinasa y se inmunotransfirieron utilizando un anticuerpo monoclonal específico de fosfotirosina. Los datos obtenidos indicaron que la HpTK5 se fosforila en tirosina. Sin embargo, la presencia de otras bandas que aparecían sistemáticamente en la inmunoprecipitación marcada con 32 P sugirió que la proteína HpTK5 estaba

únicamente parcialmente purificada y, por lo tanto, no se podía concluir que la HpTK5 era enzimáticamente activa. Para superar este problema, se generó una construcción de fusión en la que se añadió un epítipo de 8 aminoácidos (FLAP) al extremo N-terminal de HpTK5. La fusión FLAG-HpTK5 se tradujo in vivo y se inmunoprecipitó con un anticuerpo monoclonal específico de FLAG dando lugar a una única proteína de tamaño adecuado (~120kD).

- 5 Cuando se sometió a las condiciones de quinasa en presencia de ³²P-ATP, la proteína de fusión HpTK5-FLAG se marcó en la tirosina confirmando la autofosforilación de tirosina la autofosforilación de tirosina y, de este modo, la función quinasa de HpTK5.

EJEMPLO 5

PRODUCCIÓN DE ANTICUERPOS MONOCLONALES PARA HPTK5

- 10 **[0103]** Se produjeron anticuerpos monoclonales anti-HpTK5 mediante la hiperinmunización intraperitoneal de ratones BALB/c con la proteína de fusión Fc de IgG₁ Humana-dominio extracelular (ECD) de HpTK5 (producida utilizando las técnicas descritas anteriormente) en adyuvante RIBI (RIBI ImmunoChem Research, Hamilton, MT) y la fusión de esplenocitos con la línea celular de mieloma de ratón X63-Ag8.653 (Kearney et al., J. Immunol. 123:1548-1550 [1979]). Se purificaron los anticuerpos del fluido ascítico utilizando proteína A Sefarosa (Repligen Corp., Cambridge, MA) y se establecieron procedimientos de cromatografía de afinidad (Goding, J.W., J. Immunol. Methods 20:241-253 [1978]).

- 20 **[0104]** Se cribaron anticuerpos monoclonales por su capacidad de unirse al antígeno de HpTK5. Comenzando en el día 15 después de la fusión, se recogieron los sobrenadantes del cultivo de las placas de fusión y se analizó su capacidad de "capturar" específicamente HpTK5-IgG. En este ensayo ELISA, se recubrió la IgG anti-ratón de cabra en placas de microtitulación de 96 pocillos. Se añadieron los sobrenadantes de cultivo (100 µl) a los pocillos y se unieron las IgG de ratón presentes mediante los anticuerpos IgG anti-ratón de cabra. Se lavaron las placas y se añadió HpTK5-IgG o CD4-IgG (100 µl a 6 nM). Se detectó inmunoadhesina "capturada" utilizando un conjugado de peroxidasa de rábano picante anti-hu de cabra (específico de Fc) y sustrato de ortofenilén diamina. Se determinó la cuantificación de catálisis de sustrato mediante la densidad óptica a 490 nm.

- 25 **[0105]** A continuación, se cribaron los anticuerpos agonistas para utilizar las técnicas descritas en el Ejemplo 6 siguiente. Se identificaron dos anticuerpos monoclonales agonistas, uno de los cuales fue depositado con la ATCC.

EJEMPLO 6

ACTIVIDAD AGONISTA DE ANTICUERPOS MONOCLONALES PARA HPTK5

- 30 **[0106]** Se evaluaron los anticuerpos monoclonales producidos utilizando las técnicas descritas en el Ejemplo 5 por su capacidad de inducir la fosforilación de HpTK5 en células NIH3T3 transfectadas con HpK5. Se pusieron en placas las células a una densidad de 5×10^5 células/pocillo en una placa de 6 pocillos y, después de 24 horas, se les privó de suero durante 1 hora antes de añadir suero pre-inmune o anticuerpo monoclonal anti-HpTK5 (se utilizaron medios de hibridoma acondicionados sin diluir) durante 30 minutos. A continuación, se lavaron las células en PBS y se lisaron en 2X de tampón de muestra o tampón de lisis NP-40 tal y como se ha descrito anteriormente. A
- 35 continuación, se separaron los lisados crudos o los lisados de células inmunoprecipitadas mediante gradiente de SDS-PAGE al 4-12% y se analizaron mediante inmunotransferencia de antifosfotirosina tal y como se ha descrito anteriormente. Se expusieron células que expresaban HpTK5 al anticuerpo monoclonal y se separaron mediante SDS-PAGE con o sin inmunoprecipitación. El gel electrotransferido se inmunotransfirió con anticuerpo anti-fosfotirosina. Se observó una mayor fosforilación de tirosina de HpTK5 después de la exposición a anticuerpos
- 40 monoclonales mostrando un efecto de tipo agonista de la unión a anticuerpo. Por consiguiente, la interacción de HpTK5 con un anticuerpo monoclonal dirigido contra su ECD es capaz de inducir la fosforilación del dominio quinasa del mismo.

EJEMPLO 7

PRODUCCIÓN DE ANTICUERPOS POLICLONALES PARA SAL-S1

- 45 **[0107]** Se construyó un gen de la fusión dominio extracelular (ECD) de SAL-S1-Fc de IgG1 humana y se produjo la proteína de fusión tal como se ha descrito previamente en Bennett et al., J. Biol. Chem. 266:23060-23067 [1991]. Brevemente, los cebadores de PCR otk 1.41.1 (SEC ID NO: 43) y otk 1.41.2 (SEC ID NO: 44) se utilizaron en la técnica de PCR utilizando el plásmido pRK5.tk1-1.1 (SEC ID NO: 45) que contenía el ácido nucleico de SAL-S1 como plantilla para crear un fragmento de ADN que, cuando se digiere con Sall/BstEII, generó un fragmento
- 50 Sall/BstEII de 155 pb. Este fragmento de 155 pb se combinó con un fragmento Sall/HindIII de 6839 pb aislado de pRKS.tk1-1.1 y un fragmento BstEII/HindIII de 719 pb aislado de pBSSK-CH2-CH3 (Bennett et al., supra). Estos fragmentos se ligaron para crear un plásmido pRK5.tk1.igl.1 (7713 pb de tamaño) que, cuando se transfectaron en células 293, se utilizaron para producir una proteína de fusión del dominio extracelular (ECD) de SAL-S1-Fc de IgG humana. La proteína de fusión se preparó y purificó tal como se ha descrito en Bennett et al., supra. Se generaron anticuerpos policlinales en conejos blancos hembra de Nueva Zelanda contra la proteína de fusión. Brevemente, se emulsionaron 12,5 µg de proteína de fusión en 0,625 ml de PBS con 0,625 ml de adyuvante de Freund (adyuvante completo para la inyección primaria y adyuvante incompleto para todos los refuerzos). La inyección primaria y todos

los refuerzos fueron intramusculares en dos sitios y subcutáneos en múltiples sitios. Los refuerzos se realizaron en intervalos de 3 semanas con sangrados tomados 1 y 2 semanas después de cada refuerzo. La especificidad de SAL-S1 del suero de conejo inmunizado se valoró mediante análisis citométrico de flujo de 293 (ATCC CRL 1593) y COS7 (ATCC CRL 1651) transfectadas con Sal-S1 de longitud completa o vector solo (ver a continuación) utilizando una dilución 1:200 de suero pre-inmune o suero anti-SAL-S1-Fc de IgG. Se observaron desplazamientos significativos de los picos en varios clones que expresan SAL-S1 en comparación con el suero preinmune o controles transfectantes de vector solo.

EJEMPLO 8

UTILIDAD Y ACTIVIDAD AGONISTA DE ANTICUERPOS POLICLONALES SAL-S1

10 A. Inmunoprecipitación

[0108] Se utilizaron células de control 293 y COS7 además de las células 293 y COS7 transfectadas con SAL-S1 para inmunoprecipitación con suero pre-inmune o anticuerpo policlonal anti-SAL-S1-Fc de IgG. Se transfectaron las células COS7 y 293 utilizando un procedimiento de CaPO₄ tal y como describe Gorman, C. DNA Cloning, Glover D. Ed., IRL Press, Oxford, vol2: 143-190 (1985). Para la expresión transitoria, se transfectaron células 293 tal y como describe Gearing et al. EMBO 8: 3667-3676 (1989). Se marcaron células subconfluentes con 200µCi/ml de ³⁵S-metionina durante 18 horas y se lisaron en tampón de lisis (NaCl 150 mM, HEPES 50mM, pH 7,5, EGTA 1 mM, azida de Na 0,025, Triton-X 100 al 1%, MgCl₂ 1,5mM, Glicerol al 10%, fluoruro de fenilmetilsulfonilo [PMSF] 1 mM, aprotinina 10 µg/ml, leupeptina 10 µg/ml y vanadato de Na 50 µM) durante 10 min en hielo. Se centrifugó el lisato celular (12.000 X g) durante 10 min a 4°C. Después de la centrifugación, se añadió anticuerpo preinmune o policlonal al sobrenadante y se giró durante 4 horas a 4°C antes de añadir 100 µl de proteína-A Sefarosa CL4B y la solución se agitó a 4°C durante 2 horas adicionales. Se lavaron los inmunoprecipitados, se suspendieron en tampón de carga SDS/PAGE (glicerol al 10%, 2-ME al 5%, SDS al 2,3% y Tris-HCl 62,5mM pH 6,8), se calentaron hasta 95°C durante 5 min y se analizaron mediante SDS-PAGE al 7,5%.

B. Transferencia western y ensayo de fosfotirosina

[0109] Se transfirieron electroforéticamente las proteínas a una membrana de nitrocelulosa de 0,2 µm (Bio-Rad) o una membrana de difluoruro de polivinilideno de 0,45 µm (Millipore) en un tampón que contenía Tris-HCl 25 mM (pH 7,5), glicina 192 mM y metanol al 20% a 100 mA durante 2 h. Se lavaron los filtros en TBS (Tris-HCl 10 mM pH 8,0, NaCl 150 mM), se bloquearon mediante incubación en TBST (TBS con Tween-20 al 0,05%) más BSA al 5% durante toda la noche. Se lavaron los filtros cuatro veces durante 5 min cada uno en TBST y se incubaron durante 2 h con anticuerpo 4G10 anti-fosfotirosina de UBI (dilución 1:1000 en TBST). Se lavaron los filtros cuatro veces durante 5 min cada uno en TBST y se incubaron durante 1 h con fosfatasa alcalina marcada con anticuerpo secundario anti-ratón (Promega) en una dilución 1:5000 en TBST. Después de lavar cuatro veces, se reveló la transferencia durante 30-60 min en tampón AP (Tris-HCl 100mM, NaCl 100 mM, MgCl₂ 5 mM) más sustratos BCIP, NBT.

C. Ensayo de fosforilación inducida por anticuerpo

[0110] Se evaluaron los antisueros de conejo para SAL-S1-Fc de IgG por su capacidad de inducir la fosforilación de SAL-S1 en células 293 transfectadas con SAL-S1. Se pusieron en placas las células a una densidad de 5 x 10⁵ células/pocillo en una placa de 6 pocillos y, después de 24 horas, se les privó de suero durante 12 horas antes de añadir suero pre-inmune o suero inmune en una dilución 1:5 durante 30 minutos. A continuación se lavaron las células en PBS y se lisaron en tampón de muestra o tampón de lisis Triton-X tal y como se describe anteriormente. A continuación, se separaron lisatos crudos o lisatos de células inmunoprecipitadas mediante gradiente de SDS-PAGE al 8% o al 4-12% y se analizaron mediante inmunotransferencia de antifosfotirosina tal y como se describe anteriormente. Se expusieron células que expresaban SAL-S1 a antisueros y se separaron mediante SDS-PAGE con o sin inmunoprecipitación. El gel electrotransferido se inmunotransfirió con anticuerpo anti-fosfotirosina. Se observó una mayor fosforilación de tirosina de SAL-S1 después de la exposición a antisuero policlonal mostrando un efecto de tipo agonista de la unión a anticuerpo. La interacción de SAL-S1 con un anticuerpo dirigido contra su ECD induce la fosforilación

EJEMPLO 9

PRODUCCIÓN DE ANTICUERPOS MONOCLONALES PARA SAL-S1

[0111] Se produjeron anticuerpos monoclonales anti-SAL-S1 mediante la hiperinmunización de ratones BALB/c en la base de la pata con la proteína de fusión Fc de IgG₁ humana-dominio extracelular de SAL-S1 en adyuvante de RIBI (RIBI Immunochem Research, Hamilton, MT) y la fusión de linfocitos de nódulos linfáticos con la línea celular de mieloma de ratón X63-Ag8U1.

[0112] Empezando en el día 10 después de la fusión, se recogieron los sobrenadantes cultivados de las placas de fusión y se ensayó su capacidad de unirse a SAL-S1. En este ensayo ELISA, se recubrió IgG₁ SAL-S1 en placas de microtitulación de 96 pocillos. Se añadieron los sobrenadantes cultivados (100 µl) a los pocillos y se unieron los anticuerpos de ratón presentes a IgG₁ SAL-S1. Se lavaron las placas y se detectó IgG de ratón utilizando un

conjugado de peroxidasa de rábano picante con IgG de cabra anti-ratón (Fc específico sin reactividad cruzada contra Fc de IgG humano) y sustrato de ortofenilén diamina. Se determinó la cuantificación de catálisis de sustrato mediante densidad óptica a 490 nm.

- 5 **[0113]** A continuación, se evaluaron los sobrenadantes cultivados que fueron positivos en ELISA por su capacidad de unirse específicamente a 293 transfectadas con receptor de SAL-S1 y se analizaron mediante citometría de flujo. A continuación, se cribaron los anticuerpos agonistas para utilizar las técnicas descritas en el Ejemplo 10 siguiente. Se identificaron seis anticuerpos monoclonales agonistas.

EJEMPLO 10

ACTIVIDAD AGONISTA DE ANTICUERPOS MONOCLONALES PARA SAL-S1

- 10 **[0114]** Se evaluaron los anticuerpos monoclonales por su capacidad de inducir la fosforilación de SAL-S1 en células 293 transfectadas con SAL-S1. Las células se recogieron de la placa de cultivo tisular mediante tampón de ensayo y se lavaron 2 veces con el mismo tampón. Se añadieron 1×10^5 células a una placa de base U de 96 pocillos que se centrifugó y se extrajo el tampón de ensayo. Se añadieron 150 μ l de sobrenadantes de cultivo a cada pocillo seguido de incubación a 37°C durante 30 minutos, la placa se centrifugó y se extrajeron los sobrenadantes de cultivo. Se
15 añadieron 100 μ l de solución de fijación, las células se fijaron durante 30 minutos a -20°C, las células se lavaron con tampón dos veces y se tiñeron con conjugado anti-fosfotirosina con FITC durante 60 minutos a 4°C. Las células se analizaron mediante citometría de flujo (FACScan Becton Dickinson, milpitas, CA). Los seis anticuerpos monoclonales anti-SAL-S1 eran capaces de inducir la fosforilación de SAL-S1 en células 293 transfectadas con SAL-S1
- 20 Depósito de materiales

[0115] Los siguientes materiales se han depositado con la American TIPO Culture Collection, 12301 Parklawn Drive, Rockville, MD, USA (ATCC):

Hibridoma	ATCC Dep. No.	Data del depósito
Anti-HpTK5	HB 11.583	15 de marzo de 1994

- 25 **[0116]** Este depósito se realizó según lo estipulado en el Tratado de Budapest sobre el Reconocimiento Internacional del Depósito de Microorganismos a los fines del Procedimiento en Materia de Patentes y el Reglamento bajo el mismo (Tratado de Budapest). Esto asegura el mantenimiento de un cultivo viable del depósito durante 30 años a partir de la fecha del depósito. El organismo estará disponible mediante la ATCC según los términos del Tratado de Budapest, y están sujetos a un acuerdo entre Genentech, Inc. y ATCC, que asegura la disponibilidad permanente y sin restricción de la progenie del cultivo del depósito al uso público tras la concesión de
30 la respectiva patente estadounidense o tras ponerse abierta a la inspección pública de cualquier solicitud de patente estadounidense o extranjera, la que sea primera, y asegura la disponibilidad de la progenie a la persona autorizada por el Comisionado de Estados Unidos de Patentes y Marcas de acuerdo con la norma 35 USC § 122 y las normas del Comisionado según lo acordado (incluyendo 37 CFR § 1.14 con particular referencia a 886 OG 638).
- 35 **[0117]** El cesionario de la presente solicitud ha acordado que si un cultivo de los materiales en el depósito muriera o se perdiera o se destruyera cuando se cultiva en las condiciones adecuadas, los materiales serán inmediatamente reemplazados en una notificación por otros iguales. La disponibilidad del material depositado no se interpreta como una licencia para realizar la invención contraviniendo los derechos concedidos bajo la autoridad de cualquier gobierno de acuerdo con sus leyes de patente.
- 40 **[0118]** La memoria escrita anterior se considera que es suficiente para permitir a un experto en la materia realizar la invención. La presente invención no se limita en su alcance por la construcción depositada, ya que la realización depositada pretende ser una ilustración individual de ciertos aspectos de la presente invención y cualquier cultivo que sea funcionalmente equivalente están dentro del alcance de la presente invención. El depósito del material de la presente invención no constituye una admisión de que la descripción escrita contenida en la presente invención sea inadecuada para permitir la práctica de cualquier aspecto de la invención, incluyendo el modo óptimo de la misma, ni
45 se interpreta como limitante del alcance de las reivindicaciones a las ilustraciones específicas que representa. De hecho, las diversas modificaciones además de las mostradas y descritas en la presente invención serán evidentes para los expertos en la materia a partir de la descripción anterior y caen dentro del alcance de las reivindicaciones adjuntas.

LISTADO DE SECUENCIAS

[0119]

(1) INFORMACIÓN GENERAL:

(i) SOLICITANTE: Genentech, Inc.

Bennett, Brian D.

5 Goeddel, David

Lee, James M.

Matthews, William

Tsai, Siao Ping

Wood, William I.

10 (ii) TÍTULO DE LA INVENCIÓN: ANTICUERPOS AGONISTAS DE PROTEÍNA TIROSINA QUINASA

(iii) NÚMERO DE SECUENCIAS: 45

(iv) DIRECCIÓN DE CORRESPONDENCIA:

(A) DESTINATARIO: Genentech, Inc.

(B) CALLE: 460 Point San Bruno Blvd

15 (C) CIUDAD: South San Francisco

(D) ESTADO: California

(E) PAÍS: USA

(F) CP: 94080

(v) FORMA DE LECTURA POR ORDENADOR:

20 (A) TIPO DE MEDIO: disquete de 5.25 pulgadas, 360 Kb

(B) ORDENADOR: IBM PC compatible

(C) SISTEMA OPERATIVO: PC-DOS/MS-DOS

(D) SOFTWARE: patin (Genentech)

(vi) DATOS DE LA PRESENTE SOLICITUD:

25 (A) NÚMERO DE SOLICITUD:

(B) FECHA DE SOLICITUD:

(C) CLASIFICACIÓN:

(vii) DATOS DE SOLICITUD ANTERIOR:

(A) NÚMERO DE SOLICITUD: 08/222616

30 (B) FECHA DE SOLICITUD: 04-APR-1994

(viii) INFORMACIÓN ABOGADO/AGENTE:

(A) NOMBRE: Wendy M. Lee

(B) NÚMERO DE REGISTRO: 00,000

(C) NÚMERO DE REFERENCIA/EXPEDIENTE: 821P3PCT

35 (ix) INFORMACIÓN DE COMUNICACIÓN:

(A) TELÉFONO: 415/225-1994

(B) TELEFAX: 415/952-9881

(C) TELEX: 910/371-7168

(2) INFORMACIÓN PARA LA SEC ID NO:1:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

5 (A) LONGITUD: 17 bases

(B) TIPO: ácido nucleico

(C) CADENA: única

(D) TOPOLOGÍA: lineal

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

10 CGGATCCACA GNGACCT 17

(2) INFORMACIÓN PARA LA SEC ID NO:2:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 23 bases

(B) TIPO: ácido nucleico

15 (C) CADENA: única

(D) TOPOLOGÍA: lineal

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

GGAATTCCAA AGGACCAGAC GTC 23

(2) INFORMACIÓN PARA LA SEC ID NO:3:

20 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 21 bases

(B) TIPO: ácido nucleico

(C) CADENA: única

(D) TOPOLOGÍA: lineal

25 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:3:

CGGATCCATC CACAGAGATG T 21

(2) INFORMACIÓN PARA LA SEC ID NO:4:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 26 bases

30 (B) TIPO: ácido nucleico

(C) CADENA: única

(D) TOPOLOGÍA: lineal

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

GGAATTCCTT CAGGAGCCAT CCACTT 26

35 (2) INFORMACIÓN PARA LA SEC ID NO:5:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 160 bases

(B) TIPO: ácido nucleico

(C) CADENA: única

(D) TOPOLOGÍA: lineal

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

5 **GGATCCTGTG CATCAGTGAC TTAGGGCTAG GAACATTCTG CTGTCCGAAA 50**

GCGACGTGGT GAAGATCTGT GACTTTGGCC TTGCCCGGGA CATCTACAAA 100

10 **GACCCCAGCT ACGTCCGCAA GCATGCCCGG CTGCCCTGA AGTGGATGGC 150**

GCCAGAATTC 160

(2) INFORMACIÓN PARA LA SEC ID NO:6:

15 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 53 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

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

20 **Asp Pro Val His Gln Xaa Leu Arg Ala Arg Asn Ile Leu Leu Ser**
1 5 10 15

Glu Ser Asp Val Val Lys Ile Cys Asp Phe Gly Leu Ala Arg Asp
20 25 30

25 **Ile Tyr Lys Asp Pro Ser Tyr Val Arg Lys His Ala Arg Leu Pro**
35 40 45

Leu Lys Trp Met Ala Pro Glu Phe
50 53

(2) INFORMACIÓN PARA LA SEC ID NO:7:

30 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 147 bases

(B) TIPO: ácido nucleico

(C) CADENA: única

(D) TOPOLOGÍA: lineal

35 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:7:

GGATCCATTC ACAGAGACCT AGCAGCACGC AACATCCTGG TCTCAGAGGA 50

CCTGGTAACC AAGGTCAGCG ACTTTGGCCT GGCCAAAGCC GAGCGGAAGG 100

5

GGCTAGACTC AAGCCGGCTG CCCGTCAAAT GGATGGCTCC CGAATTC 147

(2) INFORMACIÓN PARA LA SEC ID NO:8:

10 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 49 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

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

15

Gly Ser Ile His Arg Asp Leu Ala Ala Arg Asn Ile Leu Val Ser
 1 5 10 15

Glu Asp Leu Val Thr Lys Val Ser Asp Phe Gly Leu Ala Lys Ala
 20 25 30

20

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

Ala Pro Glu Phe
 49

25 (2) INFORMACIÓN PARA LA SEC ID NO:9:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 149 bases

(B) TIPO: ácido nucleico

(C) CADENA: única

30 (D) TOPOLOGÍA: lineal

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

GTTGGAATTC CTTCCGGCGC CATCCATTC ACCGGCAGCT TTATTCGTG 50

TCTAGATTCA TAGATGTCTT CATTATCTAC CTTAAAACT CTGGCAAGTC 100

5 CAAAATCTGC TACTTTGTAG ATATTATGTT CACCAACGAG GACATTCCT 149

(2) INFORMACIÓN PARA LA SEC ID NO:10:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 47 aminoácidos

10 (B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:10:

Val	Gly	Ile	Pro	Ser	Gly	Ala	Ile	His	Phe	Thr	Gly	Ser	Phe	Ile
1				5					10					15

15

Ser	Cys	Leu	Asp	Ser	Met	Ser	Ser	Leu	Ser	Thr	Leu	Lys	Thr	Leu
				20					25					30

20

Ala	Ser	Pro	Lys	Ser	Ala	Thr	Leu	Ile	Leu	Cys	Ser	Pro	Thr	Arg
				35					40					45

Thr	Phe
	47

(2) INFORMACIÓN PARA LA SEC ID NO:11:

25 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 151 bases

(B) TIPO: ácido nucleico

(C) CADENA: única

(D) TOPOLOGÍA: lineal

30 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:11:

GTGCACAGGG ATCTCGCGGC TCGGAACATC CTCGTCGGGG AAAACACCCT 50

CTCGAAAGTT GGGGACTTCG GGTTAGCCAG GCTTATCAAG GAGGACGTCT 100

5

ACCTCTCCCA TGACCACAAT ATCCCCTACA AATGGATGGC CCCTGAGGGA 150

A 151

(2) INFORMACIÓN PARA LA SEC ID NO:12:

10 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 50 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal (xi)

DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:12:

15

Val His Arg Asp Leu Ala Ala Arg Asn Ile Leu Val Gly Glu Asn
1 5 10 15

Thr Leu Ser Lys Val Gly Asp Phe Gly Leu Ala Arg Leu Ile Lys
20 25 30

20

Glu Asp Val Tyr Leu Ser His Asp His Asn Ile Pro Tyr Lys Trp
35 40 45

Met Ala Pro Glu Gly
50

25 (2) INFORMACIÓN PARA LA SEC ID NO:13:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 137 bases

(B) TIPO: ácido nucleico

(C) CADENA: única

30 (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:13:

GTTACCCGAG ATCTCAAGTC CAACAACATT TTGCTGCTGC AGCCCATTGA 50

GAGTGACGAC ATGGAGCACA AGACCCTGAA GATCACCGAC TTTGGCCTGG 100

CCCGAGAGTG GCACAAAACC ACACAAATGA GTGCCGC 137

5 (2) INFORMACIÓN PARA LA SEC ID NO:14:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 45 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

10 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:14:

Val	His	Arg	Asp	Leu	Lys	Ser	Asn	Asn	Ile	Leu	Leu	Leu	Gln	Pro
1				5					10					15

Ile	Glu	Ser	Asp	Asp	Met	Glu	His	Lys	Thr	Leu	Lys	Ile	Thr	Asp
			20						25					30

15

Phe	Gly	Leu	Ala	Arg	Glu	Trp	His	Lys	Thr	Thr	Gln	Met	Ser	Ala
			35						40					45

(2) INFORMACIÓN PARA LA SEC ID NO:15:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

20 (A) LONGITUD: 211 bases

(B) TIPO: ácido nucleico

(C) CADENA: única

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:15:

25

GTCAATCGTG ACCTCGCCGC CCGAAATGTG TTGCTAGTTA CCCAACATTA 50

CGCCAAGATC AGTGATTTCG GACTTTCCAA AGCACTGCGT GCTGATGAAA 100

ACTACTACAA GGCCCAGACC CATGGAAAGT GGCCTGTCAA GTGGTACGCT 150

30

CCGGAATGCA TCAACTACTA CAAGTTCTCC AGCAAAGCG ATGTCTGGTC 200

CTTTGGAATT C 211

(2) INFORMACIÓN PARA LA SEC ID NO:16:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 70 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:16:

5
Val Asn Arg Asp Leu Ala Ala Arg Asn Val Leu Leu Val Thr Gln
1 5 10 15

His Tyr Ala Lys Ile Ser Asp Phe Gly Leu Ser Lys Ala Leu Arg
20 25 30

10 **Ala Asp Glu Asn Tyr Tyr Lys Ala Gln Thr His Gly Lys Trp Pro**
35 40 45

Val Lys Trp Tyr Ala Pro Glu Cys Ile Asn Tyr Tyr Lys Phe Ser
50 55 60

15 **Ser Lys Ser Asp Val Trp Ser Phe Gly Ile**
65 70

(2) INFORMACIÓN PARA LA SEC ID NO:17:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 6827 bases

20 (B) TIPO: ácido nucleico

(C) CADENA: única

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:17:

TTCGAGCTCG CCCGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT 50

25 **TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC 100**

TTACGGTAAA TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG 150

30 **ACGTCAATAA TGACGTATGT TCCCATAGTA ACGCCAATAG GGACTTTCCA 200**

TTGACGTCAA TGGGTGGAGT ATTTACGGTA AACTGCCAC TTGGCAGTAC 250

ATCAAGTGTA TCATATGCCA AGTACGCCCC CTATTGACGT CAATGACGGT 300
 AAATGGCCCCG CCTGGCATTG TGCCAGTAC ATGACCTTAT GGGACTTTCC 350
 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC 400
 GGTTTTGGCA GTACATCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA 450
 TTTCCAAGTC TCCACCCCAT TGACGTCAAT GGGAGTTTGT TTTGGCACCA 500
 AAATCAACGG GACTTTCCAA AATGTCGTAA CAACTCCGCC CCATTGACGC 550
 AAATGGGCGG TAGGCGTGTA CGGTGGGAGG TCTATATAAG CAGAGCTCGT 600
 TTAGTGAACC GTCAGATCGC CTGGAGACGC CATCCACGCT GTTTTGACCT 650
 CCATAGAAGA CACCGGGACC GATCCAGCCT CCGCGGCCGG GAACGGTGCA 700
 TTGGAACGCG GATTCCCCGT GCCAAGAGTG ACGTAAGTAC CGCCTATAGA 750
 GTCTATAGGC CCACTTGGCT TCGTTAGAAC GCGGCTACAA TTAATACATA 800
 ACCTTATGTA TCATACACAT ACGATTTAGG TGACACTATA GAATAACATC 850
 CACTTTGCCT TTCTCTCCAC AGGTGTCCAC TCCCAGGTCC AACTGCACCT 900
 CGGTTCTATC GATTGAATTC CCCGGGGATC CTCTAGAGAT CCCTCGACCT 950
 CGAGATCCAT TGTGCTGGCG CGGATTCTTT ATCACTGATA AGTTGGTGGG 1000
 CATATTATGT TTATCAGTGA TAAAGTGTC AAGCATGACAA AGTTGCAGCC 1050
 GAATACAGTG ATCCGTGCCG CCCTAGACCT GTTGAACGAG GTCGGCGTAG 1100
 ACGGTCTGAC GACACGCAA CTGGCGGAAC GGTTGGGGGT TCAGCAGCCG 1150
 GCGCTTTACT GGCACCTCAG GAACAAGCGG GCGCTGCTCG ACGCACTGGC 1200

CGAAGCCATG CTGGCGGAGA ATCATAGCAC TTCGGTGCCG AGAGCCGACG 1250
 ACGACTGGCG CTCATTTCTG ACTGGGAATG CCCGCAGCTT CAGGCAGGCG 1300
 CTGCTCGCCT ACCGCCAGCA CAATGGATCT CGAGGGATCT TCCATACCTA 1350
 CCAGTTCTGC GCCTGCAGGT CGCGGCCGCA CTACTIONTTG ATGTATTACT 1400
 CATATTACCA AGGAATAACT GCGGGGCACA GGGTCAGGTG CTGAAGGGAC 1450
 ATTGTGAGAA GTGACCTAGA AGGCAAGAGG TGAGCCCTCT GTCACGCTGG 1500
 CATAAGGGCC GCTTGAGGGC TCTTTGGTCA AGCAGTAACG CCAGTGTCTG 1550
 GGAAGGCACC TGTACTCAG CAGACCATGA AAGGGCGTCT CCCTTTCCTT 1600
 GGAGCAGTCA GGAACACTC TGCTCCACCA GCTTCTTGTG GGAGCCTGGA 1650
 TATTATCCAG GCCTGCCCGC AGTCATCCGG AGGCCTAACC CCTCCCTGTG 1700
 GTGCTTCAGT GGTCACTC CTTGTCCACT TTCATGCTCC TCTTGGCCTC 1750
 CTGGTTCCTC TTGGAAGTTT GTAGTAGATA GCAGAAGAAA TAGCGAAAGT 1800
 CTTAAAGTCT TTGATCTTTC TTATAAGTGC AGAGAAGAAA TGCTGACGTA 1850
 TGCTGCCTTC TCTCTCTCTG CTTAGCTAC CTGAAGCCGC TTTCTTGTCT 1900
 ATACCTGCTC TCTATCTGCT CACTCCTC CGAGGCCAGC ACCATCCCAC 1950
 TGTCTGTCTG GTTGTCCACA GAGCCTTTGT AGGTCGTTGG GGTCATGGGG 2000
 AATTCCTCAA ATGTCTTCAT CCTGGAGGAA CCACGGGTCT CAGCCCCTCT 2050
 GGCCAGGCAC CCGGAAAGG ACACCCAGTT GTAATACCTG GCGGCCAGGC 2100
 TGTGGCGCTG CAGGCTTGGC GGGCTGTCCCT CAGCGTCAGC CTGGGCGATG 2150

TGTAGGGCCA TGGTGGACAC CTGCGAGAAG CTGCCCTCTT CTGAGCTCTG 2200
 AGAGCTGCGC GGGGCCATGC AGACCTCCTC TTCCTCTTGC AGGCCCTGC 2250
 CCTGGAGCAG GTCCCCCAGG ATCTCCACCA GCTCCGAGAA TGCAGGTCTC 2300
 GCCTTGGGGT CTCCGGACCA GCAGTTCAGC ATGATGCGGC GTATGGCGGG 2350
 AGTGGCCAGC TCCGGGGCCC TCATCCTTGT GCCGTCTCTC AGCCGCTGGC 2400
 AGAACTCCTC ATTGATCTGC ACCCCAGGGT ACGGGGAGGC CCCCAGAGAG 2450
 AAGATCTCCC AGAGAAGCAC CCCAAAGGAC CACACGTAC TCTGCGTGGT 2500
 GTACACCTTG TCGAAGATGC TTTCAGGGGC CATCCACTTC AGGGGCAGCC 2550
 GGGCACTGCC CTTGCGGACG TAGTCGGGGT CTTTGTAGAT GTCCCGGGCA 2600
 AGGCCAAAGT CACAGATCTT CACCACGTCG CTTTCCGACA GCAGAATGTT 2650
 CCGAGCAGCC AGGTCTCTGT GGATGCACTT TCGGGAAGCC AGGAACTCCA 2700
 TCCCTCTGGC CACCTGGAAG CTGTAGCAGA CAAGATCTTC CATGGTCAGC 2750
 GGGCTCAGCC ACAGGTCCTC AGCTTCTTGG TCTGGAGAAG CCCGCCTCGC 2800
 TCCGCCCTCG GTCTTCGAGA ACCGCGCGAA GAGGACCCTG TCGCTGCTCC 2850
 CCGGCCGCCT CCGATCCAGC CTGGCGAGCT CCACCATGGC GCGGAAGCGT 2900
 CCGCGCTGCT CGGGAGACTT CTCCTGCGGA TGCACGAAGC TGGCTCGAGG 2950
 GCGCCCAGTC GTCCGCCGCA GAGGCGCCTC CATTCCCCCG CCGCCC GCG 3000
 CGCCCCGAG GCCGCCGCT CACCGNGCAG GGGCTGCGGC CGCGACTCTA 3050
 GAGTCGACCT GCAGAAGCTT GGCCGCCATG GCCCAACTTG TTTATTGCAG 3100

CTTATAATGG TTACAAATAA AGCAATAGCA TCACAAATTT CACAAATAAA 3150
 GCATTTTTTT CACTGCATTC TAGTTGTGGT TTGTCCAAAC TCATCAATGT 3200
 ATCTTATCAT GTCTGGATCG ATCGGGAATT AATTCGGCGC AGCACCATGG 3250
 CCTGAAATAA CCTCTGAAAG AGGAACTTGG TTAGGTACCT TCTGAGGCGG 3300
 AAAGAACCAG CTGTGGAATG TGTGTCAGTT AGGGTGTGGA AAGTCCCCAG 3350
 GCTCCCCAGC AGGCAGAAGT ATGCAAAGCA TGCATCTCAA TTAGTCAGCA 3400
 ACCAGGTGTG GAAAGTCCCC AGGCTCCCCA GCAGGCAGAA GTATGCAAAG 3450
 CATGCATCTC AATTAGTCAG CAACCATAGT CCCGCCCTA ACTCCGCCCA 3500
 TCCCGCCCCT AACTCCGCC AGTTCGCCCC ATTCTCCGCC CCATGGCTGA 3550
 CTAATTTTTT TTATTATGC AGAGGCCGAG GCCGCCTCGG CCTCTGAGCT 3600
 ATTCCAGAAG TAGTGAGGAG GCTTTTTTGG AGGCCTAGGC TTTTGCAAAA 3650
 AGCTGTTAAC AGCTTGGCAC TGGCCGTCGT TTTACAACGT CGTGACTION 3700
 AAAACCCTGG CGTTACCCAA CTTAATCGCC TTGCAGCACA TCCCCCCTTC 3750
 GCCAGCTGGC GTAATAGCGA AGAGGCCCGC ACCGATCGCC CTTCCCAACA 3800
 GTTGCGTAGC CTGAATGGCG AATGGCGCCT GATGCGGTAT TTTCTCCTTA 3850
 CGCATCTGTG CGGTATTTCA CACCGCATA GTCAAAGCAA CCATAGTACG 3900
 CGCCCTGTAG CGGCGCATT AGCGCGGCGG GTGTGGTGGT TACGCGCAGC 3950
 GTGACCCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT TCGCTTTCTT 4000
 CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC 4050

GGGGGCTCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC 4100
AAAAAACTTG ATTTGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA 4150
GACGGTTTTT CGCCCTTTGA CGTTGGAGTC CACGTTCTTT AATAGTGGAC 4200
TCTTGTTCCA AACTGGAACA ACACTCAACC CTATCTCGGG CTATTCTTTT 4250
GATTTATAAG GGATTTTGCC GATTTCTGGCC TATTGGTTAA AAAATGAGCT 4300
GATTTAACAA AAATTTAACG CGAATTTTAA CAAAATATTA ACGTTTACAA 4350
TTTTATGGTG CACTCTCAGT ACAATCTGCT CTGATGCCGC ATAGTTAAGC 4400
CAACTCCGCT ATCGCTACGT GACTGGGTCA TGGCTGCGCC CCGACACCCG 4450
CCAACACCCG CTGACGCGCC CTGACGGGCT TGTCTGCTCC CGGCATCCGC 4500
TTACAGACAA GCTGTGACCG TCTCCGGGAG CTGCATGTGT CAGAGGTTTT 4550
CACCGTCATC ACCGAAACGC GCGAGGCAGT ATTCTTGAAG ACGAAAGGGC 4600
CTCGTGATAC GCCTATTTTT ATAGGTTAAT GTCATGATAA TAATGGTTTC 4650
TTAGACGTCA GGTGGCACTT TTCGGGGAAA TGTGCGCGGA ACCCCTATTT 4700
GTTTATTTTT CTAAATACAT TCAAATATGT ATCCGCTCAT GAGACAATAA 4750
CCCTGATAAA TCTTCAATAA TATTGAAAAA GGAAGAGTAT GAGTATTCAA 4800
ACATTTCCGT GTCGCCCTTA TTCCCTTTTT GCGGCATTT TGCCTTCCTG 4850
TTTTTGCTCA CCCAGAAACG CTGGTAAAAG TAAAAGATGC TGAAGATCAG 4900
TTGGGTGCAC GAGTGGGTTA CATCGAACTG GATCTCAACA GCGGTAAGAT 4950
CCTTGAGAGT TTTCGCCCCG AAGAACGTTT TCCAATGATG AGCACTTTTA 5000

AAGTTCTGCT ATGTGGCGCG GTATTATCCC GTGATGACGC CGGGCAAGAG 5050
 CAACTCGGTC GCCGCATACA CTATTCTCAG AATGACTTGG TTGAGTACTC 5100
 ACCAGTCACA GAAAAGCATC TTACGGATGG CATGACAGTA AGAGAATTAT 5150
 GCAGTGCTGC CATAACCATG AGTGATAACA CTGCGGCCAA CTTACTTCTG 5200
 ACAACGATCG GAGGACCGAA GGAGCTAACC GCTTTTTTGC ACAACATGGG 5250
 GGATCATGTA ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA 5300
 TACCAAACGA CGAGCGTGAC ACCACGATGC CAGCAGCAAT GGCAACAACG 5350
 TTGCGCAAAC TATTAAGTGG CGAACTACTT ACTCTAGCTT CCCGGCAACA 5400
 ATTAATAGAC TGGATGGAGG CGGATAAAGT TGCAGGACCA CTTCTGCGCT 5450
 CGGCCCTTCC GGCTGGCTGG TTTATTGCTG ATAAATCTGG AGCCGGTGAG 5500
 CGTGGGTCTC GCGGTATCAT TGCAGCACTG GGGCCAGATG GTAAGCCCTC 5550
 CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC 5600
 GAAATAGACA GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA 5650
 CTGTCAGACC AAGTTTACTC ATATATACTT TAGATTGATT TAAAACTTCA 5700
 TTTTAAATTT AAAAGGATCT AGGTGAAGAT CCTTTTTGAT AATCTCATGA 5750
 CCAAATCCC TTAACGTGAG TTTTCGTTC ACTGAGCGTC AGACCCCGTA 5800
 GAAAAGATCA AAGGATCTTC TTGAGATCCT TTTTTTCTGC GCGTAATCTG 5850
 CTGCTTGCAA ACAAAAAAAC CACCGCTACC AGCGGTGGTT TGTTTGCCGG 5900
 ATCAAGAGCT ACCAACTCTT TTTCCGAAGG TAACTGGCTT CAGCAGAGCG 5950

CAGATACCAA ATACTGTCCT TCTAGTGTAG CCGTAGTTAG GCCACCACTT 6000
 CAAGAACTCT GTAGCACCGC CTACATACCT CGCTCTGCTA ATCCTGTTAC 6050
 CAGTGGCTGC TGCCAGTGGC GATAAGTCGT GTCTTACCGG GTTGGACTCA 6100
 AGACGATAGT TACCGGATAA GGCAGCAGCGG TCGGGCTGAA CCGGGGGTTC 6150
 GTGCACACAG CCCAGCTTGG AGCGAACGAC CTACACCGAA CTGAGATACC 6200
 TACAGCGTGA GCATTGAGAA AGCGCCACGC TTCCCGAAGG GAGAAAGGCG 6250
 GACAGGTATC CGGTAAGCGG CAGGGTCGGA ACAGGAGAGC GCACGAGGGA 6300
 GCTTCCAGGG GGAAACGCCT GGTATCTTTA TAGTCCTGTC GGGTTTCGCC 6350
 ACCTCTGACT TGAGCGTCGA TTTTGTGAT GCTCGTCAGG GGGGCGGAGC 6400
 CTATGGAAAA ACGCCAGCAA CGCGGCCTTT TTACGGTTCC TGGCCTTTTG 6450
 CTGGCCTTTT GCTCACATGT TCTTTCCTGC GTTATCCCCT GATTCTGTGG 6500
 ATAACCGTAT TACCGCCTTT GAGTGAGCTG ATACCGCTCG CCGCAGCCGA 6550
 ACGACCGAGC GCAGCGAGTC AGTGAGCGAG GAAGCGGAAG AGCGCCCAAT 6600
 ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCATTAA TCCAGCTGGC 6650
 ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT 6700
 GTGAGTTACC TCACTCATTA GGCACCCCAG GCTTTACTACT TTATGCTTCC 6750
 GGCTCGTATG TTGTGTGGAA TTGTGAGCGG ATAACAATTT CACACAGGAA 6800
 ACAGCTATGA CCATGATTAC GAATTAA 6827

(2) INFORMACIÓN PARA LA SEC ID NO:18:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 348 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

5 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:18:

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

	Ala	Leu	His	Ile	Ala	Gln	Ala	Asp	Ala	Glu	Asp	Ser	Pro	Pro	Ser
					260					265					270
5	Leu	Gln	Arg	His	Ser	Leu	Ala	Ala	Arg	Tyr	Tyr	Asn	Trp	Val	Ser
					275					280					285
	Phe	Pro	Gly	Cys	Leu	Ala	Arg	Gly	Ala	Glu	Thr	Arg	Gly	Ser	Ser
					290					295					300
	Arg	Met	Lys	Thr	Phe	Glu	Glu	Phe	Pro	Met	Thr	Pro	Thr	Thr	Tyr
					305					310					315
10	Lys	Gly	Ser	Val	Asp	Asn	Gln	Thr	Asp	Ser	Gly	Met	Val	Leu	Ala
					320					325					330
	Ser	Glu	Glu	Cys	Glu	Gln	Ile	Glu	Ser	Arg	Tyr	Arg	Gln	Glu	Ser
					335					340					345
15	Gly	Phe	Arg												
					348										

(2) INFORMACIÓN PARA LA SEC ID NO:19:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 7607 bases

20 (B) TIPO: ácido nucleico

(C) CADENA: única

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:19:

TTCGAGCTCG CCCGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT 50
TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC 100
TTACGGTAAA TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG 150
ACGTCAATAA TGACGTATGT TCCCATAGTA ACGCCAATAG GGACTTTCCA 200
TTGACGTCAA TGGGTGGAGT ATTTACGGTA AACTGCCCCAC TTGGCAGTAC 250
ATCAAGTGTA TCATATGCCA AGTACGCCCC CTATTGACGT CAATGACGGT 300
AAATGGCCCG CCTGGCATTG TGCCCAGTAC ATGACCTTAT GGGACTTTCC 350
TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC 400
GGTTTTGGCA GTACATCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA 450

TTTCCAAGTC TCCACCCCAT TGACGTCAAT GGGAGTTTGT TTTGGCACCA 500
 AAATCAACGG GACTTTCCAA AATGTCGTAA CAACTCCGCC CCATTGACGC 550
 AAATGGGCGG TAGGCGTGTA CGGTGGGAGG TCTATATAAG CAGAGCTCGT 600
 TTAGTGAACC GTCAGATCGC CTGGAGACGC CATCCACGCT GTTTTGACCT 650
 CCATAGAAGA CACCGGGACC GATCCAGCCT CCGCGGCCGG GAACGGTGCA 700
 TTGGAACGCG GATTCCCCGT GCCAAGAGTG ACGTAAGTAC CGCCTATAGA 750
 GTCTATAGGC CCACTTGGCT TCGTTAGAAC GCGGCTACAA TTAATACATA 800
 ACCTTATGTA TCATACACAT ACGATTTAGG TGACACTATA GAATAACATC 850
 CACTTTGCCT TTCTCTCCAC AGGTGTCCAC TCCCAGGTCC AACTGCACCT 900
 CGGTTCCTATC GATTGAATTC CCCGGGGATC CTCTAGAGAT CCCTCGACCT 950
 CGAGTCGACT TTTTTTTTTT TTTTGTAGG CCAAAGGGTA CTTCTTTTTC 1000
 TTTATTAATT ACTCAGAAGT CTAGGCCACA GCAATCTACT GTTCTCCTCT 1050

 CATTTTCCTA AACTATTTTG ATACCTATTT CTCAGACTTT ATGGGCTATT 1100
 AGACATTTCT CACATTTCCA TAGATAATAA CTCATCCGTT TTGCAACCTG 1150
 ATTCTCAATA TTAAGAGATT AAAACTAATG TATATGACTC TCAGTTGACA 1200
 CATACTGAAG TACAGAAAAA TTCCATCATT TCCTTCTGCA AAATGAAAAA 1250
 GACTTCGTTT TCTCAACAGC TGCATCATTT TTTTATGCAT AGAAAAAAT 1300
 GTGCAATTAC TCCAAGTACA ATCAAGTCAT TTAACATGGC TTTACCATCA 1350
 TTGTAGTTAC AGGATATTTT AAAAGAGAAA AAAAAATCTC AAAGCACAGG 1400

TCCTGCTGTG CAGCAAAGCA ATCAAATTCC TTCATAATAA CAGCCTGATG 1450
 GGATTCAGCA ATCTGAGGAA TAATGAATAA CCACTCTAAT CAGTAAACAG 1500
 GAAAATGCTA CAACAGTCAC TGAGTAAAAA TTGGACTATC ATCTGTTGAT 1550
 TCTCTTGATC GACATTTCAA ACAATAAATG GAAATGTAAG TATCTCTTAA 1600
 AAAGAAAAAT AACTTGGTTT AGTGTGCTTA ATTTTACCAG GCAGTGAGGA 1650
 AATTATATAT CACCTTGACT GTCCTGCAGT GTTGCCAGT CAATAAAATG 1700
 CACAAATAAT CTTTTTCATA ATACATGGCC AACTTTATCC TATCACTTGA 1750
 ATATGTCAGG ATAAACTGAT TGTGCAGTTG GTTGATAACA TTGTATTTTG 1800
 GAATGGATTA TTGAATTTG TTTTGCTACT TTATTATTTG ATATTCTTCT 1850
 CCAGTGTTCA TCTTATGAAG TTATTTGCAT CTGAATATGA AGAGTCTGTT 1900
 TCAAAATAGT CTTCAAGTTT CCAACGCAGT GTCTCAAATG TAGGTCGTTT 1950
 CTTAGGCTCT GCATTCCAGC ACTCCAACAT GATGTTGTAA AATTGCTGTG 2000
 GACAGTTGGA TGGTTGCGGA AGTCTATAGT TTTGAGCCAA CATCTGGATT 2050
 ACCTGGGCAC CTGTCATACC ACTGTAAGGC ATTTTGCCAT AAGTAATGAT 2100
 TTCATAAAGA AGGATTCCAA ATGACCATAC ATCGGACTTA ATGCTGAATT 2150
 TATTACTACG AATGGCTTCG GCGCAGTCC ACTTCACCGG CAGCTTTATT 2200
 TCGTGTCTAG ATTCATAGAT GTCTTCATTA TCTACCTTAA AACTCTGGC 2250
 AAGTCCAAAA TCTGCTACTT TGTAGATATT ATGTTACCA ACGAGGACAT 2300
 TTCTGGCAGC CAGATCTCTG TGAATGTAGT TCCGAGACTC CAGATAGGCC 2350

ATTCCAGAGG CAACCTGTGC CGCCATGTCT ACCTGTTGAG TCAGATGGAT 2400
 TTTTGATCCA GTGTCATTTT GGAGATATTC TTGCAGACTT CCATGTCTCA 2450
 TCAACTCTGT AATAATATAA ATTGGATCTT CTAAAGTGCA AACAGCATAA 2500
 AGCTGGATAA GCTTTGGATG TCTTAGGTTT TTCATTATCT GTGCCTCCCT 2550
 CAGGAAGTCA TTTGGATCCA TTGAACCTGG TTTTAATGTT TTCACTGCTA 2600
 CTGGAGTGGT ATTGTTCCAC AGACCTTCCC ATACTTCGCC AAACTGACCA 2650
 GATCCCAATC GCTTCAGAAG CTGTATGGAG TTGCGGTCTA TCTCCCATG 2700
 GTCCACGGTT TTATACGACA AATCAAATGG AGCTGGGACC TGGATCTTTA 2750
 AGCATGGTTT CCCCAGCTTG ACACACAGGC CGTCACTTGT CTTGGTGTAG 2800
 TGGCTCACAA ATTCGTTTCTG TGTTGAAAAG ATTCTTCTTC GCGTGAGAAA 2850
 AAATCCCCCT TCATCCAGTC TTTTAATTCT GTAGTGTTTT ACAACTGCTC 2900
 CATCTAAAAC TGAAAGAGAG AATTCTCCTT TTTGGCTTTC ACTTTCTCTG 2950
 ATTAGAAAGG AACCGGTCTT GTTTTCTGAA TATAATAGTT GTTTCTCTGC 3000
 ATCTGATCTT CCGATTGCTC CAAAGAACCA CGGCTCTGCC TGTAGGCTTC 3050
 TGTCCTCAGC CACGTAGTTA GAAGGAATAT AGCCTTGTAG TTGCTGACTG 3100
 GAGCCATCTC GTCTTTTCTC CAAGTGTCTG GCAAACCACC AGCCCTCATG 3150
 CAAAGTGTCC AGAACTTGAA GTTTGTCCACC TGCTCGGAAG CTCAAGTCCT 3200
 CAGCAGTCCG AGCCTGGTAA TCAAACAAAG CCACAAAGTA GTGGCCATGC 3250
 CTCTGTGACT GGGGAGAGCA AAGGGCCCCT GGATTTTCAA TCACGGTTGA 3300

CTTGTCTGCC TCCGTGGACA AACAGGGGAG ATAGGGTTCT AGGTACTCCC 3350
 AGAGCCTCTG ACAGATGTTG CTCATTGTGC CTTGGTGGGG AGAAGAGGAG 3400
 CAGGGCTTCT CCCTCTCCCC TTAGTCTCTG CGATCCACCT TATCTTCCTT 3450
 CACCAGGCAA CTTTGAAGTC AGCACCAACT CACCATACTT CGGAGAGTAT 3500
 GCAAAGTCCC GTTTCAGATC AGTCCAGCAG CTGGGTTGCA GCAAGTCCTA 3550
 CCTGGAGAGA CTTACCGGCT TGCTTTCTGT GGCTGGAGGT GCTACCCCGA 3600
 GGCAAACTG AGCAGGAGCT GGGCAGCTGC TCACTAGGAA GGTGTCTTTT 3650
 CTTCTTATCT GCTTAAGAAT CCCACAACAA AAATAAAATA AAATTAAGAAG 3700
 GGCTTTATTT AGACAAATAT CTGAGAACAG AATGGTGCCA TCTTGCCTTT 3750
 TGTCCCAATA AAAAGTTAGC AAGAGGAAGC TACTAACCCC TGGTAAAACC 3800
 TCCACGTCTT GCTTTCGCCA GGGTCGACTC GAGGGATCTT CCATACCTAC 3850
 CAGTTCGCG CCTGCAGGTC GCGGCCGCGA CTCTAGAGTC GACCTGCAGA 3900
 AGCTTGCCG CCATGGCCCA ACTTGTTTAT TGCAGCTTAT AATGGTTACA 3950
 AATAAAGCAA TAGCATCACA AATTTACAAA ATAAAGCATT TTTTTCACTG 4000
 CATCTAGTT GTGGTTTGTC CAAACTCATC AATGTATCTT ATCATGTCTG 4050
 GATCGGGAAT TAATTCGGCG CAGCACCATG GCCTGAAATA ACCTCTGAAA 4100
 GAGGAACTTG GTTAGGTACC TTCTGAGGCG GAAAGAACCA GCTGTGGAAT 4150
 GTGTGTCAGT TAGGGTGTGG AAAGTCCCCA GGCTCCCCAG CAGGCAGAAG 4200
 TATGCAAAGC ATGCATCTCA ATTAGTCAGC AACCAGGTGT GGAAAGTCCC 4250

CAGGCTCCCC AGCAGGCAGA AGTATGCAAA GCATGCATCT CAATTAGTCA 4300
 GCAACCATAG TCCCGCCCCT AACTCCGCCC ATCCCGCCCC TAACTCCGCC 4350
 CAGTTCCGCC CATTCTCCGC CCCATGGCTG ACTAATTTTT TTTATTTATG 4400
 CAGAGGCCGA GGCCGCCTCG GCCTCTGAGC TATTCCAGAA GTAGTGAGGA 4450
 GGCTTTTTTG GAGGCCTAGG CTTTGGCAAA AAGCTGTTAA CAGCTTGGCA 4500
 CTGGCCGTCG TTTTACAACG TCGTGACTGG GAAAACCCTG GCGTTACCCA 4550
 ACTTAATCGC CTTGCAGCAC ATCCCCCTTT CGCCAGCTGG CGTAATAGCG 4600
 AAGAGGCCCG CACCGATCGC CCTTCCCAAC AGTTGCGCAG CCTGAATGGC 4650
 GAATGGCGCC TGATGCGGTA TTTTCTCCTT ACGCATCTGT GCGGTATTTT 4700
 ACACCGCATA CGTCAAAGCA ACCATAGTAC GCGCCCTGTA GCGGCGCATT 4750
 AAGCGCGGCG GGTGTGGTGG TTACGCGCAG CGTGACCGCT ACACTTGCCA 4800
 GCGCCCTAGC GCCCGCTCCT TTCGCTTCTT TCCCTTCCTT TCTCGCCACG 4850
 TTCGCCGGCT TTCCCCGTCA AGCTCTAAAT CGGGGGCTCC CTTTAGGGTT 4900
 CCGATTTAGT GCTTTACGGC ACCTCGACCC CAAAAAATT GATTTGGGTG 4950
 ATGGTTCACG TAGTGGGCCA TCGCCCTGAT AGACGGTTTT TCGCCCTTTG 5000
 ACGTTGGAGT CCACGTTCTT TAATAGTGGA CTCTTGTTCC AAACCTGGAAC 5050
 AACACTCAAC CCTATCTCGG GCTATTCTTT TGATTTATAA GGGATTTTGC 5100
 CGATTTCCGC CTATTGGTTA AAAAATGAGC TGATTTAACA AAAATTTAAC 5150
 GCGAATTTTA ACAAATATT AACGTTTACA ATTTTATGGT GCACTCTCAG 5200

TACAATCTGC TCTGATGCCG CATAGTTAAG CCAGCCCCGA CACCCGCCAA 5250
CACCCGCTGA CGCGCCCTGA CGGGCTTGTC TGCTCCCGGC ATCCGCTTAC 5300
AGACAAGCTG TGACCGTCTC CGGGAGCTGC ATGTGTCAGA GGTTCACACC 5350
GTCATCACCG AAACGCGCGA GACGAAAGGG CCTCGTGATA CGCCTATTTT 5400
TATAGGTAA TGTCATGATA ATAATGGTTT CTTAGACGTC AGGTGGCACT 5450
TTTCGGGGAA ATGTGCGCGG AACCCCTATT TGTTTATTTT TCTAAATACA 5500
TTCAAATATG TATCCGCTCA TGAGACAATA ACCCTGATAA ATGCTTCAAT 5550
AATATTGAAA AAGGAAGAGT ATGAGTATTC AACATTTCCG TGTCGCCCTT 5600
ATTCCCTTTT TTGCGGCATT TTGCCTTCCT GTTTTGTCTC ACCCAGAAAC 5650
GCTGGTGAAA GTAAAAGATG CTGAAGATCA GTTGGGTGCA CGAGTGGGTT 5700
ACATCGAACT GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC 5750
GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC 5800
GGTATTATCC CGTATTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC 5850
ACTATTCTCA GAATGACTTG GTTGAGTACT CACCAGTCAC AGAAAAGCAT 5900
CTTACGGATG GCATGACAGT AAGAGAATTA TGCAGTGCTG CCATAACCAT 5950
GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC GGAGGACCGA 6000
AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT 6050
GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA 6100
CACCACGATG CCTGTAGCAA TGGCAACAAC GTTGCGCAAA CTATTAAGT 6150

ES 2 355 952 T3

GCGAACTACT TACTCTAGCT TCCC GGCAAC AATTAATAGA CTGGATGGAG 6200
 GCGGATAAAG TTGCAGGACC ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG 6250
 GTTTATTGCT GATAAATCTG GAGCCGGTGA GCGTGGGTCT CGCGGTATCA 6300
 TTGCAGCACT GGGGCCAGAT GGTAAGCCCT CCCGTATCGT AGTTATCTAC 6350
 ACGACGGGGA GTCAGGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA 6400
 GATAGGTGCC TCACTGATTA AGCATTGGTA ACTGTCAGAC CAAGTTTACT 6450
 CATATATACT TTAGATTGAT TTAAAACCTC ATTTTTAATT TAAAAGGATC 6500
 TAGGTGAAGA TCCTTTTTGA TAATCTCATG ACCAAAATCC CTTAACGTGA 6550
 GTTTCGTTC CACTGAGCGT CAGACCCCGT AGAAAAGATC AAAGGATCTT 6600
 CTTGAGATCC TTTTTTCTG CGCGTAATCT GCTGCTTGCA AACAAAAAAA 6650
 CCACCGCTAC CAGCGGTGGT TTGTTTGCCG GATCAAGAGC TACCAACTCT 6700
 TTTTCCGAAG GTAACGGCT TCAGCAGAGC GCAGATACCA AATACTGTTC 6750
 TTCTAGTGTA GCCGTAGTTA GGCCACCACT TCAAGAACTC TGTAGACCCG 6800
 CCTACATACC TCGCTCTGCT AATCCTGTTA CCAGTGGCTG CTGCCAGTGG 6850
 CGATAAGTCG TGTCTTACCG GGTTGGACTC AAGACGATAG TTACCGGATA 6900
 AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT CGTGACACACA GCCCAGCTTG 6950
 GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA 7000
 AAGCGCCACG CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG 7050
 GCAGGGTCGG AACAGGAGAG CGCACGAGGG AGCTTCCAGG GGGAAACGCC 7100

TGGTATCTTT ATAGTCCTGT CGGGTTTCGC CACCTCTGAC TTGAGCGTCG 7150

ATTTTTGTGA TGCTCGTCAG GGGGGCGGAG CCTATGGAAA AACGCCAGCA 7200

ACGCGGCCTT TTTACGGTTC CTGGCCTTTT GCTGGCCTTT TGCTCACATG 7250

TTCTTTCCTG CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCCTT 7300

TGAGTGAGCT GATACCGCTC GCCGCAGCCG AACGACCGAG CGCAGCGAGT 7350

CAGTGAGCGA GGAAGCGGAA GAGCGCCCAA TACGCAAACC GCCTCTCCCC 7400

GCGCGTTGGC CGATTCATTA ATGCAGCTGG CACGACAGGT TTCCCGACTG 7450

GAAAGCGGGC AGTGAGCGCA ACGCAATTAA TGTGAGTTAG CTCACTCATT 7500

AGGCACCCCA GGCTTTACAC TTTATGCTTC CGGCTCGTAT GTTGTGTGGA 7550

ATTGTGAGCG GATAACAATT TCACACAGGA AACAGCTATG ACATGATTAC 7600

GAATTAA 7607

(2) INFORMACIÓN PARA LA SEC ID NO:20:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 505 aminoácidos

5 (B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:20:

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

His	Glu	Gly	Trp	Trp	Phe	Ala	Arg	His	Leu	Glu	Lys	Arg	Arg	Asp	80	85	90
Gly	Ser	Ser	Gln	Gln	Leu	Gln	Gly	Tyr	Ile	Pro	Ser	Asn	Tyr	Val	95	100	105
Ala	Glu	Asp	Arg	Ser	Leu	Gln	Ala	Glu	Pro	Trp	Phe	Phe	Gly	Ala	110	115	120
Ile	Gly	Arg	Ser	Asp	Ala	Glu	Lys	Gln	Leu	Leu	Tyr	Ser	Glu	Asn	125	130	135
Lys	Thr	Gly	Ser	Phe	Leu	Ile	Arg	Glu	Ser	Glu	Ser	Gln	Lys	Gly	140	145	150
Glu	Phe	Ser	Leu	Ser	Val	Leu	Asp	Gly	Ala	Val	Val	Lys	His	Tyr	155	160	165
Arg	Ile	Lys	Arg	Leu	Asp	Glu	Gly	Gly	Phe	Phe	Leu	Thr	Arg	Arg	170	175	180
Arg	Ile	Phe	Ser	Thr	Leu	Asn	Glu	Phe	Val	Ser	His	Tyr	Thr	Lys	185	190	195
Thr	Ser	Asp	Gly	Leu	Cys	Val	Lys	Leu	Gly	Lys	Pro	Cys	Leu	Lys	200	205	210
Ile	Gln	Val	Pro	Ala	Pro	Phe	Asp	Leu	Ser	Tyr	Lys	Thr	Val	Asp	215	220	225
Gln	Trp	Glu	Ile	Asp	Arg	Asn	Ser	Ile	Gln	Leu	Leu	Lys	Arg	Leu	230	235	240
Gly	Ser	Gly	Gln	Phe	Gly	Glu	Val	Trp	Glu	Gly	Leu	Trp	Asn	Asn	245	250	255
Thr	Thr	Pro	Val	Ala	Val	Lys	Thr	Leu	Lys	Pro	Gly	Ser	Met	Asp	260	265	270
Pro	Asn	Asp	Phe	Leu	Arg	Glu	Ala	Gln	Ile	Met	Lys	Asn	Leu	Arg	275	280	285
His	Pro	Lys	Leu	Ile	Gln	Leu	Tyr	Ala	Val	Cys	Thr	Leu	Glu	Asp	290	295	300
Pro	Ile	Tyr	Ile	Ile	Thr	Glu	Leu	Met	Arg	His	Gly	Ser	Leu	Gln	305	310	315
Glu	Tyr	Leu	Gln	Asn	Asp	Thr	Gly	Ser	Lys	Ile	His	Leu	Thr	Gln	320	325	330
Gln	Val	Asp	Met	Ala	Ala	Gln	Val	Ala	Ser	Gly	Met	Ala	Tyr	Leu	335	340	345
Glu	Ser	Arg	Asn	Tyr	Ile	His	Arg	Asp	Leu	Ala	Ala	Arg	Asn	Val	350	355	360

Leu	Val	Gly	Glu	His	Asn	Ile	Tyr	Lys	Val	Ala	Asp	Phe	Gly	Leu
				365					370					375
Ala	Arg	Val	Phe	Lys	Val	Asp	Asn	Glu	Asp	Ile	Tyr	Glu	Ser	Arg
				380					385					390
His	Glu	Ile	Lys	Leu	Pro	Val	Lys	Trp	Thr	Ala	Pro	Glu	Ala	Ile
				395					400					405
Arg	Ser	Asn	Lys	Phe	Ser	Ile	Lys	Ser	Asp	Val	Trp	Ser	Phe	Gly
				410					415					420
Ile	Leu	Leu	Tyr	Glu	Ile	Ile	Thr	Tyr	Gly	Lys	Met	Pro	Tyr	Ser
				425					430					435
Gly	Met	Thr	Gly	Ala	Gln	Val	Ile	Gln	Met	Leu	Ala	Gln	Asn	Tyr
				440					445					450
Arg	Leu	Pro	Gln	Pro	Ser	Asn	Cys	Pro	Gln	Gln	Phe	Tyr	Asn	Ile
				455					460					465
Met	Leu	Glu	Cys	Trp	Asn	Ala	Glu	Pro	Lys	Glu	Arg	Pro	Thr	Phe
				470					475					480
Glu	Thr	Leu	Arg	Trp	Lys	Leu	Glu	Asp	Tyr	Phe	Glu	Thr	Asp	Ser
				485					490					495
Ser	Tyr	Ser	Asp	Ala	Asn	Asn	Phe	Ile	Arg					
				500					505					

(2) INFORMACIÓN PARA LA SEC ID NO:21:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 404 bases

(B) TIPO: ácido nucleico

5 (C) CADENA: única

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:21:

GCGGCCGCAG AGAAAGCAGA GGATGGGGCT TAGCAGCTGG CAGAGCCAGG 50

10

AGCGGGGAGG TAGCAGAAAG ACCACAAGTA CAAAGAAGTC CTGAAACTTT 100

GGTTTTGCTG CTGCAGCCCA TTGAGAGTGA CGACATGGAG CACAAGACCC 150

15

TGAAGATCAC CGACTTTGGC CTGGCCCGAG AGTGGCACAA AACCACACAA 200

ATGAGTGCCG CNGGCACCTA CNCCTGGATG GCTCCTGAGG TTATCAAGGC 250

20

CTCCACCTTC TCTAAGGGCA GTGACGCTG GAGTTTTGGG GTGCTGCTGT 300

GGGAACTGCT GACCGGGGAG NTGCCATACC GTGGCATTGA CTGCCTTGCT 350

25

GTGGCCTATG GCGTAGCTGT TAACAAGCTC AACTGCCAT CCATCCACCT 400

GGCC 404

(2) INFORMACIÓN PARA LA SEC ID NO:22:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

30 (A) LONGITUD: 3120 bases

(B) TIPO: ácido nucleico

(C) CADENA: única

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:22:

ATGAGAGCGT TGGCGCGCGA CGGCGGCCAG CTGCCGCTGC TCGTTGTTTT 50
TTCTGCAATG ATATTTGGGA CTATTACAAA TCAAGATCTG CCTGTGATCA 100
AGTGTGTTTT AATCAATCAT AAGAACAATG ATTCATCAGT GGGGAAGTCA 150
TCATCATATC CCATGGTATC AGAATCCCCG GAAGACCTCG GGTGTGCGTT 200
GAGACCCAG AGCTCAGGGA CAGTGTACGA AGCTGCCGCT GTGGAAGTGG 250
ATGTATCTGC TTCCATCACA CTGCAAGTGC TGGTCGATGC CCCAGGGAAC 300
ATTCCTGTC TCTGGGTCTT TAAGCACAGC TCCCTGAATT GCCAGCCACA 350
TTTTGATTTA CAAAACAGAG GAGTTGTTTC CATGGTCATT TTGAAAATGA 400
CAGAAACCCA AGCTGGAGAA TACCTACTTT TTATTCAGAG TGAAGCTACC 450
AATTACACAA TATTGTTTAC AGTGAGTATA AGAAATACCC TGCTTTACAC 500
ATTAAGAAGA CCTTACTTTA GAAAAATGGA AAACCAGGAC GCCCTGGTCT 550
GCATATCTGA GAGCGTTCCA GAGCGGATCC TGGAATGGGT GCTTTGCGAT 600

ES 2 355 952 T3

TCACAGGGGG AAAGCTGTAA AGAAGAAAGT CCAGCTGTTG TTAAAAAGGA 650
GGAAAAAGTG CTTTCATGAAT TATTTGGGAC GGACATAAGG TGCTGTGCCA 700
GAAATGAACT GGGCAGGGAA TGCACCAGGC TGTTCACAAT AGATCTAAAT 750
CAAACCTCTC AGACCACATT GCCACAATTA TTTCTTAAAG TAGGGGAACC 800
CTTATGGATA AGGTGCAAAG CTGTTTCATGT GAACCATGGA TTCGGGCTCA 850
CCTGGGAATT AGAAAACAAA GCACTCGAGG AGGGCAACTA CTTTGAGATG 900
AGTACCTATT CAACAAACAG AACTATGATA CGGATTCTGT TTGCTTTTGT 950
ATCATCAGTG GCAAGAAACG ACACCGGATA CTACACTTGT TCCTCTTCAA 1000
AGCATCCCAG TCAATCAGCT TTGGTTACCA TCGTAGAAAA GGGATTTATA 1050
AATGCTACCA ATTCAAGTGA AGATTATGAA ATTGACCAAT ATGAAGAGTT 1100
TTGTTTTTCT GTCAGGTTTA AAGCCTACCC ACAAATCAGA TGTACGTGGA 1150
CCTTCTCTCG AAAATCATT CCTTGTGAGC AAAAGGGTCT TGATAACGGA 1200
TACAGCATAT CCAAGTTTGT CAATCATAAG CACCAGCCAG GAGAATATAT 1250
ATTCCATGCA GAAAATGATG ATGCCCAATT TACCAAAATG TTCACGCTGT 1300
ATATAAGAAG GAAACCTCAA GTCCTCGCAG AAGCTTCGGC AAGTCAGGCG 1350
TCCTGTTTCT CGGATGGATA CCCATTACCA TCTTGGACCT GGAAGAAGTG 1400
TTCAGACAAG TCTCCCAACT GCACAGAAGA GATCACAGAA GGAGTCTGGA 1450
ATAGAAAGGC TAACAGAAAA GTGTTTGGAC AGTGGGTGTC GAGCAGTACT 1500
CTAAACATGA GTGAAGCCAT AAAAGGGTTC CTGGTCAAGT GCTGTGCATA 1550

CAATTCCTT GGCACATCTT GTGAGACGAT CCTTTTAAAC TCTCCAGGCC 1600
 CCTTCCCTTT CATCCAAGAC AACATCTCAT TCTATGCAAC AATTGGTGTT 1650
 TGTCTCCTCT TCATTGTCGT TTTAACCCCTG CTAATTTGTC ACAAGTACAA 1700
 AAAGCAATTT AGGTATGAAA GCCAGCTACA GATGGTACAG GTGACCGGAT 1750
 CCTCAGATTA TGAGTACTTC TACGTTGATT TCAGAGAATA TGAATATGAT 1800
 GTCAAATGGG AGTTTCCAAG AGAAAATTTA GAGTTTGGGA AGGTACTAGG 1850
 ATCAGGTGCT TTTGGAAAAG TGATGAACGC AACAGCTTAT GGAATTAGCA 1900
 AAACAGGAGT CTCAATCCAG GTTACCGTCA AAATGCTGAA AGAAAAAGCA 1950
 GACAGCTCTG AAAGAGAGGC ACTCATGTCA GAACTCAAGA TGATGACCCA 2000
 GCTGGGAAGC CACGAGAATA TTGTGAACCT GCTGGGGGCG TGCACACTGT 2050
 CAGGACCAAT TTACTTGATT TTTGAATACT GTTGCTATGG TGATCTTCTC 2100
 AACTATCTAA GAAGTAAAAG AGAAAATTT CACAGGACTT GGACAGAGAT 2150
 TTTCAAGGAA CACAATTTCA GTTTTTACCC CACTTTCCAA TCACATCCAA 2200
 ATTCCAGCAT GCCTGGTTCA AGAGAAGTTC AGATACACCC GGACTCGGAT 2250
 CAAATCTCAG GGCTTCATGG GAATTCATTT CACTCTGAAG ATGAAATTGA 2300
 ATATGAAAAC CAAAAAGGC TGAAGAAGA GGAGGACTTG AATGTGCTTA 2350
 CATTTGAAGA TCTTCTTTGC TTTGCATATC AAGTTGCCAA AGGAATGGAA 2400
 TTTCTGGAAT TTAAGTCGTG TGTTACAGA GACCTGGCCG CCAGGAACGT 2450
 GCTTGTACC CACGGGAAAG TGGTGAAGAT ATGTGACTTT GGATTGGCTC 2500

GAGATATCAT GAGTGATTCC AACTATGTTG TCAGGGGCAA TGCCCGTCTG 2550
CCTGTAAAAT GGATGGCCCC CGAAAGCCTG TTTGAAGGCA TCTACACCAT 2600
TAAGAGTGAT GTCTGGTCAT ATGGAATATT ACTGTGGGAA ATCTTCTCAC 2650
TTGGTGTGAA TCCTTACCCT GGCATTCCGG TTGATGCTAA CTTCTACAAA 2700
CTGATTCAAA ATGGATTTAA AATGGATCAG CCATTTTATG CTACAGAAGA 2750
AATATACATT ATAATGCAAT CCTGCTGGGC TTTTGACTCA AGGAAACGGC 2800
CATCCTTCCC TAATTTGACT TCGTTTTTAG GATGTCAGCT GGCAGATGCA 2850
GAAGAAGCGA TGTATCAGAA TGTGGATGGC CGTGTTTCGG AATGTCCTCA 2900
CACCTACCAA AACAGGCGAC CTTTCAGCAG AGAGATGGAT TTGGGGCTAC 2950
TCTCTCCGCA GGCTCAGGTC GAAGATTCGT AGAGGAACAA TTTAGTTTTA 3000
AGGACTTCAT CCCTCCACCT ATCCCTAACA GGCTGTAGAT TACCAAACA 3050
AGGTTAATTT CATCACTAAA AGAAAATCTA TTATCAACTG CTGCTTCACC 3100
AGACTTTTCT CTAGAGAGCG 3120

(2) INFORMACIÓN PARA LA SEC ID NO:23:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 3969 bases

(B) TIPO: ácido nucleico

(C) CADENA: única

5 (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:23:

TCGGCGTCCA CCCGCCAGG GAGAGTCAGA CCTGGGGGGG CGAGGGCCCC 50

10

CCAAACTCAG TTCGGATCCT ACCCGAGTGA GCGGGCGCCA TGGAGCTCCG 100

ES 2 355 952 T3

GGTGCTGCTC TGCTGGGCTT CGTTGGCCGC AGCTTTGGAA GAGACCCTGC 150
 TGAACACAAA ATTGGAAACT GCTGATCTGA AGTGGGTGAC ATTCCCTCAG 200
 GTGGACGGGC AGTGGGAGGA ACTGAGCGGC CTGGATGAGG AACAGCACAG 250
 CGTGCGCACC TACGAAGTGT GTGACGTGCA GCGTGCCCCG GGCCAGGCCC 300
 ACTGGCTTCG CACAGGTGG GTCCCACGGC GGGGCGCCGT CCACGTGTAC 350
 GCCACGCTGC GCTTCACCAT GCTCGAGTGC CTGTCCCTGC CTCGGGCTGG 400
 GCGCTCCTGC AAGGAGACCT TCACCGTCTT CTACTATGAG AGCGATGCGG 450
 ACACGGCCAC GGCCCTCACG CCAGCCTGGA TGGAGAACCC CTACATCAAG 500
 GTGGACACGG TGGCCGCGGA GCATCTCACC CGGAAGCGCC CTGGGGCCGA 550
 GGCCACCGGG AAGGTGAATG TCAAGACGCT GCGTCTGGGA CCGCTCAGCA 600
 AGGCTGGCTT CTACCTGGCC TTCCAGGACC AGGGTGCCTG CATGGCCCTG 650
 CTATCCCTGC ACCTCTTCTA CAAAAGTGC GCCCAGCTGA CTGTGAACCT 700
 GACTCGATTG CCGGAGACTG TGCCTCGGGA GCTGGTTGTG CCCGTGGCCG 750
 GTAGCTGCGT GGTGGATGCC GTCCCCGCC CTGGCCCCAG CCCCAGCCTC 800
 TACTGCCGTG AGGATGGCCA GTGGGCCGAA CAGCCGGTCA CGGGCTGCAG 850
 CTGTGCTCCG GGGTTCGAGG CAGCTGAGGG GAACACCAAG TGCCGAGCCT 900
 GTGCCCAGGG CACCTTCAAG CCCCTGTCAG GAGAAGGGTC CTGCCAGCCA 950
 TGCCAGCCA ATAGCCACTC TAACACCATT GGATCAGCCG TCTGCCAGTG 1000
 CCGGTCGGG TACTTCCGGG CACGCACAGA CCCCCGGGT GCACCCTGCA 1050

ES 2 355 952 T3

CCACCCCTCC TTCGGCTCCG CGGAGCGTGG TTTCCCGCCT GAACGGCTCC 1100
 TCCCTGCACC TGGAATGGAG TGCCCCCTG GAGTCTGGTG GCCGAGAGGA 1150
 CCTCACCTAC GCCCTCCGCT GCCGGGAGTG CCGACCCGGA GGCTCCTGTG 1200
 CGCCCTGCGG GGGAGACCTG ACTTTTGACC CCGGCCCCCG GGACCTGGTG 1250
 GAGCCCTGGG TGGTGGTTCG AGGGCTACGT CCTGACTTCA CCTATACCTT 1300
 TGAGGTCACT GCATTGAACG GGGTATCCTC CTTAGCCACG GGGCCCGTCC 1350
 CATTTGAGCC TGTCAATGTC ACCACTGACC GAGAGGTACC TCCTGCAGTG 1400
 TCTGACATCC GGGTGACGCG GTCCTCACCC AGCAGCTTGA GCCTGGCCTG 1450
 GGCTGTTCCC CGGGCACCCA GTGGGGCTGT GCTGGACTAC GAGGTCAAAT 1500
 ACCATGAGAA GGGCGCCGAG GGTCCCAGCA GCGTGCGGTT CCTGAAGACG 1550
 TCAGAAAACC GGGCAGAGCT GCGGGGGCTG AAGCGGGGAG CCAGCTACCT 1600
 GGTGCAGGTA CGGGCGCGCT CTGAGGCCGG CTACGGGCC TTCGGCCAGG 1650
 AACATCACAG CCAGACCCAA CTGGATGAGA GCGAGGGCTG GCGGGAGCAG 1700
 CTGGCCCTGA TTGCGGGCAC GGCAGTCGTG GGTGTGGTCC TGGTCCTGGT 1750
 GGTCAATTGTG GTCGCAGTTC TCTGCCTCAG GAAGCAGAGC AATGGGAGAG 1800
 AAGCAGAATA TTCGGACAAA CACGGACAGT ATCTCATCGG ACATGGTACT 1850
 AAGGTCTACA TCGACCCCTT CACTTATGAA GACCCTAATG AGGCTGTGAG 1900
 GGAATTTGCA AAAGAGATCG ATGTCTCCTA CGTCAAGATT GAAGAGGTGA 1950
 TTGGTGCAGG TGAGTTTGGC GAGGTGTGCC GGGGGCGGCT CAAGGCCCCA 2000

ES 2 355 952 T3

GGGAAGAAGG AGAGCTGTGT GGCAATCAAG ACCCTGAAGG GTGGCTACAC 2050
GGAGCGGCAG CGGCGTGAGT TTCTGAGCGA GGCCTCCATC ATGGGCCAGT 2100
TCGAGCACCC CAATATCATC CGCCTGGAGG GCGTGGTCAC CAACAGCATG 2150
CCCGTCATGA TTCTCACAGA GTTCATGGAG AACGGCGCCC TGGACTCCTT 2200
CCTGCGGCTA AACGACGGAC AGTTCACAGT CATCCAGCTC GTGGGCATGC 2250
TGCGGGGCAT CGCCTCGGGC ATGCGGTACC TTGCCGAGAT GAGCTACGTC 2300
CACCGAGACC TGGCTGCTCG CAACATCCTA GTCAACAGCA ACCTCGTCTG 2350
CAAAGTGTCT GACTTTGGCC TTTCCCGATT CCTGGAGGAG AACTCTTCCG 2400
ATCCCACCTA CACGAGCTCC CTGGGAGGAA AGATTCCCAT CCGATGGACT 2450
GCCCCGGAGG CCATTGCCTT CCGGAAGTTC ACTTCCGCCA GTGATGCCTG 2500
GAGTTACGGG ATTGTGATGT GGGAGGTGAT GTCATTTGGG GAGAGGCCGT 2550
ACTGGGACAT GAGCAATCAG GACGTGATCA ATGCCATTGA ACAGGACTAC 2600
CGGCTGCCCC CGCCCCAGA CTGTCCCACC TCCCTCCACC AGCTCATGCT 2650
GGACTGTTGG CAGAAAGACC GGAATGCCCG GCCCCGCTTC CCCCAGGTGG 2700
TCAGCGCCCT GGACAAGATG ATCCGGAACC CCGCCAGCCT CAAAATCGTG 2750
GCCCCGGGAGA ATGGCGGGGC CTCACACCCT CTCCTGGACC AGCGGCAGCC 2800
TCACTACTCA GCTTTTGGCT CTGTGGGCGA GTGGCTTCGG GCCATCAAAA 2850
TGGGAAGATA CGAAGAAAGT TTCGCAGCCG CTGGCTTTGG CTCCTTCGAG 2900
CTGGTCAGCC AGATCTCTGC TGAGGACCTG CTCCGAATCG GAGTCACTCT 2950

ES 2 355 952 T3

GGCGGGACAC CAGAAGAAAA TCTTGGCCAG TGTCCAGCAC ATGAAGTCCC 3000
AGGCCAAGCC GGAACCCCG GGTGGGACAG GAGGACCGGC CCCGCAGTAC 3050
TGACCTGCAG GAACTCCCA CCCAGGGAC ACCGCCTCCC CATTTCCTGG 3100
GGCAGAGTGG GGAATCACAG AGGCCCCCAG CCCTGTGCCC CGCTGGATTG 3150
CACTTTGAGC CCGTGGGGTG AGGAGTTGGC AATTTGGAGA GACAGGATTT 3200
GGGGGTTCTG CCATAATAGG AGGGGAAAAT CACCCCCCAG CCACCTCGGG 3250
GAACTCCAGA CCAAGGGTGA GGGCGCCTTT CCCTCAGGAC TGGGTGTGAC 3300
CAGAGGAAAA GGAAGTGCCC AACATCTCCC AGCCTCCCCA GGTGCCCCCC 3350
TCACCTTGAT GGGTGCCTTC CCGCAGACCA AAGAGAGTGT GACTCCCTTG 3400
CCAGCTCCAG AGTGGGGGGG CTGTCCCAGG GGGCAAGAAG GGGTGTGAGG 3450
GCCCAGTGAC AAAATCATTG GGGTTTGTAG TCCCAACTTG CTGCTGTGAC 3500
CACCAAATC AATCATTFFF TTCCCTTGTA AATGCCCTC CCCAGCTGC 3550
TGCCTTCATA TTGAAGGTTT TTGAGTTTG TTTTGGTCT TAATTTTTCT 3600
CCCCGTTCCC TTTTGTGTTT TTCGTTTGT TTTTCTACCG TCCTTGTCAT 3650
AACTTTGTGT TGGAGGGAAC CTGTTTCACT ATGGCCTCCT TGCCCAAGT 3700
TGAAACAGGG GCCCATCATC ATGTCTGTTT CCAGAACAGT GCCTTGGTCA 3750
TCCCACATCC CCGGACCCCG CCTGGGACCC CCAAGCTGTG TCCTATGAAG 3800
GGGTGTGGGG TGAGGTAGTG AAAAGGGCGG TAGTTGGTGG TGGAACCCAG 3850
AAACGGACGC CGGTGCTTGG AGGGGTTCTT AAATTATATT TAAAAAAGTA 3900

ACTTTTTGTA TAAATAAAAG AAAATGGGAC GTGTCCCAGC TCCAGGGGTA 3950

AAAAAAAAAA AAAAAAAAAA 3969

(2) INFORMACIÓN PARA LA SEC ID NO:24:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

5 (A) LONGITUD: 1276 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:24:

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

Val	Pro	Arg	Glu	Leu	Val	Val	Pro	Val	Ala	Gly	Ser	Cys	Val	Val	215	220	225
Asp	Ala	Val	Pro	Ala	Pro	Gly	Pro	Ser	Pro	Ser	Leu	Tyr	Cys	Arg	230	235	240
Glu	Asp	Gly	Gln	Trp	Ala	Glu	Gln	Pro	Val	Thr	Gly	Cys	Ser	Cys	245	250	255
Ala	Pro	Gly	Phe	Glu	Ala	Ala	Glu	Gly	Asn	Thr	Lys	Cys	Arg	Ala	260	265	270
Cys	Ala	Gln	Gly	Thr	Phe	Lys	Pro	Leu	Ser	Gly	Glu	Gly	Ser	Cys	275	280	285
Gln	Pro	Cys	Pro	Ala	Asn	Ser	His	Ser	Asn	Thr	Ile	Gly	Ser	Ala	290	295	300
Val	Cys	Gln	Cys	Arg	Val	Gly	Tyr	Phe	Arg	Ala	Arg	Thr	Asp	Pro	305	310	315
Arg	Gly	Ala	Pro	Cys	Thr	Thr	Pro	Pro	Ser	Ala	Pro	Arg	Ser	Val	320	325	330
Val	Ser	Arg	Leu	Asn	Gly	Ser	Ser	Leu	His	Leu	Glu	Trp	Ser	Ala	335	340	345
Pro	Leu	Glu	Ser	Gly	Gly	Arg	Glu	Asp	Leu	Thr	Tyr	Ala	Leu	Arg	350	355	360
Cys	Arg	Glu	Cys	Arg	Pro	Gly	Gly	Ser	Cys	Ala	Pro	Cys	Gly	Gly	365	370	375
Asp	Leu	Thr	Phe	Asp	Pro	Gly	Pro	Arg	Asp	Leu	Val	Glu	Pro	Trp	380	385	390
Val	Val	Val	Arg	Gly	Leu	Arg	Pro	Asp	Phe	Thr	Tyr	Thr	Phe	Glu	395	400	405
Val	Thr	Ala	Leu	Asn	Gly	Val	Ser	Ser	Leu	Ala	Thr	Gly	Pro	Val	410	415	420
Pro	Phe	Glu	Pro	Val	Asn	Val	Thr	Thr	Asp	Arg	Glu	Val	Pro	Pro	425	430	435
Ala	Val	Ser	Asp	Ile	Arg	Val	Thr	Arg	Ser	Ser	Pro	Ser	Ser	Leu	440	445	450
Ser	Leu	Ala	Trp	Ala	Val	Pro	Arg	Ala	Pro	Ser	Gly	Ala	Val	Leu	455	460	465
Asp	Tyr	Glu	Val	Lys	Tyr	His	Glu	Lys	Gly	Ala	Glu	Gly	Pro	Ser	470	475	480
Ser	Val	Arg	Phe	Leu	Lys	Thr	Ser	Glu	Asn	Arg	Ala	Glu	Leu	Arg	485	490	495

Gly	Leu	Lys	Arg	Gly	Ala	Ser	Tyr	Leu	Val	Gln	Val	Arg	Ala	Arg	
				500					505					510	
Ser	Glu	Ala	Gly	Tyr	Gly	Pro	Phe	Gly	Gln	Glu	His	His	Ser	Gln	
				515					520					525	
Thr	Gln	Leu	Asp	Glu	Ser	Glu	Gly	Trp	Arg	Glu	Gln	Leu	Ala	Leu	
				530					535					540	
Ile	Ala	Gly	Thr	Ala	Val	Val	Gly	Val	Val	Leu	Val	Leu	Val	Val	
				545					550					555	
Ile	Val	Val	Ala	Val	Leu	Cys	Leu	Arg	Lys	Gln	Ser	Asn	Gly	Arg	
				560					565					570	
Glu	Ala	Glu	Tyr	Ser	Asp	Lys	His	Gly	Gln	Tyr	Leu	Ile	Gly	His	
				575					580					585	
Gly	Thr	Lys	Val	Tyr	Ile	Asp	Pro	Phe	Thr	Tyr	Glu	Asp	Pro	Asn	
				590					595					600	
Glu	Ala	Val	Arg	Glu	Phe	Ala	Lys	Glu	Ile	Asp	Val	Ser	Tyr	Val	
				605					610					615	
Lys	Ile	Glu	Glu	Val	Ile	Gly	Ala	Gly	Glu	Phe	Gly	Glu	Val	Cys	
				620					625					630	
Arg	Gly	Arg	Leu	Lys	Ala	Pro	Gly	Lys	Lys	Glu	Ser	Cys	Val	Ala	
				635					640					645	
Ile	Lys	Thr	Leu	Lys	Gly	Gly	Tyr	Thr	Glu	Arg	Gln	Arg	Arg	Glu	
				650					655					660	
Phe	Leu	Ser	Glu	Ala	Ser	Ile	Met	Gly	Gln	Phe	Glu	His	Pro	Asn	
				665					670					675	
Ile	Ile	Arg	Leu	Glu	Gly	Val	Val	Thr	Asn	Ser	Met	Pro	Val	Met	
				680					685					690	
Ile	Leu	Thr	Glu	Phe	Met	Glu	Asn	Gly	Ala	Leu	Asp	Ser	Phe	Leu	
				695					700					705	
Arg	Leu	Asn	Asp	Gly	Gln	Phe	Thr	Val	Ile	Gln	Leu	Val	Gly	Met	
				710					715					720	
Leu	Arg	Gly	Ile	Ala	Ser	Gly	Met	Arg	Tyr	Leu	Ala	Glu	Met	Ser	
				725					730					735	
Tyr	Val	His	Arg	Asp	Leu	Ala	Ala	Arg	Asn	Ile	Leu	Val	Asn	Ser	
				740					745					750	
Asn	Leu	Val	Cys	Lys	Val	Ser	Asp	Phe	Gly	Leu	Ser	Arg	Phe	Leu	
				755					760					765	
Glu	Glu	Asn	Ser	Ser	Asp	Pro	Thr	Tyr	Thr	Ser	Ser	Leu	Gly	Gly	
				770					775					780	

Lys Ile Pro Ile Arg Trp Thr Ala Pro Glu Ala Ile Ala Phe Arg
 785 790 795
 Lys Phe Thr Ser Ala Ser Asp Ala Trp Ser Tyr Gly Ile Val Met
 800 805 810
 Trp Glu Val Met Ser Phe Gly Glu Arg Pro Tyr Trp Asp Met Ser
 815 820 825
 Asn Gln Asp Val Ile Asn Ala Ile Glu Gln Asp Tyr Arg Leu Pro
 830 835 840
 Pro Pro Pro Asp Cys Pro Thr Ser Leu His Gln Leu Met Leu Asp
 845 850 855
 Cys Trp Gln Lys Asp Arg Asn Ala Arg Pro Arg Phe Pro Gln Val
 860 865 870
 Val Ser Ala Leu Asp Lys Met Ile Arg Asn Pro Ala Ser Leu Lys
 875 880 885
 Ile Val Ala Arg Glu Asn Gly Gly Ala Ser His Pro Leu Leu Asp
 890 895 900
 Gln Arg Gln Pro His Tyr Ser Ala Phe Gly Ser Val Gly Glu Trp
 905 910 915
 Leu Arg Ala Ile Lys Met Gly Arg Tyr Glu Glu Ser Phe Ala Ala
 920 925 930
 Ala Gly Phe Gly Ser Phe Glu Leu Val Ser Gln Ile Ser Ala Glu
 935 940 945
 Asp Leu Leu Arg Ile Gly Val Thr Leu Ala Gly His Gln Lys Lys
 950 955 960
 Ile Leu Ala Ser Val Gln His Met Lys Ser Gln Ala Lys Pro Gly
 965 970 975
 Thr Pro Gly Gly Thr Gly Gly Pro Ala Pro Gln Tyr Pro Ala Gly
 980 985 990
 Thr Pro His Pro Arg Asp Thr Ala Ser Pro Phe Ser Gly Ala Glu
 995 1000 1005
 Trp Gly Leu Thr Glu Ala Pro Ser Pro Val Pro Arg Trp Ile Ala
 1010 1015 1020
 Leu Ala Arg Gly Val Arg Ser Trp Gln Phe Gly Glu Thr Gly Phe
 1025 1030 1035
 Gly Gly Ser Ala Ile Ile Gly Gly Glu Asn His Pro Pro Ala Thr
 1040 1045 1050
 Ser Gly Asn Ser Arg Pro Arg Val Arg Ala Pro Phe Pro Gln Asp
 1055 1060 1065

Trp	Val	Pro	Glu	Glu	Lys	Glu	Val	Pro	Asn	Ile	Ser	Gln	Pro	Pro
				1070					1075					1080
Gln	Val	Pro	Pro	Ser	Pro	Trp	Val	Arg	Ser	Arg	Arg	Pro	Lys	Arg
				1085					1090					1095
Val	Leu	Pro	Cys	Gln	Leu	Gln	Ser	Gly	Gly	Ala	Val	Pro	Gly	Gly
				1100					1105					1110
Lys	Lys	Gly	Cys	Gln	Gly	Pro	Val	Thr	Lys	Ser	Leu	Gly	Phe	Val
				1115					1120					1125
Val	Pro	Thr	Cys	Cys	Cys	His	His	Gln	Thr	Gln	Ser	Phe	Phe	Ser
				1130					1135					1140
Leu	Val	Asn	Ala	Pro	Pro	Pro	Ala	Ala	Ala	Phe	Ile	Leu	Lys	Val
				1145					1150					1155
Phe	Glu	Phe	Cys	Phe	Trp	Ser	Phe	Phe	Ser	Pro	Phe	Pro	Phe	Cys
				1160					1165					1170
Phe	Phe	Val	Leu	Phe	Phe	Tyr	Arg	Pro	Cys	His	Asn	Phe	Val	Leu
				1175					1180					1185
Glu	Gly	Thr	Cys	Phe	Thr	Met	Ala	Ser	Phe	Ala	Gln	Val	Glu	Thr
				1190					1195					1200
Gly	Ala	His	His	His	Val	Cys	Phe	Gln	Asn	Ser	Ala	Leu	Val	Ile
				1205					1210					1215
Pro	His	Pro	Arg	Thr	Pro	Pro	Gly	Thr	Pro	Lys	Leu	Cys	Pro	Met
				1220					1225					1230
Lys	Gly	Cys	Gly	Val	Arg	Lys	Gly	Arg	Leu	Val	Val	Glu	Pro	Arg
				1235					1240					1245
Asn	Gly	Arg	Arg	Cys	Leu	Glu	Gly	Phe	Leu	Asn	Tyr	Ile	Lys	Ser
				1250					1255					1260
Asn	Phe	Leu	Tyr	Lys	Lys	Lys	Met	Gly	Arg	Val	Pro	Ala	Pro	Gly
				1265					1270					1275
Val														
1276														

(2) INFORMACIÓN PARA LA SEC ID NO:25:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 59 aminoácidos

5 (B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:25:

	Ala	Arg	Asn	Ile	Leu	Val	Asn	Ser	Asn	Leu	Val	Cys	Lys	Val	Ser	
	1				5					10					15	
10																
	Asp	Phe	Gly	Leu	Ser	Arg	Phe	Leu	Glu	Asp	Asp	Thr	Ser	Asp	Pro	
					20					25					30	
	Thr	Tyr	Thr	Ser	Ala	Leu	Gly	Gly	Lys	Ile	Pro	Met	Arg	Trp	Thr	
15					35					40					45	
	Ala	Pro	Glu	Ala	Ile	Gln	Tyr	Arg	Lys	Phe	Ala	Ser	Ala	Ser		
					50					55				59		

(2) INFORMACIÓN PARA LA SEC ID NO:26:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

20 (A) LONGITUD: 54 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:26:

	Asn	Val	Leu	Val	Lys	Ser	Pro	Asn	His	Val	Lys	Ile	Thr	Asp	Phe	
25	1				5					10					15	
	Gly	Leu	Ala	Arg	Leu	Leu	Glu	Gly	Asp	Glu	Lys	Glu	Tyr	Asn	Ala	
					20					25					30	
	Asp	Gly	Gly	Lys	Met	Pro	Ile	Lys	Trp	Met	Ala	Leu	Glu	Cys	Ile	
30					35					40					45	
	His	Tyr	Arg	Lys	Phe	Thr	His	Gln	Ser							
					50				54							

(2) INFORMACIÓN PARA LA SEC ID NO:27:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

35 (A) LONGITUD: 54 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:27:

```

    Asn Cys Met Leu Ala Gly Asp Met Thr Val Cys Val Ala Asp Phe
      1           5           10           15
5    Gly Leu Ser Trp Lys Ile Tyr Ser Gly Ala Thr Ile Val Arg Gly
      20           25           30
    Cys Ala Ser Lys Leu Pro Val Lys Trp Leu Ala Leu Gly Ser Leu
      35           40           45
10   Ala Asp Asn Leu Tyr Thr Val His Ser
      50           54

```

(2) INFORMACIÓN PARA LA SEC ID NO:28:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 27 aminoácidos

(B) TIPO: aminoácido

15 (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:28:

```

    Asn Cys Leu Val Gly Lys Asn Tyr Thr Ile Lys Ile Ala Asp Phe
      1           5           10           15
20   Gly Met Ser Arg Asn Leu Tyr Ser Gly Asp Tyr Tyr
      20           25           27

```

(2) INFORMACIÓN PARA LA SEC ID NO:29:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 58 aminoácidos

25 (B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:29:

```

    Thr Arg Asn Ile Leu Val Glu Asn Glu Asn Arg Val Lys Ile Gly
      1           5           10           15
30   Asp Phe Gly Leu Thr Lys Val Leu Pro Gln Asp Lys Glu Tyr Tyr
      20           25           30
    Lys Val Lys Glu Pro Gly Glu Ser Pro Ile Phe Trp Tyr Ala Pro
      35           40           45
35   Glu Ser Leu Thr Glu Ser Leu Phe Ser Val Ala Ser Asp
      50           55           58

```

(2) INFORMACIÓN PARA LA SEC ID NO:30:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 58 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

5 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:30:

	Ala	Arg	Asn	Ile	Leu	Val	Asn	Ser	Asn	Leu	Val	Cys	Lys	Val	Ser
	1				5					10					15
	Asp	Phe	Gly	Met	Ser	Arg	Val	Leu	Glu	Asp	Asp	Pro	Glu	Ala	Ala
					20					25					30
10	Tyr	Thr	Thr	Arg	Gly	Gly	Lys	Ile	Pro	Ile	Arg	Trp	Thr	Ala	Pro
					35					40					45
	Glu	Ala	Ile	Ala	Tyr	Arg	Lys	Phe	Thr	Ser	Ala	Ser	Asp		
					50					55			58		

15

(2) INFORMACIÓN PARA LA SEC ID NO:31:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 4425 bases

(B) TIPO: ácido nucleico

20 (C) CADENA: única

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:31:

TCGGGTCGGA CCCACGCGCA GCGGCCGGAG ATGCAGCGGG GCGCCGCGCT 50
 GTGCCTGCGA CTGTGGCTCT GCCTGGGACT CCTGGACGGC CTGGTGAGTG 100
 GCTACTCCAT GACCCCCCG ACCTTGAACA TCACGGAGGA GTCACACGTC 150
 ATCGACACCG GTGACAGCCT GTCCATCTCC TGCAGGGGAC AGCACCCCCT 200
 CGAGTGGGCT TGGCCAGGAG CTCAGGAGGC GCCAGCCACC GGAGACAAGG 250
 ACAGCGAGGA CACGGGGGTG GTGCGAGACT GCGAGGGCAC AGACGCCAGG 300
 CCCTACTGCA AGGTGTTGCT GCTGCACGAG GTACATGCCA ACGACACAGG 350
 CAGCTACGTC TGCTACTACA AGTACATCAA GGCACGCATC GAGGGCACCA 400
 CGGCCGCCAG CTCCTACGTG TTCGTGAGAG ACTTTGAGCA GCCATTCATC 450
 AACAAAGCCTG ACACGCTCTT GGTCAACAGG AAGGACGCCA TGTGGGTGCC 500
 CTGTCTGGTG TCCATCCCCG GCCTCAATGT CACGCTGCGC TCGCAAAGCT 550
 CGGTGCTGTG GCCAGACGGG CAGGAGGTGG TGTGGGATGA CCGCGGGGC 600
 ATGCTCGTGT CCACGCCACT GCTGCACGAT GCCCTGTACC TGCAGTGCGA 650
 GACCACCTGG GGAGACCAGG ACTTCCTTTC CAACCCCTTC CTGGTGACCA 700
 TCACAGGCAA CGAGCTCTAT GACATCCAGC TGTTGCCCAG GAAGTCGCTG 750
 GAGCTGCTGG TAGGGGAGAA GCTGGTCCTG AACTGCACCG TGTGGGCTGA 800
 GTTTAACTCA GGTGTCACCT TTGACTGGGA CTACCCAGGG AAGCAGGCAG 850
 AGCGGGGTAA GTGGGTGCCG GAGCGACGCT CCCAGCAGAC CCACACAGAA 900

CTCTCCAGCA TCCTGACCAT CCACAACGTC AGCCAGCACG ACCTGGGCTC 950
 GTATGTGTGC AAGGCCAACA ACGGCATCCA GCGATTTCCG GAGAGCACCG 1000
 AGGTCATTGT GCATGAAAAT CCCTTCATCA GCGTCGAGTG GCTCAAAGGA 1050
 CCCATCCTGG AGGCCACGGC AGGAGACGAG CTGGTGAAGC TGCCCGTGAA 1100
 GCTGGCAGCG TACCCCCCGC CCGAGTTCCA GTGGTACAAG GATGGAAAGG 1150
 CACTGTCCGG GCGCCACAGT CCACATGCCC TGGTGCTCAA GGAGGTGACA 1200
 GAGGCCAGCA CAGGCACCTA CACCCTCGCC CTGTGGA ACT CCGCTGCTGG 1250
 CCTGAGGCGC AACATCAGCC TGGAGCTGGT GGTGAATGTG CCCCCCAGA 1300
 TACATGAGAA GGAGGCCTCC TCCCCCAGCA TCTACTCGCG TCACAGCCGC 1350
 CAGGCCCTCA CCTGCACGGC CTACGGGGTG CCCCTGCCTC TCAGCATCCA 1400
 GTGGCACTGG CGGCCCTGGA CACCCTGCAA GATGTTTGCC CAGCGTAGTC 1450
 TCCGGCGGCG GCAGCAGCAA GACCTCATGC CACAGTGCCG TGACTGGAGG 1500
 GCGGTGACCA CGCAGGATGC CGTGAACCCC ATCGAGAGCC TGGACACCTG 1550
 GACCGAGTTT GTGGAGGGAA AGAATAAGAC TGTGAGCAAG CTGGTGATCC 1600
 AGAATGCCAA CGTGTCTGCC ATGTACAAGT GTGTGGTCTC CAACAAGGTG 1650
 GGCCAGGATG AGCGGCTCAT CTA CTCTAT GTGACCACCA TCCCCGACGG 1700
 C TTCACCATC GAATCCAAGC CATCCGAGGA GCTACTAGAG GGCCAGCCGG 1750
 TGCTCCTGAG CTGCCAAGCC GACAGCTACA AGTACGAGCA TCTGCGCTGG 1800
 TACCGCCTCA ACCTGTCCAC GCTGCACGAT GCGCACGGGA ACCCGCTTCT 1850

ES 2 355 952 T3

GCTCGACTGC AAGAACGTGC ATCTGTTTCGC CACCCCTCTG GCCGCCAGCC 1900
 TGGAGGAGGT GGCACCTGGG GCGCGCCACG CCACGCTCAG CCTGAGTATC 1950
 CCCC GCGTCG CGCCCGAGCA CGAGGGCCAC TATGTGTGCG AAGTGCAAGA 2000
 CCGGCGCAGC CATGACAAGC ACTGCCACAA GAAGTACCTG TCGGTGCAGG 2050
 CCCTGGAAGC CCCTCGGCTC ACGCAGAACT TGACCGACCT CCTGGTGAAC 2100
 GTGAGCGACT CGCTGGAGAT GCAGTGCTTG GTGGCCGGAG CGCACGCGCC 2150
 CAGCATCGTG TGGTACAAAG ACGAGAGGCT GCTGGAGGAA AAGTCTGGAG 2200
 TCGACTTGGC GGA CTCCAAC CAGAAGCTGA GCATCCAGCG CGTGCGCGAG 2250
 GAGGATGCGG GACGCTATCT GTGCAGCGTG TGCAACGCCA AGGGCTGCGT 2300
 CAACTCCTCC GCCAGCGTGG CCGTGGAAGG CTCCGAGGAT AAGGGCAGCA 2350
 TGGAGATCGT GATCCTTGTC GGTACCGGCG TCATCGCTGT CTTCTTCTGG 2400
 GTCCTCCTCC TCCTCATCTT CTGTAACATG AGGAGGCCGG CCCACGCAGA 2450
 CATCAAGACG GGCTACCTGT CCATCATCAT GGACCCCGGG GAGGTGCCTC 2500
 TGGAGGAGCA ATGCGAATAC CTGTCCTACG ATGCCAGCCA GTGGGAATTC 2550
 CCCC GAGAGC GGCTGCACCT GGGGAGAGTG CTCGGCTACG GCGCCTTCGG 2600
 GAAGGTGGTG GAAGCCTCCG CTTTCGGCAT CCACAAGGGC AGCAGCTGTG 2650
 ACACCGTGGC CGTGAAAATG CTGAAAGAGG GCGCCACGGC CAGCGAGCAC 2700
 CGCGCGCTGA TGTCGGAGCT CAAGATCCTC ATTCACATCG GCAACCACCT 2750
 CAACGTGGTC AACCTCCTCG GGGCGTGAC CAAGCCGCAG GGCCCCCTCA 2800

ES 2 355 952 T3

TGGTGATCGT GGAGTTCTGC AAGTACGGCA ACCTCTCCAA CTCCTGCGC 2850
 GCCAAGCGGG ACGCCTTCAG CCCCTGCGCG GAGAAGTCTC CCGAGCAGCG 2900
 CGGACGCTTC CGCGCCATGG TGGAGCTCGC CAGGCTGGAT CGGAGGCGGC 2950
 CGGGGAGCAG CGACAGGGTC CTCTTCGCGC GGTTCTCGAA GACCGAGGGC 3000
 GGAGCGAGGC GGGCTTCTCC AGACCAAGAA GCTGAGGACC TGTGGCTGAG 3050
 CCCGCTGACC ATGGAAGATC TTGTCTGCTA CAGCTTCCAG GTGGCCAGAG 3100
 GGATGGAGTT CCTGGCTTCC CGAAAGTGCA TCCACAGAGA CCTGGCTGCT 3150
 CGGAACATTC TGCTGTCGGA AAGCGACGTG GTGAAGATCT GTGACTTTGG 3200
 CCTTGCCCGG GACATCTACA AAGACCCTGA CTACGTCCGC AAGGGCAGTG 3250
 CCCGGCTGCC CCTGAAGTGG ATGGCCCTG AAAGCATCTT CGACAAGGTG 3300
 TACACCACGC AGAGTGACGT GTGGTCCTTT GGGGTGCTTC TCTGGGAGAT 3350
 CTTCTCTCTG GGGGCCTCCC CGTACCCTGG GGTGCAGATC AATGAGGAGT 3400
 TCTGCCAGCG GCTGAGAGAC GGCACAAGGA TGAGGGCCCC GGAGCTGGCC 3450
 ACTCCCGCCA TACGCCGCAT CATGCTGAAC TGCTGGTCCG GAGACCCCAA 3500
 GGCAGACCT GCATTCTCGG AGCTGGTGGA GATCCTGGGG GACCTGCTCC 3550
 AGGGCAGGGG CCTGCAAGAG GAAGAGGAGG TCTGCATGGC CCCGCGCAGC 3600
 TCTCAGAGCT CAGAAGAGGG CAGCTTCTCG CAGGTGTCCA CCATGGCCCT 3650
 ACACATCGCC CAGGCTGACG CTGAGGACAG CCCGCCAAGC CTGCAGCGCC 3700
 ACAGCCTGGC CGCCAGGTAT TACAACCTGGG TGTCCCTTCC CGGGTGCTG 3750

GCCAGAGGGG CTGAGACCCG TGGTTCCTCC AGGATGAAGA CATTGAGGA 3800
 ATTCCCATG ACCCCAACGA CCTACAAAGG CTCTGTGGAC AACCAGACAG 3850
 ACAGTGGGAT GGTGCTGGCC TCGGAGGAGT TTGAGCAGAT AGAGAGCAGG 3900
 CATAGACAAG AAAGCGGCTT CAGGTAGCTG AAGCAGAGAG AGAGAAGGCA 3950
 GCATACGTCA GCATTTTCTT CTCTGCACTT ATAAGAAAGA TCAAAGACTT 4000
 TAAGACTTTC GCTATTTCTT CTGCTATCTA CTACAAACTT CAAAGAGGAA 4050
 CCAGGAGGCC AAGAGGAGCA TGAAAGTGA CAAGGAGTGT GACCACTGAA 4100
 GCACCACAGG GAGGGGTTAG GCCTCCGGAT GACTGCGGGC AGGCCTGGAT 4150
 AATATCCAGC CTCCCACAAG AAGCTGGTGG AGCAGAGTGT TCCCTGACTC 4200
 CTCCAAGGAA AGGGAGACGC CCTTTCATGG TCTGCTGAGT AACAGGTGCC 4250
 TTCCCAGACA CTGGCGTTAC TGCTTGACCA AAGAGCCCTC AAGCGGCCCT 4300
 TATGCCAGCG TGACAGAGGG CTCACCTCTT GCCTTCTAGG TCACTTCTCA 4350
 CAATGTCCCT TCAGCACCTG ACCCTGTGCC CGCCAGTTAT TCCTTGGTAA 4400
 TATGAGTAAT ACATCAAAGA GTAGT 4425

(2) INFORMACIÓN PARA LA SEC ID NO:32:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 4425 bases

(B) TIPO: ácido nucleico

5 (C) CADENA: única

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:32:

AGCCCAGCCT GGGTGC GCGT CGCCGGCCTC TACGTCGCCCG CGCGGGCGCGA 50

CACGGACGCT GACACCGAGA CGGACCCCTGA GGACCTGCCG GACCACTCAC 100

CGATGAGGTA CTGGGGGGGC TGGAACCTTGT AGTGCCTCCT CAGTGTGCAG 150

TAGCTGTGGC CACTGTCCGA CAGGTAGAGG ACGTCCCCTG TCGTGGGGGA 200

GCTCACCCGA ACCGGTCTC GAGTCTCCG CGGTCGGTGG CCTCTGTTCC 250

TGTCGCTCCT GTGCCCCAC CACGCTCTGA CGCTCCCGTG TCTGCGGTCC 300

GGGATGACGT TCCACAACGA CGACGTGCTC CATGTACGGT TGCTGTGTCC 350

GTCGATGCAG ACGATGATGT TCATGTAGTT CCGTGCGTAG CTCCCCTGGT 400

GCCGGCGGTC GAGGATGCAC AAGCACTCTC TGAAACTCGT CGGTAAGTAG 450

TTGTTCCGAC TGTGCGAGAA CCAAGTTGTCC TTCCTGCGGT ACACCCACGG 500

GACAGACCAC AGGTAGGGGC CGGAGTIACA GTGCGACGCG AGCGTTTCGA 550

GCCACGACAC CGGTCTGCCG GTCCCTCCACC ACACCCTACT GGCCGCCCCG 600

TACGAGCACA GGTGCGGTGA CGACGTGCTA CGGGACATGG ACGTCACGCT 650

CTGGTGGACC CCTCTGGTCC TGAAGGAAAG GTTGGGGAAG GACCACGTGT 700

AGTGTCCGTT GCTCGAGATA CTGTAGGTCG ACAACGGGTC CTTCAGCGAC 750

CTCGACGACC ATCCCCTCTT CGACCAGGAC TTGACGTGGC ACACCCGACT 800

CAAATTGAGT CCACAGTGGA AACTGACCCT GATGGGTCCC TTCGTCCGTC 850

TCGCCCCATT CACCCACGGG CTCGCTGCGA GGGTCGTCTG GGTGTGTCTT 900

GAGAGGTCGT AGGACTGGTA GGTGTTGCAG TCGGTCGTGC TGGACCCGAG 950

CATACACACG TTCCGGTTGT TGCCGTAGGT CGCTAAAGCC CTCTCGTGGC 1000

ES 2 355 952 T3

TCCAGTAACA CGTACTTTTA GGGAAAGTAGT CGCAGCTCAC CGAGTTTCCT 1050

GGGTAGGACC TCCGGTGCCG TCCTCTGCTC GACCACTTCG ACGGGCACTT 1100

CGACCGTCGC ATGGGGGGCG GGCTCAAGGT CACCATGTTC CTACCTTTCC 1150

GTGACAGGCC CGCGGTGTCA GGTGTACGGG ACCACGAGTT CCTCCACTGT 1200

CTCCGGTCGT GTCCGTGGAT GTGGGAGCGG GACACCTTGA GGCACGACC 1250

GGACTCCGCG TTGTAGTCGG ACCTCGACCA CCACTTACAC GGGGGGGTCT 1300

ATGTACTCTT CCTCCGGAGG AGGGGGTTCGT AGATGAGCGC AGTGTCGGCG 1350

GTCCGGGAGT GGACGTGCCG GATGCCCCAC GGGGACGGAG AGTCGTAGGT 1400

CACCGTGACC GCCGGGACCT GTGGGACGTT CTACAAACGG GTCGCATCAG 1450

AGGCCGCCGC CGTCGTCGTT CTGGAGTACG GTGTCACGGC ACTGACCTCC 1500

CGCCACTGGT GCGTCCTACG GCACTTGGGG TAGCTCTCGG ACCTGTGGAC 1550

CTGGCTCAAA CACCTCCCTT TCTTATTCTG ACACTCGTTC GACCACTAGG 1600

TCTTACGGTT GCACAGACGG TACATGTTCA CACACCAGAG GTTGTTCAC 1650

CCGGTCCTAC TCGCCGAGTA GATGAAGATA CACTGGTGGT AGGGGCTGCC 1700

GAAGTGGTAG CTTAGGTTTCG GTAGGCTCCT CGATGATCTC CCGGTCGGCC 1750

ACGAGGACTC GACGGTTCGG CTGTGATGT TCATGCTCGT AGACGCGACC 1800

ATGGCGGAGT TGGACAGGTG CGACGTGCTA CGCGTGCCCT TGGCGAAGA 1850

CGAGCTGACG TTCTTGACAG TAGACAAGCG GTGGGGAGAC CGGCGGTCGG 1900

ACCTCCTCCA CCGTGGACCC CGCGCGGTGC GGTGCGAGTC GGACTIONTAG 1950

ES 2 355 952 T3

GGGGCGCAGC GCGGGCTCGT GCTCCCGGTG ATACACACGC TTCACGTTCT 2000

GGCCGCGTCG GTACTGTTCG TGACGGTGTT CTTCATGGAC AGCCACGTCC 2050

GGGACCTTCG GGGAGCCGAG TCGTCTTGA ACTGGCTGGA GGACCACTTG 2100

CACTCGCTGA GCGACCTCTA CGTCACGAAC CACCGGCCTC GCGTGCGCGG 2150

GTCGTAGCAC ACCATGTTTC TGCTCTCCGA CGACCTCCTT TTCAGACCTC 2200

AGCTGAACCG CCTGAGGTTG GTCTTCGACT CGTAGGTCGC GCACGCGCTC 2250

CTCCTACGCC CTGCGATAGA CACGTGCGAC ACGTTGCGGT TCCCGACGCA 2300

GTTGAGGAGG CGGTCGCACC GGCACCTTCC GAGGCTCCTA TTCCCGTCGT 2350

ACCTCTAGCA CTAGGAACAG CCATGGCCGC AGTAGCGACA GAAGAAGACC 2400

CAGGAGGAGG AGGAGTAGAA GACATTGTAC TCCTCCGGCC GGGTGCGTCT 2450

GTAGTTCTGC CCGATGGACA GGTAGTAGTA CCTGGGGCCC CTCCACGGAG 2500

ACCTCCTCGT TACGCTTATG GACAGGATGC TACGGTCGGT CACCCTTAAG 2550

GGGGCTCTCG CCGACGTGGA CCCCTCTCAC GAGCCGATGC CGCGGAAGCC 2600

CTTCCACCAC CTTGAGGAGC GAAAGCCGTA GGTGTTCCCG TCGTCGACAC 2650

TGTGGCACCG GCACTTTTAC GACTTTCTCC CGCGGTGCCG GTCGCTCGTG 2700

GCGCGCGACT ACAGCCTCGA GTTCTAGGAG TAAGTGTAGC CGTTGGTGGA 2750

GTTGCACCAG TTGGAGGAGC CCCGCACGTG GTTCGGCGTC CCGGGGGAGT 2800

ACCACTAGCA CCTCAAGACG TTCATGCCGT TGGAGAGGTT GAAGGACGCG 2850

CGGTTGCCCC TGCGGAAGTC GGGGACGCGC CTCTTCAGAG GGCTCGTCGC 2900

ES 2 355 952 T3

GCCTGCGAAG GCGCGGTACC ACCTCGAGCG GTCCGACCTA GCCTCCGCCG 2950

GCCCCTCGTC GCTGTCCCAG GAGAAGCGCG CCAAGAGCTT CTGGCTCCCG 3000

CCTCGCTCCG CCCGAAGAGG TCTGGTTCTT CGACTCCTGG ACACCGACTC 3050

GGGCGACTGG TACCTTCTAG AACAGACGAT GTCGAAGGTC CACCGGTCTC 3100

CCTACCTCAA GGACCGAAGG GCTTTCACGT AGGTGTCTCT GGACCGACGA 3150

GCCTTGTAAG ACGACAGCCT TTCGCTGCAC CACTTCTAGA CACTGAAACC 3200

GGAACGGGCC CTGTAGATGT TTCTGGGACT GATGCAGGCG TTCCCGTCAC 3250

GGGCCGACGG GGACTTCACC TACCGGGGAC TTTCGTAGAA GCTGTTCAC 3300

ATGTGGTGCG TCTCACTGCA CACCAGGAAA CCCACGAAG AGACCCTCTA 3350

GAAGAGAGAC CCCC GGAGGG GCATGGGACC CCACGTCTAG TTA CTCTCA 3400

AGACGGTCCG CGACTCTCTG CCGTGTTCCT ACTCCCGGGG CCTCGACCGG 3450

TGAGGGCGGT ATGCGGCGTA GTACGACTTG ACGACCAGGC CTCTGGGGTT 3500

CCGCTCTGGA CGTAAGAGCC TCGACCACCT CTAGGACCCC CTGGACGAGG 3550

TCCCGTCCCC GGACGTCTC CTTCTCCTCC AGACGTACCG GGGCGGTCG 3600

AGAGTCTCGA GTCTTCTCCC GTCGAAGAGC GTCCACAGGT GGTACCGGGA 3650

TGTGTAGCGG GTCCGACTGC GACTCCTGTC GGGCGGTTCG GACGTCCGG 3700

TGTCGGACCG GCGGTCCATA ATGTTGACCC ACAGGAAAGG GCCACGGAC 3750

CGGTCTCCCC GACTCTGGGC ACCAAGGAGG TCCTACTTCT GTAAACTCCT 3800

TAAGGGGTAC TGGGGTTGCT GGATGTTTCC GAGACACCTG TTGGTCTGTC 3850

TGTCACCCTA CCACGACCGG AGCCTCCTCA AACTCGTCTA TCTCTCGTCC 3900

GTATCTGTTC TTTCGCCGAA GTCCATCGAC TTCGTCTCTC TCTCTTCCGT 3950

CGTATGCAGT CGTAAAAGAA GAGACGTGAA TATTCTTTCT AGTTTCTGAA 4000

5 ATTCTGAAAG CGATAAAGAA GACGATAGAT GATGTTTGAA GTTTCTCCTT 4050

GGTCCTCCGG TTCTCCTCGT ACTTTCACCT GTTCCTCACA CTGGTGACTT 4100

CGTGGTGTCC CTCCCCAATC CGGAGGCCTA CTGACGCCCG TCCGGACCTA 4150

10 TTATAGGTCG GAGGGTGTTT TTCGACCACC TCGTCTCACA AGGGACTGAG 4200

GAGGTTCCCTT TCCCTCTGCG GGAAAGTACC AGACGACTCA TTGTCCACGG 4250

15 AAGGGTCTGT GACCGCAATG ACGAACTGGT TTCTCGGGAG TTCGCCGGGA 4300

ATACGGTCGC ACTGTCTCCC GAGTGGAGAA CGGAAGATCC AGTGAAGAGT 4350

GTTACAGGGA AGTCGTGGAC TGGGACACGG GCGGTCAATA AGGAACCATT 4400

20 ATACTCATT TGTAGTTTCT CATCA 4425

(2) INFORMACIÓN PARA LA SEC ID NO:33:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1298 aminoácidos

25 (B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:33:

Met	Gln	Arg	Gly	Ala	Ala	Leu	Cys	Leu	Arg	Leu	Trp	Leu	Cys	Leu
1				5					10					15
Gly	Leu	Leu	Asp	Gly	Leu	Val	Ser	Gly	Tyr	Ser	Met	Thr	Pro	Pro
				20					25					30
Thr	Leu	Asn	Ile	Thr	Glu	Glu	Ser	His	Val	Ile	Asp	Thr	Gly	Asp
				35					40					45
Ser	Leu	Ser	Ile	Ser	Cys	Arg	Gly	Gln	His	Pro	Leu	Glu	Trp	Ala
				50					55					60

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

Thr Ala Gly Asp Glu Leu Val Lys Leu Pro Val Lys Leu Ala Ala
 350 355 360
 Tyr Pro Pro Pro Glu Phe Gln Trp Tyr Lys Asp Gly Lys Ala Leu
 365 370 375
 Ser Gly Arg His Ser Pro His Ala Leu Val Leu Lys Glu Val Thr
 380 385 390
 Glu Ala Ser Thr Gly Thr Tyr Thr Leu Ala Leu Trp Asn Ser Ala
 395 400 405
 Ala Gly Leu Arg Arg Asn Ile Ser Leu Glu Leu Val Val Asn Val
 410 415 420
 Pro Pro Gln Ile His Glu Lys Glu Ala Ser Ser Pro Ser Ile Tyr
 425 430 435
 Ser Arg His Ser Arg Gln Ala Leu Thr Cys Thr Ala Tyr Gly Val
 440 445 450
 Pro Leu Pro Leu Ser Ile Gln Trp His Trp Arg Pro Trp Thr Pro
 455 460 465
 Cys Lys Met Phe Ala Gln Arg Ser Leu Arg Arg Arg Gln Gln Gln
 470 475 480
 Asp Leu Met Pro Gln Cys Arg Asp Trp Arg Ala Val Thr Thr Gln
 485 490 495
 Asp Ala Val Asn Pro Ile Glu Ser Leu Asp Thr Trp Thr Glu Phe
 500 505 510
 Val Glu Gly Lys Asn Lys Thr Val Ser Lys Leu Val Ile Gln Asn
 515 520 525
 Ala Asn Val Ser Ala Met Tyr Lys Cys Val Val Ser Asn Lys Val
 530 535 540
 Gly Gln Asp Glu Arg Leu Ile Tyr Phe Tyr Val Thr Thr Ile Pro
 545 550 555
 Asp Gly Phe Thr Ile Glu Ser Lys Pro Ser Glu Glu Leu Leu Glu
 560 565 570
 Gly Gln Pro Val Leu Leu Ser Cys Gln Ala Asp Ser Tyr Lys Tyr
 575 580 585
 Glu His Leu Arg Trp Tyr Arg Leu Asn Leu Ser Thr Leu His Asp
 590 595 600
 Ala His Gly Asn Pro Leu Leu Leu Asp Cys Lys Asn Val His Leu
 605 610 615
 Phe Ala Thr Pro Leu Ala Ala Ser Leu Glu Glu Val Ala Pro Gly
 620 625 630

Ala Arg His Ala Thr Leu Ser Leu Ser Ile Pro Arg Val Ala Pro
635 640 645

Glu His Glu Gly His Tyr Val Cys Glu Val Gln Asp Arg Arg Ser
650 655 660

His Asp Lys His Cys His Lys Lys Tyr Leu Ser Val Gln Ala Leu
665 670 675

Glu Ala Pro Arg Leu Thr Gln Asn Leu Thr Asp Leu Leu Val Asn
680 685 690

Val Ser Asp Ser Leu Glu Met Gln Cys Leu Val Ala Gly Ala His
695 700 705

Ala Pro Ser Ile Val Trp Tyr Lys Asp Glu Arg Leu Leu Glu Glu
710 715 720

Lys Ser Gly Val Asp Leu Ala Asp Ser Asn Gln Lys Leu Ser Ile
725 730 735

Gln Arg Val Arg Glu Glu Asp Ala Gly Arg Tyr Leu Cys Ser Val
740 745 750

Cys Asn Ala Lys Gly Cys Val Asn Ser Ser Ala Ser Val Ala Val
755 760 765

Glu Gly Ser Glu Asp Lys Gly Ser Met Glu Ile Val Ile Leu Val
770 775 780

Gly Thr Gly Val Ile Ala Val Phe Phe Trp Val Leu Leu Leu Leu
785 790 795

Ile Phe Cys Asn Met Arg Arg Pro Ala His Ala Asp Ile Lys Thr
800 805 810

Gly Tyr Leu Ser Ile Ile Met Asp Pro Gly Glu Val Pro Leu Glu
815 820 825

Glu Gln Cys Glu Tyr Leu Ser Tyr Asp Ala Ser Gln Trp Glu Phe
830 835 840

Pro Arg Glu Arg Leu His Leu Gly Arg Val Leu Gly Tyr Gly Ala
845 850 855

Phe Gly Lys Val Val Glu Ala Ser Ala Phe Gly Ile His Lys Gly
860 865 870

Ser Ser Cys Asp Thr Val Ala Val Lys Met Leu Lys Glu Gly Ala
875 880 885

Thr Ala Ser Glu His Arg Ala Leu Met Ser Glu Leu Lys Ile Leu
890 895 900

Ile His Ile Gly Asn His Leu Asn Val Val Asn Leu Leu Gly Ala
905 910 915

Cys Thr Lys Pro Gln Gly Pro Leu Met Val Ile Val Glu Phe Cys
 920 925 930
 Lys Tyr Gly Asn Leu Ser Asn Phe Leu Arg Ala Lys Arg Asp Ala
 935 940 945
 Phe Ser Pro Cys Ala Glu Lys Ser Pro Glu Gln Arg Gly Arg Phe
 950 955 960
 Arg Ala Met Val Glu Leu Ala Arg Leu Asp Arg Arg Arg Pro Gly
 965 970 975
 Ser Ser Asp Arg Val Leu Phe Ala Arg Phe Ser Lys Thr Glu Gly
 980 985 990
 Gly Ala Arg Arg Ala Ser Pro Asp Gln Glu Ala Glu Asp Leu Trp
 995 1000 1005
 Leu Ser Pro Leu Thr Met Glu Asp Leu Val Cys Tyr Ser Phe Gln
 1010 1015 1020
 Val Ala Arg Gly Met Glu Phe Leu Ala Ser Arg Lys Cys Ile His
 1025 1030 1035
 Arg Asp Leu Ala Ala Arg Asn Ile Leu Leu Ser Glu Ser Asp Val
 1040 1045 1050
 Val Lys Ile Cys Asp Phe Gly Leu Ala Arg Asp Ile Tyr Lys Asp
 1055 1060 1065
 Pro Asp Tyr Val Arg Lys Gly Ser Ala Arg Leu Pro Leu Lys Trp
 1070 1075 1080
 Met Ala Pro Glu Ser Ile Phe Asp Lys Val Tyr Thr Thr Gln Ser
 1085 1090 1095
 Asp Val Trp Ser Phe Gly Val Leu Leu Trp Glu Ile Phe Ser Leu
 1100 1105 1110
 Gly Ala Ser Pro Tyr Pro Gly Val Gln Ile Asn Glu Glu Phe Cys
 1115 1120 1125
 Gln Arg Leu Arg Asp Gly Thr Arg Met Arg Ala Pro Glu Leu Ala
 1130 1135 1140
 Thr Pro Ala Ile Arg Arg Ile Met Leu Asn Cys Trp Ser Gly Asp
 1145 1150 1155
 Pro Lys Ala Arg Pro Ala Phe Ser Glu Leu Val Glu Ile Leu Gly
 1160 1165 1170
 Asp Leu Leu Gln Gly Arg Gly Leu Gln Glu Glu Glu Glu Val Cys
 1175 1180 1185
 Met Ala Pro Arg Ser Ser Gln Ser Ser Glu Glu Gly Ser Phe Ser
 1190 1195 1200

Gln Val Ser Thr Met Ala Leu His Ile Ala Gln Ala Asp Ala Glu
 1205 1210 1215

Asp Ser Pro Pro Ser Leu Gln Arg His Ser Leu Ala Ala Arg Tyr
 1220 1225 1230

5 Tyr Asn Trp Val Ser Phe Pro Gly Cys Leu Ala Arg Gly Ala Glu
 1235 1240 1245

Thr Arg Gly Ser Ser Arg Met Lys Thr Phe Glu Glu Phe Pro Met
 1250 1255 1260

Thr Pro Thr Thr Tyr Lys Gly Ser Val Asp Asn Gln Thr Asp Ser
 1265 1270 1275

10 Gly Met Val Leu Ala Ser Glu Glu Phe Glu Gln Ile Glu Ser Arg
 1280 1285 1290

His Arg Gln Glu Ser Gly Phe Arg
 1295 1298

(2) INFORMACIÓN PARA LA SEC ID NO:34:

15 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 3348 bases
- (B) TIPO: ácido nucleico
- (C) CADENA: única
- (D) TOPOLOGÍA: lineal

20 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:34:

ATGGCTGGGA TTTTCTATTT CGCCCTATTT TCGTGTCTCT TCGGGATTTG 50

CGACGCTGTC ACAGGTTCCA GGGTATACCC CGCGAATGAA GTTACCTTAT 100

TGGATTCCAG ATCTGTTTCAG GGAGAACTTG GGTGGATAGC AAGCCCTCTG 150

GAAGGAGGGT GGGAGGAAGT GAGTATCATG GATGAAAAAA ATACACCAAT 200

CCGAACCTAC CAAGTGTGCA ATGTGATGGA ACCCAGCCAG AATAACTGGC 250

TACGAACTGA TTGGATCACC CGAGAAGGGG CTCAGAGGGT GTATATTGAG 300

ATTAAATTCA CCTTGAGGGA CTGCAATAGT CTTCCGGGCG TCATGGGGAC 350

TTGCAAGGAG ACGTTTAACC TGTACTACTA TGAATCAGAC AACGACAAAG 400

AGCGTTTCAT CAGAGAGAAC CAGTTTGTC AATTGACAC CATTGCTGCT 450

ES 2 355 952 T3

GATGAGAGCT TCACCCAAGT GGACATTGGT GACAGAATCA TGAAGCTGAA 500

CACCGAGATC CGGGATGTAG GGCCATTAAG CAAAAGGGG TTTTACCTGG 550

CTTTTCAGGA TGTGGGGGCC TGCATCGCCC TGGTATCAGT CCGTGTGTTC 600

TATAAAAAGT GTCCACTCAC AGTCCGCAAT CTGGCCCAGT TTCCTGACAC 650

CATCACAGGG GCTGATACGT CTTCCCTGGT GGAAGTTCGA GGCTCCTGTG 700

TCAACAACCTC AGAAGAGAAA GATGTGCCAA AAATGTACTG TGGGGCAGAT 750

GGTGAATGGC TGGTACCCAT TGGCAACTGC CTATGCAACG CTGGGCATGA 800

GGAGCGGAGC GGAGAATGCC AAGCTTGCAA AATTGGATAT TACAAGGCTC 850

TCTCCACGGA TGCCACCTGT GCCAAGTGCC CACCCACAG CTA CTCTGTGTC 900

TGGGAAGGAG CCACCTCGTG CACCTGTGAC CGAGGCTTTT TCAGAGCTGA 950

CAACGATGCT GCCTCTATGC CCTGCACCCG TCCACCATCT GCTCCCCTGA 1000

ACTTGATTTT AAATGTCAAC GAGACATCTG TGA ACTTGGGA ATGGAGTAGC 1050

CCTCAGAATA CAGGTGGCCG CCAGGACATT TCCTATAATG TGGTATGCAA 1100

GAAATGTGGA GCTGGTGACC CCAGCAAGTG CCGACCCTGT GGAAGTGGGG 1150

TCCACTACAC CCCACAGCAG AATGGCTTGA AGACCACCAA AGGCTCCATC 1200

ACTGACCTCC TAGCTCATA C AATTACACC TTTGAAATCT GGGCTGTGAA 1250

TGGAGTGTCC AAATATAACC CTAACCCAGA CCAATCAGTT TCTGTCACTG 1300

TGACCACCAA CCAAGCAGCA CCATCATCCA TTGCTTTGGT CCAGGCTAAA 1350

GAAGTCACAA GATACAGTGT GGC ACTGGCT TGGCTGGAAC CAGATCGGCC 1400

CAATGGGGTA ATCCTGGAAT ATGAAGTCAA GTATTATGAG AAGGATCAGA 1450
 ATGAGCGAAG CTATCGTATA GTTCGGACAG CTGCCAGGAA CACAGATATC 1500
 AAAGGCCTGA ACCCTCTCAC TTCCTATGTT TTCCACGTGC GAGCCAGGAC 1550
 AGCAGCTGGC TATGGAGACT TCAGTGAGCC CTTGGAGGTT ACAACCAACA 1600
 CAGTGCCTTC CCGGATCATT GGAGATGGGG CTAACTCCAC AGTCCTTCTG 1650
 GTCTCTGTCT CGGGCAGTGT GGTGCTGGTG GTAATTCTCA TTGCAGCTTT 1700
 TGTCATCAGC CGGAGACGGA GTAAATACAG TAAAGCCAAA CAAGAAGCGG 1750
 ATGAAGAGAA ACATTTGAAT CAAGGTGTAA GAACATATGT GGACCCCTTT 1800
 ACGTACGAAG ATCCCAACCA AGCAGTGCGA GAGTTTGCCA AAGAAATTGA 1850
 CGCATCCTGC ATTAAGATTG AAAAAGTTAT AGGAGTTGGT GAATTTGGTG 1900
 AGGTATGCAG TGGGCGTCTC AAAGTGCCTG GCAAGAGAGA GATCTGTGTG 1950
 GCTATCAAGA CTCTGAAAGC TGGTTATACA GACAAACAGA GGAGAGACTT 2000
 CCTGAGTGAG GCCAGCATCA TGGGACAGTT TGACCATCCG AACATCATT 2050
 ACTTGGAAGG CGTGGTCACT AAATGTAAAC CAGTAATGAT CATAACAGAG 2100
 TACATGGAGA ATGGCTCCTT GGATGCATT CTCAGGAAAA ATGATGGCAG 2150
 ATTTACAGTC ATTCAGCTGG TGGGCATGCT TCGTGGCATT GGGTCTGGGA 2200
 TGAAGTATTT ATCTGATATG AGCTATGTGC ATCGTGATCT GGCCGCACGG 2250
 AACATCCTGG TGAACAGCAA CTTGGTCTGC AAAGTGTCTG ATTTTGGCAT 2300
 GTCCCGAGTG CTTGAGGATG ATCCGGAAGC AGCTTACACC ACCAGGGGTG 2350

GCAAGATTCC TATCCGGTGG ACTGCGCCAG AAGCAATTGC CTATCGTAAA 2400
 TTCACATCAG CAAGTGATGT ATGGAGCTAT GGAATCGTTA TGTGGGAAGT 2450
 GATGTCGTAC GGGGAGAGGC CCTATTGGGA TATGTCCAAT CAAGATGTGA 2500
 TTAAAGCCAT TGAGGAAGGC TATCGGTTAC CCCCTCCAAT GGACTGCCCC 2550
 ATTGCGCTCC ACCAGCTGAT GCTAGACTGC TGGCAGAAGG AGAGGAGCGA 2600
 CAGGCCTAAA TTTGGGCAGA TTGTCAACAT GTTGGACAAA CTCATCCGCA 2650
 ACCCCAACAG CTTGAAGAGG ACAGGGACGG AGAGCTCCAG ACCTAACACT 2700
 GCCTTGTTGG ATCCAAGCTC CCCTGAATTC TCTGCTGTGG TATCAGTGGG 2750
 CGATTGGCTC CAGGCCATTA AAATGGACCG GTATAAGGAT AACTTCACAG 2800
 CTGCTGGTTA TACCACACTA GAGGCTGTGG TGCACGTGAA CCAGGAGGAC 2850
 CTGGCAAGAA TTGGTATCAC AGCCATCACA CACCAGAATA AGATTTTGAG 2900
 CAGTGTCCAG GCAATGCGAA CCCAAATGCA GCAGATGCAC GGCAGAATGG 2950
 TTCCCGTCTG AGCCAGTACT GAATAAACTC AAAACTCTTG AAATTAGTTT 3000
 ACCTCATCCA TGCACTTTAA TTGAAGAACT GCACTTTTTT TACTTCGTCT 3050
 TCGCCCTCTG AAATTAAAGA AATGAAAAAA AAAAAACAAT ATCTGCAGCG 3100
 TTGCTTGGTG CACAGATTGC TGAAACTGTG GGGCTTACAG AAATGACTGC 3150
 CGGTCATTTG AATGAGACCT GGAACAAATC GTTCTCAGA AGTACTTTTC 3200
 TGTTCATCAC CAGTCTGTAA AATACATGTA CCTATAGAAA TAGAACACTG 3250
 CCTCTGAGTT TTGATGCTGT ATTTGCTGCC AGACACTGAG CTTCTGAGAC 3300

ATCCCTGATT CTCTCTCCAT TTGGAATTAC AACGGTCGAC GAGCTCGA 3348

(2) INFORMACIÓN PARA LA SEC ID NO:35:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 3348 bases

(B) TIPO: ácido nucleico

5 (C) CADENA: única

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:35:

TACCGACCCT AAAAGATAAA GCGGGATAAA AGCACAGAGA AGCCCTAAAC 50

GCTGCGACAG TGTCCAAGGT CCCATATGGG GCGCTTACTT CAATGGAATA 100

ACCTAAGGTC TAGACAAGTC CCTCTTGAAC CCACCTATCG TTCGGGAGAC 150

CTTCCTCCCA CCCTCCTTCA CTCATAGTAC CTACTTTTTT TATGTGGTTA 200

GGCTTGGATG GTTCACACGT TACACTACCT TGGGTCGGTC TTATTGACCG 250

ATGCTTGACT AACCTAGTGG GCTCTTCCCC GAGTCTCCCA CATATAACTC 300

TAATTTAAGT GGAACTCCCT GACGTTATCA GAAGGCCCGC AGTACCCCTG 350

AACGTTCCCTC TGCAAATTGG ACATGATGAT ACTTAGTCTG TTGCTGTTTC 400

TCGCAAAGTA GTCTCTCTTG GTCAAACAGT TTTAACTGTG GTAACGACGA 450

CTACTCTCGA AGTGGGTTC A CTGTAACCA CTGTCTTAGT ACTTCGACTT 500

GTGGCTCTAG GCCCTACATC CCGGTAATTC GTTTTTCCCC AAAATGGACC 550

GAAAAGTCCT ACACCCCCGG ACGTAGCGGG ACCATAGTCA GGCACACAAG 600

ATATTTTTCA CAGGTGAGTG TCAGGCGTTA GACCGGGTCA AAGGACTGTG 650

GTAGTGTCCC CGACTATGCA GAAGGGACCA CCTTCAAGCT CCGAGGACAC 700

ES 2 355 952 T3

AGTTGTTGAG TCTTCTCTTT CTACACGGTT TTTACATGAC ACCCCGTCTA 750
 CCACTTACCG ACCATGGGTA ACCGTTGACG GATACGTTGC GACCCGTACT 800
 CCTCGCCTCG CCTCTTACGG TTCGAACGTT TTAACCTATA ATGTTCCGAG 850
 AGAGGTGCCT ACGGTGGACA CGGTTACGG GTGGGGTGTC GATGAGACAG 900
 ACCCTTCCTC GGTGGAGCAC GTGGACACTG GCTCCGAAAA AGTCTCGACT 950
 GTTGCTACGA CGGAGATACG GGACGTGGGC AGGTGGTAGA CGAGGGGACT 1000
 TGAACTAAAG TTTACAGTTG CTCTGTAGAC ACTTGAACCT TACCTCATCG 1050
 GGAGTCTTAT GTCCACCGGC GGTCCGTGTA AGGATATTAC ACCATACGTT 1100
 CTTTACACCT CGACCACTGG GGTGTTTAC GGCTGGGACA CCTTCACCCC 1150
 AGGTGATGTG GGGTGTGTC TTACCGAACT TCTGGTGGTT TCCGAGGTAG 1200
 TGACTGGAGG ATCGAGTATG GTTAATGTGG AACTTTAGA CCCGACACTT 1250
 ACCTCACAGG TTTATATTGG GATTGGGTCT GGTTAGTCAA AGACAGTGAC 1300
 ACTGGTGGTT GGTTCGTCGT GGTAGTAGGT AACGAAACCA GGTCCGATTT 1350
 CTTCAGTGTT CTATGTCACA CCGTGACCGA ACCGACCTTG GTCTAGCCGG 1400
 GTTACCCCAT TAGGACCTTA TACTTCAGTT CATAATACTC TTCCTAGTCT 1450
 TACTCGCTTC GATAGCATAT CAAGCCTGTC GACGGTCCTT GTGTCTATAG 1500
 TTTCCGGACT TGGGAGAGTG AAGGATACAA AAGGTGCACG CTCGGTCCTG 1550
 TCGTCGACCG ATACCTCTGA AGTCACTCGG GAACCTCCAA TGTTGGTTGT 1600
 GTCACGGAAG GGCCTAGTAA CCTCTACCCC GATTGAGGTG TCAGGAAGAC 1650

ES 2 355 952 T3

CAGAGACAGA GCCCGTCACA CCACGACCAC CATTAAAGAGT AACGTCGAAA 1700
 ACAGTAGTCG GCCTCTGCCT CATTATATGTC ATTTTCGGTTT GTTCTTCGCC 1750
 TACTTCTCTT TGTAAACTTA GTTCCACATT CTTGTATACA CCTGGGGAAA 1800
 TGCATGCTTC TAGGGTTGGT TCGTCACGCT CTCAAACGGT TTCTTTAACT 1850
 GCGTAGGACG TAATTCTAAC TTTTCAATA TCCTCAACCA CTTAAACCAC 1900
 TCCATACGTC ACCCGCAGAG TTTCACGGAC CGTTCTCTCT CTAGACACAC 1950
 CGATAGTTCT GAGACTTTCG ACCAATATGT CTGTTTGTCT CCTCTCTGAA 2000
 GGACTCACTC CGGTCGTAGT ACCCTGTCAA ACTGGTAGGC TTGTAGTAAG 2050
 TGAACCTTC GCACCAGTGA TTTACATTG GTCATTACTA GTATTGTCTC 2100
 ATGTACCTCT TACCGAGGAA CCTACGTAAG GAGTCCTTTT TACTACCGTC 2150
 TAAATGTCAG TAAGTCGACC ACCCGTACGA AGCACCGTAA CCCAGACCCT 2200
 ACTTCATAAA TAGACTATAC TCGATACACG TAGCACTAGA CCGGCGTGCC 2250
 TTGTAGGACC ACTTGTCGTT GAACCAGACG TTTCACAGAC TAAAACCGTA 2300
 CAGGGCTCAC GAACTCCTAC TAGGCCTTCG TCGAATGTGG TGGTCCCCAC 2350
 CGTTCTAAGG ATAGGCCACC TGACGCGGTC TTCGTTAACG GATAGCATTT 2400
 AAGTGTAGTC GTTCACTACA TACCTCGATA CCTTAGCAAT ACACCCTTCA 2450
 CTACAGCATG CCCCTCTCCG GGATAACCCT ATACAGGTTA GTTCTACTACT 2500
 AATTTCCGTA ACTCCTTCCG ATAGCCAATG GGGGAGGTTA CCTGACGGGG 2550
 TAACGCGAGG TGGTCGACTA CGATCTGACG ACCGTCTTCC TCTCCTCGCT 2600

GTCCGGATT AAACCCGTCT AACAGTTGTA CAACCTGTTT GAGTAGGCGT 2650
 TGGGGTTGTC GAACTTCTCC TGTCCCTGCC TCTCGAGGTC TGGATTGTGA 2700
 CGGAACAACC TAGGTTTCGAG GGGACTTAAG AGACGACACC ATAGTCACCC 2750
 5 GCTAACCGAG GTCCGGTAAT TTTACCTGGC CATATTCCTA TTGAAGTGTC 2800
 GACGACCAAT ATGGTGTGAT CTCCGACACC ACGTGCACTT GGCCTCCTG 2850
 GACCGTTCTT AACCATAGTG TCGGTAGTGT GTGGTCTTAT TCTAAACTC 2900
 10 GTCACAGGTC CGTTACGCTT GGGTTTACGT CGTCTACGTG CCGTCTTACC 2950
 AAGGGCAGAC TCGGTCATGA CTTATTTGAG TTTTGAGAAC TTTAATCAA 3000
 TGGAGTAGGT ACGTGAAATT AACTTCTTGA CGTGAAAAAA ATGAAGCAGA 3050
 15 AGCGGGAGAC TTTAATTTCT TTTACTTTTT TTTTTTGTTA TAGACGTCGC 3100
 AACGAACCAC GTGTCTAACG ACTTTGACAC CCCGAATGTC TTTACTGACG 3150
 20 GCCAGTAAAC TTTACTCTGGA CCTTGTTTGA CAAAGAGTCT TCATGAAAAG 3200
 ACAAGTAGTG GTCAGACATT TTATGTACAT GGATATCTTT ATCTTGTGAC 3250
 GGAGACTCAA AACTACGACA TAAACGACGG TCTGTGACTC GAAGACTCTG 3300
 25 TAGGGACTAA GAGAGAGGTA AACCTTAATG TTGCCAGCTG CTCGAGCT 3348

(2) INFORMACIÓN PARA LA SEC ID NO:36:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1104 aminoácidos

30 (B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:36:

Met Ala Gly Ile Phe Tyr Phe Ala Leu Phe Ser Cys Leu Phe Gly
 1 5 10 15

Ile	Cys	Asp	Ala	Val	Thr	Gly	Ser	Arg	Val	Tyr	Pro	Ala	Asn	Glu
				20					25					30
Val	Thr	Leu	Leu	Asp	Ser	Arg	Ser	Val	Gln	Gly	Glu	Leu	Gly	Trp
				35					40					45
Ile	Ala	Ser	Pro	Leu	Glu	Gly	Gly	Trp	Glu	Glu	Val	Ser	Ile	Met
				50					55					60
Asp	Glu	Lys	Asn	Thr	Pro	Ile	Arg	Thr	Tyr	Gln	Val	Cys	Asn	Val
				65					70					75
Met	Glu	Pro	Ser	Gln	Asn	Asn	Trp	Leu	Arg	Thr	Asp	Trp	Ile	Thr
				80					85					90
Arg	Glu	Gly	Ala	Gln	Arg	Val	Tyr	Ile	Glu	Ile	Lys	Phe	Thr	Leu
				95					100					105
Arg	Asp	Cys	Asn	Ser	Leu	Pro	Gly	Val	Met	Gly	Thr	Cys	Lys	Glu
				110					115					120
Thr	Phe	Asn	Leu	Tyr	Tyr	Tyr	Glu	Ser	Asp	Asn	Asp	Lys	Glu	Arg
				125					130					135
Phe	Ile	Arg	Glu	Asn	Gln	Phe	Val	Lys	Ile	Asp	Thr	Ile	Ala	Ala
				140					145					150
Asp	Glu	Ser	Phe	Thr	Gln	Val	Asp	Ile	Gly	Asp	Arg	Ile	Met	Lys
				155					160					165
Leu	Asn	Thr	Glu	Ile	Arg	Asp	Val	Gly	Pro	Leu	Ser	Lys	Lys	Gly
				170					175					180
Phe	Tyr	Leu	Ala	Phe	Gln	Asp	Val	Gly	Ala	Cys	Ile	Ala	Leu	Val
				185					190					195
Ser	Val	Arg	Val	Phe	Tyr	Lys	Lys	Cys	Pro	Leu	Thr	Val	Arg	Asn
				200					205					210
Leu	Ala	Gln	Phe	Pro	Asp	Thr	Ile	Thr	Gly	Ala	Asp	Thr	Ser	Ser
				215					220					225
Leu	Val	Glu	Val	Arg	Gly	Ser	Cys	Val	Asn	Asn	Ser	Glu	Glu	Lys
				230					235					240
Asp	Val	Pro	Lys	Met	Tyr	Cys	Gly	Ala	Asp	Gly	Glu	Trp	Leu	Val
				245					250					255
Pro	Ile	Gly	Asn	Cys	Leu	Cys	Asn	Ala	Gly	His	Glu	Glu	Arg	Ser
				260					265					270
Gly	Glu	Cys	Gln	Ala	Cys	Lys	Ile	Gly	Tyr	Tyr	Lys	Ala	Leu	Ser
				275					280					285
Thr	Asp	Ala	Thr	Cys	Ala	Lys	Cys	Pro	Pro	His	Ser	Tyr	Ser	Val
				290					295					300

Trp	Glu	Gly	Ala	Thr	Ser	Cys	Thr	Cys	Asp	Arg	Gly	Phe	Phe	Arg
				305					310					315
Ala	Asp	Asn	Asp	Ala	Ala	Ser	Met	Pro	Cys	Thr	Arg	Pro	Pro	Ser
				320					325					330
Ala	Pro	Leu	Asn	Leu	Ile	Ser	Asn	Val	Asn	Glu	Thr	Ser	Val	Asn
				335					340					345
Leu	Glu	Trp	Ser	Ser	Pro	Gln	Asn	Thr	Gly	Gly	Arg	Gln	Asp	Ile
				350					355					360
Ser	Tyr	Asn	Val	Val	Cys	Lys	Lys	Cys	Gly	Ala	Gly	Asp	Pro	Ser
				365					370					375
Lys	Cys	Arg	Pro	Cys	Gly	Ser	Gly	Val	His	Tyr	Thr	Pro	Gln	Gln
				380					385					390
Asn	Gly	Leu	Lys	Thr	Thr	Lys	Gly	Ser	Ile	Thr	Asp	Leu	Leu	Ala
				395					400					405
His	Thr	Asn	Tyr	Thr	Phe	Glu	Ile	Trp	Ala	Val	Asn	Gly	Val	Ser
				410					415					420
Lys	Tyr	Asn	Pro	Asn	Pro	Asp	Gln	Ser	Val	Ser	Val	Thr	Val	Thr
				425					430					435
Thr	Asn	Gln	Ala	Ala	Pro	Ser	Ser	Ile	Ala	Leu	Val	Gln	Ala	Lys
				440					445					450
Glu	Val	Thr	Arg	Tyr	Ser	Val	Ala	Leu	Ala	Trp	Leu	Glu	Pro	Asp
				455					460					465
Arg	Pro	Asn	Gly	Val	Ile	Leu	Glu	Tyr	Glu	Val	Lys	Tyr	Tyr	Glu
				470					475					480
Lys	Asp	Gln	Asn	Glu	Arg	Ser	Tyr	Arg	Ile	Val	Arg	Thr	Ala	Ala
				485					490					495
Arg	Asn	Thr	Asp	Ile	Lys	Gly	Leu	Asn	Pro	Leu	Thr	Ser	Tyr	Val
				500					505					510
Phe	His	Val	Arg	Ala	Arg	Thr	Ala	Ala	Gly	Tyr	Gly	Asp	Phe	Ser
				515					520					525
Glu	Pro	Leu	Glu	Val	Thr	Thr	Asn	Thr	Val	Pro	Ser	Arg	Ile	Ile
				530					535					540
Gly	Asp	Gly	Ala	Asn	Ser	Thr	Val	Leu	Leu	Val	Ser	Val	Ser	Gly
				545					550					555
Ser	Val	Val	Leu	Val	Val	Ile	Leu	Ile	Ala	Ala	Phe	Val	Ile	Ser
				560					565					570
Arg	Arg	Arg	Ser	Lys	Tyr	Ser	Lys	Ala	Lys	Gln	Glu	Ala	Asp	Glu
				575					580					585

Glu	Lys	His	Leu	Asn	Gln	Gly	Val	Arg	Thr	Tyr	Val	Asp	Pro	Phe
				590					595					600
Thr	Tyr	Glu	Asp	Pro	Asn	Gln	Ala	Val	Arg	Glu	Phe	Ala	Lys	Glu
				605					610					615
Ile	Asp	Ala	Ser	Cys	Ile	Lys	Ile	Glu	Lys	Val	Ile	Gly	Val	Gly
				620					625					630
Glu	Phe	Gly	Glu	Val	Cys	Ser	Gly	Arg	Leu	Lys	Val	Pro	Gly	Lys
				635					640					645
Arg	Glu	Ile	Cys	Val	Ala	Ile	Lys	Thr	Leu	Lys	Ala	Gly	Tyr	Thr
				650					655					660
Asp	Lys	Gln	Arg	Arg	Asp	Phe	Leu	Ser	Glu	Ala	Ser	Ile	Met	Gly
				665					670					675
Gln	Phe	Asp	His	Pro	Asn	Ile	Ile	His	Leu	Glu	Gly	Val	Val	Thr
				680					685					690
Lys	Cys	Lys	Pro	Val	Met	Ile	Ile	Thr	Glu	Tyr	Met	Glu	Asn	Gly
				695					700					705
Ser	Leu	Asp	Ala	Phe	Leu	Arg	Lys	Asn	Asp	Gly	Arg	Phe	Thr	Val
				710					715					720
Ile	Gln	Leu	Val	Gly	Met	Leu	Arg	Gly	Ile	Gly	Ser	Gly	Met	Lys
				725					730					735
Tyr	Leu	Ser	Asp	Met	Ser	Tyr	Val	His	Arg	Asp	Leu	Ala	Ala	Arg
				740					745					750
Asn	Ile	Leu	Val	Asn	Ser	Asn	Leu	Val	Cys	Lys	Val	Ser	Asp	Phe
				755					760					765
Gly	Met	Ser	Arg	Val	Leu	Glu	Asp	Asp	Pro	Glu	Ala	Ala	Tyr	Thr
				770					775					780
Thr	Arg	Gly	Gly	Lys	Ile	Pro	Ile	Arg	Trp	Thr	Ala	Pro	Glu	Ala
				785					790					795
Ile	Ala	Tyr	Arg	Lys	Phe	Thr	Ser	Ala	Ser	Asp	Val	Trp	Ser	Tyr
				800					805					810
Gly	Ile	Val	Met	Trp	Glu	Val	Met	Ser	Tyr	Gly	Glu	Arg	Pro	Tyr
				815					820					825
Trp	Asp	Met	Ser	Asn	Gln	Asp	Val	Ile	Lys	Ala	Ile	Glu	Glu	Gly
				830					835					840
Tyr	Arg	Leu	Pro	Pro	Pro	Met	Asp	Cys	Pro	Ile	Ala	Leu	His	Gln
				845					850					855
Leu	Met	Leu	Asp	Cys	Trp	Gln	Lys	Glu	Arg	Ser	Asp	Arg	Pro	Lys
				860					865					870

	Phe Gly Gln Ile Val Asn Met Leu Asp Lys Leu Ile Arg Asn Pro	875	880	885
	Asn Ser Leu Lys Arg Thr Gly Thr Glu Ser Ser Arg Pro Asn Thr	890	895	900
5	Ala Leu Leu Asp Pro Ser Ser Pro Glu Phe Ser Ala Val Val Ser	905	910	915
	Val Gly Asp Trp Leu Gln Ala Ile Lys Met Asp Arg Tyr Lys Asp	920	925	930
	Asn Phe Thr Ala Ala Gly Tyr Thr Thr Leu Glu Ala Val Val His	935	940	945
10	Val Asn Gln Glu Asp Leu Ala Arg Ile Gly Ile Thr Ala Ile Thr	950	955	960
	His Gln Asn Lys Ile Leu Ser Ser Val Gln Ala Met Arg Thr Gln	965	970	975
	Met Gln Gln Met His Gly Arg Met Val Pro Val Ala Ser Thr Glu	980	985	990
15	Thr Gln Asn Ser Asn Phe Thr Ser Ser Met His Phe Asn Arg Thr	995	1000	1005
	Ala Leu Phe Leu Leu Arg Leu Arg Pro Leu Lys Leu Lys Lys Lys	1010	1015	1020
20	Lys Lys Asn Asn Ile Cys Ser Val Ala Trp Cys Thr Asp Cys Asn	1025	1030	1035
	Cys Gly Ala Tyr Arg Asn Asp Cys Arg Ser Phe Glu Asp Leu Glu	1040	1045	1050
	Gln Ile Val Ser Gln Lys Tyr Phe Ser Val His His Gln Ser Val	1055	1060	1065
25	Lys Tyr Met Tyr Leu Lys Asn Thr Ala Ser Glu Phe Cys Cys Ile	1070	1075	1080
	Cys Cys Gln Thr Leu Ser Phe Asp Ile Pro Asp Ser Leu Ser Ile	1085	1090	1095
	Trp Asn Tyr Asn Gly Arg Arg Ala Arg	1100	1104	
30				

(2) INFORMACIÓN PARA LA SEC ID NO:37:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 24 bases

(B) TIPO: ácido nucleico

35 (C) CADENA: única

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:37:

TCGGATCCAC ACGNGACTCT TGGC 24

(2) INFORMACIÓN PARA LA SEC ID NO:38:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 28 bases

5 (B) TIPO: ácido nucleico

(C) CADENA: única

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:38:

TCGGATCCAC TCAGNGACTC TTNGCNGC 28

10 (2) INFORMACIÓN PARA LA SEC ID NO:39:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 32 bases

(B) TIPO: ácido nucleico

(C) CADENA: única

15 (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:39:

CTCGAATTCC AGATAAGCGT ACCAGCACAG TC 32

(2) INFORMACIÓN PARA LA SEC ID NO:40:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

20 (A) LONGITUD: 32 bases

(B) TIPO: ácido nucleico

(C) CADENA: única

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:40:

25 CTCGAATTCC AGATATCCGT ACCATAACAG TC 32

(2) INFORMACIÓN PARA LA SEC ID NO:41:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 13 aminoácidos

(B) TIPO: aminoácido

30 (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:41:

Met Asp Tyr Lys Asp Asp Asp Asp Lys Lys Leu Ala Met
1 5 10 13

35 (2) INFORMACIÓN PARA LA SEC ID NO:42:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 54 bases

(B) TIPO: ácido nucleico

(C) CADENA: única

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:42:

5

CCGGATATCA TGGACTACAA GGACGACGAT GACAAGAAGC TTGCCATGGA 50

GCTC 54

(2) INFORMACION PARA LA SEC ID NO:43:

10 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 22 bases

(B) TIPO: ácido nucleico

(C) CADENA: única

(D) TOPOLOGÍA: lineal

15 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:43:

AGGCTGCTGG AGGAAAAGTC TG 22

(2) INFORMACIÓN PARA LA SEC ID NO:44:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 32 bases

20 (B) TIPO: ácido nucleico

(C) CADENA: única

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:44:

GGAGGGTGAC CTCCATGCTG CCCTTATCCT CG 32

25 (2) INFORMACIÓN PARA LA SEC ID NO:45:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 9108 bases

(B) TIPO: ácido nucleico

(C) CADENA: única

30 (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID NO:45:

TTCGAGCTCG CCCGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT 50
 TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC 100
 TTACGGTAAA TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG 150
 ACGTCAATAA TGACGTATGT TCCCATAGTA ACGCCAATAG GGACTTTCCA 200
 TTGACGTCAA TGGGTGGAGT ATTTACGGTA AACTGCCCAC TTGGCAGTAC 250
 ATCAAGTGTA TCATATGCCA AGTACGCCCC CTATTGACGT CAATGACGGT 300
 AAATGGCCCG CCTGGCATT A TGCCAGTAC ATGACCTTAT GGGACTTTCC 350
 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC 400
 GGTTTTGGCA GTACATCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA 450
 TTTCCAAGTC TCCACCCCAT TGACGTCAAT GGGAGTTTGT TTTGGCACCA 500
 AAATCAACGG GACTTTCCAA AATGTCGTAA CAACTCCGCC CCATTGACGC 550
 AAATGGGCGG TAGGCGTGTA CGGTGGGAGG TCTATATAAG CAGAGCTCGT 600
 TTAGTGAACC GTCAGATCGC CTGGAGACGC CATCCACGCT GTTTTGACCT 650
 CCATAGAAGA CACCGGGACC GATCCAGCCT CCGCGGCCGG GAACGGTGCA 700
 TTGGAACGCG GATTCCCCGT GCCAAGAGTG ACGTAAGTAC CGCCTATAGA 750
 GTCTATAGGC CCACCCCCTT GGCTTCGTTA GAACGCGGCT ACAATTAATA 800
 CATAACCTTA TGTATCATAC ACATACGATT TAGGTGACAC TATAGAATAA 850

CATCCACTTT GCCTTTCTCT CCACAGGTGT CCACTCCCAG GTCCAACTGC 900
 ACCTCGGTTC TATCGATTGA ATTCGCGGCC GCTCGGGTCG GACCCACGCG 950
 CAGCGGCCGG AGATGCAGCG GGGCGCCGCG CTGTGCCTGC GACTGTGGCT 1000
 CTGCCTGGGA CTCCTGGACG GCCTGGTGAG TGGCTACTCC ATGACCCCCC 1050
 CGACCTTGAA CATCACGGAG GAGTCACACG TCATCGACAC CGGTGACAGC 1100
 CTGTCCATCT CCTGCAGGGG ACAGCACCCC CTCGAGTGGG CTTGGCCAGG 1150
 AGCTCAGGAG GCGCCAGCCA CCGGAGACAA GGACAGCGAG GACACGGGGG 1200
 TGGTGCGAGA CTGCGAGGGC ACAGACGCCA GGCCCTACTG CAAGGTGTTG 1250
 CTGCTGCACG AGGTACATGC CAACGACACA GGCAGCTACG TCTGCTACTA 1300
 CAAGTACATC AAGGCACGCA TCGAGGGCAC CACGGCCGCC AGCTCCTACG 1350
 TGTTTCGTGAG AGACTTTGAG CAGCCATTCA TCAACAAGCC TGACACGCTC 1400
 TTGGTCAACA GGAAGGACGC CATGTGGGTG CCCTGTCTGG TGTCCATCCC 1450
 CGGCCTCAAT GTCACGCTGC GCTCGCAAAG CTCGGTGCTG TGGCCAGACG 1500
 GGCAGGAGGT GGTGTGGGAT GACCGGCGGG GCATGCTCGT GTCCACGCCA 1550
 CTGCTGCACG ATGCCCTGTA CCTGCAGTGC GAGACCACCT GGGGAGACCA 1600
 GGACTTCCTT TCCAACCCTT TCCTGGTGCA CATCACAGGC AACGAGCTCT 1650
 ATGACATCCA GCTGTTGCC AGGAAGTCGC TGGAGCTGCT GGTAGGGGAG 1700
 AAGCTGGTCC TGAAGTGCAC CGTGTGGGCT GAGTTTAACT CAGGTGTCAC 1750
 CTTTGACTGG GACTACCCAG GGAAGCAGGC AGAGCGGGGT AAGTGGGTGC 1800

ES 2 355 952 T3

CCGAGCGACG CTCCCAGCAG ACCCACACAG AACTCTCCAG CATCCTGACC 1850
 ATCCACAACG TCAGCCAGCA CGACCTGGGC TCGTATGTGT GCAAGGCCAA 1900
 CAACGGCATC CAGCGATTTC GGGAGAGCAC CGAGGTCATT GTGCATGAAA 1950
 ATCCCTTCAT CAGCGTCGAG TGGCTCAAAG GACCCATCCT GGAGGCCACG 2000
 GCAGGAGACG AGCTGGTGAA GCTGCCCGTG AAGCTGGCAG CGTACCCCCC 2050
 GCCCGAGTTC CAGTGGTACA AGGATGGAAA GGCCTGTCC GGGCGCCACA 2100
 GTCCACATGC CCTGGTGCTC AAGGAGGTGA CAGAGGCCAG CACAGGCACC 2150
 TACACCCTCG CCCTGTGGAA CTCCGCTGCT GGCCTGAGGC GCAACATCAG 2200
 CCTGGAGCTG GTGGTGAATG TGCCCCCCA GATACATGAG AAGGAGGCCT 2250
 CCTCCCCCAG CATCTACTCG CGTCACAGCC GCCAGGCCCT CACCTGCACG 2300
 GCCTACGGGG TGCCCCTGCC TCTCAGCATC CAGTGGCACT GGCGGCCCTG 2350
 GACACCCTGC AAGATGTTTG CCCAGCGTAG TCTCCGGCGG CGGCAGCAGC 2400
 AAGACCTCAT GCCACAGTGC CGTGACTGGA GGGCGGTGAC CACGCAGGAT 2450
 GCCGTGAACC CCATCGAGAG CCTGGACACC TGGACCGAGT TTGTGGAGGG 2500
 AAAGAATAAG ACTGTGAGCA AGCTGGTGAT CCAGAATGCC AACGTGTCTG 2550
 CCATGTACAA GTGTGTGGTC TCCAACAAGG TGGGCCAGGA TGAGCGGCTC 2600
 ATCTACTTCT ATGTGACCAC CATCCCCGAC GGCTTCACCA TCGAATCCAA 2650
 GCCATCCGAG GAGCTACTAG AGGGCCAGCC GGTGCTCCTG AGCTGCCAAG 2700
 CCGACAGCTA CAAGTACGAG CATCTGCGCT GGTACCGCCT CAACCTGTCC 2750

ES 2 355 952 T3

ACGCTGCACG ATGCGCACGG GAACCCGCTT CTGCTCGACT GCAAGAACGT 2800
 GCATCTGTTC GCCACCCCTC TGGCCGCCAG CCTGGAGGAG GTGGCACCTG 2850
 GGGCGGCCA CGCCACGCTC AGCCTGAGTA TCCCCGCGT CGGCCCCGAG 2900
 CACGAGGGCC ACTATGTGTG CGAAGTGCAA GACCGGCGCA GCCATGACAA 2950
 GCACTGCCAC AAGAAGTACC TGTCGGTGCA GGCCCTGGAA GCCCCTCGGC 3000
 TCACGCAGAA CTTGACCGAC CTCCTGGTGA ACGTGAGCGA CTCGCTGGAG 3050
 ATGCAGTGCT TGGTGGCCGG AGCGCACGCG CCCAGCATCG TGTGGTACAA 3100
 AGACGAGAGG CTGCTGGAGG AAAAGTCTGG AGTCGACTTG GCGGACTCCA 3150
 ACCAGAAGCT GAGCATCCAG CGCGTGCGCG AGGAGGATGC GGGACGCTAT 3200
 CTGTGCAGCG TGTGCAACGC CAAGGGCTGC GTCAACTCCT CCGCCAGCGT 3250
 GGCCGTGGAA GGCTCCGAGG ATAAGGGCAG CATGGAGATC GTGATCCTTG 3300
 TCGGTACCGG CGTCATCGCT GTCTTCTTCT GGGTCCTCCT CCTCCTCATC 3350
 TTCTGTAACA TGAGGAGGCC GGCCACGCA GACATCAAGA CGGGCTACCT 3400
 GTCCATCATC ATGGACCCCG GGGAGGTGCC TCTGGAGGAG CAATGCGAAT 3450
 ACCTGTCCTA CGATGCCAGC CAGTGGGAAT TCCCCGAGA GCGGCTGCAC 3500
 CTGGGGAGAG TGCTCGGCTA CGGCGCCTTC GGGAAAGGTGG TGGAAGCCTC 3550
 CGCTTTCGGC ATCCACAAGG GCAGCAGCTG TGACACCGTG GCCGTGAAAA 3600
 TGCTGAAAGA GGGCGCCACG GCCAGCGAGC ACCGCGCGCT GATGTCGGAG 3650
 CTCAAGATCC TCATTACAT CGGCAACCAC CTCAACGTGG TCAACCTCCT 3700

ES 2 355 952 T3

CGGGGCGTGC ACCAAGCCGC AGGGCCCCCT CATGGTGATC GTGGAGTTCT 3750
 GCAAGTACGG CAACCTCTCC AACTTCCTGC GCGCCAAGCG GGACGCCTTC 3800
 AGCCCCTGCG CGGAGAAGTC TCCCGAGCAG CGCGGACGCT TCCGCGCCAT 3850
 GGTGGAGCTC GCCAGGCTGG ATCGGAGGCG GCCGGGAGC AGCGACAGGG 3900
 TCCTCTTCGC GCGGTTCTCG AAGACCGAGG GCGGAGCGAG GCGGGCTTCT 3950
 CCAGACCAAG AAGCTGAGGA CCTGTGGCTG AGCCCGCTGA CCATGGAAGA 4000
 TCTTGTCTGC TACAGCTTCC AGGTGGCCAG AGGGATGGAG TTCCTGGCTT 4050
 CCCGAAAGTG CATCCACAGA GACCTGGCTG CTCGGAACAT TCTGCTGTCG 4100
 GAAAGCGACG TGGTGAAGAT CTGTGACTTT GGCCTTGCCC GGGACATCTA 4150
 CAAAGACCCT GACTACGTCC GCAAGGGCAG TGCCCGGCTG CCCCTGAAGT 4200
 GGATGGCCCC TGAAAGCATC TTCGACAAGG TGTACACCAC GCAGAGTGAC 4250
 GTGTGGTCTT TTGGGGTGCT TCTCTGGGAG ATCTTCTCTC TGGGGGCCTC 4300
 CCCGTACCCT GGGGTGCAGA TCAATGAGGA GTTCTGCCAG CGGCTGAGAG 4350
 ACGGCACAAG GATGAGGGCC CCGGAGCTGG CCACTCCC GC CATA CGCCGC 4400
 ATCATGCTGA ACTGCTGGTC CGGAGACCCC AAGGCGAGAC CTGCATTCTC 4450
 GGAGCTGGTG GAGATCCTGG GGGACCTGCT CCAGGGCAGG GGCCTGCAAG 4500
 AGGAAGAGGA GGTCTGCATG GCCCCGCGCA GCTCTCAGAG CTCAGAAGAG 4550
 GGCAGCTTCT CGCAGGTGTC CACCATGGCC CTACACATCG CCCAGGCTGA 4600
 CGCTGAGGAC AGCCCGCCAA GCCTGCAGCG CCACAGCCTG GCCGCCAGGT 4650

ES 2 355 952 T3

ATTACAAC TG GGTGTCCTTT CCCGGGTGCC TGGCCAGAGG GGCTGAGACC 4700

CGTGGTTCCT CCAGGATGAA GACATTTGAG GAATTCCCCA TGACCCCAAC 4750

GACCTACAAA GGCTCTGTGG ACAACCAGAC AGACAGTGGG ATGGTGCTGG 4800

CCTCGGAGGA GTTTGAGCAG ATAGAGAGCA GGCATAGACA AGAAAGCGGC 4850

TTCAGGTAGC TGAAGCAGAG AGAGAGAAGG CAGCATACGT CAGCATTTTC 4900

TTCTCTGCAC TTATAAGAAA GATCAAAGAC TTTAAGACTT TCGCTATTTT 4950

TTCTGCTATC TACTACAAAC TTCAAAGAGG AACCAGGAGG CCAAGAGGAG 5000

CATGAAAGTG GACAAGGAGT GTGACCACTG AAGCACCACA GGGAGGGGTT 5050

AGGCCTCCGG ATGACTGCGG GCAGGCCTGG ATAATATCCA GCCTCCCACA 5100

AGAAGCTGGT GGAGCAGAGT GTTCCCTGAC TCCTCCAAGG AAAGGGAGAC 5150

GCCCTTTCAT GGTCTGCTGA GTAACAGGTG CCTTCCCAGA CACTGGCGTT 5200

ACTGCTTGAC CAAAGAGCCC TCAAGCGGCC CTTATGCCAG CGTGACAGAG 5250

GGCTCACCTC TTGCCTTCTA GGTCACTTCT CACAATGTCC CTTCAGCACC 5300

TGACCCTGTG CCCGCCAGTT ATTCCTTGGT AATATGAGTA ATACATCAAA 5350

GAGTAGTGCG GCCCGGAATT CCCCAGGGAT CCTCTAGAGT CGACCTGCAG 5400

AAGCTTGGCC GCCATGGCCC AACTTGTTTA TTGCAGCTTA TAATGGTTAC 5450

AAATAAAGCA ATAGCATCAC AAATTTTACA AATAAAGCAT TTTTTTCACT 5500

GCATTCTAGT TGTGGTTTGT CCAAACATCAT CAATGTATCT TATCATGTCT 5550

GGATCGGGAA TTAATTCGGC GCAGCACCAT GGCCTGAAAT AACCTCTGAA 5600

ES 2 355 952 T3

AGAGGAACTT GGTTAGGTAC CTTCTGAGGC GGAAAGAACC AGCTGTGGAA 5650
 TGTGTGTCAG TTAGGGTGTG GAAAGTCCCC AGGCTCCCCA GCAGGCAGAA 5700
 GTATGCAAAG CATGCATCTC AATTAGTCAG CAACCAGGTG TGGAAAGTCC 5750
 CCAGGCTCCC CAGCAGGCAG AAGTATGCAA AGCATGCATC TCAATTAGTC 5800
 AGCAACCATA GTCCCGCCCC TAACTCCGCC CATCCCGCCC CTAACTCCGC 5850
 CCAGTTCGGC CCATTCTCCG CCCCATGGCT GACTAATTTT TTTTATTTAT 5900
 GCAGAGGCCG AGGCCGCCTC GGCCTCTGAG CTATTCCAGA AGTAGTGAGG 5950
 AGGCTTTTTT GGAGGCCTAG GCTTTTGCAA AAAGCTGTTA ACAGCTTGGC 6000
 ACTGGCCGTC GTTTTACAAC GTCGTGACTG GGAAAACCCT GGC GTTACCC 6050
 AACTTAATCG CCTTGCAGCA CATCCCCCTT TCGCCAGCTG GCGTAATAGC 6100
 GAAGAGGCCC GCACCGATCG CCCTTCCCAA CAGTTGCGCA GCCTGAATGG 6150
 CGAATGGCGC CTGATGCGGT ATTTTCTCCT TACGCATCTG TGCGGTATTT 6200
 CACACCGCAT ACGTCAAAGC AACCATAGTA CGCGCCCTGT AGCGGCGCAT 6250
 TAAGCGCGGC GGGTGTGGTG GTTACGCGCA GCGTGACCGC TACACTTGCC 6300
 AGCGCCCTAG CGCCCGCTCC TTTCGCTTTC TTCCCTTCCCT TTCTCGCCAC 6350
 GTTCGCCGTC TTCCCCGTC AAGCTCTAAA TCGGGGGCTC CCTTTAGGGT 6400
 TCCGATTTAG TGCTTTACGG CACCTCGACC CCAAAAAACT TGATTTGGGT 6450
 GATGGTTCAC GTAGTGGGCC ATCGCCCTGA TAGACGGTTT TTCGCCCTTT 6500
 GACGTTGGAG TCCACGTTCT TTAATAGTGG ACTCTTGTTT CAAACTGGAA 6550

ES 2 355 952 T3

CAACACTCAA CCCTATCTCG GGCTATTCTT TTGATTTATA AGGGATTTTG 6600
 CCGATTTTCGG CCTATTGGTT AAAAAATGAG CTGATTTAAC AAAAATTTAA 6650
 CGCGAATTTT AACAAAATAT TAACGTTTAC AATTTTATGG TGCACTCTCA 6700
 GTACAATCTG CTCTGATGCC GCATAGTTAA GCCAGCCCCG ACACCCGCCA 6750
 ACACCCGCTG ACGCGCCCTG ACGGGCTTGT CTGCTCCCGG CATCCGCTTA 6800
 CAGACAAGCT GTGACCGTCT CCGGGAGCTG CATGTGTCAG AGGTTTTTAC 6850
 CGTCATCACC GAAACGCGCG AGACGAAAGG GCCTCGTGAT ACGCCTATTT 6900
 TTATAGGTTA ATGTCATGAT AATAATGGTT TCTTAGACGT CAGGTGGCAC 6950
 TTTTCGGGGA AATGTGCGCG GAACCCCTAT TTGTTTATTT TTCTAAATAC 7000
 ATTCAAATAT GTATCCGCTC ATGAGACAAT AACCCTGATA AATGCTTCAA 7050
 TAATATTGAA AAAGGAAGAG TATGAGTATT CAACATTTCC GTGTCGCCCT 7100
 TATTCCCTTT TTTGCGGCAT TTTGCCTTCC TGTTTTTGCT CACCCAGAAA 7150
 CGCTGGTGAA AGTAAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT 7200
 TACATCGAAC TGGATCTCAA CAGCGGTAAG ATCCTTGAGA GTTTTCGCCC 7250
 CGAAGAACGT TTTCCAATGA TGAGCACTTT TAAAGTTCTG CTATGTGGCG 7300
 CGGTATTATC CCGTATTGAC GCCGGGCAAG AGCAACTCGG TCGCCGCATA 7350
 CACTATTCTC AGAATGACTT GGTGAGTAC TCACCAGTCA CAGAAAAGCA 7400
 TCTTACGGAT GGCATGACAG TAAGAGAATT ATGCAGTGCT GCCATAACCA 7450
 TGAGTGATAA CACTGCGGCC AACTTACTTC TGACAACGAT CGGAGGACCG 7500

ES 2 355 952 T3

AAGGAGCTAA CCGCTTTTTT GCACAACATG GGGGATCATG TAACTCGCCT 7550
 TGATCGTTGG GAACCGGAGC TGAATGAAGC CATAACAAAC GACGAGCGTG 7600
 ACACCACGAT GCCTGTAGCA ATGGCAACAA CGTTGCGCAA ACTATTAAC 7650
 GCGGAAC TACTCTAGC TTCCCGCAA CAATTAATAG ACTGGATGGA 7700
 GCGGATAAA GTTGCAGGAC CACTTCTGCG CTCGGCCCTT CCGGCTGGCT 7750
 GGTTTATTGC TGATAAATCT GGAGCCGGTG AGCGTGGGTC TCGCGGTATC 7800
 ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC TCCCGTATCG TAGTTATCTA 7850
 CACGACGGGG AGTCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG 7900
 AGATAGGTGC CTCACTGATT AAGCATTGGT AACTGTCAGA CCAAGTTTAC 7950
 TCATATATAC TTTAGATTGA TTTAAAACCTT CATTTTTAAT TTAAAAGGAT 8000
 CTAGGTGAAG ATCCTTTTTG ATAATCTCAT GACCAAATC CCTTAACGTG 8050
 AGTTTTTCGTT CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT 8100
 TCTTGAGATC CTTTTTTTCT GCGCGTAATC TGCTGCTTGC AAACAAAAAA 8150
 ACCACCGCTA CCAGCGGTGG TTTGTTTGCC GGATCAAGAG CTACCAACTC 8200
 TTTTCCGAA GGTAAC TGGC TTCAGCAGAG CGCAGATACC AAATACTGTT 8250
 CTTCTAGTGT AGCCGTAGTT AGGCCACCAC TTCAAGAACT CTGTAGCACC 8300
 GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT GCTGCCAGTG 8350
 GCGATAAGTC GTGTCTTACC GGGTTGACT CAAGACGATA GTTACCGGAT 8400
 AAGGCGCAGC GGTCGGGCTG AACGGGGGGT TCGTGACAC AGCCCAGCTT 8450

GGAGCGAACG ACCTACACCG AACTGAGATA CCTACAGCGT GAGCTATGAG 8500

AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA TCCGGTAAGC 8550

GGCAGGGTCG GAACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACGC 8600

CTGGTATCTT TATAGTCCTG TCGGGTTTCG CCACCTCTGA CTTGAGCGTC 8650

GATTTTTGTG ATGCTCGTCA GGGGGGCGGA GCCTATGGAA AAACGCCAGC 8700

AACGCGGCCT TTTTACGGTT CCTGGCCTTT TGCTGGCCTT TTGCTCACAT 8750

GTTCTTTCCT GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCCT 8800

TTGAGTGAGC TGATACCGCT CGCCGCAGCC GAACGACCGA GCGCAGCGAG 8850

TCAGTGAGCG AGGAAGCGGA AGAGCGCCCA ATACGCAAAC CGCCTCTCCC 8900

CGCGCGTTGG CCGATTCATT AATGCAGCTG GCACGACAGG TTTCCCGACT 8950

GGAAAGCGGG CAGTGAGCGC AACGCAATTA ATGTGAGTTA GCTCACTCAT 9000

TAGGCACCCC AGGCTTTACA CTTTATGCTT CCGGCTCGTA TGTGTGTGG 9050

AATTGTGAGC GGATAACAAT TTCACACAGG AACAGCTAT GACATGATTA 9100

CGAATTAA 9108

REIVINDICACIONES

1. Proteína quimérica que comprende una fusión del dominio extracelular de bpTK7 con una secuencia de dominio constante de inmunoglobulina.
2. Proteína quimérica según la reivindicación 1, donde la secuencia de dominio constante de inmunoglobulina es la de una inmunoglobulina IgG.
- 5 3. Ácido nucleico que codifica la proteína quimérica según la reivindicación 1.
4. Vector replicable que comprende el ácido nucleico según la reivindicación 3.
5. Célula huésped recombinante que comprende el ácido nucleico según la reivindicación 3.
- 10 6. Método de utilización de una molécula de ácido nucleico que codifica una proteína quimérica que comprende una fusión del dominio extracelular de bpTK7, con una secuencia de dominio constante de inmunoglobulina, para realizar la producción de la proteína quimérica que comprende cultivar la célula huésped según la reivindicación 5.
7. Utilización de la proteína quimérica según la reivindicación 1 en la producción de un anticuerpo que se une al dominio extracelular de bpTK7.

FIG. 1A

GGATCCTGTG CATCAGTGAC TTAGGGCTAG GAACATTCTG CTGTCCGAAA GCGACGTGGT 60
 GAAGATCTGT GACTTTGGCC TTGCCCGGA CATCTACAAA GACCCCAGCT ACGTCCGCAA 120
 GCATGCCCGG CTGCCCCCTGA AGTGGATGGC GCCAGNATTC 160

FIG. 1B

Asp Pro Val His Gln Xaa Leu Arg Ala Arg Asn Ile Leu Leu Ser Glu 15
 1 5 10
 Ser Asp Val Val Lys Ile Cys Asp Phe Gly Leu Ala Arg Asp Ile Tyr 30
 20 25
 Lys Asp Pro Ser Tyr Val Arg Lys His Ala Arg Leu Pro Leu Lys Trp 45
 35 40
 Met Ala Pro Glu Phe 50

FIG. 2A

GGATCCATTC ACAGAGACCT AGCAGCACGC AACATCCTGG TCTCAGAGGA CCTGGTAACC 60
 AAGGTCAGCG ACTTTGGCCT GGCCAAAGCC GAGCGGAAGG GGCTAGACTC AAGCCGGCTG 120
 CCCGTCAAAT GGATGGCTCC CGAATTC 147

FIG. 2B

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

FIG. 3A

GTT GGA ATT CCT TCC GGC GCC ATC CAT TTC ACC GGC AGC TTT ATT TCG 48
 Val Gly Ile Pro Ser Gly Ala Ile His Phe Thr Gly Ser Phe Ile Ser 15
 1 5 10
 TGT CTA GAT TCA TAG ATG TCT TCA TTA TTA TCT ACC TTA AAA ACT CTG GCA 96
 Cys Leu Asp Ser Met Ser Ser Leu Ser Thr Leu Lys Thr Leu Ala 30
 20 25
 AGT CCA AAA TCT GCT ACT TTG TAG ATA TTA TGT TCA CCA ACG AGG ACA 144
 Ser Pro Lys Ser Ala Thr Leu Ile Leu Cys Ser Pro Thr Arg Thr 45
 35 40
 TTCCT 149
 Phe

FIG. 3B

GTG CAC AGG GAT CTC GCG GCT CGG AAC ATC CTC GTC GGG GAA AAC ACC 48
 Val His Arg Asp Leu Ala Ala Arg Asn Ile Leu Val Gly Glu Asn Thr 15
 1 5 10
 CTC TCG AAA GTT GGG GAC TTC GGG TTA GCC AGG CTT ATC AAG GAG GAC 96
 Leu Ser Lys Val Gly Asp Phe Gly Leu Ala Arg Leu Ile Lys Glu Asp 30
 20 25
 GTC TAC CTC TCC CAT GAC CAC AAT ATC CCC TAC AAA TGG ATG GCC CCT 144
 Val Tyr Leu Ser His Asp His Asn Ile Pro Tyr Lys Trp Met Ala Pro 45
 35 40
 GAG GGA A 151
 Glu Gly 50

FIG. 3C

GTT CAC CGA GAT CTC AAG TCC AAC AAC ATT TTG CTG CTG CAG CCC ATT 48
 Val His Arg Asp Leu Lys Ser Asn Asn Ile Leu Leu Leu Gln Pro Ile 15
 1 5 10
 GAG AGT GAC GAC ATG GAG CAC AAG ACC CTG AAG ATC ACC GAC TTT GGC 96
 Glu Ser Asp Asp Met Glu His Lys Thr Leu Lys Ile Thr Asp Phe Gly 30
 20 25
 CTG GCC CGA GAG TGG CAC AAA ACC ACA CAA ATG AGT GCC GC 137
 Leu Ala Arg Glu Trp His Lys Thr Thr Gln Met Ser Ala 45
 35 40

FIG. 3D

GTC AAT CGT GAC CTC GCC GCC CGA AAT GTG TTG CTA GTT ACC CAA CAT 48
 Val Asn Arg Asp Leu Ala Ala Arg Asn Val Leu Leu Val Thr Gln His 15
 1 5 10
 TAC GCC AAG ATC AGT GAT TTC GGA CTT TCC AAA GCA CTG CGT GCT GAT 96
 Tyr Ala Lys Ile Ser Asp Phe Gly Leu Ser Lys Ala Leu Arg Ala Asp 30
 20 25
 GAA AAC TAC TAC AAG GCC CAG ACC CAT GGA AAG TGG CCT GTC AAG TGG 144
 Glu Asn Tyr Tyr Lys Ala Gln Thr His Gly Lys Trp Pro Val Lys Trp 45
 35 40
 TAC GCT CCG GAA TGC ATC AAC TAC AAC TTC TCC AGC AAA AGC GAT 192
 Tyr Ala Pro Glu Cys Ile Asn Tyr Tyr Lys Phe Ser Ser Lys Ser Asp 60
 50 55
 GTC TGG TCC TTT GGA ATT C 211
 Val Trp Ser Phe Gly Ile 70
 65

FIG. 4A

TTCGAGCTCG CCGGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT TACGGGGTCA 60
 TTAGTTCATA GCCCATATAT GGAGTTCGGC GTTACATAAC TTACGGTAAA TGGCCCCGCT 120
 GGCTGACCGC CCAACGACCC CCGCCCCATTG ACGTCAATAA TGACGTATGT TCCCATAGTA 180
 ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT ATTTACGGTA AACTGCCCCAC 240
 TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTACGCCCC CTATTGACGT CAATGACGGT 300
 AAATGGCCCC CCTGGCATTG TGCCCCAGTAC ATGACCTTAT GGGACITTTCC TACTTGGCAG 360
 TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC GGTTTTGGCA GTACATCAAT 420
 GGGCGTGGAT AGCGGTTTGA CTCACGGGGA TTTCCAAGTC TCCACCCCAT TGACGTCAAT 480
 GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA AATGTCGTAA CAACTCCGCC 540
 CCATTGACGC AAATGGGCGG TAGGCGTGTA CCGTGGGAGG TCTATATAAG CAGAGCTCGT 600
 TTAGTGAACC GTCAGATCGC CTGGAGACGC CATCCACGCT GTTTTGACCT CCATAGAAGA 660
 CACCGGGACC GATCCAGCCT CCGCGGCCGG GAACGGTGCA TTGGAACGGG GATTCGCCGT 720
 GCCAAGAGTG ACGTAAGTAC CGCCTATAGA GTCTATAGGC CCACTTGGCT TCGTTAGAAC 780
 GCGGCTACAA TTAATACATA ACCTTATGTA TCATACACAT ACCATTTAGG TGACACTATA 840
 GAATAACATC CACTTTGGCT TTCTCTCCAC AGGTGTCCAC TCCCAGGTCC AACTGCACCT 900
 CGGTTCTATC GATTGAATTC CCGGGGGATC CTCTAGAGAT CCGTCGACCT CGAGATCCAT 960
 TGTGCTGGCG CGGATTTCTTT ATCACTGATA AGTTGGTGGA CATATTATGT TTATCAGTGA 1020

FIG. 4B

TAAAGTGTCA AGCATGACAA AGTTGCAGCC GAATACAGTG ATCCGTGCCG CCTAGACCT 1080
 GTTGAACGAG GTCGGCGTAG ACGGTCTGAC GACACGGAAA CTGGCGGAAC GGTGGGSET 1140
 TCAGCAGCCG GCGCTTTACT GGCACCTCAG GAACAAGCGG GCGTGTCTCG ACGCACTGGC 1200
 CGAAGCCATG CTGGCGGAGA ATCATAGCAC TTCGGTGCCG AGAGCCGACG ACGACTGGCG 1260
 CTCATTTCTG ACTGGGAATG CCCGCAGCTT CAGGCAGGCG CTGCTCGCCT ACCGCCAGCA 1320
 CAATGGATCT CGAGGGATCT TCCATACCTA CCAGTTCTGC GCCTGCAGGT CGCGGCCGCA 1380
 CTACTCTTTG ATGTATTACT CATATTACCA AGGAATAACT GCGGGGCACA GGGTCAGGTG 1440
 CTGAAGGGAC ATTGTGAGAA GTGACCTAGA AGGCAAGAGG TGAGCCCTCT GTCACGGCTGG 1500
 CATAAAGGCC GCTTGAGGGC TCCTTTGGTCA AGCAGTAACG CCAGTGTCTG GGAAGGCACC 1560
 TGTTACTCAG CAGACCATGA AAGGGCGTCT CCCTTTCCCTT GGAGCAGTCA GGAACACTC 1620
 TGCTCCACCA GCTTCTTGTG GGAGCCTGGA TATTATCCAG GCCTGCCCCGC AGTCATCCGG 1680
 AGGCCTAACC CCTCCCTGTG GTGCTTCAGT GGTCAACTC CTTGTCCACT TTCATGCTCC 1740
 TCTTTGGCCTC CTGGTTCCTC TTGGAAGTTT GTAGTAGATA GCAGAAGAAA TAGCGAAAAGT 1800
 CTTAAAGTCT TTGATCTTTC TTATAAGTGC AGAGAAGAAA TGCTGACGTA TGCTGCCTTC 1860
 TCTCTCTCIG CTTCAGCTAC CTGAAGCCGC TTTCTTGTCT ATACCTGCTC TCTATCTGCT 1920
 CACTCTCCTC CGAGGCCAGC ACCATCCCAC TGTCTGTCTG GTTGTCCACA GAGCCTTTGT 1980
 AGGTCGTTGG GGTCAATGGG AATTCCTCAA ATGTCTTCAT CCTGGAGGAA CCACGGGTCT 2040

FIG. 4C

CAGCCCCCTCT GGCCAGGCAC CCGGGAAGG ACACCCAGTT GTAATACCTG GCGGCCAGGC 2100
 TGTGGCGCTG CAGGCTTGGC GGGCTGTCTT CAGCGTCAGC CTGGGCGATG TGTAGGGCCA 2160
 TGGTGGACAC CTGGGAGAAG CTGCCCTCTT CTGAGCTCTG AGAGCTGCGC GGGGCCATGC 2220
 AGACCTCCTC TTCCCTCTTGC AGGCCCCCTGC CCTGGAGCAG GTCCCCCAGG ATCTCCACCA 2280
 GCTCCGAGAA TGCAGGTCTC GCCTTGGGGT CTCCGGGACCA GCAGTTCAGC ATGATGGCGC 2340
 GTATGGCGGG AGTGGCCAGC TCCGGGGCCC TCATCCTTGT GCCGTCTCTC AGCCCGCTGGC 2400
 AGAACTCCTC ATTGATCTGC ACCCCAGGGT ACGGGGAGGC CCCCAGAGAG AAGATCTCCC 2460
 AGAGAAGCAC CCCAAAGGAC CACACGTAC TCTGCGTGGT GTACACCTTG TCGAAGATGC 2520
 TTTTAGGGGC CATCCACTTC AGGGGCAGCC GGGCACTGCC CTTGGGGACG TAGTGGGGGT 2580
 CTTTGTAGAT GTCCCCGGGCA AGGCCAAAGT CACAGATCTT CACCACGTCG CTTTCCGACA 2640
 GCAGAATGTT CCGAGCAGCC AGTCTCTGT GGATGCACTT TCGGGAAGCC AGGAACTCCA 2700
 TCCCTCTGGC CACCTGGAAG CTGTAGCAGA CAAGATCTTC CATGGTCAGC GGGCTCAGCC 2760
 ACAGGTCTC AGCTTCTTGG TCTGGAGAAG CCCGCCCTGC TCCGCCCTCG GTCTTCGAGA 2820
 ACCGCGCGAA GAGGACCCTG TCGCTGCTCC CCGGCCGCCT CCGATCCAGC CTGGCCGAGCT 2880
 CCACCATGGC GCGGAAGCGT CCGCGCTGCT CCGGAGACTT CTCCTGCGGA TGCACGMAAGC 2940
 TGGCTCGAGG GCGCCCAAGTC GTCCGCCGCA GAGGGCCCTC CATTCGCCCG CCGCCCGCGG 3000
 CGCCCCGCAG GCCGCCCGCT CACCGNGCAG GGGCTGGGC CGCGACTCTA GAGTCGACCT 3060

FIG. 4D

GCAGAAAGCTT GGCCGGCCATG GCCCAACTTG TTATTATGCAG CTTATAATGG TTACAAATAA 3120
 AGCAATAGCA TCACAAATTT CACAAATAAA GCATTTTTTTT CACTGCATTC TAGTTGTGGT 3180
 TTGTCCAAAC TCATCAATGT ATCTTATCAT GTCTGGATCG ATCGGGAAATT AATTCGGCGC 3240
 AGCACCATGG CCTGAAATAA CCTCTGAAAG AGGAACTTGG TTAGGTACCT TCTGAGGCGG 3300
 AAAGAAACCAG CTGTGGAATG TGTGTCAAGT AGGGTGTGGA AAGTCCCCCAG GCTCCCCAGC 3360
 AGCAGAAAGT ATGCAAAGCA TGCATCTCAA TTAGTCAGCA ACCAGGTGTG GAAAAGTCCCC 3420
 AGGCTCCCCA GCAGGCAGAA GTATGCAAG CATGCATCTC AATTAGTCAG CAACCATAGT 3480
 CCGGCCCTA ACTCCGCCCA TCCCGCCCTT AACTCCGCCC AGTTCGCCCC ATTCTCCGCC 3540
 CCATGGCTGA CTAATTTTTT TTATTTATGC AGAGGCCGAG GCCGCCCTCGG CCTCTGAGCT 3600
 ATTCCAGAAG TAGTGAGGAG GCTTTTTTGG AGGCCTAGGC TTTTGCAAAA AGCTGTTAAC 3660
 AGCTTGCCAC TGGCCGTCGT TTTACAACGT CGTGACTGGG AAAACCCCTGG CGTTACCCAA 3720
 CTTAATCGCC TTGCAGCACA TCCCCCCTTC GCCAGCTGGC GTAATAGCGA AGAGGCCCGC 3780
 ACCGATCGCC CTTCCCAACA GTTGGGTAGC CTGAATGGG AATGGGCGCT GATGCGGTAT 3840
 TTTCTCCITA CGCATCTGTG CGGTATTTCA CACCCGATAC GTCAAAGCAA CCATAGTACG 3900
 CGCCCTGTAG CGGGCGATTA AGCCGGGCGG GTGTGGTGGT TACGGCCAGC GTGACCCGTA 3960
 CACTTGCCAG CGCCCTAGCG CCCGCTCCTT TCGCTTTCTT CCGTTCCTTT CTGCCACCGT 4020
 TCGCCGGCTT TCCCCGTCAA GCTCTAAATC GGGGGCTCCC TTTAGGGTTC CGATTTAGTG 4080

FIG. 4E

CTTTACGGCA CCTCGACCCC AAAAAACTTG ATTTGGGTGA TGGTTCACGT AGTGGGCCAT 4140
 CGCCCTGATA GACGGTTTTT CGCCCTTTGA CGTTGGAGTC CACGTTCTTT AATAGTGGAC 4200
 TCTTGTTCOA AACTGGAACA ACACTCAACC CTATCTCGGG CTATTCCTTT GATTATAAG 4260
 GGATTTTGCC GATTTCCGCC TATTTGGTTAA AAAATGAGCT GATTTAACAA AAATTTAACC 4320
 CGAATTTTAA CAAAATATTA ACGTTTACAA TTTTATGGTG CACTCTCAGT ACAATCTGCT 4380
 CTGATGCCGC ATAGTTAAGC CAACTCCGCT ATCGCTACGT GACTGGGTCA TGGCTGCCGC 4440
 CCGACACCCG CCAACACCCG CTGACGGGCC CTGACGGGCT TGTCTGCTCC CGGCATCCGC 4500
 TTACAGACAA GCTGTGACCG TCTCCGGGAG CTGCATGIGT CAGAGGTTTT CACCGTCAATC 4560
 ACCGAAACCG GCGAGGCAGT ATTCTTGAAG ACGAAAGGGC CTCGTGATAC GCCTATTTTT 4620
 ATAGGTTAAT GTCATGATAA TAATGGTTTC TTAGACGTC A GGTGGCACCTT TTCGGGGAAA 4680
 TGTGCGCGGA ACCCCTATTT GTTTATTTTT CTAAATACAT TCAAAATATGT ATCCGCTCAT 4740
 GAGACAATAA CCCTGATAAA TCTTCAATA ATATTGAAA AGGAAGAGTA TGAGTATTCAA 4800
 ACATTTCCGT GTCGCCCTTA TTCCCTTTTT GCGGCAATTT TGCCTTCCTG TTTTTCCTCA 4860
 CCCAGAAACG CTGGTGAAG TAAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGGTTA 4920
 CATCGAACTG GATCTCAACA GCGGTAAGAT CCTTGAGAGT TTTCCGCCCG AAGAACGTTT 4980
 TCCAAATGATG AGCACTTTTA AAGTCTCTGCT ATGTGGCGCG GTATTATCCC GTGATGACGC 5040
 CCGGCAAGAG CAACTCGGTC GCCGCATACA CTATTCCTCAG AATGACTTGG TTGAGTACTC 5100

FIG. 4F

ACCAGTCACA GAAAAGCATC TTACGGATGG CATGACAGTA AGAGAATTAT GCAGTGCTGC 5160
 CATAACCATG AGTGATAACA CTGCGGCCAA CTTACTTCTG ACAACGATCG GAGGACCGAA 5220
 GGAGCTAACC GCCTTTTTTGC ACAACATGGG GGATCATGTA ACTCGCCTTG ATCGTTGGGA 5280
 ACCGGAGCTG AATGAAGCCA TACCAAACGA CGAGCGTGAC ACCACGATGC CAGCAGCAAT 5340
 GGCAACAACG TTGCGCAAAC TATTAACCTG CGAACTACTT ACTCTAGCTT CCCGGCAACA 5400
 ATTAATAGAC TGGATGGAGG CCGATAAAGT TGCAGGACCA CTTCTGCGCT CGGCCCTTCC 5460
 GGCTGGCTGG TTTATTGCTG ATAAATCTGG AGCCGGTGAG CGTGGGTCTC GCGGTATCAT 5520
 TGCAGCACTG GGGCCAGATG GTAAGCCCTC CCGTATCGTA GTTATCTACA CGACGGGGAG 5580
 TCAGGCAACT ATGGATGAAC GAAATAGACA GATCGCTGAG ATAGGTGCCT CACTGATTAA 5640
 GCATTGGTAA CTGTCAGACC AAGTTTACTC ATATATACTT TAGATTGATT TAAAACTTCA 5700
 TTTTAAATTT AAAAGGATCT AGGTGAAGAT CCTTTTTGAT AATCTCATGA CCAAAATCCC 5760
 TTAACGTGAG TTTTCGTTCC ACTGAGCGTC AGACCCCGTA GAAAAGATCA AAGGATCTTC 5820
 TTGAGATCCT TTTTTTCTGC GCGTAATCTG CTGCTTGCAA ACAAAAAAC CACCGCTACC 5880
 AGCGGTGGTT TGTTTGCCGG ATCAAGAGCT ACCAACTCTT TTTCCGAAGG TAACTGGCTT 5940
 CAGCAGAGCG CAGATACCAA ATACTGTCTT TCTAGTGTAG CCGTAGTTAG GCCACCCTT 6000
 CAAGAACTCT GTAGCACCGC CTACATACCT CGCTCTGCTA ATCCCTGTTAC CAGTGGCTGC 6060
 TGCCAGTGGC GATAAGTCGT GTCTTACCGG GTTGGACTCA AGACGATAGT TACCGGATAA 6120

FIG. 4G

GGCGCAGCGG TCGGGCTGAA CCGGGGGTTC GTGCACACAG CCCAGCTTGG AGCGAACGAC 6180
CTACACCGAA CTGAGATACC TACAGCGTGA GCATTGAGAA AGCGCCACGC TTCCCAGAAG 6240
GAGAAAGCGG GACAGGTATC CGGTAAGCGG CAGGGTCGGA ACAGGAGAGC GCACGAGGGA 6300
GCTTCCAGGG GGAACGCCCT GGTATCTTTA TAGTCCCTGC GGGTTTCGCC ACCTCTGACT 6360
TGAGCGTCGA TTTTGTGAT GCTCGTCAGG GGGGCGGAGC CTATGGAAA ACGCCAGCAA 6420
CGCGGCCCTT TTACGGTTCC TGGCCCTTTG CTGGCCCTTT GCTCACATGT TCTTTCCTGC 6480
GTTATCCCCCT GATTCTGTGG ATAACCGTAT TACCGCCCTT GAGTGAGCTG ATACCGCTCG 6540
CCGCAGCCGA ACGACCGAGC GCAGCGAGTC AGTGAGCGAG GAAGCGGAAG AGCGCCCAAT 6600
ACGCAAAACCG CCTCTCCCCG CGCGTTGGCC GATTCATTAA TCCAGCTGGC ACGACAGGTT 6660
TCCCAGACTG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTACC TCACTCATTA 6720
GGCACCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA TTGTGAGCGG 6780
ATAACAATTT CACACAGGAA ACAGCTATGA CCATGATTAC GAATTA 6827

FIG. 4H

1 Glu Lys Ser Pro Glu Gln Arg Gly Arg Phe Arg Ala Met Val Glu Leu 15
 5
 20 Ala Arg Leu Asp Arg Arg Pro Gly Ser Ser Asp Arg Val Leu Phe 30
 25
 35 Ala Arg Phe Ser Lys Thr Glu Gly Ala Arg Arg Ala Ser Pro Asp 45
 40
 50 Gln Glu Ala Glu Asp Leu Trp Leu Ser Pro Leu Thr Met Glu Asp Leu 60
 55
 65 Val Cys Tyr Ser Phe Gln Val Ala Arg Gly Met Glu Phe Leu Ala Ser 80
 70
 85 Arg Lys Cys Ile His Arg Asp Leu Ala Ala Arg Asn Ile Leu Leu Ser 95
 90
 100 Glu Ser Asp Val Val Lys Ile Cys Asp Phe Gly Leu Ala Arg Asp Ile 110
 105
 115 Tyr Lys Asp Pro Asp Tyr Val Arg Lys Gly Ser Ala Arg Leu Pro Leu 125
 120
 130 Lys Trp Met Ala Pro Glu Ser Ile Phe Asp Lys Val Tyr Thr Thr Gln 140
 135
 145 Ser Asp Val Trp Ser Phe Gly Val Leu Leu Trp Glu Ile Phe Ser Leu 160
 150
 165 Gly Ala Ser Pro Tyr Pro Gly Val Gln Ile Asn Glu Glu Phe Cys Gln 175
 170
 180 Arg Leu Arg Asp Gly Thr Arg Met Arg Ala Pro Glu Leu Ala Thr Pro 190
 185

FIG. 4I

Ala Ile Arg Arg Ile Met Leu Asn Cys Trp Ser Gly Asp Pro Lys Ala
 195 200 205
 Arg Pro Ala Phe Ser Glu Leu Val Glu Ile Leu Gly Asp Leu Leu Gln
 210 215 220
 Gly Arg Gly Leu Gln Glu Glu Glu Val Cys Met Ala Pro Arg Ser
 225 230 235 240
 Ser Gln Ser Ser Glu Glu Gly Ser Phe Ser Gln Val Ser Thr Met Ala
 245 250 255
 Leu His Ile Ala Gln Ala Asp Ala Glu Asp Ser Pro Pro Ser Leu Gln
 260 265 270
 Arg His Ser Leu Ala Ala Arg Tyr Tyr Asn Trp Val Ser Phe Pro Gly
 275 280 285
 Cys Leu Ala Arg Gly Ala Glu Thr Arg Gly Ser Ser Arg Met Lys Thr
 290 295 300
 Phe Glu Glu Phe Pro Met Thr Pro Thr Thr Tyr Lys Gly Ser Val Asp
 305 310 315 320
 Asn Gln Thr Asp Ser Gly Met Val Leu Ala Ser Glu Glu Cys Glu Gln
 325 330 335
 Ile Glu Ser Arg Tyr Arg Gln Glu Ser Gly Phe Arg *
 340 345

FIG. 5A

FTCCGAGCTCG CCCGACATTC ATTATTGACT AGTTATTAAT AGTAATCAAT TACGGGGTCA 60
 TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAAC TTACGGTAAA TGGCCCGCCT 120
 GGCTGACCGC CCAACGACCC CCGCCCATTC ACGTCAATAA TGACGTATGT TCCCATAGTA 180
 ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT ATTTACGGTA AACTGCCCCAC 240
 TTGGCAGTAC ATCAAGTGA TCATATGCCA AGTACGCCCC CTATTGACGT CAATGACGGT 300
 AAATGGCCCG CCTGGCATTG TGGCCAGTAC ATGACCTTAT GGGACTTTCC TACTTGGCAG 360
 TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC GGTTTTGGCA GTACATCAAT 420
 GGGCGTGGAT AGCGTTTGA CTCACGGGGA TTTCCAAGTC TCCACCCCAT TGACGTCAAT 480
 GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA AATGTCGTAA CAACTCCGCC 540
 CCATTGACCG AAATGGGCGG TAGGCCGTGA CGGTGGGAGG TCTATATAAG CAGAGCTCGT 600
 TTAGTGAACC GTCAGATCGC CTGGAGACGC CATCCACCGCT GTTTTGACCT CCATAGAAGA 660
 CACCGGGACC GATCCAGCCT CCGGGGCCGG GAACGGTGCA TTGGAACGGG GATTCGCCCGT 720
 GCCAAGAGTG ACGTAAGTAC CGCCTATAGA GTCTATAGGC CCACTTGGCT TCGTTAGAAC 780
 GCGGCTACAA TTAATACATA ACCTTATGTA TCATACACAT ACGATTTAGG TGACACTATA 840
 GAATAACATC CACTTTGCCT TTCTCTCCAC AGGTGTCCAC TCCCAGGTCC AACTGCACCT 900
 CGGTCTATC GATTGAATTC CCCGGGGATC CTCTAGAGAT CCCTCGACCT CGAGTCGACT 960
 TTTTTTTTTT TTTTTGTAGG CCAAAGGGA CTTCCTTTTC TTATTAAT ACTCAGAAGT 1020

FIG. 5B

CTAGGCCACA GCAATCTACT GTTCTCCTCT CATTTTCCTA AACTATTTTG ATACCTATTT 1080
 CTCAGACTTT ATGGGCTATT AGACATTTCT CACATTTCCA TAGATAATAA CTCATCCGTT 1140
 TTGCAACCTG ATTTCTCAATA TTAAGAGATT AAAACTAATG TATATGACTC TCAGTTGACA 1200
 CATACTGAAG TACAGAAAAA TTCCATCATT TCCTTCTGCA AAATGAAAAA GACTTCGTTT 1260
 TCTCAACAGC TGCATCATTI TTTTATGCAT AGAAAAAAAT GTGCAATTAC TCCAAGTACA 1320
 ATCAAAGTCAT TTAACATGGC TTTACCATCA TTGTAGTTAC AGGATATTTT AAAAGAGAAA 1380
 AAAAAATCTC AAAGCACAGG TCCTGCTGTG CAGCAAAGCA ATCAAAATCC TTCATAATAA 1440
 CAGCCTGATG GGATTCAGCA ATCTGAGGAA TAATGAATAA CCACTCTAAT CAGTAAACAG 1500
 GAAAAAGCTA CAACAGTCAC TGAGTAAAAA TTGGACTATC ATCTGTTGAT TCTCTTGATC 1560
 GACATTTCAA ACAATAAATG GAAATGTAAG TATCTCTTAA AAAGAAAAAT AACTTGGTTT 1620
 AGTGTGCTTA ATTTTACCAG GCAGTGAGGA AATTATATAT CACCTTGACT GTCCTGCAGT 1680
 GTTGCCCGT CAATAAAAATG CACAATAAT CTTTTTCATA ATACATGGCC AACTTTATCC 1740
 TATCACTTGA ATATGTCAGG ATAACTGAT TGTGCAGTTG GTTGATAACA TTGTATTTTG 1800
 GAATGGATTA TTTGAATTTG TTTTGTCTACT TTATTATTTG ATATTCTTCT CCAGTGTTCA 1860
 TCTTAIGAAG TTATTTGCAT CTGAATATGA AGAGTCTGTT TCAAAAATAGT CTTCAAGTTT 1920
 CCAACGCAGT GTCTCAAATG TAGGTCGTTT CTTAGGCTCT GCATTCCAGC ACTCCAACAT 1980
 GATGTTGTAA AATTGCTGTG GACAGTTGGA TGGTTGCGGA AGTCTATAGT TTTGAGCCAA 2040

FIG. 5C

CATCTGGATT ACCTGGGCAC CTGTCAATACC ACTGTAAGGC ATTTTGCCAT AAGTAATGAT 2100
 TTCATAAAGA AGGATTCCAA ATGACCATAC ATCGGACTTA ATGCTGAATT TATTACTACG 2160
 AATGGCTTCG GCGGCAGTCC ACTTCACCGG CAGCTTTATT TCGTGTCTAG ATTCATAGAT 2220
 GTCITCATTA TCTACCTTAA AAACTCTGGC AAGTCCAAAA TCTGCTACTT TGTAGATATT 2280
 ATGTTACCCA ACGAGGACAT TTCTGGCAGC CAGATCTCTG TGAATGTAGT TCCGAGACTC 2340
 CAGATAGGCC ATTCCAGAGG CAACCTGTGC CGCCATGTCT ACCTGTTGAG TCAGATGGAT 2400
 TTTTGTATCCA GTGTCAATTT GGAGATATTC TTGCAGACTT CCATGTCTCA TCAACTCTGT 2460
 AATAATATAA ATTGGATCTT CTAAGTGCA AACAGCATAA AGCTGGATAA GCCTTGGATG 2520
 TCTTAGGTTT TTCATTATCT GTGCCCTCCCT CAGGAAGTCA TTTGGATCCA TTGAACCTGG 2580
 TTTTAAATGTT TTCACTGCTA CTGGAGTGGT ATTGTTCCAC AGACCTTCCC ATACTTCGCC 2640
 AAACTGACCA GATCCCAATC GCTTCAGAAG CTGTATGGAG TTGGGGTCTA TCTCCCATTG 2700
 GTCCACGGTT TTATACGACA AATCAAAATGG AGCTGGGACC TGGATCTTTA AGCATGGTTT 2760
 CCCCAGCTTG ACACACAGGC CGTCACTTGT CTTGGTGTAG TGGCTCACAA ATTTCGTTTCCAG 2820
 TGTGAAAAG ATTCTTCTTC GCGTGAGAAA AAATCCCCCT TCATCCAGTC TTTTAAATCT 2880
 GTAGTGTTTT ACAACTGCTC CATCTAAAAC TGAAGAGAG AATTCTCCTT TTTGGCTTTTC 2940
 ACTTTCTCTG ATTAGAAAGG AACCGGTCTT GTTTTCTGAA TATAAATAGTT GTTTCTCTGC 3000
 ATCTGATCTT CCGATTGCTC CAAAGAACCA CGGCTCTGCC TGTAGGCTTC TGTCCCTCAGC 3060

FIG. 5D

CACGTAGTTA GAAGGAATAT AGCCTTGTAG TTGCTGACTG GAGCCATCTC GTCTTTTCTC 3120
 CAAGTGTCTG GCAAACCACC AGCCCTCATG CAAAGTGTCC AGAACTTGAA GTTTGTCAAC 3180
 TGCTCGGAAG CTCAAAGTCTT CAGCAGTCCG AGCCTGGTAA TCAAACAAG CCACAAAGTA 3240
 GTGGCCATGC CTCTGTGACT GGGGAGAGCA AAGGGCCCTT GGATTTTCAA TCACGGTTGA 3300
 CTTGTCTGCC TCCGTGGACA AACAGGGGAG ATAGGGTTCT AGGTACTCCC AGAGCCTCTG 3360
 ACAGATGTTG CTCATTGTGC CTTGGTGGG AGAAGAGGAG CAGGGCTTCT CCTCTCCCC 3420
 TTAGTCTCTG CGATCCACCT TATCTTCCTT CACCAGGCAA CTTTGAAGTC AGCACCAACT 3480
 CACCATACTT CGGAGAGTAT GCAAAGTCCC GTTTCAGATC AGTCCAGCAG CTGGGTTGCA 3540
 GCAAGTCCTA CCTGGAGAGA CTTACCCGGCT TGCCTTCTGT GGCTGGAGGT GCTACCCCGA 3600
 GGCAAAACTG AGCAGGAGCT GGGCAGCTGC TCACTAGGAA GGTGTCITTTT CTTCTTATCT 3660
 GCTTAAGAAT CCCACAACAA AAATAAATA AAATAAAG GCCTTTATTT AGACAAATAT 3720
 CTGAGAACAG AATGGTGCCA TCTTGCCTTT TGTCCCAATA AAAAGTTAGC AAGAGGAAGC 3780
 TACTAACCCC TGGTAAAACC TCCACGTCCTT GCTTTCGCCA GGGTCGACTC GAGGGATCCT 3840
 CCATACCTAC CAGTTCCTGG CCTGCAGGTC GCGGCCCGGA CTCTAGAGTC GACCTGCAGA 3900
 AGCTTGGCCG CCATGGCCCA ACTTGTATT TGCAGCTTAT AATGGTTACA AATAAAGCAA 3960
 TAGCATCACA AATTTACAA ATAAAGCATT TTTTCACTG CATCTAGTT GTGGTTTGTG 4020
 CAAACTCATC AATGTATCTT ATCATGTCTG GATCGGGAAT TAATTCGGCG CAGCACCATG 4080

FIG. 5E

4140 GCCTGAATA ACCTCTGAAA GAGGAACTTG GTTAGGTACC TTCTGAGGCG GAAAGAACCA
 4200 GCTGTGGAAT GTGTGTCAGT TAGGTTGTGG AAAGTCCCCA GGCTCCCCAG CAGGCAGAAG
 4260 TATGCAAAGC ATGCATCTCA ATTAGTCAGC AACCAGGTGT GGAAGTCCC CAGGCTCCCC
 4320 AGCAGGCAGA AGTATGCAAA GCATGCATCT CAATTAGTCA GCNACCATAG TCCCGCCCTT
 4380 AACTCCGCCC ATCCCCGCCC TAACTCCGCC CAGTTCCGCC CAFTCTCCGC CCCATGGCTG
 4440 ACTAATTTTT TTTATTTATG CAGAGGCCGA GGCCGCCCTCG GCCTCTGAGC TATTCAGAA
 4500 GTAGTGAGGA GGCTTTTTG GAGGCCTAGG CTTTTGCMAA AAGCTGTTAA CAGCTTGGCA
 4560 CTGGCCGTCG TTTTACAACG TCGTGACTGG GAAAACCCCTG GCGTTACCCA ACTTAATCGC
 4620 CTTGCAGCAC ATCCCCCTTT CGCCAGCTGG CGTAATAGCG AAGAGGCCCG CACCGATCGC
 4680 CCTTCCCAAC AGTTGGCCAG CCTGAATGGC GAATGGGCC TGATGGGGTA TTTTCTCCTT
 4740 ACGCATCTGT GCGGTATTC ACACCGCATA CGTCAAAGCA ACCATAGTAC GCGCCCTGTA
 4800 GCGGCGCATT AAGCGGGCGG GGTGTGTGG TTACGGCCAG CGTGACCCGT AACTTTGCCA
 4860 GCGCCCTAGC GCCCCCTCCT TTGCGTTTCT TCCCCTCCTT TCTCGCCACG TTCGCCCCGT
 4920 TTCCCCGTCA AGCTCTAAAT CGGGGGCTCC CTTTAGGTT CCGATTTAGT GCTTTACGGC
 4980 ACCTCGACCC CAAAAAATT GATTTGGGTG ATGGTTACAG TAGTGGGCCA TCGCCCTGAT
 5040 AGACGGTTTT TCGCCCTTTG ACGTTGGAGT CCACGTTCTT TAATAGTGA CTCCTTGTTC
 5100 AAACCTGGAAC AACACTCAAC CCTATCTCGG GCTATTCTTT TGATTATAA GGGATTTTGC

FIG. 5F

CGATTTCCGGC CTATTGGTTA AAAAATGAGC TGATTTAACA AAAATTTAAC GCGAATTTTA 5160
 ACAAATATTT AACGTTTACA ATTTTATGGT GCACTCTCAG TACAATCTGC TCTGATGCCG 5220
 CATAGTTAAG CCAGCCCCCGA CACCCGCCAA CACCCGCTGA CCGCCCTGTA CGGGCTTGTG 5280
 TGCTCCCGGC ATCCGGCTTAC AGACAAGCTG TGACCGTCTC CGGGAGCTGC ATGTGTCAGA 5340
 GGTTTTCACC GTCATCACCG AAACGGCCGA GACGAAAGGG CCTCGTGATA CGCCTATTTT 5400
 TATAGGTTAA TGCATGATA ATAATGGTTT TGTTTATTTT TCTAAATACA TTCAAATATG TATCCGCTCA 5460
 ATGTGCCGGG AACCCCTATT TGTTTATTTT TCTAAATACA TTCAAATATG TATCCGCTCA 5520
 TGAGACAAATA ACCCTGATAA ATGCTTCAAT AATATTTGAAA AAGGAAGAGT ATGAGTATTC 5580
 AACATTTCCG TGTGCCCTT ATTCCTTTT TTGCCGCATT TTGCCITCCT GTTTTGTGCTC 5640
 ACCCAGAAAC GCTGGTGAAA GTAAAAGATG CTGAAGATCA GTTGGGTGCA CGAGTGGGTT 5700
 ACATCGAACT GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT 5760
 TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCCG GGTATTATCC CGTATTGACG 5820
 CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG GTTGAGTACT 5880
 CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAAATTA TGCAGTGCIG 5940
 CCATAACCAT GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACCATC GGAGGACCGA 6000
 AGGAGCTAAC CGCTTTTGTG CACAACATGG GGGATCATGT AACTCGCCTT GATCGTTGGG 6060
 AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG CCTGTAGCAA 6120

FIG. 5G

TGGCAACAAC GTTGGCAAA CTATTAAGTGGCGAACTACT TACTCTAGCT TCCCGGCAAC 6180
 AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC ACTTCTGCGC TCGGCCCTTC 6240
 CGGCTGGCTG GTTTATTGCT GATAAATCTG GAGCCGGTGA GCGTGGGTCT CGCGGTATCA 6300
 TTGCAGCACT GGGGCCAGAT GGTAAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGGA 6360
 GTCAGGCAAC TATGGATGAA CGAATAGAC AGATCGCTGA GATAGGTGCC TCACTGATTA 6420
 AGCATTGGTA ACTGTCAGAC CAAGTTTACT CATATATACT TTAGATTGAT TTAAAACCTTC 6480
 ATTTTAAAT TAAAAGGATC TAGGTGAAGA TCCTTTMTGA TAATCTCATG ACCAAAATCC 6540
 CTTAACGTGA GTTTTCGTTT CACTGAGCGT CAGACCCCGT AGAAAAGATC AAAGGATCTT 6600
 CTTGAGATCC TTTTTTTCTG CGCGTAATCT GCTGCTTGCA AACAAAAAAA CCACCCGCTAC 6660
 CAGCGGTGGT TTGTTTGCCG GATCAAGAGC TACCAACTCT TTTTCCGAAG GTAACCTGGCT 6720
 TCAGCAGAGC GCAGATACCA AATACTGTTT TTCTAGTGTA GCCGTAGTTA GGCCACCACT 6780
 TCAAGAATCT TGTAGCACCG CCTACATACC TCGCTCTGCT AATCCTGTTA CCAGTGGCTG 6840
 CTGCCAGTGG CGATAAGTCC TGTCTTACCG GGTGGACTC AAGACGATAG TTACCCGGATA 6900
 AGGCCAGCG GTCCGGCTGA ACGGGGGGT CGTGCACACA GCCCAGCTTG GAGCGAACGA 6960
 CCTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA AAGGCCACG CTTCCCGAAG 7020
 GGAGAAAGGC GGACAGGTAT CCGGTAAGCG GCAGGGTCCG AACAGGAGAG CGCACGAGGG 7080
 AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCTTGT CCGGTTTCCG CACCTCTGAC 7140

FIG. 5H

TTGAGCGTCG ATTTTGTGA TGCTCGTCAG GGGGGCGGAG CCTATGGAAA AACGCCAGCA 7200
 ACGCGGCCCTT TTTACGGTTC CTGGCCCTTT GCTGGCCTTT TGCTCACATG TTCTTTCCCTG 7260
 CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCCTT TGAGTGAGCT GATACCGCTC 7320
 GCCGCAGCCG AACGACCGAG CGCAGCGAGT CAGTGAGCGA GGAAGCGGAA GAGCGCCCAA 7380
 TACGCAAACC GCCTCTCCCC GCGCGTTGGC CGATTCATTA ATGCAGCTGG CACGACAGGT 7440
 TTCCCGACTG GAAAGCGGGC AGTGAGCGCA ACGCAATTAA TGTGAGTTAG CTCACTCATT 7500
 AGGCACCCCA GGCCTTACAC TTTATGCTTC CGGCTCGTAT GTTGTGTGGA ATTGTGAGCC 7560
 GATAACAATT TCACACAGGA AACAGCTATG ACATGATTAC GAATTAA 7607

FIG. 5I

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

FIG. 5J

Tyr Thr Lys Thr Ser Asp Gly Leu Cys Val Lys Leu Gly Lys Pro Cys
 195 200 205
 Leu Lys Ile Gln Val Pro Ala Pro Phe Asp Leu Ser Tyr Lys Thr Val
 210 215 220
 Asp Gln Trp Glu Ile Asp Arg Asn Ser Ile Gln Leu Leu Lys Arg Leu
 225 230 235 240
 Gly Ser Gly Gln Phe Gly Glu Val Trp Glu Gly Leu Trp Asn Asn Thr
 245 250 255
 Thr Pro Val Ala Val Lys Thr Leu Lys Pro Gly Ser Met Asp Pro Asn
 260 265 270
 Asp Phe Leu Arg Glu Ala Gln Ile Met Lys Asn Leu Arg His Pro Lys
 275 280 285
 Leu Ile Gln Leu Tyr Ala Val Cys Thr Leu Glu Asp Pro Ile Tyr Ile
 290 295 300
 Ile Thr Glu Leu Met Arg His Gly Ser Leu Gln Glu Tyr Leu Gln Asn
 305 310 315 320
 Asp Thr Gly Ser Lys Ile His Leu Thr Gln Gln Val Asp Met Ala Ala
 325 330 335
 Gln Val Ala Ser Gly Met Ala Tyr Leu Glu Ser Arg Asn Tyr Ile His
 340 345 350
 Arg Asp Leu Ala Ala Arg Asn Val Leu Val Gly Glu His Asn Ile Tyr
 355 360 365
 Lys Val Ala Asp Phe Gly Leu Ala Arg Val Phe Lys Val Asp Asn Glu
 370 375 380

FIG. 5K

Asp Ile Tyr Glu Ser Arg His Glu Ile Lys Leu Pro Val Lys Trp Thr
 385 390 395 400
 Ala Pro Glu Ala Ile Arg Ser Asn Lys Phe Ser Ile Lys Ser Asp Val
 405 410 415
 Trp Ser Phe Gly Ile Leu Leu Tyr Glu Ile Ile Thr Tyr Gly Lys Met
 420 425 430
 Pro Tyr Ser Gly Met Thr Gly Ala Gln Val Ile Gln Met Leu Ala Gln
 435 440 445
 Asn Tyr Arg Leu Pro Gln Pro Ser Asn Cys Pro Gln Gln Phe Tyr Asn
 450 455 460
 Ile Met Leu Glu Cys Trp Asn Ala Glu Pro Lys Glu Arg Pro Thr Phe
 465 470 475 480
 Glu Thr Leu Arg Trp Lys Leu Glu Asp Tyr Phe Glu Thr Asp Ser Ser
 485 490 495
 Tyr Ser Asp Ala Asn Asn Phe Ile Arg *
 500 505

FIG. 6

GCGGCCGAG AGAAGCAGA GGATGGGGCT TAGCAGCTGG CAGAGCCAGG AGCGGGGAGG 60
 TAGCAGAAAG ACCACAAGTA CAAAGAAGTC CTGAACCTTT GGTTTTGCTG CTGCAGCCCA 120
 TTGAGAGTGA CGACATGGAG CACAAGACCC TGAAGATCAC CGACTTTGGC CTGGCCCGAG 180
 AGTGGACAA AACACACAA ATGAGTGCCG CNGGCACCTA CNCCTGGATG GCTCCTGAGG 240
 TTATCAAGGC CTCACCTTC TCTAAGGGCA GTGACGTCTG GAGTTTTGGG GTGCTGCTGT 300
 GGGAACTGCT GACCGGGGAG NTGCCATACC GTGGCATTGA CTGCCTTGCT GTGGCCTATG 360
 GCGTAGCTGT TAACAAGCTC ACACTGCCAT CCATCCACCT GGCC 404

FIG. 7A

ATGAGAGCGT TGGCGGCGGA CGGCGGCCAG CTGCCGCTGC TCGTTGTTTT TTCTGCAATG 60
 ATATTTGGGA CTATTACAAA TCAAGATCTG CCTGTGATCA AGTGTGTTTT AATCAATCAT 120
 AAGAACAATG ATTCAATCAGT GGGGAAGTCA TCATCATATC CCATGGTATC AGAATCCCCG 180
 GAAGACCTCG GGTGTGGGTT GAGACCCCGAG AGCTCAGGGA CAGTGTACGA AGCTGCCGCT 240
 GTGGAAGTGG ATGTATCTGC TTCCATCACA CTGCAAGTGC TGGTCCGATGC CCCAGGGGAC 300
 ATTTCCCTGC TCTGGGTCTT TAAGCACAGC TCCCTGAATT GCCAGCCACA TTTTGATTTA 360
 CAAAAACAGAG GAGTTGTTC CATCGTCAAT TTGAAAATGA CAGAAACCCA AGCTGGAGAA 420
 TACCCTACTTT TTATTCAGAG TGAAGCTACC AATTACACAA TATTGTTTAC AGTGAGTATA 480
 AGAANTACCC TGCTTTACAC ATTAAGAAGA CCTTACTTTA GAAAAATGGA AAACCCAGGAC 540
 GCCCTGGTCT GCATATCTGA GAGCGTTCCA GAGGGATCC TGGAAATGGT GCTTTGGGAT 600
 TCACAGGGGG AAAGCTGTAA AGAAGAAAGT CCAGCTGTTG TTA AAAAAGGA GGA AAAAGTG 660
 CTTCAATGAAT TATTTGGGAC GGACATAAGG TGCTGTGCCA GAAATGAACT GGGCAGGGAA 720
 TGCACCAGGC TGTTACAAAT AGATCTAAAT CAAACTCCTC AGACCACATT GCCACAATTA 780
 TTTCTTAAAG TAGGGGAACC CTTATGGATA AGGTGAAAG CTGTTTCATGT GAAACCATGA 840
 TTCGGGCTCA CCTGGGAATT AGAAAACAAA GCACTCGAGG AGGGCAACTA CTTTGAGATG 900
 AGTACCTATT CAACAAACAG AACTATGATA CGGATTCCTGT TTGCTTTTTGT ATCATCAGTG 960
 GCAAGAAACG ACACCGGATA CTACACTTGT TCCTCTTCAA AGCATCCCAG TCAATCAGCT 1020
 TTGGTTACCA TCGTAGAAAA GGGATTTATA AATGCTACCA ATTCAAGTGA AGATTATGAA 1080

FIG. 7B

ATTGACCAAT ATGAAGAGTT TTGTTTTTCT GTCCAGTTTA AAGCCTACCC ACAAAATCAGA 1140
 TGTACGTGGA CCTTCTCTCG AAAATCAATTT CCTTGTGAGC AAAAGGGTCT TGATAACGGA 1200
 TACAGCATAT CCAAGTTTTG CAATCATRAG CACCAGCCAG GAGAATATAT ATTCCATGCA 1260
 GAAAATGATG ATGCCCAATT TACCAAAATG TTCACGGCTGT ATATAAGAAG GAAACCTCAA 1320
 GTCCCTCGCAG AAGCTTCGGC AAGTCAGGGC TCCTGTTTCT CGGATGGATA CCCATTACCA 1380
 TCTTGGACCT GGAAGAAGTG TTCAGACAAG TCTCCCACT GCACAGAAGA GATCACAGAA 1440
 GGAGTCTGGA ATAGAAAGGC TAACAGAAA GTGTTGGAC AGTGGGTGC GAGCAGTACT 1500
 CTAACATGA GTGAAGCCAT AAAAGGGTTC CTGGTCAAGT GCTGTGCATA CAATTCCCTT 1560
 GGCACATCTT GTGAGACGAT CCTTTTAAAC TCTCCAGGCC CCTTCCCTTT CATCCAAGAC 1620
 AACATCTCAT TCTATGCAAC AATTGGTGT TGTCTCCTCT TCATTTGTCGT TTTAAACCTG 1680
 CTAATTTGTC ACAGTACAA AAAGCAATTT AGGTATGAAA GCCAGCTACA GATGGTACAG 1740
 GTGACCCGGAT CCTCAGATTA TGAGTACTTC TACGTTGATT TCAGAGAATA TGAATATGAT 1800
 GTCAAATGGG AGTTTCCAAG AGAAAATTTA GAGTTTGGGA AGGTACTAGG ATCAGGTGCT 1860
 TTTGGAAAAG TGATGAACGC AACAGCTTAT GGAATTAGCA AAACAGGAGT CTCAAATCCAG 1920
 GTTACCGTCA AAATGCTGAA AGMAAAGCA GACAGCTCTG AAAGAGAGGC ACTCATGTCA 1980
 GAACTCAAGA TGATGACCCA GCTGGGAAGC CACGAGAATA TTGTGAACCT GCTGGGGGGC 2040
 TGCACACTGT CAGGACCAAT TTACTTTGATT TTTGAATACT GTTGCTATGG TGATCTTCTC 2100
 AACTATCTAA GAAGTAAAG AGAAAATTT CACAGGACTT GGACAGAGAT TTTCAAGGAA 2160

FIG. 7C

CACAATTTCA GTTTTTACC CACTTTCCAA TCACATCCAA ATTCCAGCAT GCCTGGTTCA 2220
 AGAGAAGTTC AGATACACCC GGACTCGGAT CAAATCTCAG GGCTTCATGG GAATTCATTT 2280
 CACTCTGAAG ATGAAATTGA ATATGAAAAC CAAAAAGGC TGGAAAGA GAAGGACTTG 2340
 AATGTGCTTA CATTGGAAGA TCTTCTTTGC TTGTCATATC AAGTTGCCAA AGGAATGGAA 2400
 TTTCIGGAAT TTAAGTCGTG TGTTACACAGA GACCCTGGCCG CCAGGAACGT GCTTGTCAAC 2460
 CACGGGAAG TGGTGAAGAT ATGTGACTTT GGATTTGGCTC GAGATATCAT GAGTGATTC 2520
 AACTATGTTG TCAGGGGCAA TGCCCGTCTG CCTGTAAAAT GGATGGCCCC CGAAAGCCCTG 2580
 TTTGAAGGCA TCTACACCAT TAAGAGTGAT GTCTGGTTCAT ATGGAATATT ACTGTGGGAA 2640
 ATCTTCTCAC TTGGTGTGAA TCCTTACCCT GGCAITCCGG TTGATGCTAA CTTCTACAAA 2700
 CTGATTCAAA ATGGATTTAA AATGGATCAG CCATTTTATG CTACAGAAGA AATATACATT 2760
 ATAAATGCAAT CCTGCTGGC TMTTGACTCA AGAAAACGGC CATCCCTCCC TAATTTGACT 2820
 TCGTTTTTAG GATGTCAGCT GGCAGATGCA GAAGAAGCGA TGTATCAGAA TGTGGATGGC 2880
 CGTGTTTCGG AATGTCCTCA CACCTACCAA AACAGGGGAC CTTTCAGCAG AGAGATGGAT 2940
 TTGGGGCTAC TCTCTCCGCA GGCTCAGSTC GAAGATTCGT AGAGGAACAA TTTAGTTTTA 3000
 AGGACTTCAT CCCTCCACCT ATCCCTAACA GGCTGTAGAT TACCAAACA AGGTTAATTT 3060
 CATCACTAAA AGAAAATCTA TTATCAACTG CTGCTTCACC AGACTTTTCT CTAGAGAGCG 3120

FIG. 8A

TCGGGTCCA CCGCCCGAG GAGAGTCAGA CCTGGGGGG CGAGGGCCCC CCAAACCTCAG 60
 TTCGGATCCT ACCCGAGTGA GCGGGGGCC ATG GAG CTC CGG GTG CTG CTC TGC 113
 Met Glu Leu Arg Val Leu Leu Cys 5
 TGG GCT TCG TTG GCC GCA GCT TTG GAA GAG ACC CTG CTG AAC ACA AAA 161
 Trp Ala Ser Leu Ala Ala Leu Glu Thr Leu Leu Asn Thr Lys 20
 10
 TTG GAA ACT GCT GAT CTG AAG TGG GTG ACA TTC CCT CAG GTG GAC GGG 209
 Leu Glu Thr Ala Asp Leu Lys Trp Val Thr Phe Pro Gln Val Asp Gly 40
 25
 CAG TGG GAG GAA CTG AGC GGC CTG GAT GAG GAA CAG CAC AGC GTG CGC 257
 Gln Trp Glu Glu Leu Ser Gly Leu Asp Glu Glu His Ser Val Arg 55
 45
 ACC TAC GAA GTG TGT GAC GTG CAG CGT GCC CCG GGC CAG GCC CAC TGG 305
 Thr Tyr Glu Val Cys Asp Val Gln Arg Ala Pro Gly Gln Ala His Trp 70
 60
 CTT CGC ACA GGT TGG GTC CCA CGG CGG GGC GGC GTC CAC GTG TAC GCC 353
 Leu Arg Thr Gly Trp Val Pro Arg Arg Gly Ala Val His Val Tyr Ala 85
 75
 ACG CTG CGC TTC ACC ATG CTC GAG TGC CTG TCC CTG CCT CGG GCT GGG 401
 Thr Leu Arg Phe Thr Met Leu Glu Cys Leu Ser Leu Pro Arg Ala Gly 100
 90
 CGC TCC TGC AAG GAG ACC TTC ACC GTC TTC TAC TAT GAG AGC GAT GCG 449
 Arg Ser Cys Lys Glu Thr Phe Thr Val Phe Tyr Tyr Glu Ser Asp Ala 120
 105
 GAC ACG GCC ACG GCC CTC ACG CCA GCC TGG ATG GAG AAC CCC TAC ATC 497
 Asp Thr Ala Thr Ala Leu Thr Pro Ala Trp Met Glu Asn Pro Tyr Ile 135
 125
 130

FIG. 8B

AAG GTG GAC ACG GTG GCC GCG GAG CAT CTC ACC CGG AAG CGC CCT GGG 545
 Lys Val Asp Thr Val Ala Ala Glu His Leu Thr Arg Lys Arg Pro Gly 150
 140
 GCC GAG GCC ACC GGG AAG GTG AAT GTC AAG ACG CTG CGT CTG GGA CCG 593
 Ala Glu Ala Thr Gly Lys Val Asn Val Lys Thr Leu Arg Leu Gly Pro 165
 155
 CTC AGC AAG GCT GGC TTC TAC CTG GCC TTC CAG GAC CAG GGT GCC TGC 641
 Leu Ser Lys Ala Gly Phe Tyr Leu Ala Phe Gln Asp Gln Gly Ala Cys 180
 170
 ATG GCC CTG CTA TCC CTG CAC CTC TTC TAC AAA AAG TGC GCC CAG CTG 689
 Met Ala Leu Leu Ser Leu His Leu Phe Tyr Lys Cys Ala Gln Leu 200
 185
 ACT GTG AAC CTG ACT CGA TTC CCG GAG ACT GTG CCT CGG GAG CTG GTT 737
 Thr Val Asn Leu Thr Arg Phe Pro Glu Thr Val Pro Arg Glu Leu Val 210
 205
 GTG CCC GTG GCC GGT AGC TGC GTG GAT GCC GTC CCC GCC CCT GGC 785
 Val Pro Val Ala Ala Gly Ser Cys Val Val Asp Ala Val Pro Ala Pro Gly 230
 220
 CCC AGC CCC AGC CTC TAC TGC CGT GAG GAT GGC CAG TGG GCC GAA CAG 833
 Pro Ser Pro Ser Leu Tyr Cys Arg Glu Asp Gly Gln Trp Ala Glu Gln 245
 235
 CCG GTC ACG GCC TGC AGC TGT GCT CCG GGG TTC GAG GCA GCT GAG GGG 881
 Pro Val Thr Gly Cys Ser Cys Ala Pro Gly Phe Glu Ala Ala Glu Gly 260
 250
 AAC ACC AAG TGC CGA GCC TGT GCC CAG GGC ACC TTC AAG CCC CTG TCA 929
 Asn Thr Lys Cys Arg Ala Cys Ala Gln Gly Thr Phe Lys Pro Leu Ser 280
 265 270 275

FIG. 8C

GGA GAA GGG TCC TGC CAG CCA TGC CCA GCC AAT AGC CAC TCT AAC ACC 977
 Gly Glu Gly Ser Cys Gln Pro Cys Pro Ala Asn Ser His Ser Asn Thr 295
 285
 ATT GGA TCA GCC GTC TGC CAG TGC CGC GTC GGG TAC TTC CGG GCA CGC 1025
 Ile Gly Ser Ala Val Cys Gln Cys Arg Val Gly Tyr Phe Arg Ala Arg 310
 300
 ACA GAC CCC CGG GGT GCA CCC TGC ACC ACC CCT CCT TCG GCT CCG CGG 1073
 Thr Asp Pro Arg Gly Ala Pro Cys Thr Thr Pro Pro Ser Ala Pro Arg 325
 315
 AGC GTG GTT TCC CGC CTG AAC GGC TCC TCC CTG CAC CTG GAA TGG AGT 1121
 Ser Val Val Ser Arg Leu Asn Gly Ser Ser Leu His Leu Glu Trp Ser 340
 335
 GCC CCC CTG GAG TCT GGT GGC CGA GAG GAC CTC ACC TAC GCC CTC CGC 1169
 Ala Pro Leu Glu Ser Gly Gly Arg Glu Asp Leu Thr Tyr Ala Leu Arg 360
 345
 TGC CGG GAG TGC CGA CCC GGA GGC TCC TGT GCG CCC TGC GGG GGA GAC 1217
 Cys Arg Glu Cys Arg Pro Gly Gly Ser Cys Ala Pro Cys Gly Gly Asp 375
 365
 CTG ACT TTT GAC CCC GGC CCC CGG GAC CTG GTG GAG CCC TGG GTG GTG 1265
 Leu Thr Phe Asp Pro Gly Pro Arg Asp Leu Val Glu Pro Trp Val Val 390
 380
 GTT CGA GGG CTA CGT CCT GAC TTC ACC TAT ACC TTT GAG GTC ACT GCA 1313
 Val Arg Gly Leu Arg Pro Asp Phe Thr Tyr Thr Phe Glu Val Thr Ala 405
 395
 TTG AAC GGG GTA TCC TCC TTA GCC ACG GGC CCC GTC CCA TTT GAG CCT 1361
 Leu Asn Gly Val Ser Ser Leu Ala Thr Gly Pro Val Pro Phe Glu Pro 420
 410
 415

FIG. 8D

GTC AAT GTC ACC ACT GAC CGA GAG GTA CCT CCT GCA GTG TCT GAC ATC 1409
 Val Asn Val Thr Thr Asp Arg Glu Val Pro Pro Ala Val Ser Asp Ile 440
 425 430

CCG GTG ACG CCG TCC TCA CCC AGC AGC TTG AGC CTG GCC TGG GCT GTT 1457
 Arg Val Thr Arg Ser Ser Pro Ser Leu Ser Leu Ala Trp Ala Val 455
 445 450

CCC CCG GCA CCC AGT GGG GCT GTG CTG GAC TAC GAG GTC AAA TAC CAT 1505
 Pro Arg Ala Pro Ser Gly Ala Val Leu Asp Tyr Glu Val Lys Tyr His 470
 460 465

GAG AAG GCC GCG GAG GGT CCC AGC AGC GTG CCG TTC CTG AAG ACG TCA 1553
 Glu Lys Gly Ala Glu Gly Pro Ser Ser Val Arg Phe Leu Lys Thr Ser 485
 475 480

GAA AAC CCG GCA GAG CTG CCG GGG CTG AAG CCG GGA GCC AGC TAC CTG 1601
 Glu Asn Arg Ala Glu Leu Arg Gly Leu Lys Arg Gly Ala Ser Tyr Leu 495
 490 500

GTG CAG GTA CCG GCG TCT GAG GCC GCG TAC GGG CCC TTC GGC CAG 1649
 Val Gln Val Arg Ala Arg Ser Glu Ala Gly Tyr Gly Pro Phe Gly Gln 515
 505 510

GAA CAT CAC AGC CAG ACC CAA CTG GAT GAG AGC GAG GGC TGG CCG GAG 1697
 Glu His His Ser Gln Thr Gln Leu Asp Glu Ser Glu Gly Trp Arg Glu 535
 525 530

CAG CTG GCC CTG ATT GCG GCG ACG GCA GTC GTG GGT GTG GTC CTG GTC 1745
 Gln Leu Ala Leu Ile Ala Gly Thr Ala Val Val Gly Val Val Leu Val 545
 540 550

CTG GTG GTC ATT GTG GTC GCA GTT CTC TGC CTC AGG AAG CAG AGC AAT 1793
 Leu Val Val Ile Val Val Ala Val Leu Cys Leu Arg Lys Gln Ser Asn 565
 555 560

FIG. 8E

1841 GGG AGA GAA GCA GAA GAA TAT TCG GAC AAA CAC GGA CAG TAT CTC ATC GGA
 Gly Arg Glu Ala Glu Tyr Ser Asp Lys His Gly Gln Tyr Leu Ile Gly
 570 575 580

1889 CAT GGT ACT AAG GTC TAC ATC GAC CCC TTC ACT TAT GAA GAC CCT AAT
 His Gly Thr Lys Val Tyr Ile Asp Pro Phe Thr Tyr Glu Asp Pro Asn
 585 590 595 600

1937 GAG GCT GTG AGG GAA TTT GCA AAA GAG ATC GAT GTC TCC TAC GTC AAG
 Glu Ala Val Arg Glu Phe Ala Lys Glu Ile Asp Val Ser Tyr Val Lys
 605 610 615

1985 ATT GAA GAG GTG ATT GGT GCA GGT GAG TTT GGC GAG GTG TGC CGG GGG
 Ile Glu Glu Val Ile Gly Ala Glu Phe Gly Glu Val Cys Arg Gly
 620 625 630

2033 CCG CTC AAG GCC CCA GGG AAG AAG GAG AGC TGT GTG GCA ATC AAG ACC
 Arg Leu Lys Ala Pro Gly Lys Lys Glu Ser Cys Val Ala Ile Lys Thr
 635 640 645

2081 CTG AAG GGT GGC TAC ACG GAG CCG CAG CCG CGT GAG TTT CTG AGC GAG
 Leu Lys Gly Gly Tyr Thr Glu Arg Gln Arg Arg Glu Phe Leu Ser Glu
 650 655 660

2129 GCC TCC ATC ATG GGC CAG TTC GAG CAC CCC AAT ATC ATC CGC CTG GAG
 Ala Ser Ile Met Gly Gln Phe Glu His Pro Asn Ile Ile Arg Leu Glu
 665 670 675 680

2177 GGC GTG GTC ACC AAC AGC ATG CCC GTC ATG ATT CTC ACA GAG TTC ATG
 Gly Val Val Thr Asn Ser Met Pro Val Met Ile Leu Thr Glu Phe Met
 685 690 695

2225 GAG AAC GGC GCC CTG GAC TCC TTC CTG CGG CTA AAC GAC GGA CAG TTC
 Glu Asn Gly Ala Leu Asp Ser Phe Leu Arg Leu Asn Asp Gly Gln Phe
 700 705 710

FIG. 8F

ACA GTC ATC CAG CTC GTG GGC ATG CTG CGG GGC ATC GCC TCG GGC ATG 2273
 Thr Val Ile Gln Leu Val Gly Met Leu Arg Gly Ile Ala Ser Gly Met 725
 715 720
 CGG TAC CTT GCC GAG ATG AGC TAC GTC CAC CGA GAC CTG GCT GCT CGC 2321
 Arg Tyr Leu Ala Glu Met Ser Tyr Val His Arg Asp Leu Ala Ala Arg 740
 730 735
 AAC ATC CTA GTC AAC AGC AAC CTC GTC TGC AAA GTG TCT GAC TTT GGC 2369
 Asn Ile Leu Val Asn Ser Asn Leu Val Cys Lys Val Ser Asp Phe Gly 760
 745 750 755
 CTT TCC CGA TTC CTG GAG GAG AAC TCT TCC GAT CCC ACC TAC ACG AGC 2417
 Leu Ser Arg Phe Leu Glu Glu Asn Ser Ser Asp Pro Thr Tyr Thr Ser 775
 765 770
 TCC CTG GGA GGA AAG ATT CCC ATC CGA TGG ACT GCC CCG GAG GCC ATT 2465
 Ser Leu Gly Gly Lys Ile Pro Ile Arg Trp Thr Ala Pro Glu Ala Ile 790
 780 785
 GCC TTC CGG AAG TTC ACT TCC GCC AGT GAT GCC TGG AGT TAC GGG ATT 2513
 Ala Phe Arg Lys Phe Thr Ser Ala Ser Asp Ala Trp Ser Tyr Gly Ile 805
 795 800
 GTG ATG TGG GAG GTG ATG TCA TTT GGG GAG AGG CCG TAC TGG GAC ATG 2561
 Val Met Trp Glu Val Met Ser Phe Gly Glu Arg Pro Tyr Trp Asp Met 820
 810 815
 AGC AAT CAG GAC GTG ATC AAT GCC ATT GAA CAG GAC TAC CGG CTG CCC 2609
 Ser Asn Gln Asp Val Ile Asn Ala Ile Glu Gln Asp Tyr Arg Leu Pro 840
 825 830 835
 CCG CCC CCA GAC TGT CCC ACC TCC CTC CAC CAG CTC ATG CTG GAC TGT 2657
 Pro Pro Pro Asp Cys Pro Thr Ser Leu His Gln Leu Met Leu Asp Cys 850
 845 850 855

FIG. 8G

TGG CAG AAA GAC CGG AAT GCC CGG CCC CGC TTC CCC CAG GTG GTC AGC 2705
 Trp Gln Lys Asp Arg Asn Ala Arg Pro Arg Phe Pro Gln Val Val Ser 860
 870
 GCC CTG GAC AAG ATG ATC CGG AAC CCC GCC AGC CTC AAA ATC GTG GCC 2753
 Ala Leu Asp Lys Met Ile Arg Asn Pro Ala Ser Leu Lys Ile Val Ala 875
 880
 CGG GAG AAT GGC GGG GCC TCA CAC CCT CTC CTG GAC CAG CGG CAG CCT 2801
 Arg Glu Asn Gly Gly Ala Ser His Pro Leu Leu Asp Gln Arg Gln Pro 890
 895
 CAC TAC TCA GCT TTT GGC TCT GTG GGC GAG TGG CTT CGG GCC ATC AAA 2849
 His Tyr Ser Ala Phe Gly Ser Val Gly Glu Trp Leu Arg Ala Ile Lys 900
 905
 ATG GGA AGA TAC GAA GAA AGT TTC GCA GCC GCT GGC TTT GGC TCC TTC 2897
 Met Gly Arg Tyr Glu Glu Ser Phe Ala Ala Ala Gly Phe Gly Ser Phe 910
 915
 920
 GAG CTG GTC AGC CAG ATC TCT GCT GAG GAC CTG CTC CGA ATC GGA GTC 2945
 Glu Leu Val Ser Gln Ile Ser Ala Glu Asp Leu Leu Arg Ile Gly Val 930
 935
 940
 ACT CTG GCG GGA CAC CAG AAG AAA ATC TTG GCC AGT GTC CAG CAC ATG 2993
 Thr Leu Ala Gly His Gln Lys Lys Ile Leu Ala Ser Val Gln His Met 945
 950
 955
 AAG TCC CAG GCC AAG CCG GGA ACC CCG GGT GGG ACA GGA GGA CCG GCC 3041
 Lys Ser Gln Ala Lys Pro Gly Thr Pro Gly Gly Thr Gly Gly Pro Ala 960
 965
 970
 CCG CAG TAC TGA CCT GCA GGA ACT CCC CAC CCC AGG GAC ACC GCC TCC 3089
 Pro Gln Tyr * Pro Ala Gly Thr Pro His Pro Arg Asp Thr Ala Ser 975
 980
 985
 990
 1000

FIG. 8H

3137 CCA TTT TCC GGG GCA GAG TGG GGA CTC ACA GAG GCC CCC AGC CCT GTG
 Pro Phe Ser Gly Ala Glu Trp Gly Leu Thr Glu Ala Pro Ser Pro Val
 1005 1010 1015
 3185 CCC CGC TGG ATT GCA CTT TGA GCC CGT GGG GTG AGG AGT TGG CAA TTT
 Pro Arg Trp Ile Ala Leu * Ala Arg Gly Val Arg Ser Trp Gln Phe
 1020 1025 1030
 3233 GGA GAG ACA GGA TTT GGG GGT TCT GCC ATA ATA GGA GGG GAA AAT CAC
 Gly Glu Thr Gly Phe Gly Gly Ser Ala Ile Ile Gly Gly Glu Asn His
 1035 1040 1045
 3281 CCC CCA GCC ACC TCG GGG AAC TCC AGA CCA AGG GTG AGG GCG CCT TTC
 Pro Pro Ala Thr Ser Gly Asn Ser Arg Pro Arg Val Arg Ala Pro Phe
 1050 1055 1060
 3329 CCT CAG GAC TGG GTG TGA CCA GAG GAA AAG GAA GTG CCC AAC ATC TCC
 Pro Gln Asp Trp Val * Pro Glu Glu Lys Glu Val Pro Asn Ile Ser
 1065 1070 1075
 3377 CAG CCT CCC CAG GTG CCC CCC TCA CCT TGA TGG GTG CGT TCC CGC AGA
 Gln Pro Pro Gln Val Pro Pro Ser Pro * Trp Val Arg Ser Arg Arg
 1085 1090 1095
 3425 CCA AAG AGA GTG TGA CTC CCT TGC CAG CTC CAG AGT GGG GGG GCT GTC
 Pro Lys Arg Val * Leu Pro Cys Gln Leu Gln Ser Gly Ala Val
 1100 1105 1110
 3473 CCA GGG GGC AAG AAG GGG TGT CAG GGC CCA GTG ACA AAA TCA TTG GGG
 Pro Gly Gly Lys Lys Gly Cys Gln Gly Pro Val Thr Lys Ser Leu Gly
 1115 1120 1125
 3521 TTT GTA GTC CCA ACT TGC TGC TGT CAC CAC CAA ACT CAA TCA TTT TTT
 Phe Val Val Pro Thr Cys Cys Cys His His Gln Thr Gln Ser Phe Phe
 1130 1135 1140

FIG. 8 I

TCC CTT GTA AAT GCC CCT CCC CCA GCT GCT GCC TTC ATA TTG AAG GTT 3569
 Ser Leu Val Asn Ala Pro Pro Ala Ala Ala Phe Ile Leu Lys Val 1145
 1150

TTT GAG TTT TGT TTT TGG TCT TAA TTT TTC TCC CCG TTC CCT TTT TGT 1160
 Phe Glu Phe Cys Phe Trp Ser * Phe Phe Ser Pro Phe Pro Phe Cys 1175
 1165

TTC TTC GTT TTG TTT TTC TAC CGT CCT TGT CAT AAC TTT GTG TTG GAG 3617
 Phe Phe Val Leu Phe Phe Tyr Arg Pro Cys His Asn Phe Val Leu Glu 1180
 1185

GGA ACC TGT TTC ACT ATG GCC TCC TTT GCC CAA GTT GAA ACA GGG GCC 3665
 Gly Thr Cys Phe Thr Met Ala Ser Phe Ala Gln Val Glu Thr Gly Ala 1190
 1195

CAT CAT CAT GTC TGT TTC CAG AAC AGT GCC TTG GTC ATC CCA CAT CCC 3713
 His His His Val Cys Phe Gln Asn Ser Ala Leu Val Ile Pro His Pro 1200
 1205

CGG ACC CCG CCT GGG ACC CCC AAG CTG TGT CCT ATG AAG GGG TGT GGG 3761
 Arg Thr Pro Pro Gly Thr Pro Lys Leu Cys Pro Met Lys Gly Cys Gly 1210
 1215

GTG AGG TAG TGA AAA GGG CCG TAG TTG GTG GAA CCC AGA AAC GGA 3809
 Val Arg * * Lys Gly Arg * Leu Val Val Glu Pro Arg Asn Gly 1220
 1235

CGC CGG TGC TTG GAG GGG TTC TTA AAT TAT ATT TAA AAA AGT AAC TTT 3857
 Arg Arg Cys Leu Glu Gly Phe Leu Asn Tyr Ile * Lys Ser Asn Phe 1240
 1245

TTG TAT AAA TAA AAG AAA ATG GGA CGT GTC CCA GCT CCA GGG GTA 3905
 Leu Tyr Lys * Lys Lys Met Gly Arg Val Pro Ala Pro Gly Val 1250
 1255

AAAAAAAAAAAA AAAAAAAAAA 3950
 3969

FIG. 9

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

FIG. 10

Asn Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr Asp Phe Gly
 1 5 10 15
 Leu Ala Arg Leu Leu Glu Gly Asp Glu Lys Glu Tyr Asn Ala Asp Gly
 20 25 30
 Gly Lys Met Pro Ile Lys Trp Met Ala Leu Glu Cys Ile His Tyr Arg
 35 40 45
 Lys Phe Thr His Gln Ser
 50

FIG. 11

Asn Cys Met Leu Ala Gly Asp Met Thr Val Cys Val Ala Asp Phe Gly
 1 5 10 15
 Leu Ser Trp Lys Ile Tyr Ser Gly Ala Thr Ile Val Arg Gly Cys Ala
 20 25 30
 Ser Lys Leu Pro Val Lys Trp Leu Ala Leu Gly Ser Leu Ala Asp Asn
 35 40 45
 Leu Tyr Thr Val His Ser
 50

FIG. 12

Asn Cys Leu Val Gly Lys Asn Tyr Thr Ile Lys Ile Ala Asp Phe Gly
 1 5 10 15
 Met Ser Arg Asn Leu Tyr Ser Gly Asp Tyr Tyr
 20 25

FIG. 13

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

FIG. 14

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

FIG. 15A

```

1 TCGGGTCGGA CCCACGCGCA GCGGCCGGAG ATGCAGCGGG GCGCCGCGCT GTGCCTGCGA
  AGCCCAGCCT GGGTGCGCGT CGCCGGCCTC TACGTGCCCC CGCGGCGCGA CACGGACGCT
1 M Q R G A A L C L R

61 CTGTGGCTCT GCCTGGGACT CCTGGACGGC CTGGTGAGTG GCTACTCCAT GACCCCCCGG
  GACACCGAGA CGGACCCTGA GGACCTGCCG GACCACTCAC CGATGAGGTA CTGGGGGGGC
11 L W L C L G L L D G L V S G Y S M T P P

121 ACCTTGAACA TCACGGAGGA GTCACACGTC ATCGACACCG GTGACAGCCT GTCCATCTCC
  TGGAACTTGT AGTGCCTCCT CAGTGTGCAG TAGCTGTGGC CACTGTCCGA CAGGTAGAGG
31 T L N I T E E S H V I D T G D S L S I S

181 TGCAGGGGAC AGCACCCCCT CGAGTGGGCT TGGCCAGGAG CTCAGGAGGC GCCAGCCACC
  ACGTCCCCTG TCGTGGGGGA GCTCACCCGA ACCGGTCCTC GAGTCCTCCG CGGTGGGTGG
51 C R G Q H P L E W A W P G A Q E A P A T

241 GGAGACAAGG ACAGCGAGGA CACGGGGGTG GTGCGAGACT GCGAGGGCAC AGACGCCAGG
  CCTCTGTTC TGTGCTCCT GTGCCCCAC CACGCTCTGA CGCTCCCGTG TCTGCGGTCC
71 G D K D S E D T G V V R D C E G T D A R

301 CCCTACTGCA AGGTGTTGCT GCTGCACGAG GTACATGCCA ACGACACAGG CAGCTACGTC
  GGGATGACGT TCCACAACGA CGACGTGCTC CATGTACGGT TGCTGTGTCC GTCGATGCAG
91 P Y C K V L L L H E V H A N D T G S Y V

361 TGCTACTACA AGTACATCAA GGCACGCATC GAGGGCACCA CGGCCGCCAG CTCCTACGTC
  ACGATGATGT TCATGTAGTT CCGTGCCTAG CTCCCGTGGT GCCGGCGGTC GAGGATGCAC
111 C Y Y K Y I K A R I E G T T A A S S Y V

421 TTCGTGAGAG ACTTTGAGCA GCCATTCATC AACAAAGCCTG ACACGCTCTT GGTCAACAGG
  AAGCACTCTC TGAAACTCGT CGGTAAGTAG TTGTTCCGAC TGTGCGAGAA CCAGTTGTCC
131 F V R D F E Q P F I N K P D T L L V N R

481 AAGGACGCCA TGTGGGTGCC CTGTCTGGTG TCCATCCCCG GCCTCAATGT CACGCTGCGC
  TTCCTGCGGT ACACCCACGG GACAGACCAC AGGTAGGGGC CGGAGTTACA GTGCGACGCG
151 K D A M W V P C L V S I P G L N V T L R

541 TCGCAAAGCT CGGTGTCTGTG GCCAGACGGG CAGGAGGTGG TGTGGGATGA CCGGCGGGGC
  AGCGTTTCGA GCCACGACAC CGGTCTGCCC GTCCCTCCACC ACACCCTACT GGCCGCCCCG
171 S Q S S V L W P D G Q E V V W D D R R G

601 ATGCTCGTGT CCACGCCACT GCTGCACGAT GCCCTGTACC TGCAGTGCGA GACCACCTGG
  TACGAGCACA GGTGCGGTGA CGACGTGCTA CGGGACATGG ACGTCACGCT CTGGTGGACC
191 M L V S T P L L H D A L Y L Q C E T T W

661 GGAGACCAGG ACTTCCTTTC CAACCCCTTC CTGGTGCACA TCACAGGCAA CGAGCTCTAT
  CCTCTGGTCC TGAAGGAAAG GTTGGGGAAG GACCACGTGT AGTGTCCGTT GCTCGAGATA
211 G D Q D F L S N P F L V H I T G N E L Y

```


FIG. 15B

721 GACATCCAGC TGTGGCCAG GAAGTCGCTG GAGCTGCTGG TAGGGGAGAA GCTGGTCCTG
 CTGTAGGTCG ACAACGGGTC CTTTCAGCGAC CTCGACGACC ATCCCTCTTT CGACCAGGAC
 231 D I Q L L P R K S L E L L V G E K L V L

781 AACTGCACCG TGTGGCTGA GTTAACTCA GGTGTCACCT TTGACTGGGA CTACCCAGGG
 TTGACGTGGC ACACCCGACT CAAATTGAGT CCACAGTGGG AACTGACCCT GATGGGTCCC
 251 N C T V W A E F N S G V T F D W D Y P G

841 AAGCAGGCAG AGCGGGGTAA GTGGGTGCC GAGCGACGCT CCCAGCAGAC CCACACAGAA
 TTCGTCCGTC TCGCCCATC CACCCACGGG CTCGCTGCGA GGGTCGTCTG GGTGTGCTTT
 271 K Q A E R G K W V P E R R S Q Q T H T E

901 CTCTCCAGCA TCCTGACCAT CCACAACGTC AGCCAGCAGC ACCTGGGCTC GTATGTGTGC
 GAGAGTTCGT AGGACTGGTA GGTGTTCAGT TCGGTCTGTC TGGACCCGAG CATAACACAG
 291 L S S I L T I H N V S Q H D L G S Y V C

961 AAGGCCAACA ACGGCATCCA GCGATTTCCG GAGAGCACCG AGGTCATTGT GCATGAAAT
 TTCCGGTTGT TGCCGTAGGT CGCTAAAGCC CTCTCGTGGC TCCAGTAACA CGTACTTTTA
 311 K A N N G I Q R F R E S T E V I V H E N

1021 CCCTTCATCA GCGTCGAGTG GCTCAAAGGA CCCATCCTGG AGGCCACGGC AGGAGACGAG
 GGAAGTAGT CGCAGCTCAC CGAGTTTCCT GGGTAGGACC TCCGGTGCCG TCCTCTGCTC
 331 P F I S V E W L K G P I L E A T A G D E

1081 CTGGTGAAGC TGCCCGTGAA GCTGGCAGCG TACCCCCCGC CCGAGTTCCA GTGGTACAAG
 GACCACTTCG ACGGGCACTT CGACCGTCGC ATGGGGGGCG GGCTCAAGGT CACCATGTTT
 351 L V K L P V K L A A Y P P P E F Q W Y K

1141 GATGGAAAGG CACTGTCCGG GCGCCACAGT CCACATGCC TGGTGTCAA GGAGGTGACA
 CTACCTTTCC GTGACAGGCC CGCGGTGTCA GGTGTACGGG ACCACGAGTT CCTCCACTGT
 371 D G K A L S G R H S P H A L V L K E V T

1201 GAGGCCAGCA CAGGCACCTA CACCCTCGCC CTGTGGAAct CCGCTGCTGG CCTGAGGCGC
 CTCCGGTTCGT GTCCGTGGAT GTGGGAGCGG GACACCTTGA GGCACGACC GGACTCCGCG
 391 E A S T G T Y T L A L W N S A A G L R R

1261 AACATCAGCC TGGAGCTGGT GGTGAATGTG CCCCCCAGA TACATGAGAA GGAGGCCTCC
 TTGTAGTCCG ACCTCGACCA CCACTTACAC GGGGGGGTCT ATGTACTCTT CCTCCGGAGG
 411 N I S L E L V V N V P P Q I H E K E A S

1321 TCCCCAGCA TCTACTCGCG TCACAGCCGC CAGGCCCTCA CCTGCACGGC CTACGGGGTG
 AGGGGGTTCG AGATGAGCGC AGTGTGCGCG GTCCGGGAGT GGACGTGCCG GATGCCCCAC
 431 S P S I Y S R H S R Q A L T C T A Y G V

1381 CCCCTGCCTC TCAGCATCCA GTGGCACTGG CCGCCCTGGA CACCCTGCAA GATGTTTGCC
 GGGGACGGAG AGTCGTAGGT CACCGTGACC GCCGGGACCT GTGGGACGTT CTACAAACGG
 451 P L P L S I Q W H W R P W T P C K M F A

1441 CAGCGTAGTC TCCGGCGGCG GCAGCAGCAA GACCTCATGC CACAGTCCCG TGAAGTGGAG
 GTCGCATCAG AGGCCGCCGC CGTCGTCTGT CTGGAGTACG GTGTACGGC ACTGACCTCC
 471 Q R S L R R R Q Q Q D L M P Q C R D W R

FIG. 15C

```

1501 GCGGTGACCA CGCAGGATGC CGTGAACCCC ATCGAGAGCC TGGACACCTG GACCGAGTTT
      CGCCACTGGT GCGTCTACG GCACTTGGGG TAGCTCTCGG ACCTGTGGAC CTGGCTCAAA
491 A V T T Q D A V N P I E S L D T W T E F

1561 GTGGAGGGAA AGAATAAGAC TGTGAGCAAG CTGGTGATCC AGAATGCCAA CGTGCTGCCC
      CACCTCCCTT TCTTATCTG ACACTCGTTC GACCACTAGG TCTTACGGTT GCACAGACGG
511 V E G K N K T V S K L V I Q N A N V S A

1621 ATGTACAAGT GTGTGGTCTC CAACAAGGTG GGCCAGGATG AGCGGCTCAT CTACTTCTAT
      TACATGTTCA CACACCAGAG GTTGTTCCAC CCGTCTTAC TCGCCGAGTA GATGAAGATA
531 M Y K C V V S N K V G Q D E R L I Y F Y

1681 GTGACCACCA TCCCCGACGG CTTACCATC GAATCCAAGC CATCCGAGGA GCTACTAGAG
      CACTGGTGGT AGGGGCTGCC GAAGTGGTAG CTTAGGTTCG GTAGGCTCCT CGATGACTCTC
551 V T T I P D G F T I E S K P S E E L L E

1741 GGCCAGCCGG TGCTCCTGAG CTGCCAAGCC GACAGCTACA AGTACGAGCA TCTGCGCTGG
      CCGGTCCGCC ACGAGGACTC GACGGTTCGG CTGTGATGT TCATGCTCGT AGACGCGACC
571 G Q P V L L S C Q A D S Y K Y E H L R W

1801 TACCGCCTCA ACCTGTCCAC GCTGCACGAT GCGCACGGGA ACCCGTTTCT GCTCGACTGC
      ATGGCGGAGT TGGACAGGTG CGACTGCTA GCGGTGCCCT TGGGCGAAGA CGAGCTGACG
591 Y R L N L S T L H D A H G N P L L L D C

1861 AAGAACGTGC ATCTGTTCGC CACCCCTCTG GCCGCCAGCC TGGAGGAGGT GGCACCTGGG
      TTCTTGCACG TAGACAAGCG GTGGGGAGAC CGGCGGTCGG ACCTCCTCCA CCGTGGACCC
611 K N V H L F A T P L A A S L E E V A P G

1921 GCGCGCCACG CCACGCTCAG CCTGAGTATC CCCC GCGTCG CGCCCGAGCA CGAGGGCCAC
      CGCGCGGTGC GGTGCGAGTC GGACTCATAG GGGGCGCAGC GCGGGCTCGT GCTCCCGGTG
631 A R H A T L S L S I P R V A P E H E G H

1981 TATGTGTGCG AAGTGC AAGA CCGGCGCAGC CATGACAAGC ACTGCCACAA GAAGTACCTG
      ATACACACGC TTCACGTTCT GGCCGCGTGC GTACTGTTTCG TGACGGTGTTC TTTCATGGAC
651 Y V C E V Q D R R S H D K H C H K K Y L

2041 TCGGTGACAG CCCTGGAAGC CCCTCGGCTC ACGCAGA AACT TGACCGACCT CCTGGTGAAC
      AGCCACGTCC GGGACCTTCG GGGAGCCGAG TGCCTCTTGA ACTGGCTGGA GGACCACTTG
671 S V Q A L E A P R L T Q N L T D L L V N

2101 GTGAGCGACT CGCTGGAGAT GCAGTGCTTG GTGGCCGGAG CGCACGCGCC CAGCATCGTG
      CACTCGCTGA GCGACCTCTA CGTCACGAAC CACCGGCCCTC GCGTGCGCGG GTCGTAGCAC
691 V S D S L E M Q C L V A G A H A P S I V

2161 TGGTACAAAG ACGAGAGGCT GCTGGAGGAA AAGTCTGGAG TCGACTTGGC GGACTCCAAC
      ACCATGTTTC TGCTCTCCGA CGACCTCCTT TTCAGACCTC AGCTGAACCG CCTGAGGTTG
711 W Y K D E R L L E E K S G V D L A D S N

2221 CAGAAGCTGA GCATCCAGCG CGTGCGCGAG GAGGATGCGG GACGCTATCT GTGCAGCGTG
      GTCTTCGACT CGTAGGTCGC GCACGCGCTC CTCCTACGCC CTGCGATAGA CACGTGCGAC
731 Q K L S I Q R V R E E D A G R Y L C S V

```

FIG. 15D

2281 TGCAACGCCA AGGGCTGCGT CAACTCCTCC GCCAGCGTGG CCGTGAAGG CTCCGAGGAT
 ACGTTGCGGT TCCCGACGCA GTTGAGGAGG CGGTCGCACC GGCACCTTCC GAGGCTCCTA
 751 C N A K G C V N S S A S V A V E G S E D

2341 AAGGGCAGCA TGGAGATCGT GATCCTTGTC GGTACCGGCG TCATCGCTGT CTTCTTCTGG
 TTCCCGTCGT ACCTCTAGCA CTAGGAACAG CCATGGCCGC AGTAGCGACA GAAGAAGACC
 771 K G S M E I V I L V G T G V I A V F F W

2401 GTCCTCCTCC TCCTCATCTT CTGTAACATG AGGAGGCCGG CCCACGCAGA CATCAAGACG
 CAGAGGAGG AGGAGTAGAA GACATTGTAC TCCTCCGGCC GGGTGCCTCT GTAGTTCTGC
 791 V L L L L I F C N M R R P A H A D I K T

2461 GGCTACCTGT CCATCATCAT GGACCCCGGG GAGGTGCCTC TGGAGGAGCA ATGCGAATAC
 CCGATGACACA GGTAGTAGTA CCTGGGGCCC CTCCACGGAG ACCTCCCTGT TACGCTTATG
 811 G Y L S I I M D P G E V P L E E Q C E Y

2521 CTGTCTTACG ATGCCAGCCA GTGGGAATTC CCCCAGAGAG GGCTGCACCT GGGGAGAGTG
 GACAGGATGC TACGGTCGGT CACCCTTAAG GGGGCTCTCG CCGACGTGGA CCCCTCTCAC
 831 L S Y D A S Q W E F P R E R L H L G R V

2581 CTCGGCTACG GCGCCTTCGG GAAGGTGGTG GAAGCCTCCG CTTTCGGCAT CCACAAGGGC
 GAGCCGATGC CGCGGAAGCC CTTCCACCAC CTTCCGGAGG GAAAGCCGTA GGTGTTCCCG
 851 L G Y G A F G K V V E A S A F G I H K G

2641 AGCAGCTGTG ACACCGTGGC CGTGAAAATG CTGAAAGAGG GCGCCACGGC CAGCGAGCAC
 TCGTCGACAC TGTGGCACCG GCACTTTTAC GACTTTCTCC CGCGGTGCGG GTCGCTCGTG
 871 S S C D T V A V K M L K E G A T A S E H

2701 CGCGCGCTGA TGTCCGAGCT CAAGATCCTC ATTCACATCG GCAACCACCT CAACGTGGTC
 GCGCGCGACT ACAGCCTCGA GTTCTAGGAG TAAGTGTAGC CGTTGGTGGG GTTGCACCAG
 891 R A L M S E L K I L I H I G N H L N V V

2761 AACCTCCTCG GGGCGTGCAC CAAGCCGCAG GGCCCCCTCA TGGTGATCGT GGAGTTCTGC
 TTGGAGGAGC CCCGCACGTG GTTCGGCGTC CCGGGGGAGT ACCACTAGCA CCTCAAGACG
 911 N L L G A C T K P Q G P L M V I V E F C

2821 AAGTACGGCA ACCTCTCCAA CTTCTGCGC GCCAAGCGGG ACGCCTTCAG CCCCTGCGCG
 TTCATCCCGT TGGAGAGGTT GAAGGACGCG CGTTTCGCCC TCGGGAAGTC GGGGACGCGC
 931 K Y G N L S N F L R A K R D A F S P C A

2881 GAGAAGTCTC CCGAGCAGCG CGGACGCTTC CGCGCCATGG TGGAGCTCGC CAGGCTGGAT
 CTCTTCAGAG GGCTCGTTCG CCTGCGAAG GCGCGGTACC ACCTCGAGCG GTCCGACCTA
 951 E K S P E Q R G R F R A M V E L A R L D

2941 CGGAGGCGGC CGGGGAGCAG CGACAGGGTC CTCTTCGCGC GGTTCCTCGAA GACCGAGGGC
 GCCTCCGCCG GCCCCTCGTC GCTGTCCCAG GAGAAGCGCG CCAAGAGCTT CTGGCTCCCG
 971 R R R P G S S D R V L F A R F S K T E G

3001 GGAGCGAGGC GGGCTTCTCC AGACCAAGAA GCTGAGGACC TGTGGCTGAG CCCGCTGACC
 CCTCGCTCCG CCCGAAGAGG TCTGTTCTTT CGACTCCTGG ACACCGACTC GGGCGACTGG
 991 G A R R A S P D Q E A E D L W L S P L T

FIG. 15E

3061 ATGGAAGATC TTGTCTGCTA CAGCTTCCAG GTGGCCAGAG GGATGGAGTT CCTGGCTTCC
 TACCTTCTAG AACAGACGAT GTCGAAGGTC CACCGGTCTC CCTACCTCAA GGACCGAAGG
 1011 M E D L V C Y S F Q V A R G M E F L A S
 3121 CGAAAGTGCA TCCACAGAGA CCTGGCTGCT CGGAACATTC TGCTGTGCGA AAGCGACGTG
 GCTTTCACGT AGGTGTCTCT GGACCGACGA GCCTTGTAAG ACGACAGCCT TTCGCTGCAC
 1031 R K C I H R D L A A R N I L L S E S D V
 3181 GTGAAGATCT GTGACTTTGG CCTTGCCCGG GACATCTACA AAGACCCTGA CTACGTCCGC
 CACTTCTAGA CACTGAAACC GGAACGGGCC CTGTAGATGT TTCTGGGACT GATGCAGGCG
 1051 V K I C D F G L A R D I Y K D P D Y V R
 3241 AAGGGCAGTG CCCGGCTGCC CCTGAAGTGG ATGGCCCCTG AAAGCATCTT CGACAAGGTG
 TTCGCTCAC GGGCCGACGG GGACTTCACC TACCGGGGAC TTTCGTAGAA GCTGTTCAC
 1071 K G S A R L P L K W M A P E S I F D K V
 3301 TACACCACGC AGAGTGACGT GTGGTCCTTT GGGGTGCTTC TCTGGGAGAT CTTCTCTCTG
 ATGTGGTGCG TCTCACTGCA CACCAGGAAA CCCCACGAAG AGACCCTCTA GAAGAGAGAC
 1091 Y T T Q S D V W S F G V L L W E I F S L
 3361 GGGGCCTCCC CGTACCCTGG GGTGCAGATC AATGAGGAGT TCTGCCAGCG GCTGAGAGAC
 CCCCAGAGGG GCATGGGACC CCACGTCTAG TTACTCCTCA AGACGGTCCG CACTCTCTG
 1111 G A S P Y P G V Q I N E E F C Q R L R D
 3421 GGACAAGGA TGAGGGCCCC GGAGCTGGCC ACTCCCAGCA TACGCCGAT CATGCTGAAC
 CCGTGTTCCT ACTCCCGGGG CCTCGACCGG TGAGGGCGGT ATGCGGCGTA GTACGACTTG
 1131 G T R M R A P E L A T P A I R R I M L N
 3481 TGCTGGTCCG GAGACCCAA GCGAGACCT GCATTCTCGG AGCTGGTGA GATCCTGGGG
 ACGACCAGGC CTCTGGGGTT CCGCTCTGGA CGTAAGAGCC TCGACCACCT CTAGGACCCC
 1151 C W S G D P K A R P A F S E L V E I L G
 3541 GACCTGCTCC AGGGCAGGGG CCTGCAAGAG GAAGAGGAGG TCTGCATGGC CCCGCGCAGC
 CTGGACGAGG TCCCGTCCCC GGACGTTCTC CTCTCCTCC AGACGTACCG GGGCGCGTCCG
 1171 D L L Q G R G L Q E E E E V C M A P R S
 3601 TCTCAGAGCT CAGAAGAGGG CAGCTTCTCG CAGGTGTCCA CCATGGCCCT ACACATCGCC
 AGAGTCTCGA GTCTTCTCCC GTCGAAGAGC GTCCACAGGT GGTACCGGA TGTGTAGCGG
 1191 S Q S S E E G S F S Q V S T M A L H I A
 3661 CAGGCTGACG CTGAGGACAG CCCGCCAAGC CTGCAGCGCC ACAGCCTGGC CGCCAGGTAT
 GTCCGACTGC GACTCCTGTC GGGCGGTTCG GACGTCGCGG TGTCGGACCG GCGGTCCATA
 1211 Q A D A E D S P P S L Q R H S L A A R Y
 3721 TACAACGGG TGTCCTTTCC CGGGTGCCTG GCCAGAGGGG CTGAGACCCG TGGTTCCTCC
 ATGTTGACCC ACAGGAAAGG GCCCACGGAC CGGTCTCCCC GACTCTGGGC ACCAAGGAGG
 1231 Y N W V S F P G C L A R G A E T R G S S
 3781 AGGATGAAGA CATTTGAGGA ATTCCCCATG ACCCCAACGA CCTACAAAGG CTCTGTGGAC
 TCCTACTTCT GTAAACTCCT TAAGGGGTAC TGGGGTTGCT GGATGTTTCC GAGACACCTG
 1251 R M K T F E E F P M T P T T Y K G S V D

FIG. 15F

3841 AACCAGACAG ACAGTGGGAT GGTGCTGGCC TCGGAGGAGT TTGAGCAGAT AGAGAGCAGG
 TTGGTCTGTC TGTCACCCTA CCACGACCGG AGCCTCCTCA AACTCGTCTA TCCTCGTCC
 1271 N Q T D S G M V L A S E E F E Q I E S R
 3901 CATAGACAAG AAAGCGGCTT CAGGTAGCTG AAGCAGAGAG AGAGAAGGCA GCATACGTCA
 GTATCTGTTC TTTCGCCGAA GTCCATCGAC TTCGTCTCTC TCTCTTCCGT CGTATGCAGT
 1291 H R Q E S G F R O
 3961 GCATTTTCTT CTCTGCACCT ATAAGAAAGA TCAAAGACTT TAAGACTTTC GCTATTTCTT
 CGTAAAAGAA GAGACGTGAA TATCTTTTCT AGTTTCTGAA ATTCTGAAAG CGATAAAGAA
 4021 CTGCTATCTA CTACAAACTT CAAAGAGGAA CCAGGAGGCC AAGAGGAGCA TGAAAGTGG
 GACGATAGAT GATGTTTGAA GTTCTCCTT GGTCCCTCCG TTCTCCTCGT ACTTTCACCT
 4081 CAAGGAGTGT GACCACTGAA GCACCACAGG GAGGGGTTAG GCCTCCGGAT GACTGCGGGC
 GTTCCTCACA CTGGTGACTT CGTGGTGTCC CTCCCAATC CGGAGGCCTA CTGACGCCCC
 4141 AGGCCTGGAT AATATCCAGC CTCCCACAAG AAGCTGGTGG AGCAGAGTGT TCCCTGACTC
 TCCGGACCTA TTATAGGTCG GAGGGTGTTC TTCGACCACC TCGTCTCACA AGGGACTGAG
 4201 CTCCAAGGAA AGGGAGACGC CCTTTCATGG TCTGCTGAGT AACAGGTGCC TTCCCAGACA
 GAGGTTCTTT TCCCTCTGCG GGAAAGTACC AGACGACTCA TTGTCCACGG AAGGGTCTGT
 4261 CTGGCGTTAC TGCTTGACCA AAGAGCCCTC AAGCGGCCCT TATGCCAGCG TGACAGAGGG
 GACCGCAATG ACGAACTGGT TTCTCGGGAG TTCGCCGGA ATACGGTCCG ACTGTCTCCC
 4321 CTCACCTCTT GCCTTCTAGG TCACTTCTCA CAATGTCCCT TCAGCACCTG ACCCTGTGCC
 GAGTGGAGAA CGGAAGATCC AGTGAAGAGT GTTACAGGGA AGTCGTGGAC TGGGACACGG
 4381 CGCCAGTTAT TCCTTGGTAA TATGAGTAAT ACATCAAAGA GTAGT
 GCGGTCAATA AGGAACCATT ATACTCATT TGTAGTTTCT CATCA

FIG. 16A

```

1 ATGGCTGGGA TTTTCTATTT CGCCCTATTT TCGTGTCTCT TCGGGATTTG
  TACCGACCCT AAAAGATAAA GCGGGATAAA AGCACAGAGA AGCCCTAAAC
1 MetAlaGlyI lePheTyrPh eAlaLeuPhe SerCysLeuP heGlyIleCy
  CGACGCTGTC ACAGGTTCCA GGGTATACCC CGCGAATGAA GTTACCTTAT
  GCTGCGACAG TGTCCAAGGT CCCATATGGG GCGCTTACTT CAATGGAATA
  sAspAlaVal ThrGlySerA rgValTyrPr oAlaAsnGlu ValThrLeuLeu

101 TGGATTCCAG ATCTGTTTCAG GGAGAACTTG GGTGGATAGC AAGCCCTCTG
  ACCTAAGGTC TAGACAAGTC CCTCTTGAAC CCACCTATCG TTCGGGAGAC
  35 AspSerAr gSerValGln GlyGluLeuG lyTrpIleAl aSerProLeu
  GAAGGAGGGT GGGAGGAAGT GAGTATCATG GATGAAAAAA ATACACCAAT
  CTTCTCCCA CCCTCCTTCA CTCATAGTAC CTACTTTTTT TATGTGGTTA
  GluGlyGlyT rpGluGluVa lSerIleMet AspGluLysA snThrProIle

201 CCGAACCTAC CAAGTGTGCA ATGTGATGGA ACCCAGCCAG AATAACTGGC
  GGCTTGGATG GTTCACACGT TACACTACCT TGGGTCGGTC TTATTGACCG
  68 ArgThrTyr GlnValCysA snValMetGl uProSerGln AsnAsnTrpL
  TACGAACTGA TTGGATCACC CGAGAAGGGG CTCAGAGGGT GTATATTGAG
  ATGCTTGACT AACCTAGTGG GCTCTTCCCC GAGTCTCCCA CATATAACTC
  euArgThrAs pTrpIleThr ArgGluGlyA laGlnArgVa lTyrIleGlu

301 ATTAAATTC A CCTTGAGGGA CTGCAATAGT CTTCCGGGCG TCATGGGGAC
  TAATTTAAGT GGA ACTCCCT GACGTTATCA GAAGGCCCGC AGTACCCCTG
101 IleLysPheT hrLeuArgAs pCysAsnSer LeuProGlyV alMetGlyTh
  TTGCAAGGAG ACGTTTAACC TGTACTACTA TGAATCAGAC AACGACAAAG
  AACGTTCCCTC TGCAAATTGG ACATGATGAT ACTTAGTCTG TTGCTGTTTC
  rCysLysGlu ThrPheAsnL euTyrTyrTy rGluSerAsp AsnAspLysGlu

```

FIG. 16B

401 AGCGTTTCAT CAGAGAGAAC CAGTTTGTCA AAATTGACAC CATTGCTGCT
 TCGCAAAGTA GTCTCTCTTG GTCAAACAGT TTTAACTGTG GTAACGACGA
 135 ArgPheIl eArgGluAsn GlnPheValL ysIleAspTh rIleAlaAla
 GATGAGAGCT TCACCCAAGT GGACATTGGT GACAGAATCA TGAAGCTGAA
 CTACTCTCGA AGTGGGTTC A CTGTCTTAGT ACTTCGACTT
 AspGluSerP heThrGlnVa lAspIleGly AspArgIleM etLysLeuAsn
 501 CACCGAGATC CGGGATGTAG GGCCATTAAG CAAAAAGGGG TTTTACCTGG
 GTGGCTCTAG GCCCTACATC CCGGTAATTC GTTTTTCCCC AAAATGGACC
 168 ThrGluIle ArgAspValG lyProLeuSe rLysLysGly PheTyrLeuA
 CTTTTCAGGA TGTGGGGGCC TGCATCGCCC TGGTATCAGT CCGTGTGTTC
 GAAAAGTCCT ACACCCCCGG ACGTAGCGGG ACCATAGTCA GGCACACAAG
 laPheGlnAs pValGlyAla CysIleAlaL euValSerVa lArgValPhe
 601 TATAAAAAGT GTCCACTCAC AGTCCGCAAT CTGGCCCAGT TTCCTGACAC
 ATATTTTTCA CAGGTGAGTG TCAGGCGTTA GACCGGGTCA AAGGACTGTG
 201 TyrLysLysC ysProLeuTh rValArgAsn LeuAlaGlnP heProAspTh
 CATCACAGGG GCTGATACGT CTTCCCTGGT GGAAGTTCGA GGCTCCTGTG
 GTAGTGTCCC CGACTATGCA GAAGGGACCA CCTTCAAGCT CCGAGGACAC
 rIleThrGly AlaAspThrS erSerLeuVa lGluValArg GlySerCysVal
 701 TCAACAACCTC AGAAGAGAAA GATGTGCCAA AAATGTACTG TGGGGCAGAT
 AGTTGTTGAG TCTTCTCTTT CTACACGGTT TTTACATGAC ACCCCGTCTA
 235 AsnAsnSe rGluGluLys AspValProL ysMetTyrCy sGlyAlaAsp
 GGTGAATGGC TGGTACCCAT TGGCAACTGC CTATGCAACG CTGGGCATGA
 CCACTTACCG ACCATGGGTA ACCGTTGACG GATACGTTGC GACCCGTACT
 GlyGluTrpL euValProIl eGlyAsnCys LeuCysAsnA laGlyHisGlu
 801 GGAGCGGAGC GGAGAATGCC AAGCTTGCAA AATTGGATAT TACAAGGCTC
 CCTCGCCTCG CCTCTTACGG TTCGAACGTT TTAACCTATA ATGTTCCGAG
 268 GluArgSer GlyGluCysG lnAlaCysLy sIleGlyTyr TyrLysAlaL
 TCTCCACGGA TGCCACCTGT GCCAAGTGCC CACCCCACAG CTACTCTGTC
 AGAGGTGCCT ACGGTGGACA CGGTTACGG GTGGGGTGTC GATGAGACAG
 euSerThrAs pAlaThrCys AlaLysCysP roProHisSe rTyrSerVal

FIG. 16C

```

901 TGGGAAGGAG CCACCTCGTG CACCTGTGAC CGAGGCTTTT TCAGAGCTGA
    ACCCTTCCTC GGTGGAGCAC GTGGACACTG GCTCCGAAAA AGTCTCGACT
301 TrpGluGlyA laThrSerCy sThrCysAsp ArgGlyPheP heArgAlaAs
    CAACGATGCT GCCTCTATGC CCTGCACCCG TCCACCATCT GCTCCCCTGA
    GTTGCTACGA CGGAGATACG GGACGTGGGC AGGTGGTAGA CGAGGGGACT
    pAsnAspAla AlaSerMetP roCysThrAr gProProSer AlaProLeuAsn
1001 ACTTGATTTT AAATGTCAAC GAGACATCTG TGAAC TTGGA ATGGAGTAGC
    TGAAC TAAAG TTTACAGTTG CTCTGTAGAC ACTTGAACCT TACCTCATCG
    335 LeuIleSe rAsnValAsn GluThrSerV alAsnLeuGl uTrpSerSer
    CCTCAGAATA CAGGTGGCCG CCAGGACATT TCCTATAATG TGGTATGCAA
    GGAGCTTTAT GTCCACCGGC GTCCTGTAA AGGATATTAC ACCATACGTT
    ProGlnAsnT hrGlyGlyAr gGlnAspIle SerTyrAsnV alValCysLys
1101 GAAATGTGGA GCTGGTGACC CCAGCAAGTG CCGACCCTGT GGAAGTGGGG
    CTTTACACCT CGACCACTGG GGTGCTTCAC GGCTGGGACA CCTTCACCCC
    368 LysCysGly AlaGlyAspP roSerLysCy sArgProCys GlySerGlyV
    TCCACTACAC CCCACAGCAG AATGGCTTGA AGACCACCAA AGGCTCCATC
    AGGTGATGTG GGGTGTGTCGTC TTACCGAACT TCTGGTGGTT TCCGAGGTAG
    alHisTyrTh rProGlnGln AsnGlyLeuL ysThrThrLy sGlySerIle
1201 ACTGACCTCC TAGCTCATAC CAATTACACC TTTGAAATCT GGGCTGTGAA
    TGACTGGAGG ATCGAGTATG GTTAATGTGG AAAC TTAGA CCCGACACTT
    401 ThrAspLeuL euAlaHisTh rAsnTyrThr PheGluIleT rpAlaValAs
    TGGAGTGTCC AAATATAACC CTAACCCAGA CCAATCAGTT TCTGTCACTG
    ACCTCACAGG TTTATATTGG GATTGGGTCT GGT TAGTCAA AGACAGTGAC
    nGlyValSer LysTyrAsnP roAsnProAs pGlnSerVal SerValThrVal
1301 TGACCACCAA CCAAGCAGCA CCATCATCCA TTGCTTTGGT CCAGGCTAAA
    ACTGGTGGTT GGTTCGTCGT GGTAGTAGGT AACGAAACCA GGTCCGATTT
    435 ThrThrAs nGlnAlaAla ProSerSerI leAlaLeuVa lGlnAlaLys
    GAAGTCACAA GATACAGTGT GGC ACTGGCT TGGCTGGAAC CAGATCGGCC
    CTT CAGTGT CTATGTCACA CCGTGACCGA ACCGACCTTG GTCTAGCCGG
    GluValThra rgTyrSerVa lAlaLeuAla TrpLeuGluP roAspArgPro

```


FIG. 16D

```

1401 CAATGGGGTA ATCCTGGAAT ATGAAGTCAA GTATTATGAG AAGGATCAGA
    GTTACCCCAT TAGGACCTTA TACTTCAGTT CATAATACTC TTCCTAGTCT
468  AsnGlyVal IleLeuGluT yrGluVally sTyrTyrGlu LysAspGlnA
    ATGAGCGAAG CTATCGTATA GTTCGGACAG CTGCCAGGAA CACAGATATC
    TACTCGCTTC GATAGCATAT CAAGCCTGTC GACGGTCCTT GTGTCTATAG
    snGluArgSe rTyrArgIle ValArgThra laAlaArgAs nThrAspIle
1501 AAAGGCCTGA ACCCTCTCAC TTCCTATGTT TTCCACGTGC GAGCCAGGAC
    TTTCCGGACT TGGGAGAGTG AAGGATACAA AAGGTGCACG CTCGGTCCTG
501  LysGlyLeuA snProLeuTh rSerTyrVal PheHisValA rgAlaArgTh
    AGCAGCTGGC TATGGAGACT TCAGTGAGCC CTTGGAGGTT ACAACCAACA
    TCGTCGACCG ATACCTCTGA AGTCACTCGG GAACCTCCAA TGTTGGTTGT
    rAlaAlaGly TyrGlyAspP heSerGluPr oLeuGluVal ThrThrAsnTh:
1601 CAGTGCCTTC CCGGATCATT GGAGATGGGG CTAACTCCAC AGTCCTTCTG
    GTCACGGAAG GGCCTAGTAA CCTCTACCCG GATTGAGGTG TCAGGAAGAC
535  ValProSe rArgIleIle GlyAspGlyA laAsnSerTh rValLeuLeu
    GTCTCTGTCT CGGGCAGTGT GGTGCTGGTG GTAATTCTCA TTGCAGCTTT
    CAGAGACAGA GCCCGTCACA CCACGACCAC CATTAAAGAGT AACGTCGAAA
    ValSerValS erGlySerVa lValLeuVal ValIleLeuI leAlaAlaPhe
1701 TGTCATCAGC CGGAGACGGA GTAAATACAG TAAAGCCAAA CAAGAAGCGG
    ACAGTAGTCG GCCTCTGCCT CATTATGTC ATTTTCGGTTT GTTCTTCGCC
568  ValIleSer ArgArgArgS erLysTyrSe rLysAlaLys GlnGluAlaA
    ATGAAGAGAA ACATTTGAAT CAAGGTGTAA GAACATATGT GGACCCCTTT
    TACTTCTCTT TGTAACCTTA GTTCCACATT CTTGTATACA CCTGGGGAAA
    spGluGluLy sHisLeuAsn GlnGlyValA rgThrTyrVa lAspProPhe

```

FIG. 16E

```

1801 ACGTACGAAG ATCCCAACCA AGCAGTGCGA GAGTTTGCCA AAGAAATTGA
      TGCATGCTTC TAGGGTTGGT TCGTCACGCT CTCAAACGGT TTCTTTAACT
601  ThrTyrGluA spProAsnGl nAlaValArg GluPheAlaL ysGluIleAs
      CGCATCCTGC ATTAAGATTG AAAAAGTTAT AGGAGTTGGT GAATTTGGTG
      GCGTAGGACG TAATTCTAAC TTTTTC AATA TCCTCAACCA CTTAAACCAC
      pAlaSerCys ileLysIleG luLysValIl eGlyValGly GluPheGlyGlu
1901 AGGTATGCAG TGGGCGTCTC AAAGTGCCTG GCAAGAGAGA GATCTGTGTG
      TCCATACGTC ACCCGCAGAG TTTCACGGAC CGTTCTCTCT CTAGACACAC
635  ValCysSe rGlyArgLeu LysValProG lyLysArgGl uIleCysVal
      GCTATCAAGA CTCTGAAAGC TGGTTATACA GACAAACAGA GGAGAGACTT
      CGATAGTTCT GAGACTTTCG ACCAATATGT CTGTTTGTCT CCTCTCTGAA
      AlaIleLysT hrLeuLysAl aGlyTyrThr AspLysGlnA rgArgAspPhe
2001 CCTGAGTGAG GCCAGCATCA TGGGACAGTT TGACCATCCG AACATCATTC
      GGACTCACTC CGGTCG TAGT ACCCTGTCAA ACTGGTAGGC TTGTAGTAAG
668  LeuSerGlu AlaSerIleM etGlyGlnPh eAspHisPro AsnIleIleH
      ACTTGAAGG CGTGGTCACT AAATGTAAAC CAGTAATGAT CATAACAGAG
      TGAACCTTCC GCACCAGTGA TTTACATTTG GTCATTACTA GTATTGTCTC
      isLeuGluGl yValValThr LysCysLysP roValMetIl eIleThrGlu
2101 TACATGGAGA ATGGCTCCTT GGATGCATTC CTCAGGAAAA ATGATGGCAG
      ATGTACCTCT TACCGAGGAA CCTACGTAAG GAGTCCTTTT TACTACCGTC
701  TyrMetGluA snGlySerLe uAspAlaPhe LeuArgLysA snAspGlyAr
      ATTTACAGTC ATTCAGCTGG TGGGCATGCT TCGTGGCATT GGGTCTGGGA
      TAAATGTCAG TAAGTCGACC ACCCGTACGA AGCACCATA CCCAGACCCT
      gPheThrVal ileGlnLeuV alGlyMetLe uArgGlyIle GlySerGlyMet
2201 TGAAGTATTT ATCTGATATG AGCTATGTGC ATCGTGATCT GGCCGCACGG
      ACTTCATAAA TAGACTATAC TCGATACACG TAGCACTAGA CCGGCGTGCC
735  LysTyrLe uSerAspMet SerTyrValH isArgAspLe uAlaAlaArg
      AACATCCTGG TGAACAGCAA CTTGGTCTGC AAAGTGTCTG ATTTTGGCAT
      TTGTAGGACC ACTTGTCGTT GAACCAGACG TTTCACAGAC TAAAACCGTA
      AsnIleLeuV alAsnSerAs nLeuValCys LysValSerA spPheGlyMet

```

FIG. 16F

2301 GTCCCGAGTG CTTGAGGATG ATCCGGAAGC AGCTTACACC ACCAGGGGTG
 CAGGGCTCAC GAACTCCTAC TAGGCCTTCG TCGAATGTGG TGGTCCCCAC
 768 SerArgVal LeuGluAspA spProGluAl aAlaTyrThr ThrArgGlyG
 GCAAGATTCC TATCCGGTGG ACTGCGCCAG AAGCAATTGC CTATCGTAAA
 CGTTCTAAGG ATAGGCCACC TGACCGGGTC TTCGTAAACG GATAGCATTT
 lyLysIlePr oIleArgTrp ThrAlaProG luAlaIleAl aTyrArgLys
 2401 TTCACATCAG CAAGTGATGT ATGGAGCTAT GGAATCGTTA TGTGGGAAGT
 AAGTGTAAGT GTTCACTACA TACCTCGATA CCTTAGCAAT ACACCCTTCA
 801 PheThrSerA laSerAspVa lTrpSerTyr GlyIleValM etTrpGluVa
 GATGTCGTAC GGGGAGAGGC CCTATTGGGA TATGTCCAAT CAAGATGTGA
 CTACAGCATG CCCCTCTCCG GGATAACCCT ATACAGGTTA GTTCTACACT
 lMetSerTyr GlyGluArgP roTyrTrpAs pMetSerAsn GlnAspValIle
 2501 TTAAAGCCAT TGAGGAAGGC TATCGGTTAC CCCCTCCAAT GGACTGCCCC
 AATTTCCGTA ACTCCTTCCG ATAGCCAATG GGGGAGGTTA CCTGACGGGG
 835 LysAlaIleGluGluGly TyrArgLeuP roProProMe tAspCysPro
 ATTGCGCTCC ACCAGCTGAT GCTAGACTGC TGGCAGAAGG AGAGGAGCGA
 TAACGCGAGG TGGTCGACTA CGATCTGACG ACCGTCTTCC TCTCCTCGCT
 IleAlaLeuH isGlnLeuMe tLeuAspCys TrpGlnLysG luArgSerAsp
 2601 CAGGCCTAAA TTTGGGCAGA TTGTCAACAT GTTGGACAAA CTCATCCGCA
 GTCCGGATTT AAACCCGTCT AACAGTTGTA CAACCTGTTT GAGTAGGCGT
 868 ArgProLys PheGlyGlnI leValAsnMe tLeuAspLys LeuIleArgA
 ACCCAACAG CTTGAAGAGG ACAGGGACGG AGAGCTCCAG ACCTAACACT
 TGGGGTTGTC GAACTTCTCC TGTCCTGCC TCTCGAGGTC TGGATTGTGA
 snProAsnSe rLeuLysArg ThrGlyThrG luSerSerAr gProAsnThr

FIG. 16G

2701 GCCTTGTTGG ATCCAAGCTC CCCTGAATTC TCTGCTGTGG TATCAGTGGG
 CGGAACAACC TAGGTTTCGAG GGGACTTAAG AGACGACACC ATAGTCACCC
 901 AlaLeuLeuA spProSerSe rProGluPhe SerAlaValV alSerValGl
 CGATTGGCTC CAGGCCATTA AAATGGACCG GTATAAGGAT AACTTCACAG
 GCTAACCGAG GTCCGGTAAT TTTACCTGGC CATATTCCCTA TTGAAGTGTC
 yAspTrpLeu GlnAlaIleL ysMetAspAr gTyrLysAsp AsnPheThrAla
 2801 CTGCTGGTTA TACCACACTA GAGGCTGTGG TGCACGTGAA CCAGGAGGAC
 GACGACCAAT ATGGTGTGAT CTCCGACACC ACGTGCACCTT GGTCCCTCCTG
 935 AlaGlyTy rThrThrLeu GluAlaValV alHisValAs nGlnGluAsp
 CTGGCAAGAA TTGGTATCAC AGCCATCACA CACCAGAATA AGATTTTGAG
 GACCGTTCTT AACCATAGTG TCGGTAGTGT GTGGTCTTAT TCTAAAACCTC
 LeuAlaArgI leGlyIleTh rAlaIleThr HisGlnAsnL ysIleLeuSer
 2901 CAGTGTCAG GCAATGCGAA CCCAAATGCA GCAGATGCAC GGCAGAATGG
 GTCACAGGTC CGTTACGCTT GGGTTTACGT CGTCTACGTG CCGTCTTACC
 968 SerValGln AlaMetArgT hrGlnMetGl nGlnMetHis GlyArgMetV
 TTCCCGTCTG AGCCAGTACT GAATAAACTC AAAACTCTTG AAATTAGTTT
 AAGGGCAGAC TCGGTCATGA CTTATTTGAG TTTTGAGAAC TTTAATCAAA
 alProValOp *AlaSerThr GluOc*ThrG lnAsnSerOp *AsnAm*Phe
 3001 ACCTCATCCA TGCACCTTAA TTGAAGAACT GCACTTTTTT TACTTCGTCT
 TGGAGTAGGT ACGTGAAATT AACTTCTTGA CGTGAAAAAA ATGAAGCAGA
 1001 ThrSerSerM etHisPheAs nOp*ArgThr AlaLeuPheL euLeuArgLe
 TCGCCCTCTG AAATTAAAGA AATGAAAAAA AAAAAACAAT ATCTGCAGCG
 AGCGGGAGAC TTTAATTTCT TTTACTTTTTT TTTTTTGTTA TAGACGTCGC
 uArgProLeu LysLeuLysL ysOp*LysLy sLysAsnAsn IleCysSerVal

FIG. 16H

```

3101 TTGCTTGGTG CACAGATTGC TGAAACTGTG GGGCTTACAG AAATGACTGC
AACGAACCAC GTGTCTAACG ACTTTGACAC CCCGAATGTC TTTACTGACG
1035 AlaTrpCv sThrAspCvs Op*AsnCvsG lvAlaTvrAr gAsnAspCvs
CGGTCATTTG AATGAGACCT GGAACAAATC GTTCTCAGA AGTACTTTTC
GCCAGTAAAC TTAGCTCTGGA CCTTGTTTAG CAAAGAGTCT TCATGAAAAG
ArqSerPheG luOp*AspLe uGluGlnIle ValSerGlnL ysTyrPheSer
3201 TGTTTCATCAC CAGTCTGTAA AATACATGTA CCTATAGAAA TAGAACACTG
ACAAGTAGTG GTCAGACATT TTATGTACAT GGATATCTTT ATCTTGTGAC
1068 ValHisHis GlnSerValL ysTyrMetTy rLeuAm*Lys Am*AsnThrA
CCTCTGAGTT TTGATGCTGT ATTTGCTGCC AGACACTGAG CTTCTGAGAC
GGAGACTCAA AACTACGACA TAAACGACGG TCTGTGACTC GAAGACTCTG
laSerGluPh eOp*CysCys IleCysCysG lnThrLeuSe rPheOp*Asp
3301 ATCCCTGATT CTCTCTCCAT TTGGAATTAC AACGGTCGAC GAGCTCGA
TAGGGACTAA GAGAGAGGTA AACCTTAATG TTGCCAGCTG CTCGAGCT
1101 IleProAspS erLeuSerIl eTrpAsnTyr AsnGlyArgA rgAlaArg

```

REFERENCIAS CITADAS EN LA DESCRIPCIÓN

Esta lista de referencias citadas por el solicitante está prevista únicamente para ayudar al lector y no forma parte del documento de patente europea. Aunque se ha puesto el máximo cuidado en su realización, no se pueden excluir errores u omisiones y la OEP declina cualquier responsabilidad al respecto.

Documentos de patente citados en la descripción

- WO 9315201 A [0006]
- US 4745055 A [0052]
- EP 256654 A [0052]
- EP 120694 A [0052]
- EP 125023 A [0052]
- EP 255694 A [0052]
- EP 266663 A [0052]
- WO 8803559 A [0052]
- US 4444878 A [0052]
- WO 8803565 A [0052]
- EP 68763 A [0052]
- EP 0125023 A [0056]
- EP 173494 A [0056]
- US 4816567 A, Cabilly [0063] [0064]
- WO 9109968 A [0064]
- EP 452508 A [0064]
- WO 9116927 A [0064]
- WO 08222616 B [0119]

Documentos no procedentes de patentes citados en la descripción

- Ullrich et al. Cell, 1990, vol. 61, 203 [0002]
- Hirai et al. Science, 1987, vol. 238, [0003]
- Letwin et al. Oncogene, 1988, vol. 3, 621-678 [0003]
- Lhotak et al. Mol. Cell. Biol., 1993, vol. 13, 7071-7079 [0003]
- Lindberg et al. Mol. cell. Biol., 6316-6324 [0003]
- Bohme et al. Oncogene, 1993, vol. 8, 2857-2862 [0003]
- Wicks et al. Proc. Natl. Acad. Sci. USA., 1992, vol. 89, 1611-1615 [0003]
- Pasquale et al. Cell Regulation, 523-534 [0003]
- Sajjadi et al. New Biol., 1991, vol. 3, 769-778 [0003]
- Wicks et al. Proc. Natl. Acad. Sci. USA, 1992, vol. 89, 1611-1615 [0003]
- Lhotak et al. Mol. Cell. Bio., 1991, vol. 11, 2496-2502 [0003]
- Gilardi-Hebenstreit et al. Oncogene, 1992, vol. 7, 2499-2506 [0003]
- Lai et al. Neuron, 1991, vol. 6, 691-704 [0003]
- Sajjadi et al. Oncogene, 1993, vol. 8, 1807-1813 [0003]
- Maisonpierre et al. Oncogene, 3277-3288 [0003]
- Ashman et al. Journal of Cellular Physiology, 1994, vol. 158, 545-554 [0005]
- The Journal of Biological Chemistry, 1992, vol. 267, 26166-26171 [0007]
- Matthews. Cell, 1991, vol. 65, 1143 [0027]
- Wilks. Proc. Natl. Acad. Sci. USA, 1989, vol. 86, 1603 [0027]
- Mullis et al. Cold Spring Harbor Symp. Advan. Biol., 1986, vol. 51, 263 [0027]
- Faulkner et al. Nature, 1982, vol. 398, 286 [0052]
- Morrison. J. Immun., 1979, vol. 123, 793 [0052]
- Köhler et al. Proc. Nat'l. Acad. Sci. USA, 1980, vol. 77, 2197 [0052]
- Raso et al. Cancer Res, 1981, vol. 41, 2073 [0052]
- Morrison et al. Ann. Rev. Immunol., 1984, vol. 2, 239 [0052]
- Morrison. Science, 1985, vol. 229, 1202 [0052]
- Morrison et al. Proc. Nat'l. Acad. Sci. USA, 1984, vol. 81, 6851 [0052]
- Kabat et al. Sequences of Immunological Interest. National Institutes of Health, 1987 [0054]

- Munro. Nature, 13 December 1984, vol. 312 [0056]
- Neuberger et al. Nature, 13 December 1984, vol. 312 [0056]
- Sharon et al. Nature, 24 May 1984, vol. 309 [0056]
- Morrison et al. Proc. Nat'l. Acad. Sci. USA, 1984, vol. 81, 6851-6855 [0056]
- Morrison et al. Science, 1985, vol. 229, 1202-1207 [0056]
- Boulianne et al. Nature, 13 December 1984, vol. 312, 643-646 [0056]
- Capon et al. Nature, 1989, vol. 337, 525-531 [0056]
- Traunecker et al. Nature, 1989, vol. 339, 68-70 [0056]
- Adams et al. Biochemistry, 1980, vol. 19, 2711-2719 [0057]
- Gough et al. Biochemistry, 1980, vol. 19, 2702-2710 [0057]
- Dolby et al. P.N.A.S. USA, 1980, vol. 22, 6027-6031 [0057]
- Rice et al. P.N.A.S. USA, 1982, vol. 79, 7862-7865 [0057]
- Falkner et al. Nature, 1982, vol. 298, 286-288 [0057]
- Morrison et al. Ann. Rev. Immunol., 1984, vol. 2, 239-256 [0057]
- Lyman et al. Cell, 1993, vol. 75, 1157-1167 [0058]
- Bennett et al. J. Biol. Chem., 1991, vol. 266 (34), 23060-23067 [0058]
- Kohler ; Milstein. Eur. J. Immunol., 1976, vol. 6, 511 [0060]
- Hammerling et al. Monoclonal Antibodies and T-Cell Hybridomas. Elsevier, 1981, 563-681 [0060]
- Cote et al. Monoclonal Antibodies and Cancer Therapy. Alan R. Liss, 1985, 77 [0060]
- Boerner et al. J. Immunol., 1991, vol. 147 (1), 86-95 [0060]
- Mage ; Lamoyi. Monoclonal Antibody Production Techniques and Applications. Marcel Dekker, Inc, 1987, 79-97 [0063]
- Remington's Pharmaceutical Sciences. Mack Publishing Company, 1975 [0074]
- Hanks et al. Science, 1988, vol. 241, 42-52 [0081]
- Wilks. Proc. Nat. Acad. Sci.. USA, 1989, vol. 86, 1603-1607 [0081]
- Matthews et al. Cell, 1991, vol. 65, 1143-1152 [0081]
- Kozak. J. Cell Biol., 1991, vol. 115, 887-903 [0083]
- Baer et al. Curr. Opin. Oncol., 1992, vol. 4, 24-32 [0088]
- Bennett et al. J. Biol. Chem., 1991, vol. 266, 23060-23067 [0093] [0107]
- Sigel et al. Methods Enzymol., 1983, vol. 93, 3-12 [0093]
- Gorman et al. DNA Prot. Engineer. Tech., 1990, vol. 2, 3-10 [0096]
- Kearney et al. J. Immunol., 1979, vol. 123, 1548-1550 [0103]
- Goding, J.W. J. Immunol. Methods, 1978, vol. 20, 241-253 [0103]
- Gorman, C. DNA Cloning. IRL Press, 1985, vol. 2, 143-190 [0108]
- Gearing et al. EMBO, 1989, vol. 8, 3667-3676 [0108]