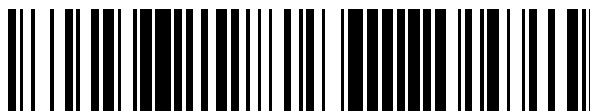


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 704 697**

51 Int. Cl.:

C12Q 1/37 (2006.01)

C12N 9/10 (2006.01)

C12N 9/50 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **05.08.2014 PCT/US2014/049805**

87 Fecha y número de publicación internacional: **12.02.2015 WO15021058**

96 Fecha de presentación y número de la solicitud europea: **05.08.2014 E 14752728 (7)**

97 Fecha y número de publicación de la concesión europea: **03.10.2018 EP 3030671**

54 Título: **Proteínas genomodificadas con un sitio de escisión de proteasa**

30 Prioridad:

05.08.2013 US 201361862363 P
02.05.2014 US 201461987518 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
19.03.2019

73 Titular/es:

GREENLIGHT BIOSCIENCES, INC. (100.0%)
200 Boston Avenue Suite 3100
Medford, MA 02155, US

72 Inventor/es:

BLAKE, WILLIAM JEREMY y
CUNNINGHAM, DREW S.

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 704 697 T3

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

DESCRIPCIÓN

Proteínas genomodificadas con un sitio de escisión de proteasa

5 Antecedentes de la invención

La ingeniería metabólica permite la producción de compuestos mediante manipulación de reacciones bioquímicas (por ejemplo, rutas biosintéticas) en una célula. No obstante, la producción de determinados compuestos puede entrar en conflicto con objetivos celulares esenciales. Por ejemplo, el desvío de nutrientes y energía para la producción de un compuesto puede dar como resultado una escasez de esos sustratos y cofactores para la producción de biomasa. El organismo genomodificado puede evolucionar alejándose de la producción del compuesto de interés o crecer de forma subóptima. Para abordar este problema, se han desarrollado sistemas sin células para la producción *in vitro* de compuestos mediante la expresión coordinada de proteínas en una ruta biosintética. Una salvedad de sistemas de bioproducción tanto *in vivo* como *in vitro* es que muchas proteínas clave que desvían el flujo de una ruta biosintética también son importantes o incluso esenciales para el crecimiento celular.

La supresión o inactivación de estas proteínas es con frecuencia difícil o imposible porque hacerlo da como resultado crecimiento o viabilidad celular reducida. Un modo de inactivar proteínas es mediante inactivación mediada por proteasa. Puede conseguirse inactivación mediada por proteasa de una proteína diana mediante la incorporación de un sitio de reconocimiento de proteasa en la secuencia de aminoácidos primaria de la proteína diana. El sitio de reconocimiento de proteasa puede incorporarse en la secuencia primaria de modo que la proteína resultante sea activa en ausencia de una proteasa que escinde el sitio de reconocimiento e inactiva en presencia de la proteasa. Dichas proteínas diana genomodificadas o recombinantes son particularmente útiles para la síntesis sin células de compuestos de interés.

El documento WO2006090385 describe inhibidores de proteína o péptido recombinantes de serina proteasa NS3 del virus de la hepatitis C (VHC). Baum *et al.* (Proc. Natl. Acad. Sci. USA, Vol. 87, págs. 10023-10027, diciembre de 1990) describe proteínas beta-galactosidasa recombinantes que comprenden "casetes de escisión" insertados con sitios de escisión de proteasa del virus de la inmunodeficiencia humana (VIH).

30 Sumario de la invención

La invención es como se define en las reivindicaciones adjuntas. Se proporcionan en el presente documento enzimas recombinantes que pueden inactivarse de forma selectiva durante la producción sin células *in vitro* de un compuesto. Se consigue inactivación selectiva de enzimas recombinantes introduciendo, entre dos codones (por ejemplo, dos codones adyacentes) en un gen que codifica la enzima recombinante, al menos una secuencia de nucleótidos que codifica una secuencia de reconocimiento de proteasa. Antes de la producción sin células *in vitro* de un compuesto, la proteasa afín (es decir, una proteasa que reconoce y escinde específicamente la secuencia de reconocimiento de proteasa en la proteína recombinante) se introduce, o activa, en el sistema sin células de modo que puede escindir la enzima recombinante, inactivándola de este modo. También se proporcionan en el presente documento métodos de producción y exploración de enzimas recombinantes que pueden inactivarse selectivamente y que conservan un nivel de actividad comparable a la de las enzimas de tipo silvestre.

En algunos aspectos de la invención, se proporcionan en el presente documento proteínas de fosfoglucosa isomerasa (Pgi) recombinantes con al menos una (o una) secuencia de reconocimiento de proteasa localizada entre dos aminoácidos (por ejemplo, dos aminoácidos adyacentes) de la proteína nativa. En otros aspectos, se proporcionan ácidos nucleicos que codifican las proteínas fosfoglucosa isomerasa recombinantes.

En algunas realizaciones, la proteína Pgi comprende la secuencia de la SEQ ID NO: 17. En determinadas realizaciones, la proteína Pgi puede comprender una secuencia que es 90 %, 95 %, 98 % o 99 % homóloga de la SEQ ID NO: 25. En algunas realizaciones, el ácido nucleico que codifica la proteína Pgi comprende la secuencia de la SEQ ID NO: 1. En determinadas realizaciones, el ácido nucleico que codifica la proteína Pgi puede comprender una secuencia que es 90 %, 95 %, 98 % o 99 % homóloga de la SEQ ID NO: 9.

Se desvelan en el presente documento proteínas de fosfotransacetilasa (Pta) recombinantes con al menos una (o una) secuencia de reconocimiento de proteasa localizada entre dos aminoácidos (por ejemplo, dos aminoácidos adyacentes) de la proteína nativa. También se desvelan en el presente documento ácidos nucleicos que codifican las proteínas de fosfotransacetilasa recombinantes.

La proteína Pta puede comprender la secuencia de la SEQ ID NO: 48. En determinadas realizaciones, la proteína Pta puede comprender una secuencia que es 90 %, 95 %, 98 % o 99 % homóloga de la SEQ ID NO: 48. El ácido nucleico que codifica la proteína Pta puede comprender la secuencia de la SEQ ID NO: 47. El ácido nucleico que codifica la proteína Pta puede comprender una secuencia que es 90 %, 95 %, 98 % o 99 % homóloga de la SEQ ID NO: 47.

Se desvelan en el presente documento proteínas de transcetolasa A (TktA) recombinantes con al menos una (o una) secuencia de reconocimiento de proteasa localizada entre dos aminoácidos (por ejemplo, dos aminoácidos adyacentes) de la proteína nativa. También se desvelan en el presente documento ácidos nucleicos que codifican las proteínas de transcetolasa A recombinantes.

La proteína TktA puede comprender la secuencia de la SEQ ID NO: 63. La proteína TktA puede comprender una secuencia que es 90 %, 95 %, 98 % o 99 % homóloga de la SEQ ID NO: 63. El ácido nucleico que codifica la proteína TktA puede comprender la secuencia de la SEQ ID NO: 57. El ácido nucleico que codifica la proteína TktA puede comprender una secuencia que es 90 %, 95 %, 98 % o 99 % homóloga de la SEQ ID NO: 57.

En algunas realizaciones, al menos una secuencia de reconocimiento de proteasa es una secuencia de reconocimiento de proteasa reconocida por una proteasa seleccionada del grupo que consiste en alanina carboxipeptidasa, astacina de *Armillaria mellea*, leucil aminopeptidasa bacteriana, procoagulante de cáncer, catepsina B, clostripaina, alanil aminopeptidasa de citosol, elastasa, endo-proteinasa Arg-C, enteroquinasa, gastricsina, gelatinasa, Gly-X carboxipeptidasa, glicil endopeptidasa, proteasa 3C de rinovirus humano, hipodermina C, serina endopeptidasa específica de IgA, leucil aminopeptidasa, leucil endopeptidasa, lysC, pro-X carboxipeptidasa lisosómica, lisil aminopeptidasa, metionil aminopeptidasa, mixobacteria, nardilisina, endopeptidasa E pancreática, picornaína 2A, picornaína 3C, proendopeptidasa, prolil aminopeptidasa, proproteína convertasa I, proproteína convertasa II, ruselisina, sacaropepsina, semenogelasa, activador de T-plasminógeno, trombina, calicreína tisular, virus del grabado del tabaco (TEV), togavirina, triptofanil aminopeptidasa, activador de U-plasminógeno, V8, venombina A, venombina AB y Xaa-pro aminopeptidasa.

En algunas realizaciones, al menos una secuencia de reconocimiento de proteasa es una secuencia de reconocimiento de proteasa reconocida por proteasa 3C de rinovirus humano. En algunas realizaciones, la secuencia de aminoácidos de la secuencia de reconocimiento de proteasa reconocida por proteasa 3C de rinovirus humano comprende la secuencia de la SEQ ID NO: 38. En algunas realizaciones, la secuencia de ácido nucleico de la secuencia de reconocimiento de proteasa reconocida por proteasa 3C de rinovirus humano comprende la secuencia de la SEQ ID NO: 37.

En algunas realizaciones, al menos una secuencia de reconocimiento de proteasa de la proteína Pgi se localiza después del aminoácido 410, 526, 527, 528, 529, 530, 531 o 532 de la secuencia de la SEQ ID NO: 17. En algunas realizaciones, al menos una secuencia de reconocimiento de proteasa de la proteína Pgi se localiza después del aminoácido 526 de la secuencia de la SEQ ID NO: 17. En algunas realizaciones, al menos una secuencia de reconocimiento de proteasa del ácido nucleico que codifica la proteína Pgi se localiza después de los codones 410, 526, 527, 528, 529, 530, 531 o 532 de la secuencia de la SEQ ID NO: 1. En algunas realizaciones, la secuencia de reconocimiento de proteasa del ácido nucleico que codifica la proteína Pgi se localiza después del codón 526 de la secuencia de la SEQ ID NO: 1.

Al menos una secuencia de reconocimiento de proteasa de la proteína Pta se puede localizar después del aminoácido 381, 382, 387 o 409 de la secuencia de la SEQ ID NO: 48. Al menos una secuencia de reconocimiento de proteasa de la proteína Pta se puede localizar después del aminoácido 381 de la secuencia de la SEQ ID NO: 48. Al menos una secuencia de reconocimiento de proteasa del ácido nucleico que codifica la proteína Pta se puede localizar después de los codones 381, 382, 387 o 409 de la secuencia de la SEQ ID NO: 47. La secuencia de reconocimiento de proteasa del ácido nucleico que codifica la proteína Pta se puede localizar después del codón 381 de la secuencia de la SEQ ID NO: 47.

Al menos una secuencia de reconocimiento de proteasa de la proteína TktA se puede localizar después del aminoácido 635, 636, 637, 638 o 640 de la secuencia de la SEQ ID NO: 63. Al menos una secuencia de reconocimiento de proteasa de la proteína TktA se puede localizar después del aminoácido 637 de la secuencia de la SEQ ID NO: 63. Al menos una secuencia de reconocimiento de proteasa del ácido nucleico que codifica la proteína TktA se puede localizar después de los codones 635, 636, 637, 638 o 640 de la secuencia de la SEQ ID NO: 57. La secuencia de reconocimiento de proteasa del ácido nucleico que codifica la proteína TktA se puede localizar después del codón 637 de la secuencia de la SEQ ID NO: 57.

En algunas realizaciones, al menos una secuencia de reconocimiento de proteasa de la proteína Pgi se puede localizar en una región C terminal de la proteína. Al menos una secuencia de reconocimiento de proteasa de la proteína Pta se puede localizar en una región central o una región C terminal de la proteína. Al menos una secuencia de reconocimiento de proteasa de la proteína TktA se puede localizar en una región central o una región C terminal de la proteína. En algunas realizaciones, al menos una secuencia de reconocimiento de proteasa de la proteína Pgi se puede localizar en una región de bucle expuesta a disolvente de la proteína. Al menos una secuencia de reconocimiento de proteasa de la proteína Pta se puede localizar en una región de bucle expuesta a disolvente de la proteína. Al menos una secuencia de reconocimiento de proteasa de la proteína TktA se puede localizar en una región de bucle expuesta a disolvente de la proteína.

Algunos aspectos de la invención proporcionan vectores que comprenden un ácido nucleico que codifica una proteína Pgi recombinante. Se desvelan en el presente documento vectores que comprenden un ácido nucleico que

codifica una proteína Pta recombinante. También se desvelan en el presente documento vectores que comprenden un ácido nucleico que codifica una proteína TktA recombinante. En algunas realizaciones, un vector puede ser un vector de clonación o un vector de expresión. En algunas realizaciones, un vector puede ser un plásmido, un fósido, un fagémido, un genoma vírico o un cromosoma artificial. En determinadas realizaciones, un vector es un plásmido.

Otros aspectos de la invención proporcionan células que comprenden una cualquiera de las proteínas, ácidos nucleicos o vectores como se exponen en las reivindicaciones adjuntas y se describen en el presente documento. En algunas realizaciones, la célula es una célula bacteriana, una célula fúngica, una célula de mamífero o una célula vegetal. En algunas realizaciones, la célula es una célula bacteriana. En algunas realizaciones, la célula bacteriana es una célula de *Escherichia coli*.

Aspectos adicionales de la invención proporcionan células que expresan proteínas recombinantes con un sitio de escisión de proteasa como se describe en el presente documento. En determinadas realizaciones, la proteína recombinante es una proteína Pgi recombinante. Se contempla donde la proteína recombinante es una proteína Pta recombinante. En otras realizaciones determinadas, la proteína recombinante es una proteína TktA recombinante. También se contempla donde la célula es una célula bacteriana, una célula fúngica, una célula de mamífero o una célula vegetal. En determinadas realizaciones, la célula es una célula bacteriana tal como, por ejemplo, una célula de *Escherichia coli*.

También se proporcionan lisados de cualquiera de las células descritas en el presente documento.

También se proporcionan en el presente documento proteínas recombinantes con al menos una secuencia de reconocimiento de proteasa localizada entre dos aminoácidos (por ejemplo, dos aminoácidos adyacentes) de la proteína nativa en una región de bucle expuesta a disolvente de la proteína, en donde al menos una secuencia de reconocimiento de proteasa es escindida por una proteasa afín con especificidad de secuencia de reconocimiento individual y en donde la actividad de proteína recombinante en presencia de la proteasa afín se reduce en al menos 30 %, al menos 40 %, al menos 50 %, al menos 75 %, al menos 90 %, al menos 99 % o al menos 99,9 % (o es aproximadamente 70 %, aproximadamente 60 %, aproximadamente 50 %, aproximadamente 25 %, aproximadamente 10 %, aproximadamente 1 % o aproximadamente 0,1 %) de la actividad de proteína recombinante en ausencia de la proteasa afín. En algunas realizaciones, la actividad de la proteína recombinante en ausencia de la proteasa afín es suficiente para permitir el crecimiento de tipo silvestre de la célula. En algunas realizaciones, la actividad de la proteína recombinante en presencia de la proteasa afín se reduce en al menos 50 % en comparación con una proteína de control de tipo silvestre y en donde la actividad de la proteína recombinante en ausencia de la proteasa afín es al menos 80 % en comparación con una proteína de control de tipo silvestre. En algunas realizaciones, la proteína recombinante, en ausencia de la proteasa afín, mantiene una tasa de crecimiento celular de al menos 75 % de una tasa de crecimiento celular de tipo silvestre (por ejemplo, tasa de crecimiento de una célula sin la proteína recombinante).

También se proporcionan métodos de genomodificación de proteínas recombinantes (por ejemplo, proteínas Pgi recombinantes y/o proteínas Pta recombinantes y/o proteínas TktA recombinantes) con una secuencia de reconocimiento de proteasa. Los métodos pueden comprender las etapas de: (a) transformar células con una pluralidad de variantes de ácido nucleico, en donde cada variante de ácido nucleico contiene una secuencia de nucleótidos que codifica una proteína recombinante con al menos una secuencia de reconocimiento de proteasa localizada entre dos aminoácidos (por ejemplo, dos aminoácidos adyacentes) de la proteína nativa y un ácido nucleico con una secuencia promotora inducible unida operativamente con una secuencia de nucleótidos que codifica una proteasa afín; (b) cultivar las células en condiciones no inductoras en medio que impide el crecimiento celular de proteínas recombinantes inactivas y recoger células con una tasa de crecimiento comparable a células de control de tipo silvestre; y (c) cultivar las células en condiciones que inducen la expresión de la proteasa afín y recoger células que no crecen o que tienen una tasa de crecimiento reducida en comparación con células de control de tipo silvestre. En algunas realizaciones, las células se transforman con el ácido nucleico con una secuencia promotora inducible unida operativamente con una secuencia de nucleótidos que codifica una proteasa afín después de la etapa (b) y antes de la etapa (c). Los métodos comprenden además secuenciar las variantes de ácido nucleico de las células recogidas en la etapa (c) para identificar la localización de al menos una secuencia de reconocimiento de proteasa.

Se contempla en el presente documento donde la proteína de tipo silvestre (por ejemplo, proteína Pgi y/o proteína Pta y/o proteína TktA) es necesaria para crecimiento celular y el genoma de la célula carece de una copia cromosómica del gen de tipo silvestre que codifica la proteína de tipo silvestre.

Se proporcionan en el presente documento métodos para genomodificar una pluralidad de variantes de ácido nucleico que codifican proteínas recombinantes. Los métodos pueden comprender insertar al menos una secuencia que codifica al menos una secuencia de reconocimiento de proteasa después de cada codón de un ácido nucleico que codifica una proteína diana para producir una pluralidad de variantes de ácido nucleico que codifican proteínas recombinantes, en donde cada proteína recombinante tiene una secuencia de reconocimiento de proteasa en una localización única en su secuencia de aminoácidos primaria. Los métodos pueden comprender además (a)

transformar células con la pluralidad de variantes de ácido nucleico y un ácido nucleico con una secuencia promotora inducible unida operativamente con una secuencia de nucleótidos que codifica una proteasa afín; (b) cultivar las células en condiciones no inductoras en medio que impide el crecimiento celular de proteínas recombinantes inactivas y recoger células con una tasa de crecimiento normal; y (c) cultivar las células en condiciones que inducen la expresión de la proteasa afín y recoger células que no crecen o que tienen una tasa de crecimiento reducida. En algunas realizaciones, las células se transforman con el ácido nucleico con una secuencia promotora inducible unida operativamente con una secuencia de nucleótidos que codifica una proteasa afín después de la etapa (b) y antes de la etapa (c). Los métodos comprenden además secuenciar las variantes de ácido nucleico de las células recogidas en la etapa (c) para identificar la localización de la secuencia de reconocimiento de proteasa.

En otros aspectos de la invención, se proporcionan en el presente documento pluralidades heterogéneas de variantes de ácido nucleico, en donde cada variante de ácido nucleico codifica una proteína recombinante que se modifica para incluir al menos una secuencia de reconocimiento de proteasa localizada entre dos aminoácidos (por ejemplo, dos aminoácidos adyacentes) de la secuencia nativa. En algunas realizaciones, una única secuencia de reconocimiento de proteasa se localiza entre dos aminoácidos de la proteína nativa.

En el presente documento también se proporcionan poblaciones de células heterogéneas, en donde cada célula de la población comprende una variante de ácido nucleico y en donde cada variante de ácido nucleico codifica una proteína recombinante que se modifica para incluir al menos una secuencia de reconocimiento de proteasa localizada entre dos aminoácidos (por ejemplo, dos aminoácidos adyacentes) de la secuencia nativa; y un ácido nucleico que codifica una proteasa afín unida operativamente a un promotor inducible. En algunas realizaciones, una única secuencia de reconocimiento de proteasa se localiza entre dos aminoácidos de la proteína nativa.

Breve descripción de los dibujos

No se pretende que los dibujos adjuntos estén dibujados a escala. Por razones de claridad, no puede marcarse cada componente en cada dibujo.

La Figura 1A muestra un esquema de un método ejemplar para preparar una proteína recombinante con una secuencia de reconocimiento de proteasa insertando la secuencia de reconocimiento de proteasa entre dos aminoácidos de la proteína recombinante.

La Figura 1B muestra un esquema de otro método ejemplar para preparar una proteína recombinante con una secuencia de reconocimiento de proteasa reemplazando aminoácidos nativos de la proteína recombinante con una secuencia de reconocimiento de proteasa.

La Figura 2 muestra un esquema de una modificación de uno de los métodos de la invención, donde una secuencia de reconocimiento de proteasa parcial se inserta entre dos aminoácidos de la proteína de tipo silvestre o codones de modo que la secuencia de reconocimiento completa se reconstituye en el producto final.

La Figura 3 es un diagrama del plásmido pGLA042.

La Figura 4 es un diagrama del plásmido pGLC217.

La Figura 5 muestra un gráfico de actividad proteica de fosfoglucosa isomerasa (Pgi) en lisados celulares en función del tiempo.

La Figura 6 muestra una estructura cristalina de Pgi que indica una localización de inserción de sitio de proteasa.

La Figura 7 es un diagrama del plásmido pGLC089.

La Figura 8 es un diagrama del plásmido pGLC221.

La Figura 9 es una tabla que resume la cepa y datos de tasa de crecimiento (izquierda) y un gráfico de actividad de Pgi con y sin inducción de proteasa (derecha).

La Figura 10A muestra un gráfico de secreción de acetato en variantes de Pta con y sin inducción de proteasa.

La Figura 10B muestra un gráfico de secreción de lactato en variantes de Pta con y sin inducción de proteasa.

La Figura 11 muestra un gráfico de actividad proteica de fosfotransacetilasa (Pta) en lisados celulares con y sin inducción de proteasa.

La Figura 12 muestra una estructura cristalina de un dímero proteico de transcetolasa A (TktA). Los bucles destacados por una línea blanca discontinua (centro superior y centro izquierda) ilustran la localización en la que el motivo de secuencia de reconocimiento de proteasa de rinovirus humano (RVH) se insertó para permitir la desactivación mediada por proteasa de la proteína TktA.

Breve descripción del listado de secuencias

SEQ ID NO: 1 es una secuencia de nucleótidos de un gen *pgi* de tipo silvestre.

SEQ ID NO: 2 es una secuencia de nucleótidos de una variante del gen *pgi* con una secuencia de reconocimiento de 3C de rinovirus humano (RVH) insertada después del codón 108.

SEQ ID NO: 3 es una secuencia de nucleótidos de una variante del gen *pgi* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 109.

SEQ ID NO: 4 es una secuencia de nucleótidos de una variante del gen *pgi* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 110.

- SEQ ID NO: 5 es una secuencia de nucleótidos de una variante del gen *pgi* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 138.
- SEQ ID NO: 6 es una secuencia de nucleótidos de una variante del gen *pgi* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 410.
- 5 SEQ ID NO: 7 es una secuencia de nucleótidos de un variante del gen *pgi* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 524.
- SEQ ID NO: 8 es una secuencia de nucleótidos de una variante del gen *pgi* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 525.
- 10 SEQ ID NO: 9 es una secuencia de nucleótidos de una variante del gen *pgi* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 526.
- SEQ ID NO: 10 es una secuencia de nucleótidos de una variante del gen *pgi* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 527.
- SEQ ID NO: 11 es una secuencia de nucleótidos de una variante del gen *pgi* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 528.
- 15 SEQ ID NO: 12 es una secuencia de nucleótidos de una variante del gen *pgi* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 529.
- SEQ ID NO: 13 es una secuencia de nucleótidos de una variante del gen *pgi* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 530.
- 20 SEQ ID NO: 14 es una secuencia de nucleótidos de una variante del gen *pgi* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 531.
- SEQ ID NO: 15 es una secuencia de nucleótidos de una variante del gen *pgi* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 532.
- SEQ ID NO: 16 es una secuencia de nucleótidos de una variante del gen *pgi* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 545.
- 25 SEQ ID NO: 17 es una secuencia de aminoácidos de una proteína Pgi de tipo silvestre.
- SEQ ID NO: 18 es una secuencia de aminoácidos de una proteína Pgi codificada por SEQ ID NO: 2.
- SEQ ID NO: 19 es una secuencia de aminoácidos de una proteína Pgi codificada por SEQ ID NO: 3.
- SEQ ID NO: 20 es una secuencia de aminoácidos de una proteína Pgi codificada por SEQ ID NO: 4.
- SEQ ID NO: 21 es una secuencia de aminoácidos de una proteína Pgi codificada por SEQ ID NO: 5.
- 30 SEQ ID NO: 22 es una secuencia de aminoácidos de una proteína Pgi codificada por SEQ ID NO: 6.
- SEQ ID NO: 23 es una secuencia de aminoácidos de una proteína Pgi codificada por SEQ ID NO: 7.
- SEQ ID NO: 24 es una secuencia de aminoácidos de una proteína Pgi codificada por SEQ ID NO: 8.
- SEQ ID NO: 25 es una secuencia de aminoácidos de una proteína Pgi codificada por SEQ ID NO: 9.
- SEQ ID NO: 26 es una secuencia de aminoácidos de una proteína Pgi codificada por SEQ ID NO: 10.
- 35 SEQ ID NO: 27 es una secuencia de aminoácidos de una proteína Pgi codificada por SEQ ID NO: 11.
- SEQ ID NO: 28 es una secuencia de aminoácidos de una proteína Pgi codificada por SEQ ID NO: 12.
- SEQ ID NO: 29 es una secuencia de aminoácidos de una proteína Pgi codificada por SEQ ID NO: 13.
- SEQ ID NO: 30 es una secuencia de aminoácidos de una proteína Pgi codificada por SEQ ID NO: 14.
- SEQ ID NO: 31 es una secuencia de aminoácidos de una proteína Pgi codificada por SEQ ID NO: 15.
- 40 SEQ ID NO: 32 es una secuencia de aminoácidos de una proteína Pgi codificada por SEQ ID NO: 16.
- SEQ ID NO: 33 es una secuencia de nucleótidos de una proteasa 3C de RVH con codones optimizados.
- SEQ ID NO: 34 es una secuencia de aminoácidos de una proteasa 3C de RVH con codones optimizados.
- SEQ ID NO: 35 es una secuencia de nucleótidos de una proteasa 3C de RVH con codones optimizados con una secuencia líder de OmpA.
- 45 SEQ ID NO: 36 es una secuencia de aminoácidos de una proteasa 3C de RVH con codones optimizados con una secuencia líder de OmpA.
- SEQ ID NO: 37 es una secuencia de nucleótidos de una secuencia de reconocimiento de proteasa 3C de RVH.
- SEQ ID NO: 38 es una secuencia de aminoácidos de una secuencia de reconocimiento de proteasa 3C de RVH.
- 50 SEQ ID NO: 39 es una secuencia de aminoácidos de una secuencia de reconocimiento de proteasa 3C de RVH parcial.
- SEQ ID NO: 40 es una secuencia de aminoácidos de una secuencia de reconocimiento de proteasa 3C de RVH parcial.
- SEQ ID NO: 41 es una secuencia de aminoácidos de una secuencia de reconocimiento de proteasa 3C de RVH parcial.
- 55 SEQ ID NO: 42 es una secuencia de aminoácidos de una secuencia de reconocimiento de proteasa 3C de RVH parcial.
- SEQ ID NO: 43 es una secuencia de aminoácidos de una secuencia de reconocimiento de proteasa 3C de RVH parcial.
- SEQ ID NO: 44 es una secuencia de aminoácidos de una secuencia de reconocimiento de proteasa 3C de RVH parcial.
- 60 SEQ ID NO: 45 es una secuencia de aminoácidos de una secuencia de reconocimiento de proteasa 3C de RVH parcial.
- SEQ ID NO: 46 es una secuencia de aminoácidos de una secuencia líder de OmpA N-terminal.
- SEQ ID NO: 47 es una secuencia de nucleótidos de un gen *pta* de tipo silvestre.
- 65 SEQ ID NO: 48 es una secuencia de aminoácidos de una proteína Pta de tipo silvestre.

SEQ ID NO: 49 es una secuencia de nucleótidos de un variante del gen *pta* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 381.

SEQ ID NO: 50 es una secuencia de aminoácidos de una proteína Pta codificada por SEQ ID NO: 49.

SEQ ID NO: 51 es una secuencia de nucleótidos de un variante del gen *pta* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 382.

SEQ ID NO: 52 es una secuencia de aminoácidos de una proteína Pta codificada por SEQ ID NO: 51.

SEQ ID NO: 53 es una secuencia de nucleótidos de un variante del gen *pta* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 387.

SEQ ID NO: 54 es una secuencia de aminoácidos de una proteína Pta codificada por SEQ ID NO: 53.

SEQ ID NO: 55 es una secuencia de nucleótidos de un variante del gen *pta* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 409.

SEQ ID NO: 56 es una secuencia de aminoácidos de una proteína Pta codificada por SEQ ID NO: 55.

SEQ ID NO: 57 es una secuencia de nucleótidos de un gen *tktA* de tipo silvestre.

SEQ ID NO: 58 es una secuencia de nucleótidos de un variante del gen *tktA* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 635.

SEQ ID NO: 59 es una secuencia de nucleótidos de un variante del gen *tktA* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 636.

SEQ ID NO: 60 es una secuencia de nucleótidos de un variante del gen *tktA* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 637.

SEQ ID NO: 61 es una secuencia de nucleótidos de un variante del gen *tktA* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 638.

SEQ ID NO: 62 es una secuencia de nucleótidos de un variante del gen *tktA* con una secuencia de reconocimiento de 3C de RVH insertada después del codón 640.

SEQ ID NO: 63 es una secuencia de aminoácidos de una proteína TktA tipo silvestre.

SEQ ID NO: 64 es una secuencia de aminoácidos de una proteína TktA codificada por SEQ ID NO: 58.

SEQ ID NO: 65 es una secuencia de aminoácidos de una proteína TktA codificada por SEQ ID NO: 59.

SEQ ID NO: 66 es una secuencia de aminoácidos de una proteína TktA codificada por SEQ ID NO: 60.

SEQ ID NO: 67 es una secuencia de aminoácidos de una proteína TktA codificada por SEQ ID NO: 61.

SEQ ID NO: 68 es una secuencia de aminoácidos de una proteína TktA codificada por SEQ ID NO: 62.

SEQ ID NO: 69 es una secuencia de nucleótidos de una secuencia de reconocimiento de proteasa de RVH.

SEQ ID NO: 70 es una secuencia de aminoácidos de una proteína Pgi codificada por SEQ ID NO: 71.

SEQ ID NO: 71 es una secuencia de nucleótidos de una proteína Pgi.

SEQ ID NO: 72 es una secuencia de aminoácidos de una proteína Pgi-HRV-I001 codificada por SEQ ID NO: 73.

SEQ ID NO: 73 es una secuencia de nucleótidos de un variante del gen *pgi-hrv-I001* con una secuencia de reconocimiento de proteasa de RVH insertada después del primer codón ilustrado en la Figura 1A.

SEQ ID NO: 74 es una secuencia de aminoácidos de una proteína Pgi-HRV-I002 codificada por SEQ ID NO: 75.

SEQ ID NO: 75 es una secuencia de nucleótidos de un variante del gen *pgi-hrv-I002* con una secuencia de reconocimiento de proteasa de RVH insertada después del segundo codón ilustrado en la Figura 1A.

SEQ ID NO: 76 es una secuencia de aminoácidos de una proteína Pgi-HRV-I003 codificada por SEQ ID NO: 77.

SEQ ID NO: 77 es una secuencia de nucleótidos de un variante del gen *pgi-hrv-I003* con una secuencia de reconocimiento de proteasa de RVH insertada después del tercer codón ilustrado en la Figura 1A.

SEQ ID NO: 78 es una secuencia de aminoácidos de una proteína Pgi-HRV-R001 codificada por SEQ ID NO: 79.

SEQ ID NO: 79 es una secuencia de nucleótidos de un variante del gen *pgi-hrv-R001* con una sustitución de secuencia de reconocimiento de proteasa de RVH después del primer codón ilustrado en la Figura 2A.

SEQ ID NO: 80 es una secuencia de aminoácidos de una proteína Pgi-HRV-R002 codificada por SEQ ID NO: 81.

SEQ ID NO: 81 es una secuencia de nucleótidos de un variante del gen *pgi-hrv-R002* con una sustitución de secuencia de reconocimiento de proteasa de RVH después del segundo codón ilustrado en la Figura 2A.

SEQ ID NO: 82 es una secuencia de aminoácidos de una proteína Pgi-HRV-R003 codificada por SEQ ID NO: 83.

SEQ ID NO: 83 es una secuencia de nucleótidos de un variante del gen *pgi-hrv-R003* con una sustitución de secuencia de reconocimiento de proteasa de RVH después del tercer codón ilustrado en la Figura 2A.

SEQ ID NO: 84 es una secuencia de aminoácidos de una proteína Pgi-HRV-I005 codificada por SEQ ID NO: 85.

SEQ ID NO: 85 es una secuencia de nucleótidos de un variante del gen *pgi-hrv-I005* con una inserción de secuencia de reconocimiento de proteasa de RVH antes de una prolina.

SEQ ID NO: 86 es una secuencia de aminoácidos de una proteína Pgi-HRV-I015 codificada por SEQ ID NO: 87.

SEQ ID NO: 87 es una secuencia de nucleótidos de un variante del gen *pgi-prv-I015* con una inserción de secuencia de reconocimiento de proteasa después de una leucina.

Descripción detallada de determinadas realizaciones de la invención

Muchas proteínas clave en rutas biosintéticas son importantes para el crecimiento celular. La supresión o inactivación de estas proteínas es con frecuencia difícil o imposible porque hacerlo da como resultado crecimiento o viabilidad celular reducida, lo que hace que las células sean insatisfactorias para producir un compuesto de interés. La invención aborda este problema de inhibición del crecimiento celular proporcionando proteínas recombinantes (por ejemplo, enzimas) que están activas durante el crecimiento celular e inactivas durante la producción sin células *in vitro* de un compuesto de interés. Las proteínas recombinantes genomodificadas por los métodos proporcionados en el presente documento tienen una secuencia de reconocimiento de proteasa localizada selectivamente en su

secuencia de aminoácidos primaria de modo que, a pesar de la presencia de la secuencia de reconocimiento, la actividad de la proteína recombinante es suficiente para permitir el crecimiento de tipo silvestre de la célula. Las proteínas recombinantes pueden inactivarse de forma selectiva mediante la introducción, expresión y/o activación de una proteasa afín, que escinde la proteína diana recombinante específicamente en la secuencia de reconocimiento de proteasa, haciendo de este modo a la proteína diana recombinante inactiva (o con actividad reducida). Por lo tanto, las proteínas recombinantes de la invención son útiles para genomodificar y/o alterar rutas biosintéticas para producir un compuesto de interés.

Las proteínas recombinantes y métodos proporcionados en el presente documento son útiles para genomodificar y alterar rutas metabólicas en una célula y un sistema sin células. Por ejemplo, la dirección a proteasa de las enzimas Pgi recombinantes de la invención permite el control del flujo de carbono entre la glucólisis y la ruta de pentosa fosfato en reacciones sin células sin alterar la función de esta enzima clave durante el crecimiento celular. Por lo tanto, la invención proporciona un modo de controlar el flujo metabólico mediante la manipulación de enzimas de rutas claves, incluyendo la manipulación para mantener o alterar la actividad celular de enzimas de rutas claves o enzimas competitivas.

Para producir proteínas recombinantes con una secuencia de reconocimiento de proteasa que tienen un nivel de actividad que permite el crecimiento de tipo silvestre de la célula, puede generarse una biblioteca de variantes de ácido nucleico, en donde cada variante de ácido nucleico codifica una proteína recombinante con al menos una secuencia de reconocimiento de proteasa localizada entre dos aminoácidos (por ejemplo, dos aminoácidos adyacentes) de la secuencia primaria de tipo silvestre de la proteína. A partir de esta biblioteca de variantes, se producen proteínas recombinantes y se exploran para determinar la actividad proteica en ausencia de proteasa e inactividad proteica en presencia de proteasa. Una población de células se transforma en primer lugar con la biblioteca de variantes de ácido nucleico. Para seleccionar proteínas recombinantes expresadas activas, las células se cultivan en medio que impide el crecimiento de proteínas recombinantes inactivas. Se supone que las células que no presentan defectos de crecimiento contienen una proteína recombinante activa, a pesar de la introducción de la secuencia de reconocimiento de proteasa, y se seleccionan para caracterización adicional. Para seleccionar células que contienen una proteína recombinante que puede inactivarse, puede introducirse en las células un ácido nucleico con una secuencia de promotor inducible unida operativamente con una secuencia de nucleótidos que codifica una proteasa afín. El ácido nucleico que codifica la proteasa afín inducible puede introducirse antes o después de la selección de proteína recombinante activa. Las células que contienen supuestamente proteína recombinante activa se cultivan después en condiciones que inducen la expresión de una proteasa afín. Se supone que las células que no crecen, o que crecen escasamente (por ejemplo, presentan defectos de crecimiento), contienen la proteína recombinante inactiva deseada y se seleccionan para caracterización y/o uso adicional. Las variantes de ácido nucleico de las células seleccionadas pueden amplificarse y secuenciarse para identificar la localización de la secuencia de reconocimiento de proteasa.

Proteínas y rutas metabólicas

Una "proteína" o "proteína de tipo silvestre", como se usa en el presente documento, se refiere a una molécula compuesta de aminoácidos unidos entre sí por enlaces peptídicos en una cadena lineal. Un aminoácido "nativo", como se usa en el presente documento, se refiere a un aminoácido en la secuencia de aminoácidos primaria de una proteína de tipo silvestre (es decir, no un aminoácido modificado o mutado). Una "proteína diana", como se usa en el presente documento, se refiere a una proteína de tipo silvestre de interés (es decir, no una proteína recombinante) o una proteína para genomodificar con una secuencia de reconocimiento de proteasa como se describe en el presente documento. Una "proteína recombinante", como se usa en el presente documento, se refiere a una proteína procedente de ácido nucleico recombinante, que se forma artificialmente combinando ácido nucleico de diferentes fuentes. En algunas realizaciones, proteínas recombinantes de la invención difieren entre sí por que la localización de la secuencia de reconocimiento de proteasa individual es única de cada proteína recombinante. Por ejemplo, una proteína recombinante puede tener una secuencia de reconocimiento de proteasa localizada después del primer aminoácido de la secuencia de aminoácidos primaria, otra proteína recombinante puede tener una secuencia de reconocimiento de proteasa localizada después del segundo aminoácido de la secuencia de aminoácidos primaria, pero otra proteína recombinante puede tener una secuencia de reconocimiento de proteasa localizada después del tercer aminoácido de la secuencia de aminoácidos primaria, y así sucesivamente. Por lo tanto, una pluralidad de proteínas recombinantes es habitualmente una pluralidad heterogénea.

Las proteínas recombinantes de la invención pueden usarse para genomodificar rutas metabólicas o una secuencia de reacciones bioquímicas catalizadas por enzimas. Los ejemplos de rutas metabólicas que pueden genomodificarse de acuerdo con la invención incluyen, sin limitación, las implicadas en el metabolismo de carbohidratos, el metabolismo de lípidos, el metabolismo de aminoácidos y el metabolismo de energía. En algunas realizaciones, la ruta metabólica es glucólisis. En algunas realizaciones, la ruta metabólica es metabolismo de exceso de acetato. En algunas realizaciones, la ruta metabólica es la ruta de pentosa fosfato.

Fosfoglucosa isomerasa (Pgi)

En algunas realizaciones, la proteína diana es una enzima fosfoglucosa isomerasa (Pgi), por ejemplo, una enzima Pgi de *Escherichia coli* (*E. coli*). Esta enzima cataliza la interconversión de glucosa-6-fosfato y fructosa-6-fosfato y es la primera etapa comprometida en glucólisis. La inactivación de Pgi inhibe el crecimiento celular; sin embargo, la actividad de Pgi da como resultado el desvío de la glucosa a la ruta de glucólisis, lo que a su vez da como resultado una escasez de glucosa para producción sin células de compuestos de interés que proceden de ribosa. Un ácido nucleico que contiene el gen *pgi* que codifica enzima Pgi puede modificarse por cualquiera de los métodos proporcionados en el presente documento o conocidos en la técnica para generar una variante que comprende una secuencia de reconocimiento de proteasa. En algunas realizaciones, la secuencia de reconocimiento de proteasa usada es una secuencia de reconocimiento de proteasa 3C de rinovirus humano (RVH) (por ejemplo, SEQ ID NO: 37, SEQ ID NO: 38), aunque la invención no se limita a esta. En algunas realizaciones, la secuencia de reconocimiento de 3C de RVH se inserta en fase después de cada codón del gen *pgi*. En algunas realizaciones, la secuencia de reconocimiento de 3C de RVH se inserta después de cada codón del gen *pgi*, excluyendo el primer y/o último codón. En algunas realizaciones, la secuencia de reconocimiento de 3C de RVH se inserta después de los aminoácidos 2-5, 9, 24-25, 33-36, 58-59, 85-96, 105-111, 113-115, 137-141, 143-144, 146, 173-176, 196, 250-251, 254, 366-370, 398-399, 410-414, 447-451, 477, 526-532 o 545. En algunas realizaciones, la secuencia de reconocimiento de 3C de RVH se inserta después de al menos un, o cada, codón de las regiones de bucle expuestas a disolvente de la proteína Pgi.

En algunas realizaciones, las proteínas Pgi recombinantes de la invención contienen una secuencia de reconocimiento de 3C de RVH localizada después del aminoácido 108, 109, 110, 138, 410, 524, 525, 526, 527, 528, 529, 530, 531, 532 o 545. En algunas realizaciones, las variantes de Pgi de la invención comprenden una secuencia de aminoácidos seleccionada entre las SEQ ID NO: 18-32.

En algunas realizaciones, las variantes de ácido nucleico de *pgi* (por ejemplo, genes) de la invención contienen una secuencia de reconocimiento de 3C de RVH localizada después del codón 108, 109, 110, 138, 410, 524, 525, 526, 527, 528, 529, 530, 531, 532 o 545. En algunas realizaciones, las variantes de *pgi* de la invención comprenden una secuencia de nucleótidos seleccionada entre las SEQ ID NO: 2-16.

En algunas realizaciones, la secuencia de reconocimiento de 3C de RVH se inserta entre codones no adyacentes del gen *pgi*. En algunas realizaciones, la secuencia de reconocimiento de 3C de RVH reemplaza codones nativos del gen *pgi*. Por ejemplo, en algunas realizaciones, los ocho codones de la secuencia de reconocimiento de 3C de RVH (es decir, CTG GAA GTG CTG TTT CAG GGT CCG; SEQ ID NO: 37) pueden reemplazar ocho codones contiguos del gen *pgi*.

Fosfotransacetilasa (Pta)

También se contempla una proteína diana de enzima fosfotransacetilasa (Pta), por ejemplo, una enzima Pta de *Escherichia coli* (*E. coli*). Esta enzima cataliza la interconversión reversible de acetil-CoA y acetil fosfato. Un ácido nucleico que contiene el gen *pta* que codifica enzima Pta puede modificarse por cualquiera de los métodos proporcionados en el presente documento o conocidos en la técnica para generar una variante que comprende una secuencia de reconocimiento de proteasa. La secuencia de reconocimiento de proteasa usada puede ser una secuencia de reconocimiento de proteasa 3C de rinovirus humano (RVH) (por ejemplo, SEQ ID NO: 37, SEQ ID NO: 38). La secuencia de reconocimiento de 3C de RVH se puede insertar en fase después de cada codón del gen *pta*. La secuencia de reconocimiento de 3C de RVH se puede insertar después de cada codón del gen *pta*, excluyendo el primer y/o último codón. La secuencia de reconocimiento de 3C de RVH se puede insertar después de al menos un, o cada, codón de las regiones de bucle expuestas a disolvente de la proteína Pta.

Las proteínas Pta recombinantes pueden contener una secuencia de reconocimiento de 3C de RVH localizada después del aminoácido 381, 382, 387 o 409. Las variantes de Pgi pueden comprender una secuencia de aminoácidos seleccionada entre las SEQ ID NO: 50, SEQ ID NO: 52, SEQ ID NO: 54 y SEQ ID NO: 56.

Las variantes de ácido nucleico de *pta* (por ejemplo, genes) pueden contener una secuencia de reconocimiento de 3C de RVH localizada después del codón 381, 382, 387 o 409. Las variantes de *pta* pueden comprender una secuencia de nucleótidos seleccionada entre las SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53 y SEQ ID NO: 55.

La secuencia de reconocimiento de 3C de RVH se puede insertar entre codones no adyacentes del gen *pta*. La secuencia de reconocimiento de 3C de RVH puede reemplazar codones nativos del gen *pta*. Por ejemplo, los ocho codones de la secuencia de reconocimiento de 3C de RVH (es decir, CTG GAA GTG CTG TTT CAG GGT CCG; SEQ ID NO: 37) pueden reemplazar ocho codones contiguos del gen *pta*.

Transcetolasa A (TktA)

También se contempla una proteína diana de enzima transcetolasa A (TktA), por ejemplo, una enzima TktA de *Escherichia coli* (*E. coli*). TktA, junto con transcetolasa B (TktB) catalizan dos reacciones de transferencia de cetol en

ruta de pentosa fosfato. Un ácido nucleico que contiene el gen *tktA* que codifica enzima TktA puede modificarse por cualquiera de los métodos proporcionados en el presente documento o conocidos en la técnica para generar una variante que comprende una secuencia de reconocimiento de proteasa. La secuencia de reconocimiento de proteasa usada puede ser una secuencia de reconocimiento de proteasa 3C de rinovirus humano (RVH) (por ejemplo, SEQ ID NO: 37, SEQ ID NO: 38). La secuencia de reconocimiento de 3C de RVH se puede insertar en fase después de cada codón del gen *tktA*. La secuencia de reconocimiento de 3C de RVH se puede insertar después de cada codón del gen *tktA*, excluyendo el primer y/o último codón. La secuencia de reconocimiento de 3C de RVH se puede insertar después de al menos un, o cada, codón de las regiones de bucle expuestas a disolvente de la proteína TktA.

Las proteínas TktA recombinantes pueden contener una secuencia de reconocimiento de 3C de RVH localizada después del aminoácido 635, 636, 637, 638 o 640. Las variantes de TktA pueden comprender una secuencia de aminoácidos seleccionada entre las SEQ ID NO: 64, SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 67 y SEQ ID NO: 68.

Las variantes de ácido nucleico de *tktA* (por ejemplo, genes) pueden contener una secuencia de reconocimiento de 3C de RVH localizada después del codón 635, 636, 637, 638 o 640. Las variantes de *tktA* pueden comprender una secuencia de nucleótidos seleccionada entre las SEQ ID NO: 58, SEQ ID NO: 59, SEQ ID NO: 60, SEQ ID NO: 61 y SEQ ID NO: 62.

La secuencia de reconocimiento de 3C de RVH se puede insertar entre codones no adyacentes del gen *tktA*. La secuencia de reconocimiento de 3C de RVH puede reemplazar codones nativos del gen *tktA*. Por ejemplo, los ocho codones de la secuencia de reconocimiento de 3C de RVH (es decir, CTG GAA GTG CTG TTT CAG GGT CCG; SEQ ID NO: 37) pueden reemplazar ocho codones contiguos del gen *tktA*.

Proteasas y secuencias de reconocimiento afines

Las proteínas de la invención pueden inactivarse mediante una cualquiera de una diversidad de proteasas que escinden en secuencias de reconocimiento específicas. Como se usa en el presente documento, "secuencia de reconocimiento de proteasa", en el contexto de una proteína, se refiere a una secuencia de aminoácidos que es reconocida y escindida por una proteasa afín. En el contexto de un ácido nucleico que codifica una proteína, una "secuencia de reconocimiento de proteasa" se refiere a una secuencia que codifica la secuencia de aminoácidos reconocida y escindida por una proteasa afín. Como se usa en el presente documento, "proteasa afín" se refiere a una proteasa que escinde y de este modo inactiva una proteína diana recombinante (por ejemplo, enzima). Las proteasas afines que pueden usarse en el presente documento incluyen las que tienen una secuencia de reconocimiento específica, individual, lo que significa que las proteasas escinden en o adyacente a una secuencia específica de uno o más aminoácidos. Por ejemplo, la proteasa 3C de rinovirus humano es altamente específica para la secuencia de reconocimiento Leu-Glu-Val-Leu-Phe-Gln-Gly-Pro (SEQ ID NO: 38). La proteasa reconoce esta secuencia y escinde después del resto de glutamina. La proteasa 3C de rinovirus humano habitualmente no reconoce ni escinde otras secuencias de reconocimiento aunque todas las proteasas son algo promiscuas y pueden reconocer y escindir otros sitios pero a una tasa muy reducida. En algunas realizaciones, las proteínas de la invención se preparan con una secuencia de reconocimiento de proteasa 3C de rinovirus humano genomodificada.

Otros ejemplos de proteasas que pueden usarse de acuerdo con la invención incluyen, sin limitación, alanina carboxipeptidasa, astacina de *Armillaria mellea*, leucil aminopeptidasa bacteriana, procoagulante de cáncer, cathepsina B, clostripaina, alanil aminopeptidasa de citosol, elastasa, endoproteinasa Arg-C, enteroquinasa, gastricsina, gelatinasa, Gly-X carboxipeptidasa, glicil endopeptidasa, proteasa 3C de rinovirus humano, hipodermina C, serina endopeptidasa específica de IgA, leucil aminopeptidasa, leucil endopeptidasa, lysC, pro-X carboxipeptidasa lisosómica, lisil aminopeptidasa, metionil aminopeptidasa, mixobacteria, nardilisina, endopeptidasa E pancreática, picornaína 2A, picornaína 3C, proendopeptidasa, prolil aminopeptidasa, proproteína convertasa I, proproteína convertasa II, ruselisina, sacaropepsina, semenogelasa, activador de T-plasminógeno, trombina, caliceína tisular, virus del grabado del tabaco (TEV), togavirina, triptofanil aminopeptidasa, activador de U-plasminógeno, V8, venombina A, venombina AB y Xaa-pro aminopeptidasa (véase, Rawlings, S. D., *et al.*, Handbook of Proteolytic Enzymes, Academic Press, 2013, Science, Elsevier Ltd.). Otras proteasas pueden usarse de acuerdo con la invención.

Ácidos nucleicos

Se contemplan ácidos nucleicos que codifican las proteínas recombinantes (por ejemplo, proteínas Pgi recombinantes y/o proteínas Pta recombinantes y/o proteínas TktA recombinantes) descritas en el presente documento. Un "ácido nucleico", como se usa en el presente documento, se refiere a al menos dos nucleótidos (por ejemplo, adenina, timina, citosina, guanina, uracilo) unidos covalentemente entre sí. Un ácido nucleico de la invención contendrá en general enlaces fosfodiéster. Un ácido nucleico puede ser monocatenario (mc) o bicatenario (bc), ADN o ARN. En algunas realizaciones, el ácido nucleico está en forma de ADNc. En algunas realizaciones, el ácido nucleico está en forma de ADN genómico. Un "codón", como se usa en el presente documento, se refiere a un conjunto de tres nucleótidos adyacentes que codifican un aminoácido. Los codones de la invención se definen y numeran por el nucleótido inicial a partir del que comienza la traducción.

En algunas realizaciones, se preparan variantes de ácido nucleico bicatenario lineal (por ejemplo, ADN) de acuerdo con la invención. En algunos casos, las variantes de ácido nucleico bicatenarias lineales comprenden una secuencia génica variante que codifica una proteína recombinante con una secuencia de reconocimiento de proteasa así como al menos 30 pares de bases (pb) de nucleótidos de secuencia adicional cadena arriba del codón de inicio y al menos 30 pares de bases de nucleótidos de secuencia adicional cadena abajo del codón de terminación del gen, en donde cada secuencia adicional es homóloga del locus génico de tipo silvestre del genoma de la célula en la que se transformará el ácido nucleico. Como se usa en el presente documento, "gen de tipo silvestre" se refiere al gen de tipo silvestre que codifica la proteína de tipo silvestre que corresponde a la proteína recombinante con al menos un (o un) sitio de reconocimiento de proteasa. Por ejemplo, si la proteína diana es Pgi y la célula que se transforma es *E. coli*, el ácido nucleico contendrá una variante génica que codifica Pgi con al menos una secuencia de reconocimiento de proteasa, al menos 30 pb de secuencia adicional cadena arriba del codón de inicio de la variante génica y homóloga del locus *pgi* del genoma de *E. coli* y al menos 30 pb de secuencia adicional cadena abajo del codón de inicio de la variante génica y homóloga del locus *pgi* del genoma de *E. coli*. La secuencia adicional, en algunos casos, facilita la recombinación de la variante génica con la copia de tipo silvestre cromosómica del gen.

La invención abarca vectores que comprenden una variante de ácido nucleico como se expone en las reivindicaciones. Un "vector", como se usa en el presente documento, puede ser cualquiera de varios ácidos nucleicos en los que puede insertarse una secuencia o secuencias deseadas mediante restricción y ligamiento para transporte entre diferentes ambientes genéticos o para expresión en una célula. Los vectores están compuestos habitualmente de ADN aunque también están disponibles vectores de ARN. Los ejemplos de vectores de acuerdo con la invención incluyen, sin limitación, plásmidos, fósidos, fagémidos, genomas víricos y cromosomas artificiales. En algunas realizaciones, se proporciona una variante de ácido nucleico de la invención en un vector de clonación recombinante. En algunas realizaciones, una variante de ácido nucleico de la invención se expresa en un vector de expresión recombinante.

Un vector de clonación de la invención es capaz de replicar de forma autónoma o integrado en el genoma de una célula. Un vector de clonación tiene una secuencia de restricción por endonucleasa en la que el vector puede cortarse de una manera que puede determinarse y en la que puede ligarse una secuencia de ADN deseada de modo que el nuevo vector recombinante conserva su capacidad para replicar en una célula. En el caso de plásmidos, puede producirse replicación de la secuencia deseada muchas veces a medida que aumenta el número de copias del plásmido en una célula tal como una bacteria o solamente una única vez por célula antes de que la célula se reproduzca por mitosis. En el caso de fagos, puede producirse replicación de forma activa durante una fase lítica o de forma pasiva durante una fase lisógena.

Un vector de expresión de la invención es uno en el que puede insertarse una secuencia codificante de ADN deseada mediante restricción y ligamiento de modo que se una de forma operativa con secuencias reguladoras y puede expresarse como un transcrito de ARN.

Como se usa en el presente documento, se dice que una secuencia codificante y secuencias reguladoras (por ejemplo, secuencias promotoras) están unidas "operativamente" cuando estén unidas de forma covalente de modo que se coloque la expresión o transcripción de la secuencia codificante bajo la influencia o el control de las secuencias reguladoras (por ejemplo, de modo que la secuencia reguladora "conduzca" el inicio de la transcripción y/o la expresión de la secuencia codificante). Si las secuencias codificantes van a traducirse a una proteína funcional, se considera que dos secuencias de ADN están unidas operativamente si la inducción de un promotor en las secuencias 5' reguladoras da como resultado la transcripción de la secuencia de codificación y si la naturaleza del enlace entre las dos secuencias de ADN no (1) da como resultado la introducción de una mutación de desplazamiento de marco, (2) interfiere con la capacidad de la región promotora de dirigir la transcripción de las secuencias codificantes o (3) interfiere con la capacidad del transcrito de ARN correspondiente para traducirse en una proteína. Por lo tanto, una región promotora estaría unida operativamente a una secuencia codificante si la región promotora puede efectuar la transcripción de esa secuencia de ADN de tal manera que el transcrito resultante pueda traducirse en la proteína o el polipéptido deseados.

Los vectores de la invención pueden comprender una secuencia marcadora para su uso en la identificación de células que se han transformado o transfectado o no con el vector. Los marcadores incluyen, por ejemplo, genes que codifican proteínas que aumentan o reducen la resistencia o sensibilidad a antibióticos (por ejemplo, genes de resistencia a ampicilina, genes de resistencia a kanamicina, genes de resistencia a neomicina, genes de resistencia a tetraciclina y genes de resistencia a cloranfenicol) u otros compuestos, genes que codifican enzimas con actividades detectables mediante ensayos convencionales conocidos en la técnica (por ejemplo, β -galactosidasa, luciferasa o fosfatasa alcalina) y genes que afectan visiblemente al fenotipo de células transformadas o transfectadas, hospedadores, colonias o placas (por ejemplo, proteína verde fluorescente). En algunas realizaciones, los vectores usados en el presente documento tienen capacidad de replicación autónoma y expresión de los productos génicos estructurales en los segmentos de ADN con los que están unidos de forma operativa.

Cuando un ácido nucleico que codifica cualquiera de las proteínas de la invención se expresa en una célula, puede usarse una diversidad de secuencias de control de la transcripción para dirigir su expresión. Por ejemplo, un ácido

nucleico de la invención puede contener un promotor, un potenciador y/o un terminador. Como alternativa, el vector en el que se inserta el ácido nucleico puede contener dichas secuencias reguladoras.

Un "promotor", como se usa en el presente documento, se refiere a una región de control de una secuencia de ácido nucleico en la que se controlan el inicio y la tasa de transcripción del resto de una secuencia de ácido nucleico. Un promotor también puede contener subregiones en las que pueden unirse proteínas y moléculas reguladoras, tales como ARN polimerasa y otros factores de transcripción. Los promotores pueden ser constitutivos, inducibles, activables, reprimibles, específicos de tejido o cualquier combinación de los mismos. Un promotor conduce la expresión o conduce la transcripción de la secuencia de ácido nucleico que regula. Un promotor puede ser uno asociado de forma natural con un gen o una secuencia, como puede obtenerse aislando las secuencias 5' no codificantes localizadas cadena arriba del segmento codificante y/o exón de un gen o una secuencia dada. Dicho promotor puede denominarse "endógeno".

En algunas realizaciones, un segmento de ácido nucleico codificante puede situarse bajo el control de un promotor recombinante o heterólogo, que se refiere a un promotor que normalmente no está asociado con la secuencia de ácido nucleico codificada en su ambiente natural. Un potenciador recombinante o heterólogo se refiere a un potenciador no asociado normalmente con una secuencia de ácido nucleico en su ambiente natural. Dichos promotores o potenciadores pueden incluir promotores o potenciadores de otros genes; promotores o potenciadores aislados de cualquier otra célula procariota, vírica o eucariota; y promotores o potenciadores sintéticos que no son "de origen natural" tales como, por ejemplo, los que contienen diferentes elementos de diferentes regiones reguladoras de la transcripción y/o mutaciones que alteran la expresión mediante métodos de ingeniería genética que se conocen en la técnica. Además de producir secuencias de ácido nucleico de promotores y potenciadores de forma sintética, pueden producirse secuencias usando clonación recombinante y/o tecnología de amplificación de ácido nucleico, incluyendo reacción en cadena de la polimerasa (PCR). Asimismo, pueden usarse secuencias de control que dirigen la transcripción y/o expresión de secuencias en orgánulos no nucleares tales como mitocondrias, cloroplastos y similares, de acuerdo con la invención.

Un "promotor inducible", como se usa en el presente documento, es uno que se caracteriza iniciando o potenciando la actividad transcripcional cuando está en presencia de, influido por o en contacto con un inductor o agente inductor.

Un "inductor" o "agente inductor" puede ser endógeno o un compuesto o proteína normalmente exógeno que se administra de tal manera que sea activo en la inducción de la actividad de la transcripción del promotor inducible.

Los promotores inducibles adecuados para su uso de acuerdo con la invención incluyen cualquier promotor inducible descrito en el presente documento o conocido por un experto habitual en la materia. Los ejemplos de promotores inducibles incluyen, sin limitación, promotores química/bioquímicamente regulados y físicamente regulados tales como promotores regulados por isopropil β -D-1-tiogalactopiranosido (IPTG), promotores regulados por alcohol, promotores regulados por tetraciclina (por ejemplo, promotores sensibles a anhidrotetraciclina (aTc) y otros sistemas de promotores sensibles a tetraciclina, que incluyen una proteína represora de tetraciclina (tetR), una secuencia operadora de tetraciclina (tetO) y una proteína de fusión transactivadora de tetraciclina (tTA)), promotores regulados por esteroides (por ejemplo, promotores basados en el receptor de glucocorticoides de rata, receptor de estrógenos humano, receptores de ecdisona de polilla y promotores de la superfamilia del receptor de esteroides/retinoides/tiroides), promotores regulados por metales (por ejemplo, promotores procedentes de genes de metalotioneína (proteínas que se unen y secuestran iones metálicos) de levadura, ratón y ser humano), promotores regulados por patogenicidad (por ejemplo, inducidos por ácido salicílico, etileno o benzotiadiazol (BTH)), promotores inducidos por temperatura/calor (por ejemplo, promotores de choque térmico) y promotores regulados por luz (por ejemplo, promotores sensibles a la luz de células vegetales).

Los promotores inducibles adecuados para su uso de acuerdo con la invención pueden actuar en organismos hospedadores tanto procariotas como eucariotas. En algunas realizaciones, se usan promotores inducibles de mamífero. Los ejemplos de promotores inducibles de mamífero para su uso en el presente documento incluyen, sin limitación, PAct:PAIR de tipo promotor, PART, PBIT, PCR5, PCTA, PETR, PNIC, PPIP, PROP, PSPA/PSCA, PTET, PTtgR, PRep:PCuO de tipo promotor, PETR ON8, PNIC, PPIR ON, PSCA ON8, PTetO, PUREX8, PHyb:tetO7 - ETR8 -PhCMVmin de tipo promotor, tetO7 -PIR3 -ETR8 -PhCMVmin y scbR8 -PIR3 - PhCMVmin. En algunas realizaciones, pueden usarse promotores inducibles de otros organismos, así como promotores sintéticos diseñados para actuar en un hospedador procariota o eucariota. Los ejemplos de promotores inducibles no de mamífero para su uso en el presente documento incluyen, sin limitación, promotores de Lentivirus (por ejemplo, EF α , CMV, Sinapsina I humana (hSynI), CaMKII α , hGFAP y TPH-2) y promotores de virus adenoasociados (por ejemplo, CaMKII α (AAV5), hSynI (AAV2), hThy1 (AAV5), fSST (AAV1), hGFAP (AAV5, AAV8), MBP (AAV8), SST (AAV2)). Una característica funcional importante de los promotores inducibles contemplados es su capacidad de inducción por exposición a un inductor aplicado de forma externa.

Un promotor inducible adecuado para su uso de acuerdo con la invención puede ser inducido por (o reprimido por) una o más condiciones fisiológicas, tales como cambios en el pH, temperatura, radiación, presión osmótica, gradientes salinos, unión a superficie celular y la concentración de uno o más agentes inductores extrínsecos o intrínsecos. El inductor o agente inductor extrínseco puede comprender, sin limitación, aminoácidos y análogos de

aminoácidos, sacáridos y polisacáridos, ácidos nucleicos, activadores y represores de la transcripción de proteínas, citocinas, toxinas, compuestos a base de petróleo, compuestos que contienen metales, sales, iones, análogos de sustratos enzimáticos, hormonas o combinaciones de los mismos. Otros promotores inducibles pueden usarse de acuerdo con la invención.

En algunas realizaciones de la invención, un promotor puede usarse o no junto con un "potenciador", que se refiere a una secuencia reguladora de acción en *cis* implicada en la activación de la transcripción de una secuencia de ácido nucleico cadena abajo del promotor. Un potenciador puede ser uno asociado de forma natural con una secuencia de ácido nucleico, localizado cadena abajo o cadena arriba de esa secuencia. El potenciador puede localizarse en cualquier localización funcional antes o después del promotor y/o el ácido nucleico codificado.

Un "terminador" o una "secuencia terminadora", como se usa en el presente documento, es una secuencia de ácido nucleico que provoca que se detenga la transcripción. Un terminador puede ser unidireccional o bidireccional. Está comprendido por una secuencia de ADN implicada en terminación específica de un transcrito de ARN por una ARN polimerasa. Una secuencia terminadora evita la activación transcripcional de secuencias de ácido nucleico cadena abajo por promotores cadena arriba. Por lo tanto, en determinadas realizaciones, se contempla un terminador que finaliza la producción de un transcrito de ARN.

El tipo más habitualmente usado de terminador es un terminador directo. Cuando se coloca cadena abajo de una secuencia de ácido nucleico que habitualmente se transcribe, un terminador de la transcripción directo provocará que se anule la transcripción. En algunas realizaciones, pueden usarse terminadores de la transcripción bidireccionales, que habitualmente provocan que la transcripción termine en la cadena tanto directa como inversa. En algunas realizaciones, pueden usarse terminadores de la transcripción inversos, que habitualmente terminan la transcripción solamente en la cadena inversa.

En sistemas procariotas, los terminadores habitualmente quedan en dos categorías (1) terminadores rho-independientes y (2) terminadores rho-dependientes. Los terminadores rho-dependientes generalmente están compuestos de una secuencia palindrómica que forma un tallo-bucle rico en pares de bases G-C seguido de varias bases T. Los promotores adecuados para su uso de acuerdo con la invención incluyen cualquier terminador de la transcripción descrito en el presente documento o conocido por un experto habitual en la materia. Los ejemplos de terminadores incluyen, sin limitación, las secuencias de terminación de genes tales como, por ejemplo, el terminador de hormona del crecimiento bovina y secuencias de terminación víricas tales como, por ejemplo, el terminador de SV40, *spy*, *yejM*, *secG-leuU*, *thrLABC*, *rrnB T1*, *hisLGDCBHAFI*, *metZWV*, *rmC*, *xapR*, *aspA*, y terminador de *arcA*. En algunas realizaciones, la señal de terminación puede ser una secuencia que no puede transcribirse o traducirse, tal como las resultantes de un truncamiento de secuencia. Otros terminadores pueden usarse de acuerdo con la invención.

En algunas realizaciones, se optimizan los codones de los ácidos nucleicos para expresión mejorada de las proteínas recombinantes de la invención. La optimización de codones, también denominada uso codónico preferente, se refiere a diferencias en la frecuencia de aparición de codones sinónimos en ADN codificante.

Células

Se contempla cualquier tipo de célula, incluyendo una célula procariota y una eucariota, que expresa de forma recombinante las proteínas proporcionadas en el presente documento. En algunas realizaciones, la célula es una célula bacteriana. En algunas realizaciones, la célula bacteriana es una célula de una bacteria del género *Escherichia*. En algunas realizaciones, la célula bacteriana es una célula de *Escherichia coli* (*E. coli*). En algunas realizaciones, la célula es una célula fúngica, tal como, por ejemplo, una célula de levadura (por ejemplo, una célula de *Saccharomyces cerevisiae*). En algunas realizaciones, la célula es una célula de mamífero o una célula vegetal. Debería apreciarse que algunas células para su uso de acuerdo con la invención no contienen la copia cromosómica de tipo silvestre de un gen que codifica la proteína de tipo silvestre (por ejemplo, la proteína de tipo silvestre correspondiente a la proteína recombinante con una secuencia de reconocimiento de proteasa).

Las células proporcionadas en el presente documento, en algunas realizaciones, son células procariotas que pueden transformarse con cualquiera de los ácidos nucleicos de la invención. La transformación y la transfección son procesos por los que se introduce material genético exógeno en una célula procariota y en una célula eucariota, respectivamente. Puede conseguirse transformación mediante electroporación o por medios químicos. Las células para transformar están típicamente en un estado de competencia. Por lo tanto, en algunas realizaciones, las células proporcionadas en el presente documento son células electrocompetentes o químicamente competentes. Se conocen en la técnica una diversidad de células electrocompetentes y químicamente competentes y pueden usarse de acuerdo con la invención.

En algunas realizaciones, las células son células de *Escherichia coli* (*E. coli*) tales como, por ejemplo, células de *E. coli* JW3985-1 (Coli Genetic Shock Center; CHSC n.º 10867; Mol. Sys. Biol. 2:2006-08, 2006). Pueden usarse otras líneas celulares disponibles en el mercado y no disponibles en el mercado de acuerdo con la invención.

Las células de la invención pueden comprender marcadores seleccionables. Los marcadores seleccionables incluyen, sin limitación, genes que codifican proteínas que aumentan o reducen la resistencia o sensibilidad a antibióticos (por ejemplo, genes de resistencia a ampicilina, genes de resistencia a kanamicina, genes de resistencia a neomicina, genes de resistencia a tetraciclina y genes de resistencia a cloranfenicol) u otros compuestos, genes que codifican enzimas con actividades detectables mediante ensayos convencionales conocidos en la técnica (por ejemplo, β -galactosidasa, luciferasa o fosfatasa alcalina) y genes que afectan visiblemente al fenotipo de células transformadas o transfectadas, hospedadores, colonias o placas (por ejemplo, proteína verde fluorescente). Otros marcadores seleccionables pueden usarse de acuerdo con la invención.

10 Construcción de bibliotecas

Pueden usarse métodos desvelados en el presente documento para construir una biblioteca de las variantes de ácido nucleico proporcionadas en el presente documento. El diseño de biblioteca puede utilizar dos secuencias de nucleótidos - una que codifica la secuencia de aminoácidos primaria de la proteína diana y una que codifica la secuencia de reconocimiento de proteasa de la proteasa que se usará para inactivación de la proteína recombinante de la invención. La secuencia de reconocimiento de proteasa puede "avanzarse" a lo largo de la secuencia anterior en uno de dos métodos (FIG. 1A y 1B).

En un método, la secuencia de reconocimiento de proteasa puede insertarse después de múltiples codones de la secuencia de nucleótidos que codifica la proteína diana, produciendo de este modo una pluralidad de variantes de ácido nucleico, en donde cada variante de ácido nucleico contiene la secuencia de reconocimiento de proteasa localizada en una posición única entre dos codones nativos (FIG. 1A). En otro método, la secuencia de reconocimiento de proteasa puede reemplazar un número equivalente de nucleótidos en la secuencia que codifica la proteína diana, produciendo de este modo una pluralidad de variantes de ácido nucleico, en donde cada variante de ácido nucleico contiene la secuencia de reconocimiento de proteasa en lugar de un tramo equivalente de nucleótidos nativos (FIG. 1B).

En algunas realizaciones, la secuencia de reconocimiento de proteasa puede insertarse después de cada codón de la secuencia de ácido nucleico que codifica la proteína diana, produciendo de este modo una pluralidad de variantes de ácido nucleico, en donde cada variante de ácido nucleico contiene la secuencia de reconocimiento de proteasa localizada en una posición única entre dos codones nativos (por ejemplo, dos codones nativos adyacentes). En algunas realizaciones, la secuencia de reconocimiento de proteasa puede insertarse después de cada codón de la secuencia de ácido nucleico excluyendo el primer y/o último codón. Como alternativa, en algunas realizaciones, la secuencia de reconocimiento de proteasa puede insertarse después de cada segundo codón, después de cada tercer codón, después de cada cuarto codón, después de cada quinto codón, después de cada décimo codón o después de cada vigésimo codón. En algunas realizaciones, la secuencia de reconocimiento de proteasa puede insertarse aleatoriamente. En algunas realizaciones, la secuencia de reconocimiento de proteasa puede insertarse en una región particular del ácido nucleico, tal como, por ejemplo, la región N terminal o la región C terminal. En algunas realizaciones, la secuencia de reconocimiento de proteasa puede reemplazar codones contiguos de la secuencia de ácido nucleico que codifica la proteína diana. La "región N terminal" de una proteína, como se usa en el presente documento, puede hacer referencia al tramo de 100, 90, 80, 70, 60, 50, 40, 30, 20 o 10 aminoácidos localizados adyacentes al aminoácido 5' terminal. La "región C terminal" de una proteína, como se usa en el presente documento, puede hacer referencia al tramo de 100, 90, 80, 70, 60, 50, 40, 30, 20 o 10 aminoácidos localizados adyacentes al aminoácido 3' terminal. En cada realización, debe entenderse que cada variante de ácido nucleico contiene al menos una (o una) secuencia de reconocimiento de proteasa.

En algunas realizaciones, en donde la estructura de una proteína diana se conoce o puede predecirse, la secuencia de reconocimiento de proteasa puede insertarse en regiones correspondientes a regiones de bucle expuestas a disolvente de la proteína. Se ha descubierto que, en algunos casos, estas regiones de bucle expuestas a disolvente son tolerantes a la inserción de una secuencia de reconocimiento de proteasa que es fácilmente escindible. Por lo tanto, se proporcionan en el presente documento métodos para construir una biblioteca de ácidos nucleicos que contienen una secuencia de reconocimiento de proteasa en regiones correspondientes a regiones de bucle expuestas a disolvente de la proteína. Dichos métodos ahorran tiempo y el coste de preparar proteínas con una secuencia de reconocimiento de proteasa. En algunas realizaciones, puede ser ventajoso usar estrategias de construcción de bibliotecas basadas en la reacción en cadena de la polimerasa (PCR) que impiden la síntesis de las secuencias de ácido nucleico de tipo silvestre.

Debe entenderse que, en algunas realizaciones, puede ser ventajoso incorporar tanta de la secuencia de reconocimiento de proteasa como sea necesario para producir una secuencia de reconocimiento de longitud completa. Por ejemplo, si la secuencia de reconocimiento de proteasa comienza con una leucina y la secuencia se inserta después de una leucina, solamente puede insertarse una secuencia de reconocimiento parcial de modo que no se repita la leucina (FIG. 2). Asimismo, si el último aminoácido en la secuencia de reconocimiento de proteasa es una prolina y la secuencia de reconocimiento de proteasa se inserta antes de una prolina, solamente puede insertarse una secuencia de reconocimiento de proteasa parcial de modo que no se duplique la prolina. Por lo tanto, una proteína o ácido nucleico con una secuencia de reconocimiento de proteasa localizada entre dos aminoácidos o codones nativos (por ejemplo, dos aminoácidos o codones nativos adyacentes), respectivamente, abarca proteínas y

ácidos nucleicos con secuencias de reconocimiento de proteasa parciales insertadas entre dos aminoácidos o codones nativos de modo que la secuencia de reconocimiento completa se reconstituya en el producto final.

Construcción de cepas

Las variantes de ácido nucleico de la invención pueden transformarse en células recombinantes (por ejemplo, células bacterianas) para explorar con respecto a proteínas recombinantes óptimas (por ejemplo, activas e inactivables). Debe entenderse que las células usadas para exploración no son necesariamente las células usadas para expresar una proteína recombinante óptima para el fin de genomodificar, por ejemplo, una ruta metabólica de interés.

En algunas realizaciones, el genoma de las células puede modificarse para (a) suprimir o mutar la copia cromosómica de tipo silvestre (o endógena) de un gen que codifica la proteína diana y/o (b) incluir un medio para inducir expresión de proteasa afín citoplasmática. Esto último puede conseguirse añadiendo un gen que codifica la proteasa afín con un promotor inducible al genoma celular, o proporcionando un gen que codifica la proteasa afín con un promotor inducible en un vector, tal como, por ejemplo, un plásmido. Como alternativa, en algunas realizaciones, las células pueden carecer completamente de la proteasa afín, que puede después añadirse en una etapa de exploración/selección posterior. En algunas realizaciones, la proteasa afín se añade en forma purificada.

En algunas realizaciones, la célula recombinante se modifica para carecer de una copia cromosómica funcional del gen de tipo silvestre (es decir, el gen de tipo silvestre que codifica la proteína de tipo silvestre que corresponde a la proteína recombinante con el sitio de reconocimiento de proteasa) y se transforma con un plásmido que contiene variantes de ácido nucleico de la invención. Sin quedar ligado a teoría alguna, la supresión de la copia de tipo silvestre cromosómica del gen de la célula permite la complementación y ayuda a minimizar el fondo cuando el ácido nucleico se inserta mediante métodos de recombinación de baja eficacia (por ejemplo, cuando el crecimiento celular debido a la presencia de un gen de tipo silvestre compensatorio representa un falso positivo). En algunas realizaciones, la inclusión de un marcador seleccionable (por ejemplo, un marcador de resistencia a antibiótico) en las células, por ejemplo, en vector episómico que contiene una variante de ácido nucleico, puede reducir la tasa de falsos positivos.

En algunas realizaciones, las células pueden modificarse para suprimir genes de tipo silvestre que codifican proteínas con funciones similares a las de la proteína diana. Por ejemplo, en algunas realizaciones, se suprimen copias cromosómicas de genes que codifican isozimas de una enzima diana (es decir, enzimas que proporcionan función similar) de las células para minimizar el fondo en la etapa o las etapas de exploración/selección.

En algunas realizaciones, las células recombinantes se modifican para contener un sistema de recombinasa inducible tal como, por ejemplo, al menos un ácido nucleico que contiene los genes del sistema de recombinasa de fago lambda (λ) *gamma* (γ), *beta* (β) y *exo*. Por lo tanto, en algunas realizaciones, se usan métodos de genomodificación por recombinación (o ingeniería genética mediada por recombinación) para modificar las células recombinantes de la invención. Dichos sistemas de recombinación homólogos pueden usarse para introducir o suprimir copias cromosómicas de genes de tipo silvestre del genoma celular. También se contemplan y pueden usarse en el presente documento otros métodos de genomodificación por recombinación. La invención también contempla el uso de enzimas de restricción y ligasas para combinar secuencias de ácido nucleico en un orden especificado (Strachan, T., *et al.*, Human Molecular Genetics, Capítulo 4, Garland Science, Nueva York, 1999).

Selección de actividad proteica

Las células recombinantes que expresan las variantes de ácido nucleico de la invención pueden cultivarse en medio selectivo en ausencia de una proteasa funcional para permitir la recuperación de proteínas recombinantes codificadas por las variantes de ácido nucleico. Por ejemplo, en algunas realizaciones, la actividad de la proteína diana puede ser necesaria para el crecimiento celular. Si la inserción de una secuencia de reconocimiento de proteasa afín afecta de forma adversa a la actividad de la proteína recombinante, después, supuestamente, las células presentarán defectos de crecimiento tales como, por ejemplo, una tasa de crecimiento reducida. En consecuencia, en esta etapa de exploración/selección, solamente se seleccionan las células con una tasa de crecimiento normal (o sin defectos de crecimiento) para caracterización adicional. Una "tasa de crecimiento normal", como se usa en el presente documento, se refiere a una tasa de crecimiento que es comparable a células de tipo silvestre de control. En algunas realizaciones, se considera que una célula tiene una "tasa de crecimiento normal" si su tasa de crecimiento está dentro de aproximadamente 15 % de la tasa de crecimiento de una célula de control de tipo silvestre (por ejemplo, célula con una variante de ácido nucleico/proteína recombinante de la invención). Por ejemplo, puede considerarse que una célula tiene una tasa de crecimiento normal si su tasa de crecimiento está dentro del 50 %, 40 %, 30 %, 25 %, 20 %, 15 %, 14 %, 13 %, 12 %, 11 %, 10 %, 9 %, 8 %, 7 %, 6 %, 5 %, 4 %, 3 %, 2 % o 1 % de la tasa de crecimiento de una célula de control de tipo silvestre. Una célula con un "defecto de crecimiento", como se usa en el presente documento, puede referirse a una célula que no consigue crecer o que tiene una tasa de crecimiento reducida mayor del 10 %, mayor del 15 %, mayor del 20 % o mayor del 25 % en comparación con la tasa de crecimiento de una célula de control de tipo silvestre.

El medio de crecimiento selectivo que puede usarse de acuerdo con la invención, en algunas realizaciones, puede depender de características particulares de las proteínas recombinantes, tales como la función específica de la proteína recombinante activa. Por ejemplo, cuando se cultiva proteína Pgi recombinante en medio mínimo en el que la única fuente de carbono disponible es glucosa, las células que contienen una variante de Pgi recombinante activa crecen bien, mientras que las células que contienen una variante de Pgi recombinante inactiva crecen mal. En algunas realizaciones, el medio selectivo usado puede depender del sustrato de la variante de Pgi recombinante. En algunas realizaciones, se usa un enfoque de "rescate" para producir condiciones de crecimiento selectivas, en las que la actividad de la proteína recombinante, que es necesaria para el crecimiento celular, se suprime del genoma de la célula (por ejemplo, se suprimen o mutan gen o genes que codifican la proteína o las proteínas de tipo silvestre), y después se introducen variantes de *pgi* de ácido nucleico en la célula. Las variantes de ácido nucleico que están activas deberían rescatar el crecimiento celular (por ejemplo, las células crecen) y las que están inactivas no deberían rescatar el crecimiento celular (por ejemplo, las células no crecen).

Selección de inactivación proteica

Las células que no presentan defectos de crecimiento se cultivan después en condiciones selectivas que inducen expresión de proteasa afín. Esta etapa permite la recuperación de células que presentan defectos de crecimiento.

Las células que presentan defectos de crecimiento portan supuestamente proteínas recombinantes que se inactivan en presencia de una proteasa afín funcional (y son activas en ausencia de una proteasa afín funcional). Estas células defectuosas en crecimiento se recuperan después y las variantes de ácido nucleico contenidas en las mismas se secuencian para caracterización adicional.

La caracterización adicional puede implicar la expresión de variantes de ácido nucleico seleccionadas en células que carecen de proteasa afín funcional. El crecimiento de estas células pueden después caracterizarse y los lisados producirse y recogerse. Los lisados pueden después ensayarse *in vitro* para pérdida de actividad proteica recombinante. Dichos ensayos pueden utilizar ensayos de actividad proteica tras la incubación con o sin proteasa afín purificada, exógena. Se conocen en la técnica diversos ensayos de actividad proteica, cualquiera de los cuales pueden usarse de acuerdo con la invención. El ensayo de actividad proteica seleccionado dependerá del tipo de proteína. En algunas realizaciones, la proteína recombinante que se inactiva más completa y rápidamente cuando se expone a proteasa puede seleccionarse para su uso adicional en, por ejemplo, genomodificación de rutas metabólicas de interés.

Las condiciones de inducción de proteasa afín dependerán del tipo de sistema promotor inducible elegido para conducir la expresión de la proteasa afín y se conocen en la técnica. Por ejemplo, isopropil β -D-1-tiogalactopiranosido (IPTG) puede añadirse a un sistema sin células *in vitro* para activar un promotor sensible a IPTG unido operativamente a la proteasa afín.

Se ilustran estos y otros aspectos por los siguientes ejemplos no limitantes.

Ejemplos

Ejemplo 1 - Enzimas fosfoglucosa isomerasa de *Escherichia coli*

La enzima fosfoglucosa isomerasa (Pgi) de *Escherichia coli* (*E. coli*) cataliza la interconversión de glucosa-6-fosfato y fructosa-6-fosfato y es la primera etapa comprometida en glucólisis. La dirección a proteasa de esta enzima, sin alterar la función/actividad de esta enzima clave durante el crecimiento celular, permite controlar el flujo de carbono entre glucólisis y la ruta de pentosa fosfato en reacciones sin células.

Construcción de bibliotecas de variantes de *Pgi*

Se diseñó una biblioteca de ADN bicatenario lineal de 562 miembros y se construyó mediante reacción en cadena de la polimerasa (PCR), donde la secuencia del gen *pgi* de *E. coli* nativa (SEQ ID NO: 1) se modificó para incluir una secuencia de nucleótidos (SEQ ID NO: 37) que codifica la secuencia de reconocimiento de proteasa de ocho aminoácidos (SEQ ID NO: 38) de la proteasa 3C de rinovirus humano (RVH) (FIG. 1A y 1B). 547 miembros de la biblioteca contenían genes *pgi* mutantes con nucleótidos que codificaban la secuencia de reconocimiento de proteasa insertada después de cada uno de los 549 codones en el gen *pgi* de tipo silvestre (excluyendo los primer y último codones). Se crearon miembros adicionales de la biblioteca reemplazando la secuencia del gen *gpi* de tipo silvestre con una secuencia de nucleótidos que codifica los ocho aminoácidos de la secuencia de reconocimiento de proteasa. Estos miembros contenían reemplazos en el gen de tipo silvestre en 15 localizaciones diferentes que comienzan con los números de codones 244, 245, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 461, 462 (donde los números de codones corresponden al primer codón de la secuencia de reemplazo). Además de la secuencia codificante del gen, cada miembro de la biblioteca también contenía ramas de homología de 50 pb (por ejemplo, secuencia adicional de 50 pb cadena arriba del codón de inicio del gen y 50 pb cadena abajo del codón de terminación del gen que son homólogos del locus *pgi* de tipo silvestre del genoma de *E. coli*). Si la secuencia LEVLFQGP (SEQ ID NO: 38) fuera a insertarse después de los aminoácidos L, LE o LEV, solamente EVLFQGP

(SEQ ID NO: 39), VLFQGP (SEQ ID NO: 40) o LFQGP (SEQ ID NO: 41) se insertaron, respectivamente. De manera similar, si la secuencia fuera a insertarse antes de los aminoácidos P, GP o QGP, solamente LEVLFFQG (SEQ ID NO: 42), LEVLFFQ (SEQ ID NO: 43) o LEVLFF (SEQ ID NO: 44) se insertaron, respectivamente. Además, si la inserción (o el reemplazo) fue entre los aminoácidos LP, por ejemplo, solamente EVLFFQG (SEQ ID NO: 45) se insertó (o reemplazó).

Diseño de cepas

JW3985-1 de *E. coli* (Coli Genetic Stock Center; CGSC n.º 10867) de la colección Keio (Mol. Syst. Biol. 2006;2:2006-08) se seleccionó como la cepa para la exploración de biblioteca de Pgi. Esta cepa contiene un marcador de resistencia a kanamicina (KanR) en lugar del gen *pgi*. Para preparar la cepa para su uso en la exploración, se realizaron varias modificaciones. En primer lugar, se retiró KanR usando pCP20, que se obtuvo de *E. coli* BT340 (CGSC n.º 7629), empleando el método descrito por Datsenko y Wanner (Proc Natl Acad Sci USA. 6 Jun 2000;97(12):6640-45). El locus de *pgi* resultante de la cepa contenía las primeras tres bases y las últimas 21 bases del gen *pgi*, con una secuencia cicatriz corta entre medias. Dos plásmidos (pGLA042 y pGLC217; FIG. 3 y 4) se cotransformaron en esta cepa para crear la cepa final que se usó en la exploración de Pgi (GL12-085) mediante el enfoque de selección de alto rendimiento agrupado (descrito posteriormente). pGLA042 se creó a partir de pKD46 (obtenido de *E. coli* BW25113; GCSC n.º 7739) reemplazando el sistema de expresión inducible por arabinosa de pKD46 con el sistema de expresión de λ cl857-pR inducible por temperatura. Este cambio permite la expresión inducible por temperatura de los genes del sistema de recombinasa de fago λ γ , β , y *exo* de pGLA042. pGLC217 es un plásmido de bajo número de copias que proporciona expresión inducible por arabinosa de la proteasa 3C de RVH (con codones optimizados para expresión en *E. coli*; SEQ ID NO: 34), cuya traducción es facilitada por un sitio de unión a ribosoma fuerte. También se produjo una cepa que carece de pGLC217 (GL12-052) para su uso en enfoque de selección y ensayo individual (véase posteriormente).

Enfoque de selección y ensayo individual

El locus cromosómico de *pgi* en GL12-052 se recombinó con un subconjunto de 76 miembros de la biblioteca de ADN bicatenario, lineal, descrita anteriormente. Este subconjunto contenía la secuencia de reconocimiento de proteasa en regiones de bucle accesibles al disolvente de Pgi, como se predijo por su estructura cristalina (banco de datos de proteínas ID: 3NBU). Los miembros de bibliotecas de Pgi resultantes portaron inserciones de secuencias de reconocimiento de proteasa después de las siguientes posiciones en la secuencia de aminoácidos primaria de Pgi de tipo silvestre: 2-5, 9, 24-25, 33-36, 58-59, 85-96, 105-111, 113-115, 137-141, 143-144, 146, 173-176, 196, 250-251, 254, 366-370, 398-399, 410-414, 447-451, 477, 526-532.

GL12-052 se cultivó a 30 °C en LB de baja salinidad (caldo de lisogenia) (NaCl 0,5X) hasta una densidad óptica (DO) de 0,5. El cultivo se transfirió a un baño de agua a 42 °C y se agitó durante 15 minutos para inducir el sistema de recombinasa de pGLA042. Las células inducidas se hicieron electrocompetentes siguiendo métodos convencionales y se transformaron con la biblioteca de variantes de ácido nucleico. Cada miembro de biblioteca (o variante de ácido nucleico) se transformó individualmente (25 μ l de células y 250 ng de miembro de biblioteca), o en subconjuntos de tres miembros, y se recuperó en 1 ml de LB de baja salinidad durante más de 1 h a 30 °C. Se sembraron placas de transformaciones recuperadas en medio de agar M9 complementado con glucosa 1 % (M9G). Las placas se incubaron a 30 °C durante 1,5-2 días. Las colonias resultantes representaron miembros de bibliotecas que contenían moléculas de Pgi activas a pesar de la inclusión de la secuencia de reconocimiento de proteasa. La región de Pgi de los genomas de estos miembros de las bibliotecas se amplificó por PCR y se secuenció. Después se verificaron las secuencias de las cepas en cultivos en matraces de agitación a pequeña escala con medio M9G a 37 °C para determinar sus tasas de crecimiento, proporcionando de este modo un método *in vivo* para evaluar la influencia de la inserción de secuencia de reconocimiento de proteasa en la actividad de Pgi. Las cepas con tasas de crecimiento en un intervalo de 15 % de la tasa de crecimiento de tipo silvestre se avanzaron para un segundo ciclo de exploración para determinar la susceptibilidad a inactivación de proteasa; se avanzaron 41 del subconjunto original de 76 miembros.

Para evaluar la susceptibilidad a proteasa, se crearon y ensayaron lisados clarificados *in vitro* para determinar la actividad de Pgi en presencia o ausencia de proteasa 3C de RVH exógena. Se dejaron crecer cultivos de M9G hasta una DO de 2, se sedimentaron (8000xg, 8 min., 4 °C), se lavaron (10 ml de PBS IX, 4 °C), se resuspendieron (12 ml de Tris-HCl 100 mM, pH 7,5, 4 °C), se lisaron (homogeneizador AVESTIN® Emulsiflex C3 a 103,39 MPa (15.000 psi)) y se clarificaron (22.000xg, 15 min, 4 °C). Se trataron lisados clarificados (100 μ l) con $\pm$ 10 unidades de proteasa 3C de RVH exógena (ACCELAGEN™ H0101S) durante 4 horas a 37 °C y se ensayaron para determinar la actividad de Pgi. La actividad de Pgi se ensayó acoplando con glucosa-6-fosfato deshidrogenasa (G6PDH) y después de una reducción de nicotinamida adenina dinucleótido fosfato (NADP+) a una absorbancia de 340 nm durante cinco minutos a 37 °C. Las reacciones contenían Tris-HCl 100 mM (pH 7,5), MgSO₄ 8 mM, fructosa-6-fosfato 5 mM, NADP+ 1 mM, albúmina sérica bovina 0,25 mg/ml, 2,5 unidades de G6PDH purificada de *Leuconostoc mesenteroides* (MEGAZYME® E-GPDH5) y muestras de lisado/proteasa al 30 por ciento en volumen.

En última instancia, se seleccionaron 12 secuencias de genes *pgi* únicas, que contenían la secuencia de reconocimiento de 3C de RVH insertada después de los siguientes codones: 108, 109, 110, 138, 410, 526, 527, 528,

529, 530, 531 y 532. Estos 12 miembros del subconjunto de 76 miembros original mostraron tasas de crecimiento en medio M9G en un intervalo de 15 % de tipo silvestre y se inactivaron significativamente tras exposición a proteasa exógena (Tabla 1). La FIG. 5 muestra un tratamiento de ciclo temporal más largo con proteasa para un subconjunto de la biblioteca. Basándose en datos obtenidos de los experimentos descritos anteriormente, la variante de Pgi óptima contenía la secuencia de reconocimiento de 3C de RVH después del aminoácido 526 (también conocida como Pgi-HRV-I526 del gen *pgi-HRV-I526*).

Tabla 1: Comparación de tasas de crecimiento celular y actividad de Pgi.

Secuencia de reconocimiento de proteasa insertada después del aminoácido n.º	μ (h ⁻¹)	μ / tipo silvestre	Actividad de Pgi (μ mol/min/mg de proteína total)		Actividad / Actividad tipo silvestre	
			- proteasa	+ proteasa	- proteasa	+ proteasa
Δ <i>pgi</i>	0,15	0,22	0	0		
tipo silvestre	0,68	1	2,31	2,37	1	1
108	0,60	0,88	2,16	2,03	0,94	0,86
109	0,71	1,04	2,54	1,86	1,10	0,78
110	0,70	1,03	2,13	1,99	0,92	0,84
138	0,58	0,86	0,62	0,37	0,27	0,16
410	0,59	0,87	1,16	0,96	0,50	0,41
526	0,66	0,97	1,92	0,44	0,83	0,19
527	0,70	1,03	1,93	1,42	0,84	0,60
528	0,65	0,96	1,04	0,63	0,45	0,26
529	0,66	0,97	1,48	1,19	0,64	0,50
530	0,68	1,00	0,77	0,29	0,33	0,12
531	0,62	0,91	1,29	0,67	0,56	0,28
532	0,69	1,01	1,15	0,52	0,50	0,22

Nota: Valor de ensayo de actividad de control de Δ *pgi* pequeño (0,089) restado de todas las actividades de Pgi indicadas

Enfoque de selección de alto rendimiento agrupada

El locus cromosómico de *pgi* en GL12-085 se recombinó con la biblioteca agrupada de 562 miembros (a concentraciones equimolares), usando los métodos descritos anteriormente. La biblioteca celular resultante se sembró en placas en medio de agar M9 que carecía de arabinosa y complementado con 34 μ g/ml de cloranfenicol y glucosa 1 % (M9CG). Las células se sembraron en placas para obtener cobertura 5X de la biblioteca, lo que sumaba 11 placas con aproximadamente 250 células/placa, permitiendo de este modo siembra en placas repetida más fácil en la siguiente etapa del método. Estas placas de bibliotecas agrupadas se incubaron a 37 °C durante 1,5-2 días. Las colonias que representan miembros de bibliotecas que proporcionan Pgi activo se sembraron posteriormente en placas repetidas tanto en medio M9GC como en medio de agar M9 complementado con cloranfenicol 34 μ g/ml, glucosa 1 % y arabinosa 2 % (M9CGA). Las placas repetidas se incubaron a 37 °C durante 1,5-2 días. Todas las colonias que estaban presentes en placas de M9CG pero no presentes en placas de M9CGA se recuperaron para análisis adicional.

La región de Pgi de los genomas de estas colonias se amplificó por PCR y se secuenció. Se identificaron siete secuencias de genes *pgi* únicas, que contenían la secuencia de reconocimiento de 3C de RVH insertada después de los siguientes codones: 524, 525, 526, 528, 529, 531 y 545. También se identificaron miembros 526, 528, 529 y 531 del subconjunto de 76 miembros usando el ensayo de exploración individual, descrito anteriormente.

Estructura cristalina de proteínas

Casi todas de las 15 variantes de Pgi obtenidas del individuo y exploraciones agrupadas se mapean en regiones de bucle expuestas a disolvente de la estructura cristalina publicada para Pgi (banco de datos de proteínas ID: 3NBU). Además, las variantes 526-532 corresponden a una región de bucle que precede a la hélice C-terminal de Pgi y las variantes 524-525 corresponden al extremo C-terminal de otra hélice, que posee actividad catalítica (FIG. 6). La escisión mediada por proteasa de una secuencia de reconocimiento insertada en la región 524-532 escindiría por lo tanto la hélice C-terminal y puede dar como resultado la malformación de la hélice catalítica precedente. La retirada de la hélice C-terminal es probablemente perjudicial porque Pgi es un dímero y esta hélice ayuda a "sujetar" el dímero unido.

Demostración de inactivación mediada por proteasa eficaz de *Pgi* diana

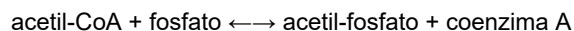
El gen *pgi* de *E. coli* BL21 (DE3) se reemplazó con *pgi-HRV-I526* (SEQ ID NO: 9). Esta cepa (GL12-116) se transformó individualmente con tres plásmidos: pACYCDuet-1, pGLC089 (FIG. 7) y pGLC221 (FIG. 8). pACYCDuet-1 es un plásmido de control de vector vacío de número de copias bajo, mientras que pGLC089 y pGLC221 son capaces de expresar proteasa 3C de RVH (con codones optimizados para expresión en *E. coli*) de un promotor T7 cuando se induce con isopropil β-D-1 tiogalactopiranosido (IPTG). El gen de proteasa de pGLC089 contiene secuencia adicional de modo que la proteasa 3C de RVH resultante posee un líder de OmpA N-terminal (MKKTAIAIAVALAGFATVAQA) (SEQ ID NO: 46) que secuestra la proteasa en el periplasma, mientras que el gen de proteasa de pGLC221 carece de dicho líder y se expresa en el citoplasma. Estas cepas se cultivaron en un medio de glucosa definido a 37 °C hasta fase semilogarítmica y se indujeron con IPTG 0,8 mM durante 2 horas. Se crearon lisados clarificados y se ensayaron posteriormente para determinar la actividad de Pgi (como se ha descrito anteriormente). La Tabla 2 muestra la tasa de crecimiento de cada cepa antes de la inducción y la actividad de Pgi medida en el lisado clarificado. Cuando se expresa proteasa 3C de RVH de forma citoplasmática, la tasa de crecimiento desciende en 40 % en comparación con la cepa que carece de proteasa, supuestamente debido a expresión filtrante de la proteasa antes de sobreexpresión inducida por IPTG.

Tabla 2: Comparación

Secuencia líder de OmpA para inmovilización periplásmica	Preinducción μ(h ⁻¹)	Actividad de Pgi (μmol/min/mg de proteína total)
sin proteasa	0,62	1,3
N.º	0,38	-0,01
Sí	0,55	0,01

Ejemplo 2 - Enzimas fosfotransacetilasa de *Escherichia coli*

La enzima fosfotransacetilasa (Pta) (también denominada fosfato acetiltransferasa) de *Escherichia coli* (*E. coli*) cataliza la primera reacción comprometida de metabolismo de exceso de acetato:



Se produce exceso de acetato en cultivos aeróbicos alimentados con glucosa rápidamente crecientes de *E. coli*. La acumulación de acetato secretado en medio de producción limita la tasa de crecimiento, la densidad de crecimiento y la producción de proteína recombinante, que es un problema típico en fermentación industrial. Las cepas con supresión de actividad de Pta crecen habitualmente 15-20 % más lento que sus homólogos de tipo silvestre y, aunque reducen la secreción de acetato, no resuelven el problema del exceso, ya que las cepas en su lugar secretan lactato y piruvato con efectos deletéreos similares. La dirección a proteasa de Pta en procesos de producción sin células previene la acumulación de acetato y desplaza el flujo de carbono al ciclo de ácido tricarboxílico, permitiendo al mismo tiempo que las cepas crezcan a tasas de crecimiento máximas más rápidas que sus homólogos con supresión de pta.

Construcción de bibliotecas de variantes de Pta

Se diseñó una biblioteca de ADN bicatenario lineal de 200 miembros y se construyó mediante reacción en cadena de la polimerasa (PCR), donde la secuencia del gen *pta* de *E. coli* nativa (SEQ ID NO: 47) se modificó para incluir una secuencia de nucleótidos (SEQ ID NO: 37) que codifica la secuencia de reconocimiento de proteasa de ocho aminoácidos (SEQ ID NO: 38) de la proteasa 3C de rinovirus humano (RVH). La secuencia de nucleótidos que codifica el motivo de reconocimiento de proteasa se insertó después de los siguientes codones en el gen *pta* de tipo silvestre: 350, 380-388, 401-405, 409-415, 426-431, 434-438, 446-465, 475-483, 490-495, 502-508, 511-518, 526-538, 544-549, 552-563, 577-586, 589-603, 615-620, 626-627, 629-632, 639-650, 653-660, 669-674, 681-687, 689-698, 709-713. Además de la secuencia codificante del gen, cada miembro de la biblioteca también contenía ramas de homología de 50 pares de bases (pb) (por ejemplo, secuencia adicional de 50 pb cadena arriba del codón de inicio del gen y 50 pb cadena abajo del codón de terminación del gen que son homólogas del locus *pta* de tipo silvestre del genoma de *E. coli*). Si la secuencia LEVLFQGP (SEQ ID NO: 38) fuera a insertarse después de los aminoácidos L, LE o LEV, solamente EVLFQGP (SEQ ID NO: 39), VLFQGP (SEQ ID NO: 40) o LFQGP (SEQ ID NO: 41) se insertaron, respectivamente. De manera similar, si la secuencia fuera a insertarse antes de los aminoácidos P, GP o QGP, solamente LEVLFQG (SEQ ID NO: 42), LEVLFQ (SEQ ID NO: 43) o LEVLF (SEQ ID NO: 44) se insertaron, respectivamente. Además, si la inserción fue entre los aminoácidos LP, por ejemplo, solamente EVLFQG (SEQ ID NO: 45) se insertó.

Diseño de cepas

Para la exploración de bibliotecas de Pta, se creó una versión modificada de JW2294-1 de *E. coli* (Coli Genetic Stock Center; CGSC n.º 9844) de la colección Keio (Mol. Syst. Biol. 2006; 2:2006-08). Para preparar la cepa para su uso en la exploración, el genoma de JW2294-1 se modificó de dos maneras, usando métodos descritos previamente en

el Ejemplo 1 para fosfoglucosa isomerasa. En primer lugar, el marcador de KanR, localizado en lugar del gen *pta*, se retiró, dejando un locus de *pta* que contenía las primeras tres bases y las últimas 21 bases del gen *pta*, con una secuencia cicatriz corta entre medias. En segundo lugar, el gen que codifica la acetil-CoA sintetasa (es decir, *acs*) se reemplazó con KanR, creando de este modo una cepa que carece de la capacidad de crecer en acetato como una única fuente de carbono y restaura la resistencia a kanamicina. El plásmido de recombinasa descrito previamente (pGLA042; Figura 3) se transformó en esta cepa para crear la cepa de exploración final (GL13-052), que se usó en un enfoque de selección y ensayo individual.

Enfoque de selección y ensayo individual

El locus cromosómico de *pta* en GL13-052 se recombinó individualmente con la biblioteca de ADN bicatenario, lineal, de 200 miembros, descrita anteriormente, que coloca la secuencia de reconocimiento de proteasa de RVH en regiones de bucle accesibles al disolvente predichas del dominio catalítico C-terminal de Pta. Como la estructura cristalina del Pta de *E. coli* aún no se ha determinado, las regiones de bucle anteriormente mencionadas se predijeron realizando un alineamiento de secuencias de aminoácidos del dominio catalítico C-terminal del Pta de *E. coli* con el de enzimas heterólogas con estructuras cristalinas publicadas (es decir, banco de datos de proteínas ID: 1R5J y 2AF3).

Se dejó crecer un cultivo de 2 litros (l) de GL13-052 hasta una densidad óptica a 600 nm de 0,5 a 30 °C en LB de baja salinidad (extracto de levadura 5 g/l, triptona 10 g/l, NaCl 5 g/l) que contiene kanamicina 10 µg/ml y carbenicilina 50 µg/ml. El cultivo se transfirió a un baño de agua a 42 °C y se agitó durante 30 minutos para inducir el sistema de recombinasa de pGLA042. Las células inducidas se hicieron electrocompetentes siguiendo métodos convencionales y se congelaron instantáneamente alícuotas de 50 µl en nitrógeno líquido y posteriormente se almacenaron a -80 °C antes de la recombinación con miembros de bibliotecas.

Cada ADN miembro de la biblioteca se transformó individualmente (25 µl de células y 50 ng de ADN miembro de la biblioteca) y se recuperó en 1 ml de LB de baja salinidad durante más de 3 horas a 30 °C. Se sembraron transformaciones recuperadas en placas en medio de agar M9 complementado con acetato 1 % y kanamicina 10 µg/ml y las placas se incubaron a 37 °C durante 2-3 días. Los acontecimientos de recombinación que produjeron colonias fueron representativos de miembros de biblioteca que contenían enzimas de Pta activas a pesar de la inclusión del motivo de secuencia de reconocimiento de proteasa. La región de Pta de los genomas de estos miembros de las bibliotecas se amplificó por PCR y se secuenció. Después se hicieron competentes y se transformaron cepas con secuencia verificada con pGLC217 para permitir la expresión inducible por arabinosa de proteasa de RVH. Las cepas también se transformaron con un plásmido de control (pGLC219), que es idéntico de pGLC217 excepto que el gen que codifica la proteasa de RVH se reemplazó con un gen que codifica la proteasa del virus del grabado del tabaco (TEV). Esta proteasa no escindiría variantes de Pta.

Para evaluar la susceptibilidad a proteasa *in vivo*, las tasas de crecimiento y secreción de acetato se midieron para cepas que contenían variantes de Pta viables que portaban pGLC217 o pGLC219 que se cultivaron a 37 °C en medio mínimo M9 que contenía glucosa 0,5 %, kanamicina 10 µg/ml y cloranfenicol 34 µg/ml. Estos cultivos se dejaron crecer ± arabinosa 2 % para sobreexpresar la proteasa. Por lo tanto, la tasa de crecimiento de cada variante de Pta viable se examinó en cuatro condiciones: (1) expresión de proteasa de TEV no inducida, (2) expresión de proteasa de TEV inducida por arabinosa, (3) expresión de proteasa de RVH no inducida y (4) expresión de proteasa de RVH inducida por arabinosa. Las condiciones (3) y (4) permiten que la desactivación de Pta mediada por proteasa se evalúe mediante la tasa de crecimiento, mientras que las condiciones (1) y (2) proporcionan un control para la carga metabólica ocasionada por el mantenimiento de plásmido y expresión de proteínas inducida por arabinosa. Como control positivo, se incluyó una cepa con *pta* de tipo silvestre en el estudio, mientras que una cepa con supresión de *pta* actuó como el control negativo.

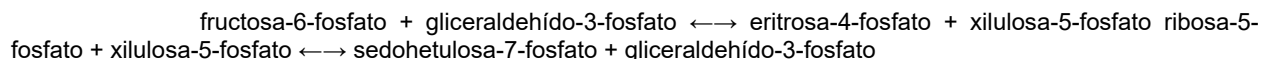
Basándose en fenotipos de crecimiento y secreción de acetato en comparación con las cepas de control, en última instancia se seleccionaron cuatro variantes de Pta únicas de la biblioteca de 200 miembros. Estas variantes contenían el motivo de escisión de proteasa insertado después del aminoácido 381, 382, 387 y 409 (SEQ ID NO: 50, 52, 54, 56, codificadas respectivamente por SEQ ID NO: 49, 51, 53, 55). La Figura 10A muestra los datos de secreción de acetato y la Figura 10B muestra los datos de secreción de lactato para las variantes seleccionadas en comparación con el Pta de tipo silvestre y controles de *pta*. El control de Pta de tipo silvestre muestra un fenotipo de alto acetato y bajo lactato, mientras que el control de *pta* muestra el fenotipo opuesto. Por lo tanto, si una variante de Pta particular es susceptible a desactivación mediada por proteasa de RVH, debería mostrar alto acetato y bajo lactato en ausencia de inducción de proteasa de RVH, mostrando al mismo tiempo lo opuesto cuando se induce la proteasa de RVH. En los controles de proteasa de TEV, el acetato debería ser alto y el lactato debería ser bajo independientemente de la inducción, ya que la secuencia de reconocimiento de proteasa de TEV no está presente en las variantes de Pta. De manera clara, las variantes 381, 382, 387 y 409 cumplen estos criterios.

Para demostrar adicionalmente que estas variantes de Pta eran susceptibles a proteasa de RVH, se realizaron ensayos de actividad *in vitro*. La Figura 11 muestra la actividad de cada variante de Pta en presencia o ausencia de proteasa de RVH proporcionada de forma exógena después de incubación durante 30 minutos a 37 °C. En todos los casos, la actividad de Pta se redujo drásticamente tras la incubación con proteasa de RVH. La Variante 381 (SEQ ID

NO: 50) mostró el mejor equilibrio entre tener actividad casi de tipo silvestre y ser particularmente susceptible a desactivación mediada por proteasa.

Ejemplo 3 - Enzimas transcetolasa de *Escherichia coli*

Las isozimas transcetolasa mayor y menor (TktA y TktB, respectivamente) de *Escherichia coli* (*E. coli*) catalizan dos reacciones de transferencia de cetol reversibles en la ruta de pentosa fosfato:



La actividad transcetolasa asegura un aporte adecuado de eritrosa-4-fosfato, un metabolito de carbono central clave necesario para la producción de los tres aminoácidos aromáticos así como varias vitaminas. Las cepas que carecen de actividad transcetolasa requieren complementación con compuestos procedentes de eritrosa-4-fosfato, tales como aminoácidos aromáticos y vitaminas. Incluso con dicha complementación, que puede ser costosa, el crecimiento hasta altas densidades celulares en fermentaciones es difícil. La transcetolasa también conecta la ruta de pentosa fosfato con la glucólisis, desviando pentosa fosfatos de la producción de nucleótidos excesiva cuando el flujo de glucosa es alto. La dirección a proteasa de transcetolasa, sin alterar significativamente su función crítica durante el crecimiento celular, beneficia la producción de moléculas procedentes de pentosa en reacciones sin células, ya que se evita que el flujo de ruta de pentosa fosfato alto se desvíe a la glucólisis. La estructura cristalina de transcetolasa se muestra en la Figura 12.

Construcción de bibliotecas de variante A de transcetolasa

Se diseñó una biblioteca de ADN bicatenario lineal de 200 miembros y se construyó mediante reacción en cadena de la polimerasa (PCR), donde la secuencia del gen (*tktA*) de transcetolasa A de *E. coli* nativa (SEQ ID NO: 57) se modificó para incluir una secuencia de nucleótidos (SEQ ID NO: 37) que codifica la secuencia de reconocimiento de proteasa de ocho aminoácidos (SEQ ID NO: 38) de la proteasa de RVH. La secuencia de nucleótidos que codifica el motivo de reconocimiento de proteasa se insertó después de los siguientes codones en el gen *tktA* de tipo silvestre: 22-24, 43-55, 78-83, 88-110, 138-148, 172-175, 185-192, 196-200, 208-210, 218-220, 233-238, 245-257, 261-287, 294-296, 331-336, 350-354, 371-372, 388-403, 484-487, 508-511, 523-529, 544-551, 573-577, 591-593, 601-607, 624-627, 633-640, 648-651. Además de la secuencia codificante del gen, cada miembro de la biblioteca contenía también ramas de homología de 50 pares de bases (pb) (por ejemplo, 50 pb de secuencia adicional cadena arriba del codón de inicio del gen y 50 pb de secuencia adicional cadena abajo del codón de terminación del gen, cada secuencia de 50 pb homóloga del locus *tktA* de tipo silvestre del genoma de *E. coli*). Para inserción de la secuencia LEVLFQGP (SEQ ID NO: 38) después de los aminoácidos L, LE o LEV, solamente EVLFQGP (SEQ ID NO: 39), VLFQGP (SEQ ID NO: 40) o LFQGP (SEQ ID NO: 41) se insertaron, respectivamente. De manera similar, para inserción de la secuencia LEVLFQGP (SEQ ID NO: 38) antes de los aminoácidos P, GP o QGP, solamente LEVLFQG (SEQ ID NO: 42), LEVLFQ (SEQ ID NO: 43) o LEVLF (SEQ ID NO: 44) se insertaron, respectivamente. Además, para inserción de LEVLFQGP (SEQ ID NO: 38) entre los aminoácidos LP, por ejemplo, solamente EVLFQG (SEQ ID NO: 45) se insertó.

Diseño de cepas

Para la exploración de bibliotecas de TktA, se creó una versión modificada de JW5478-1 de *E. coli* (Coli Genetic Stock Center; CGSC n.º 11606) de la colección Keio (Mol. Syst. Biol. 2006; 2:2006-08). Para preparar la cepa para su uso en la exploración, el genoma de JW5478-1 se modificó de dos maneras usando métodos descritos en el Ejemplo 1. En primer lugar, el marcador de KanR, localizado en lugar del gen *tktA*, se retiró, dejando un locus de *tktA* que contenía las primeras 3 bases y las últimas 21 bases del gen *tktA*, con una secuencia cicatriz corta entre medias. En segundo lugar, el gen que codifica la transcetolasa menor (*tktB*) se reemplazó con KanR, creando de este modo una cepa que carece de ninguna actividad transcetolasa y restaura la resistencia a kanamicina. El plásmido de recombinasa, pGLA042, (FIG. 3) se transformó en esta cepa para crear la cepa de exploración final, GL13-050, que se usó en un enfoque de selección y ensayo individual.

Enfoque de selección y ensayo individual

El locus cromosómico de *tktA* en GL13-050 se recombinó individualmente con la biblioteca de ADN bicatenario, lineal, de 200 miembros, descrita anteriormente, que coloca la secuencia de reconocimiento de proteasa de RVH en regiones de bucle accesibles al disolvente de TktA, como se predijo por su estructura cristalina (banco de datos de proteínas ID: 1QGD) (véase FIG. 11).

Se dejó crecer un cultivo de 2 litros (l) de GL13-050 hasta una densidad óptica a 600 nm de 0,5 a 30 °C en un medio Vogel Bonner E (mVBE) modificado (sales mínimas VBE IX, glucosa al 0,4 %, casaminoácidos al 2 %, triptófano 1 mM, 2,3-dihidroxibenzoato 0,25 mM, p-aminobenzoato 30 µM, p-hidroxibenzoato 30 µM, piridoxina-HCl 5 µM, kanamicina 10 µg/ml). El cultivo se transfirió a un baño de agua a 42 °C y se agitó durante 30 minutos para inducir el sistema de recombinasa de pGLA042. Las células inducidas se hicieron electrocompetentes siguiendo métodos

convencionales y se congelaron instantáneamente alícuotas de 50 µl en nitrógeno líquido y posteriormente se almacenaron a -80 °C antes de la recombinación con miembros de bibliotecas.

Cada ADN miembro de la biblioteca se transformó individualmente (25 µl de células y 50 ng de ADN miembro de la biblioteca) y se recuperó en 1 ml de - LB de baja salinidad durante más de 3 horas a 30 °C. Se sembraron transformantes recuperados en placas en medio de agar M9 complementado con glucosa 1 % y kanamicina 10 µg/ml y las placas se incubaron a 37 °C durante 2-3 días. Las transformaciones/recombinaciones que produjeron colonias fueron representativas de miembros de bibliotecas que contenían moléculas de TktA activas a pesar de la inclusión del motivo de secuencia de reconocimiento de proteasa. La región de TktA de los genomas de estos miembros de bibliotecas se amplificó por PCR y se secuenció. Después se hicieron competentes y se transformaron cepas con secuencia verificada con pGLC217 para permitir la expresión inducible por arabinosa de proteasa de RVH. También se transformaron cepas con un plásmido de control, pGLC219, que es casi idéntico a pGLC217, con la diferencia de que el gen que codifica la proteasa de RVH se reemplazó con un gen que codifica la proteasa del virus del grabado del tabaco (TEV). Esta proteasa no escindiría variantes de TktA.

Para evaluar la susceptibilidad a proteasa *in vivo*, las tasas de crecimiento se midieron para cepas que contenían variantes de TktA viables que portaban pGLC217 o pGLC219 que se cultivaron a 37 °C en medio mínimo M9 que contenía glucosa 0,5 %, kanamicina 10 µg/ml y cloranfenicol 34 µg/ml. Estos cultivos se dejaron crecer ± arabinosa 2 % para sobreexpresar la proteasa. Por lo tanto, la tasa de crecimiento de cada variante de TktA viable se examinó en cuatro condiciones: (1) expresión de proteasa de TEV no inducida, (2) expresión de proteasa de TEV inducida por arabinosa, (3) expresión de proteasa de RVH no inducida y (4) expresión de proteasa de RVH inducida por arabinosa. Las condiciones (3) y (4) permitieron que la desactivación de TktA mediada por proteasa de RVH se evalué mediante la tasa de crecimiento, mientras que las condiciones (1) y (2) proporcionaron un control para la carga metabólica ocasionada por el mantenimiento de plásmido y expresión de proteínas inducida por arabinosa.

Se seleccionaron cinco variantes de TktA únicas de la biblioteca de 200 miembros. Estas variantes seleccionadas contenían el motivo de escisión de proteasa de RVH insertado después del aminoácido 635, 636, 637, 638 y 640 (SEQ ID NO: 64-SEQ ID NO: 68, codificadas respectivamente por SEQ ID NO: 58-SEQ NO: 62). Como se muestra en la tabla 3, las variantes seleccionadas fueron capaces de apoyar el crecimiento en medio de glucosa mínimo cuando no se indujo expresión de proteasa de RVH pero su crecimiento se alteró significativamente cuando se indujo proteasa de RVH. Como muestran los controles de proteasa de TEV, la alteración del crecimiento ocasionada mediante inducción de proteasa no era simplemente el resultado de la complejidad de la sobreexpresión de proteínas. Por lo tanto, estas variantes de TktA fueron susceptibles a desactivación mediada por proteasa de RVH *in vivo*.

Tabla 3. Tasas de crecimiento de variantes de TktA dirigibles a proteasa en medio de glucosa mínimo

Secuencia de reconocimiento de proteasa insertada después del aminoácido n.º	proteasa de TEV		proteasa de RVH	
	- inductor	+ inductor	- inductor	+ inductor
tipo silvestre	0,55	0,53	0,60	0,57
635	0,52	0,52	0,58	0,47
636	0,39	0,42	0,43	0,17
637	0,46	0,46	0,40	0,11
638	0,18	0,18	0,13	0,08
640	0,48	0,52	0,55	0,34

Estructura cristalina de proteínas

Según la estructura cristalina publicada para TktA (banco de datos de proteínas ID: 1QGD), las cinco variantes de TktA obtenidas de la exploración (Tabla 3) se mapean en el bucle que precede inmediatamente a la hélice C-terminal (FIG. 11). TktA es activo como un dímero y este bucle aparece en una interfaz de dimerización. Sin quedar ligado a teoría alguna, la escisión de este bucle por proteasa de RVH probablemente altere la capacidad de la TktA para dimerizar. De hecho, la dimerización de transcetolasa es limitante de la tasa en la formación de la enzima activa (J. Biol. Chem. 1981; 256:4877-83).

Todas las definiciones, como se definen y usan en el presente documento, deberían entenderse como preferentes sobre las definiciones de diccionario y/o los significados ordinarios de los términos definidos.

Debería entenderse que los artículos indefinidos "un" y "una", como se usan en el presente documento en la memoria descriptiva y en las reivindicaciones, a menos que se indique claramente lo contrario, significan "al menos uno".

Debería entenderse que la expresión "y/o", como se usa en el presente documento en la memoria descriptiva y en las reivindicaciones, significa "uno o ambos" de los elementos unidos de este modo, es decir, elementos que están presentes conjuntamente en algunos casos y presentes por separado en otros casos. Múltiples elementos

enumerados con "y/o" deberían interpretarse de la misma manera, es decir, "uno o más" de los elementos unidos de este modo. Pueden estar presentes opcionalmente otros elementos distintos de los elementos identificados específicamente por la cláusula "y/o", bien relacionados o no relacionados con los elementos identificados específicamente. Por lo tanto, como un ejemplo no limitante, una referencia a "A y/o B", cuando se usa junto con lenguaje abierto tal como "que comprende" puede hacer referencia, en una realización, solamente a A (opcionalmente incluyendo elementos distintos de B); en otra realización, solamente a B (opcionalmente incluyendo elementos distintos de A); en otra realización más, tanto a A como a B (opcionalmente incluyendo otros elementos); etc.

Como se usa en el presente documento en la memoria descriptiva y en las reivindicaciones, debería entenderse que "o" tiene el mismo significado que "y/o" como se ha definido anteriormente. Por ejemplo, cuando se separan artículos en una lista, "o" o "y/o" se interpretarán como inclusivos, es decir, la inclusión de al menos uno, pero también incluyendo más de uno, de varios o una lista de elementos y, opcionalmente, artículos no enumerados adicionales. Solamente términos que indiquen claramente lo contrario, tales como "solamente uno de" o "exactamente uno de", o, cuando se use en las reivindicaciones, "que consiste en", se referirán a la inclusión de exactamente un elemento de varios o una lista de elementos. En general, el término "o" como se usa en el presente documento solamente se interpretará como indicativo de alternativas exclusivas (es decir, "uno o el otro pero no ambos") cuando está precedido de expresiones de exclusividad, tales como "bien", "uno de", "solamente uno de", o "exactamente uno de". "Que consiste esencialmente en", cuando se usa en las reivindicaciones, tendrá su significado habitual usado en el campo de la ley de patentes.

Como se usa en el presente documento en la memoria descriptiva y en las reivindicaciones, debería entenderse que la expresión "al menos uno", en referencia a una lista de uno o más elementos, significa al menos un elemento seleccionado de uno cualquiera o más de los elementos en la lista de elementos, pero no necesariamente incluyendo al menos uno de todos y cada uno de los elementos enumerados específicamente en la lista de elementos y no excluyendo cualquier combinación de elementos en la lista de elementos. Esta definición también permite que puedan estar presentes opcionalmente elementos distintos de los elementos identificados específicamente en la lista de elementos a la que se refiere la expresión "al menos uno", bien relacionados o no relacionados con los elementos identificados específicamente. Por lo tanto, como ejemplo no limitante, "al menos uno de A y B" (o, de forma equivalente, "al menos uno de A o B", o, de forma equivalente "al menos uno de A y/o B") pueden referirse, en una realización, a al menos un, opcionalmente incluyendo más de uno, A, sin B presente (y opcionalmente incluyendo elementos distintos B); en otra realización, a al menos un, opcionalmente incluyendo más de uno, B, sin A presente (y opcionalmente incluyendo elementos distintos A); en otra realización más, a al menos un, opcionalmente incluyendo más de uno, A, y al menos un, opcionalmente incluyendo más de uno, B (y opcionalmente incluyendo otros elementos); etc.

También se debería entender que, a menos que se indique claramente lo contrario, en cualquier método desvelado en el presente documento que incluya más de una etapa o un acto, el orden de las etapas o los actos del método no está necesariamente limitado al orden en el que se enumeran las etapas o los actos del método.

En las reivindicaciones, así como en la memoria descriptiva anterior, debe entenderse que todas las expresiones transicionales tales como "que comprende", "que incluye", "que porta", "que tiene", "que contiene", "que implica", "que mantiene", "compuesto de", y similares son abiertas, es decir, significan incluyendo, pero sin limitación. Solamente las expresiones "que consiste en" y "que consiste esencialmente en" serán expresiones transicionales cerradas o semicerradas, respectivamente, como se expone en el Manual de la oficina de patente de los Estados Unidos de procedimientos de examen de patentes, Sección 2111.03.

LISTADO DE SECUENCIAS

SEQ ID NO:1

ATGAAAAACATCAATCCAACGCAGACCGCTGCCTGGCAGGCACTACAGAAACACTTCGATGA
AATGAAAGACGTTACGATCGCCGATCTTTTTGCTAAAGACGGCGATCGTTTTTCTAAGTTCT
CCGCAACCTTCGACGATCAGATGCTGGTGGATTACTCCAAAAACCGCATCACTGAAGAGACG
CTGGCGAAATTACAGGATCTGGCGAAAGAGTGCGATCTGGCGGGCGCGATTAAAGTCGATGTT
CTCTGGCGAGAAGATCAACCGCACTGAAAACCGCGCCGTGCTGCACGTAGCGCTGCGTAACC
GTAGCAATACCCCGATTTTGGTTGATGGCAAAGACGTAATGCCGGAAGTCAACGCGGTGCTG
GAGAAGATGAAAACCTTCTCAGAAGCGATTATTTCCGGTGAGTGGAAGGTTATACCGGCAA
AGCAATCACTGACGTAGTGAACATCGGGATCGGCGGTTCTGACCTCGGCCCATAACATGGTGA
CCGAAGCTCTGCGTCCGTACAAAAACCACTGAACATGCACCTTGTCTTCTAACGTCGATGGG
ACTCACATCGCGGAAGTGCTGAAAAAAGTAAACCCGGAAACCAACGCTGTTCTTGGTAGCATC
TAAAACCTTCACCACTCAGGAACTATGACCAACGCCCATAGCGCGCGTGACTGGTTCCTGA
AAGCGGCAGGTGATGAAAAACACGTTGCAAAACACTTTGCGGCGCTTCCACCAATGCCAAA
GCCGTTGGCGAGTTTGGTATTGATACTGCCAACATGTTTCGAGTTCTGGGACTGGGTGGCGG
CCGTTACTCTTTGTGGTCAGCGATTGGCCTGTCGATTGTTCTCTCCATCGGCTTTGATAACT
TCGTTGAACTGCTTTCCGGCGCACACGCGATGGACAAGCATTTCTCCACCACGCCTGCCGAG
AAAAACCTGCCTGTACTGCTGGCGCTGATTGGCATCTGGTACAACAATTTCTTTGGTGCGGA
AACTGAAGCGATTCTGCCGTATGACCAGTATATGCACCGTTTCGCGGCGTACTTCCAGCAGG
GCAATATGGAGTCCAACGGTAAGTATGTTGACCGTAACGGTAACGTTGTGGATTACCAGACT
GGCCCATTATCTGGGGTGAACCAGGCACTAACGGTCAGCACGCGTTCTACCAGCTGATCCA
CCAGGGAACCAAAATGGTACCGTGCGATTTTCATCGCTCCGGCTATCACCCATAACCCGCTCT
CTGATCATCACCAGAACTGCTGTCTAACTTCTTCGCCCAGACCGAAGCGCTGGCGTTTGGT
AAATCCCGCGAAGTGTTGAGCAGGAATATCGTGATCAGGGTAAAGATCCGGCAACGCTTGA
CTACGTGGTGCCGTTCAAAGTATTGAAGGTAACCGCCCGACCAACTCCATCCTGCTGCGTG
AAATCACTCCGTTTCAAGCTGGGTGCGTTGATTGCGCTGTATGAGCACAAAATCTTTACTCAG
GGCGTGATCCTGAACATCTTCACCTTCGACCAGTGGGGCGTGGAAGTGGGTAAACAGCTGGC
GAACCGTATTCTGCCAGAGCTGAAAGATGATAAAGAAATCAGCAGCCACGATAGCTCGACCA
ATGGTCTGATTAAACCGCTATAAAGCGTGGCGCGGTTAA

5

SEQ ID NO:2

ATGAAAAACATCAATCCAACGCAGACCGCTGCCTGGCAGGCACTACAGAAACACTTCGATGAAATGAA
 AGACGTTACGATCGCCGATCTTTTTGCTAAAGATGGTGATCGTTTTCTAAGTTCTCCGCAACCTTCG
 ACGATCAGATGCTGGTGGATTACTCCAAAAACCGCATCACTGAAGAGACGCTGGCGAAATTACAGGAT
 CTGGCGAAAGAGTGCGATCTGGCGGGCGCGATTAAAGTCGATGTTCTCTGGCGAGAAGATCAACCGCAC
 TGAAAACCGCGCCGTGCTGCACGTAGCGCTGCGTAACCGTAGCAATACCCCGCTGGAAGTGCTGTTTC
 AGGGTCCGATTTTGGTTGATGGCAAAGACGTAATGCCGGAAGTCAACGCGGTGCTGGAGAAGATGAAA
 ACCTTCTCAGAAGCGATTATTTCCGGTGAGTGGAAGGTTATACCGGCAAAGCAATCACTGACGTAGT
 GAACATCGGGATCGGCGGTTCTGACCTCGGCCCATACATGGTGACCGAAGCTCTGCGTCCGTACAAAA
 ACCACCTGAACATGCACTTTGTTTCTAACGTCGATGGGACTCACATCGCGGAAGTGCTGAAAAAAGTA
 AACCCGGAACACGCTGTTCTTGGTAGCATCTAAACCTTCACCACTCAGGAACTATGACCAACGC
 CCATAGCGCGCTGACTGGTTCCTGAAAGCGGCAGGTGATGAGAAGCACGTTGCAAAACACTTTGCGG
 CGCTTTCCACCAATGCCAAAGCCGTTGGCGAGTTTGGTATTGATACTGCCAACATGTTTCGAGTTCTGG
 GACTGGGTGGCGGCCGTTACTCTTTGTGGTCAGCGATTGGCCTGTCGATTGTTCTCTCCATCGGCTT
 TGATAACTTCGTTGAACTGCTTTCCGGCGCACACGCGATGGACAAGCATTTCTCCACCACGCCTGCCG
 AGAAAAACCTGCCTGTACTGCTGGCGCTGATTGGCATCTGGTACAACAATTTCTTTGGTGCGGAACT
 GAAGCGATTCTGCCGTATGACCAGTATATGCACCGTTTCGCGGCGTACTTCAGCAGGGCAATATGGA
 GTCCAACGGTAAGTATGTTGACCGTAACGGTAACGTTGTGGATTACCAGACTGGCCCGATTATCTGGG
 GTGAACCAGGCCTAACGGTCAGCACGCTTCTACCAGCTGATCCACCAGGGAACCAAAATGGTACCG

TGCGATTTTCATCGCTCCGGCTATCACCCATAACCCGCTCTCTGATCATCACCAGAACTGCTGTCTAA
 CTTCTTCGCCCAGACCGAAGCGCTGGCGTTTGGTAAATCCCGCGAAGTGTTGAGCAGGAATATCGTG
 ATCAGGGTAAAGATCCGGCAACGCTTGACTACGTGGTGCCGTTCAAAGTATTCGAAGGTAACCGCCG
 ACCAACTCCATCCTGCTGCGTGAAATCACTCCGTTTCAGCCTGGGTGCGTTGATTGCGCTGTATGAGCA
 CAAAATCTTTACTCAGGGCGTGATCCTGAACATCTTCACCTTCGACCAGTGGGGCGTGGAAGTGGGT
 AACAGCTGGCGAACCGTATTCTGCCAGAGCTGAAAGATGATAAAGAAATCAGCAGCCACGATAGCTCG
 ACCAATGGTCTGATTAACCGCTATAAAGCGTGGCGCGGTTaa

5

SEQ ID NO:3

ATGAAAAACATCAATCCAACGCAGACCGCTGCCTGGCAGGCACTACAGAAACACTTCGATGA
 AATGAAAGACGTTACGATCGCCGATCTTTTTGCTAAAGATGGTGATCGTTTTCTAAGTTCT
 CCGCAACCTTCGACGATCAGATGCTGGTGGATTACTCCAAAAACCGCATCACTGAAGAGACG
 CTGGCGAAATTACAGGATCTGGCGAAAGAGATGCGATCTGGCGGGCGCGATTAAAGTCGATGTT
 CTCTGGCGAGAAGATCAACCGCACTGAAAACCGCGCCGTGCTGCACGTAGCGTGCCTGCGTAACC
 GTAGCAATACCCGATTCTGGAAGTGCTGTTTCAGGGTCCGTTGGTTGATGGCAAAGACGTA
 ATGCCGGAAGTCAACGCGGTGCTGGAGAAGATGAAAACCTTCTCAGAAGCGATTATTTCCGG
 TGAGTGGAAGGTTATACCGGCAAAGCAATCACTGACGTAGTGAACATCGGGATCGGCGGTT
 CTGACCTCGGCCCATACATGGTGACCGAAGCTCTGCGTCCGTACAAAAACCACTGAACATG
 CACTTTGTTTCTAACGTCGATGGGACTCACATCGCGGAAGTGCTGAAAAAAGTAAACCCGGA
 AACCACGCTGTTCTTGGTAGCATCTAAACCTTCACCACTCAGGAACTATGACCAACGCCC
 ATAGCGCGCGTGACTGGTTCCTGAAAGCGGCAGGTGATGAGAAGCACGTTGCAAAACACTTT
 GCGGCGCTTTCCACCAATGCCAAAGCCGTTGGCGAGTTTGGTATTGATACTGCCAACATGTT
 CGAGTTCTGGGACTGGGTGGCGGCCGTTACTCTTTGTGGTCAGCGATTGGCCTGTCGATTG
 TTCTCTCCATCGGCTTTGATAACTTCGTTGAACTGCTTTCCGGCGCACACGCGATGGACAAG
 CATTTCTCCACCACGCTGCCGAGAAAAACCTGCCTGTACTGCTGGCGCTGATTGGCATCTG
 GTACAACAATTTCTTTGGTGCGGAACTGAAGCGATTCTGCCGTATGACCAGTATATGCACC
 GTTTCGCGGCGTACTTCCAGCAGGGCAATATGGAGTCCAACGGTAAGTATGTTGACCGTAAC
 GGTAACGTTGTGGATTACCAGACTGGCCCGATTATCTGGGGTGAACCAGGCACTAACGGTCA
 GCACGCGTTCTACCAGCTGATCCACCAGGGAACCAAAATGGTACCGTGCGATTTTCATCGCTC
 CGGCTATCACCCATAACCCGCTCTCTGATCATCACCAGAACTGCTGTCTAACTTCTTCGCC
 CAGACCGAAGCGCTGGCGTTTGGTAAATCCCGCGAAGTGGTTGAGCAGGAATATCGTGATCA
 GGGTAAAGATCCGGCAACGCTTGACTACGTGGTGCCGTTCAAAGTATTGCAAGGTAACCGCC
 CGACCAACTCCATCCTGCTGCGTGAAATCACTCCGTTTCAGCCTGGGTGCGTTGATTGCGCTG
 TATGAGCACAAAACTTTACTCAGGGCGTGATCCTGAACATCTTCACCTTCGACCAGTGGGG
 CGTGGAAGTGGGTAAACAGCTGGCGAACCGTATTCTGCCAGAGCTGAAAGATGATAAAGAAA
 TCAGCAGCCACGATAGCTCGACCAATGGTCTGATTAACCGCTATAAAGCGTGGCGCGGTTaa

SEQ ID NO:4

ATGAAAAACATCAATCCAACGCAGACCGCTGCCTGGCAGGCACTACAGAAACACTTCGATGA
AATGAAAGACGTTACGATCGCCGATCTTTTTGCTAAAGATGGTGATCGTTTTTCTAAGTTCT
CCGCAACCTTCGACGATCAGATGCTGGTGGATTACTCCAAAAACCGCATCACTGAAGAGACG
CTGGCGAAATTACAGGATCTGGCGAAAGAGTGCGATCTGGCGGGCGCGATTAAAGTCGATGTT
CTCTGGCGAGAAGATCAACCGCACTGAAAACCGCGCCGTGCTGCACGTAGCGCTGCGTAACC
CTAGCAATACCCCGATTTTGAAGTCCTCTTTACGGTCCGGTTGATGGCAAAGACGTAATC
CCGGAAGTCAACGCGGTGCTGGAGAAGATGAAAACCTTCTCAGAAGCGATTATTTCCGGTGA
GTGGAAGGTTATACCGGCAAAGCAATCACTGACGTAGTGAACATCGGGATCGGCGGTTCTG
ACCTCGGCCCATAACATGGTGACCGAAGCTCTGCGTCCGTACAAAAACCACCTGAACATGCAC
TTTGTCTTAACGTCGATGGGACTCACATCGCGGAAGTGCTGAAAAAGTAAACCCGGAAC
CACGCTGTTCTTGGTAGCATCTAAAACCTTCACCACTCAGGAACTATGACCAACGCCCATA

GCGCGCGTGACTGGTTCCTGAAAGCGGCAGGTGATGAGAAACACGTTGCAAAACACTTTGCG
GCGCTTTCCACCAATGCCAAAGCCGTTGGCGAGTTTGGTATTGATACTGCCAACATGTTCTGA
GTTCTGGGACTGGGTTGGCGGCCGTTACTCTTTGTGGTCAGCGATTGGCCTGTCGATTGTTCT
TCTCCATCGGCTTTGATAACTTCGTTGAACTGCTTCCGGCGCACACGCGATGGACAAGCAT
TTCTCCACCACGCTGCCGAGAAAAACCTGCCTGTACTGCTGGCGCTGATTGGCATCTGGTA
CAACAATTTCTTTGGTGCGGAACTGAAGCGATTCTGCCGTATGACCAGTATATGCACCGTT
TCGCGGCGTACTTCCAGCAGGGCAATATGGAGTCCAACGGTAAGTATGTTGACCGTAACGGT
AACGTTGTGGATTACAGACTGGCCCGATTATCTGCGGTGAACAGGCACCTAACGGTCAGCA
CGCGTTCTACCAGCTGATCCACCAGGGAACCAAAATGGTACCGTGCGATTTTCATCGCTCCGG
CTATCACCCATAACCCGCTCTCTGATCATCACCAGAACTGCTGTCTAACTTCTTCGCCCAG
ACCGAAGCGCTGGCGTTTGGTAAATCCCGCGAAGTGGTTGAGCAGGAATATCGTGATCAGGG
TAAAGATCCGGCAACGCTTGACTACGTGGTGCCGTTCAAAGTATTCGAAGGTAACCGCCCGA
CCAATCCATCTCTGCTGCGTGAAATCACTCCGTTACGCTGGGTGCGTTGATTGCGCTGTAT
GAGCACAAAATCTTTACTCAGGGCGTGATCCTGAACATCTTCACCTTCGACCAGTGGGGCGT
GGAAGTGGGTAAACAGCTGGCGAACCGTATTCTGCCAGAGCTGAAAGATGATAAAGAAATCA
GCAGCCACGATAGCTCGACCAATGGTCTGATTAACCGCTATAAAGCGTGGCGCGGTtaa

5

SEQ ID NO:5

ATGAAAAACATCAATCCAACGCAGACCGCTGCCTGGCAGGCACTACAGAAACACTTCGATGA
AATGAAAGACGTTACGATCGCCGATCTTTTTGCTAAAGATGGTGATCGTTTTTCTAAGTTCT
CCGCAACCTTCGACGATCAGATGCTGGTGGATTACTCCAAAAACCGCATCACTGAAGAGACG
CTGGCGAAATTACAGGATCTGGCGAAAGAGTGCGATCTGGCGGGCGCGATTAAAGTCGATGTT
CTCTGGCGAGAAGATCAACCGCACTGAAAACCGCGCCGTGCTGCACGTAGCGCTGCGTAACC
GTAGCAATACCCCGATTTTGGTTGATGGCAAAGACGTAATGCCGGAAGTCAACGCGGTGCTG
GAGAAGATGAAAACCTTCTCAGAAGCGATTATTTCCGGTGAGCTGGAAGTGCTGTTTCAGGG
TCCGTGGAAAGGTTATACCGGCAAAGCAATCACTGACGTAGTGAACATCGGGATCGGCGGTT
CTGACCTCGGCCCATAACATGGTGACCGAAGCTCTGCGTCCGTACAAAAACCACCTGAACATG
CACTTTGTTTCTAACGTCGATGGGACTCACATCGCGGAAGTGCTGAAAAAGTAAACCCGGA
AACCACGCTGTTCTTGGTAGCATCTAAAACCTTCACCACTCAGGAACTATGACCAACGCCC
ATAGCGCGCGTGACTGGTTCCGAAAGCGGCAGGTGATGAGAAGCACGTTGCAAAACACTTT
GCGGCGCTTTCCACCAATGCCAAAGCCGTTGGCGAGTTTGGTATTGATACTGCCAACATGTT
CGAGTTCTGGGACTGGGTTGGCGGCCGTTACTCTTTGTGGTCAGCGATTGGCCTGTCGATTG
TTCTCTCCATCGGCTTTGATAACTTCGTTGAACTGCTTTCCGGCGCACACGCGATGGACAAG
CATTTCTCCACCACGCTGCCGAGAAAAACCTGCCTGTACTGCTGGCGCTGATTGGCATCTG
GTACAACAATTTCTTTGGTGCGGAACTGAAGCGATTCTGCCGTATGACCAGTATATGCACC
GTTTCGCGGCGTACTTCCAGCAGGGCAATATGGAGTCCAACGGTAAGTATGTTGACCGTAAC
GGTAACGTTGTGGATTACCAGACTGGCCGATTATCTGGGGTGAACCAGGCACCTAACGGTCA
GCACGCGTTCTACCAGCTGATCCACCAGGGAACCAAAATGGTACCGTGCGATTTTCATCGCTC
CGGCTATCACCCATAACCCGCTCTCTGATCATCACCAGAACTGCTGTCTAACTTCTTCGCC
CAGACCGAAGCGCTGGCGTTTGGTAAATCCCGCGAAGTGGTTGAGCAGGAATATCGTGATCA
GGGTAAAGATCCGGCAACGCTTGACTACGTGGTGCCGTTCAAAGTATTCGAAGGTAACCGCC
CGACCAACTCCATCTCTGCTGCGTGAAATCACTCCGTTACGCTGGGTGCGTTGATTGCGCTG
TATGAGCACAAAATCTTTACTCAGGGCGTGATCCTGAACATCTTCACCTTCGACCAGTGGGG
CGTGGAACTGGGTAAACAGCTGGCGAACCGTATTCTGCCAGAGCTGAAAGATGATAAAGAAA
TCAGCAGCCACGATAGCTCGACCAATGGTCTGATTAACCGCTATAAAGCGTGGCGCGGTtaa

10

SEQ ID NO:6

ATGAAAAACATCAATCCAACGCAGACCGCTGCCTGGCAGGCACTACAGAAACACTTCGATGA
AATGAAAGACGTTACGATCGCCGATCTTTTGGCTAAAGATGGTGATCGTTTTTCTAAGTTCT

CCGCAACCTTCGACGATCAGATGCTGGTGGATTACTCCAAAAACCGCATCACTGAAGAGACG
CTGGCGAAATTACAGGATCTGGCGAAAGAGTGCGATCTGGCGGGCGCGATTAAAGTCGATGTT
CTCTGGCGAGAAGATCAACCGCACTGAAAAACCGCGCCGTGCTGCACGTAGCGCTGCGTAACC
GTAGCAATACCCCGATTTTGGTTGATGGCAAAGACGTAATGCCGGAAGTCAACGCGGTGCTG
GAGAAGATGAAAACCTTCTCAGAAGCGATTATTTCCGGTGAGTGGAAGGTTATACCGGCAA
AGCAATCACTGACGTAGTGAACATCGGGATCGGCGGTTCTGACCTCGGCCCATACATGGTGA
CCGAAGCTCTGCGTCCGTACAAAAACCACCTGAACATGCACCTTTGTTTCTAACGTCGATGGG
ACTCACATCGCGGAAGTGCTGAAAAAAGTAAACCCGGAACACGCTGTTCTTGGTAGCATC
TAAAACCTTCACCACTCAGGAAACTATGACCAACGCCCATAGCGCGCGTGACTGGTTCCTGA
AAGCGGCAGGTGATGAGAAGCACGTTGCAAAACACTTTGCGGCGCTTTCCACCAATGCCAAA
GCCGTTGGCGAGTTTGGTATTGATACTGCCAACATGTTTCGAGTTCTGGGACTGGGTTGGCGG
CCGTTACTCTTTTGGTCAGCGATTGGCCTGTCGATTGTTCTCTCCATCGGCTTTGATAACT
TCGTTGAACTGCTTTCCGGCGCACACGCGATGGACAAGCATTTCTCCACCACGCTTCCGAG
AAAAACCTGCCTGTACTGCTGGCGCTGATTGGCATCTGGTACAACAATTTCTTTGGTGCGGA
AACTGAAGCGATTCTGCCGTATGACCAGTATATGCACCGTTTCGCGGCGTACTTCCAGCAGG
GCAATATGGAGTCCAACGGTAAGTATGTTGACCGTAACGGTAACGTTGTGGATTACCAGACT
GGCCCGATTATCTGGGGTGAACCAGGCACTAACGGTCAGCACGCGTTCTACCAGCTGATCCA
CCAGGGAACCAAAATGGTACCGTGCGATTTCATCGCTCCGGCTATCACCCATCTGGAAGTGC
TGTTTCAGGGTCCGAACCCGCTCTCTGATCATCACCAGAACTGCTGTCTAACTTCTTCGCC
CAGACCGAAGCGCTGGCGTTTGGTAAATCCCGCGAAGTGTTGAGCAGGAATATCGTGATCA
GGGTAAAGATCCGGCAACGCTTGAATACGTGGTGCCGTTCAAAGTATTGCAAGGTAACCGCC
CGACCAACTCCATCCTGCTGCGTGAAATCACTCCGTTTCAGCCTGGGTGCGTTGATTGCGCTG
TATGAGCACAAAATCTTTACTCAGGGCGTGATCCTGAACATCTTCACCTTCGACCAGTGGGG
CGTGGAAGTGGGTAAACAGCTGGCGAACCGTATTCTGCCAGAGCTGAAAGATGATAAAGAAA
TCAGCAGCCACGATAGCTCGACCAATGGTCTGATTAACCGCTATAAAGCGTGGCGCGGTtaa

5 SEQ ID NO:7

ATGAAAAACATCAATCCAACGCAGACCGCTGCCTGGCAGGCACTACAGAAACACTTCGATGA
AATGAAAGACGTTACGATCGCCGATCTTTTGGCTAAAGATGGTGATCGTTTTTCTAAGTTCT
CCGCAACCTTCGACGATCAGATGCTGGTGGATTACTCCAAAAACCGCATCACTGAAGAGACG
CTGGCGAAATTACAGGATCTGGCGAAAGAGTGCGATCTGGCGGGCGCGATTAAAGTCGATGTT
CTCTGGCGAGAAGATCAACCGCACTGAAAAACCGCGCCGTGCTGCACGTAGCGCTGCGTAACC
GTAGCAATACCCCGATTTTGGTTGATGGCAAAGACGTAATGCCGGAAGTCAACGCGGTGCTG
GAGAAGATGAAAACCTTCTCAGAAGCGATTATTTCCGGTGAGTGGAAGGTTATACCGGCAA
AGCAATCACTGACGTAGTGAACATCGGGATCGGCGGTTCTGACCTCGGCCCATACATGGTGA
CCGAAGCTCTGCGTCCGTACAAAAACCACCTGAACATGCACCTTTGTTTCTAACGTCGATGGG
ACTCACATCGCGGAAGTGCTGAAAAAAGTAAACCCGGAACACGCTGTTCTTGGTAGCATC
TAAAACCTTCACCACTCAGGAACTATGACCAACGCCCATAGCGCGCGTGACTGGTTCCTGA
AAGCGGCAGGTGATGAGAAGCACGTTGCAAAACACTTTGCGGCGCTTTCCACCAATGCCAAA
GCCGTTGGCGAGTTTGGTATTGATACTGCCAACATGTTTCGAGTTCTGGGACTGGGTTGGCGG
CCGTTACTCTTTTGGTGGTCAGCGATTGGCCTGTCGATTGTTCTCTCCATCGGCTTTGATAACT
TCGTTGAACTGCTTTCCGGCGCACACGCGATGGACAAGCATTTCTCCACCACGCTTCCGAG
AAAAACCTGCCTGTACTGCTGGCGCTGATTGGCATCTGGTACAACAATTTCTTTGGTGCGGA
AACTGAAGCGATTCTGCCGTATGACCAGTATATGCACCGTTTCGCGGCGTACTTCCAGCAGG
GCAATATGGAGTCCAACGGTAAGTATGTTGACCGTAACGGTAACGTTGTGGATTACCAGACT
GGCCCGATTATCTGGGGTGAACCAGGCACTAACGGTCAGCACGCGTTCTACCAGCTGATCCA
CCAGGGAACCAAAATGGTACCGTGCGATTTCATCGCTCCGGCTATCACCCATAACCCGCTCT
CTGATCATCACCAGAACTGCTGTCTAACTTCTTCGCCCAGACCGAAGCGCTGGCGTTTGGT
AAATCCCGCGAAGTGGTTGAGCAGGAATATCGTGATCAGGGTAAAGATCCGGCAACGCTTGA
CTACGTGGTGCCGTTCAAAGTATTCGAAGGTAACCGCCGACCAACTCCATCCTGCTGCGTG

AAATCACTCCGTTTCAGCCTGGGTGCGTTGATTGCGCTGTATGAGCACAAAATCTTTACTCAG
GGCGTGATCCTGAACATCTTCACCTTCGACCAGTGGGGCGTGGAAGTGGGTAAACAGCTGGC
GAACCGTATCTGCCAGAGCTGGAAGTGCTGTTTCAGGGTCCGAAAGATGATAAAGAAATCA
GCAGCCACGATAGCTCGACCAATGGTCTGATTAACCGCTATAAAGCGTGGCGCGGTtaa

SEQ ID NO:8

ATGAAAAACATCAATCCAACGCAGACCGCTGCCTGGCAGGCACTACAGAAACACTTCGATGA
 AATGAAAGACGTTACGATCGCCGATCTTTTTTGCTAAAGATGGTGATCGTTTTTCTAAGTTCT
 CCGCAACCTTCGACGATCAGATGCTGGTGGATTACTCCAAAAACCGCATCACTGAAGAGACG
 CTGGCGAAATTACAGGATCTGGCGAAAGAGTGCGATCTGGCGGGCGCGATTAAAGTCGATGTT
 CTCTGGCGAGAAGATCAACCGCACTGAAAACCGCGCCGTGCTGCACGTAGCGCTGCGTAACC
 GTAGCAATACCCCGATTTTGGTTGATGGCAAAGACGTAATGCCGGAAGTCAACGCGGTGCTG
 GAGAAGATGAAAACCTTCTCAGAAGCGATTATTTCCGGTGAGTGGAAAGGTTATACCGGCAA
 AGCAATCACTGACGTAGTGAACATCGGGATCGGCGGTTCTGACCTCGGCCCATACATGGTGA
 CCGAAGCTCTGCGTCCGTACAAAAACCACCTGAACATGCACTTTGTTTCTAACGTCGATGGG
 ACTCACATCGCGGAAGTGCTGAAAAAAGTAAACCCGGAACCACGCTGTTCTTGGTAGCATC
 TAAAACCTTCACCACTCAGGAACTATGACCAACGCCCATAGCGCGCGTGACTGGTTCCTGA
 AAGCGGCAGGTGATGAGAAGCACGTTGCAAAACACTTTGCGGCGCTTTCACCAATGCCAAA
 GCCGTTGGCGAGTTTGGTATTGATACTGCCAACATGTTTCGAGTTCTGGGACTGGGTGGCGG
 CCGTTACTCTTTGTGGTCAGCGATTGGCCTGTCGATTGTTCTCTCCATCGGCTTTGATAACT
 TCGTTGAACTGCTTTCCGGCGCACACGCGATGGACAAGCATTTCTCCACCACGCTGCCGAG
 AAAACCTGCCTGTACTGCTGGCGCTGATTGGCATCTGGTACAACAATTTCTTTGGTGCGGA
 AACTGAAGCGATTCTGCCGTATGACCAGTATATGCACCGTTTCGCGGCGTACTTCCAGCAGG
 GCAATATGGAGTCCAACGGTAAGTATGTTGACCGTAACGGTAACGTTGTGGATTACCAGACT
 GGCCCGATTATCTGGGGTGAACCAGGCACTAACGGTCAGCACGCGTTCTACCAGCTGATCCA
 CCAGGGAACCAAAATGGTACCGTGCGATTTTCATCGCTCCGGCTATCACCCATAACCCGCTCT
 CTGATCATCACCCAGAACTGCTGTCTAATTCTTCGCCCAGACCGAAGCGCTGGCGTTTGGT
 AAATCCCGCGAAGTGGTTGAGCAGGAATATCGTGATCAGGGTAAAGATCCGCGAACGCTTGA
 CTACGTGGTGCCGTTCAAAGTATTCGAAGGTAAACGCCCGACCAACTCCATCCTGCTGCGTG
 AAATCACTCCGTTTCAGCCTGGGTGCGTTGATTGCGCTGTATGAGCACAAAATCTTTACTCAG
 GGCGTGATCCTGAACATCTTCACCTTCGACCAGTGGGGCGTGGAAGTGGGTAAACAGCTGGC
 GAACCGTATTCTGCCAGAGCTGAAACTGGAAGTGCTGTTTCAGGGTCCGGATGATAAAGAAA
 TCAGCAGCCACGATAGCTCGACCAATGGTCTGATTAACCGCTATAAAGCGTGGCGCGGTtaa

5

SEQ ID NO:9

ATGAAAAACATCAATCCAACGCAGACCGCTGCCTGGCAGGCACTACAGAAACACTTCGATGA
 AATGAAAGACGTTACGATCGCCGATCTTTTTTGCTAAAGATGGTGATCGTTTTTCTAAGTTCT
 CCGCAACCTTCGACGATCAGATGCTGGTGGATTACTCCAAAAACCGCATCACTGAAGAGACG
 CTGGCGAAATTACAGGATCTGGCGAAAGAGTGCGATCTGGCGGGCGCGATTAAAGTCGATGTT
 CTCTGGCGAGAAGATCAACCGCACTGAAAACCGCGCCGTGCTGCACGTAGCGCTGCGTAACC
 GTAGCAATACCCCGATTTTGGTTGATGGCAAAGACGTAATGCCGGAAGTCAACGCGGTGCTG
 GAGAAGATGAAAACCTTCTCAGAAGCGATTATTTCCGGTGAGTGGAAAGGTTATACCGGCAA
 AGCAATCACTGACGTAGTGAACATCGGGATCGGCGGTTCTGACCTCGGCCCATACATGGTGA
 CCGAAGCTCTGCGTCCGTACAAAAACCACCTGAACATGCACTTTGTTTCTAACGTCGATGGG
 ACTCACATCGCGGAAGTGCTGAAAAAAGTAAACCCGGAACCACGCTGTTCTTGGTAGCATC
 TAAAACCTTCACCACTCAGGAACTATGACCAACGCCCATAGCGCGCGTGACTGGTTCCTGA
 AAGCGGCAGGTGATGAGAAGCACGTTGCAAAACACTTTGCGGCGCTTTCACCAATGCCAAA
 GCCGTTGGCGAGTTTGGTATTGATACTGCCAACATGTTTCGAGTTCTGGGACTGGGTGGCGG
 CCGTTACTCTTTGTGGTCAGCGATTGGCCTGTCGATTGTTCTCTCCATCGGCTTTGATAACT

TCGTTGAACTGCTTTCCGGCGCACACGCGATGGACAAGCATTCTCCACCACGCCTGCCGAG
 AAAAACCTGCCTGTACTGCTGGCGCTGATTGGCATCTGGTACAACAATTTCTTTGGTGCGGA
 AACTGAAGCGATTCTGCCGTATGACCAGTATATGCACCGTTTCGCGGCGTACTTCCAGCAGG
 GCAATATGGAGTCCAACGTAAGTATGTTGACCGTAACGGTAACGTTGTGGATTACCAGACT
 GGCCCATTATCTGGGGTGAACCAGGCACTAACGGTCAGCACGCGTTCTACCAGCTGATCCA
 CCAGGGAACCAAAATGGTACCGTGCGATTTTCATCGCTCCGGCTATCACCCATAACCCGCTCT
 CTGATCATCACCAGAACTGCTGTCTAACTTCTTCGCCCAGACCGAAGCGCTGGCGTTTGGT
 AAATCCCGCGAAGTGGTTGAGCAGGAATATCGTGATCAGGGTAAAGATCCGGCAACGCTTGA
 CTACGTGGTGCCGTTCAAAGTATTCTGAAGGTAACCGCCCGACCAACTCCATCCTGCTGCGTG
 AAATCACTCCGTTTCCAGCCTGGGTGCGTTGATTGCGCTGTATGAGCACAAAATCTTTACTCAG
 GGCGTGATCCTGAACATCTTCACCTTCGACCAGTGGGGCGTGGAAGTGGGTAAACAGCTGGC
 GAACCGTATTCTGCCAGAGCTGAAAGATCTGGAAGTGCTGTTTCAGGGTCCGGATAAAGAAA
 TCAGCAGCCACGATAGCTCGACCAATGGTCTGATTAAACCGCTATAAAGCGTGGCGCGGTtaa

SEQ ID NO:10

ATGAAAAACATCAATCCAACGCAGACCGCTGCCTGGCAGGCACTACAGAAACACTTCGATGA
 AATGAAAGACGTTACGATCGCCGATCTTTTTGCTAAAGATGGTGATCGTTTTTCTAAGTTCT
 CCGCAACCTTCGACGATCAGATGCTGGTGGATTACTCCAAAAACCGCATCACTGAAGAGACG
 CTGGCGAAATTACAGGATCTGGCGAAAGAGTGCGATCTGGCGGGCGCGATTAAAGTCGATGTT
 CTCTGGCGAGAAGATCAACCGCACTGAAAACCGCGCCGTGCTGCACGTAGCGCTGCGTAACC
 GTAGCAATACCCGATTTTGGTTGATGGCAAAGACGTAATGCCGGAAGTCAACGCGGTGCTG
 GAGAAGATGAAAACCTTCTCAGAAGCGATTATTTCCGGTGAGTGGAAGGTTATACCGGCAA
 AGCAATCACTGACGTAGTGAACATCGGGATCGGCGGTTCTGACCTCGGCCCATAACATGGTGA
 CCGAAGCTCTGCGTCCGTACAAAAACCACCTGAACATGCACTTTGTTTCTAACGTCGATGGG
 ACTCACATCGCGGAAGTGCTGAAAAAAGTAAACCCGGAAACCACGCTGTTCTTGGTAGCATC
 TAAAACCTTCACCACTCAGGAAACTATGACCAACGCCCATAGCGCGCGTGACTGGTTCCTGA
 AAGCGGCAGGTGATGAGAAGCACGTTGCAAAACACTTTGCGGCGCTTTCACCAATGCCAAA
 GCCGTTGGCGAGTTTGGTATTGATACTGCCAACATGTTTCGAGTTCTGGGACTGGGTGGCGG
 CCGTTACTCTTTGTGGTCAGCGATTGGCCTGTCGATTGTTCTCTCCATCGGCTTTGATAACT
 TCGTTGAACTGCTTTCCGGCGCACACGCGATGGACAAGCATTCTCCACCACGCCTGCCGAG
 AAAAACCTGCCTGTACTGCTGGCGCTGATTGGCATCTGGTACAACAATTTCTTTGGTGCGGA
 AACTGAAGCGATTCTGCCGTATGACCAGTATATGCACCGTTTCGCGGCGTACTTCCAGCAGG
 GCAATATGGAGTCCAACGTAAGTATGTTGACCGTAACGGTAACGTTGTGGATTACCAGACT
 GGCCCATTATCTGGGGTGAACCAGGCACTAACGGTCAGCACGCGTTCTACCAGCTGATCCA
 CCAGGGAACCAAAATGGTACCGTGCGATTTTCATCGCTCCGGCTATCACCCATAACCCGCTCT
 CTGATCATCACCAGAACTGCTGTCTAACTTCTTCGCCCAGACCGAAGCGCTGGCGTTTGGT
 AAATCCCGCGAAGTGGTTGAGCAGGAATATCGTGATCAGGGTAAAGATCCGGCAACGCTTGA
 CTACGTGGTGCCGTTCAAAGTATTCTGAAGGTAACCGCCCGACCAACTCCATCCTGCTGCGTG
 AAATCACTCCGTTTCCAGCCTGGGTGCGTTGATTGCGCTGTATGAGCACAAAATCTTTACTCAG
 GGCGTGATCCTGAACATCTTCACCTTCGACCAGTGGGGCGTGGAAGTGGGTAAACAGCTGGC
 GAACCGTATTCTGCCAGAGCTGAAAGATGATCTGGAAGTGCTGTTTCAGGGTCCGAAAGAAA
 TCAGCAGCCACGATAGCTCGACCAATGGTCTGATTAAACCGCTATAAAGCGTGGCGCGGTtaa

ES 2 704 697 T3

SEQ ID NO:11

ATGAAAAACATCAATCCAACGCAGACCGCTGCCTGGCAGGCACTACAGAAACACTTCGATGA
AATGAAAGACGTTACGATCGCCGATCTTTTTTGCTAAAGATGGTGATCGTTTTTCTAAGTTCT
CCGCAACCTTCGACGATCAGATGCTGGTGGATTACTCCAAAAACCGCATCACTGAAGAGACG
CTGGCGAAATTACAGGATCTGGCGAAAGAGTGCGATCTGGCGGGCGCGATTAAAGTCGATGTT
CTCTGGCGAGAAGATCAACCGCACTGAAAACCGCGCCGTGCTGCACGTAGCGCTGCGTAACC

GTAGCAATACCCCGATTTTGGTTGATGGCAAAGACGTAATGCCGGAAGTCAACGCGGTGCTG
GAGAAGATGAAAACCTTCTCAGAAGCGATTATTTCCGGTGAGTGGAAAGGTTATACCGGCAA
AGCAATCACTGACGTAGTGAACATCGGGATCGGCGGTTCTGACCTCGGCCCATACATGGTGA
CCGAAGCTCTGCGTCCGTACAAAAACCACCTGAACATGCACTTTGTTTCTAACGTCGATGGG
ACTCACATCGCGGAAGTGCTGAAAAAAGTAAACCCGGAACACGCTGTTCTTGGTAGCATC
TAAACCTTCACCACTCAGGAACTATGACCAACGCCCATAGCGCGCGTGACTGGTTCCTGA
AAGCGGCAGGTGATGAGAAGCACGTTGCAAAACACTTTGCGGCGCTTCCACCAATGCCAAA
GCCGTTGGCGAGTTTGGTATTGATACTGCCAACATGTTTCGAGTTCTGGGACTGGGTTGGCGG
CCGTTACTCTTTGTGGTCAGCGATTGGCCTGTCGATTGTTCTCTCCATCGGCTTTGATAACT
TCGTTGAACTGCTTTCCGGCGCACACGCGATGGACAAGCATTCTCCACCACGCTGCCGAG
AAAAACCTGCCTGTACTGCTGGCGCTGATTGGCATCTGGTACAACAATTTCTTTGGTGCGGA
AACTGAAGCGATTCTGCCGTATGACCAGTATATGCACCGTTTTCGCGGCGTACTTCCAGCAGG
GCAATATGGAGTCCAACGGTAAGTATGTTGACCGTAACGGTAACGTTGTGGATTACCAGACT
GGCCCCGATTATCTGGGGTGAACCAGGCACCTAACGGTCAGCACGCGTTCTACCAGCTGATCCA
CCAGGGAACCAAAATGGTACCGTGCGATTTCATCGCTCCGGCTATCACCATAACCCGCTCT
CTGATCATCACCAGAACTGCTGTCTAACTTCTTCGCCCAGACCGAAGCGCTGGCGTTTGGT
AAATCCCGCGAAGTGGTTGAGCAGGAATATCGTGATCAGGGTAAAGATCCGGCAACGCTTGA
CTACGTGGTGCCGTTCAAAGTATTGGAAGGTAACCGCCCGACCAACTCCATCCTGCTGCGTG
AAATCACTCCGTTTCAGCCTGGGTGCGTTGATTGCGCTGTATGAGCACAAAATCTTTACTCAG
GGCGTGATCCTGAACATCTTCACCTTCGACCAGTGGGGCGTGGAAGTGGGTAAACAGCTGGC
GAACCGTATTCTGCCAGAGCTGAAAGATGATAAACTGGAAGTGCTGTTTCAGGGTCCGGAAA
TCAGCAGCCACGATAGCTCGACCAATGGTCTGATTAACCGCTATAAAGCGTGGCGCGGTtaa

SEQ ID NO:12

5

ATGAAAAACATCAATCCAACGCAGACCGCTGCCTGGCAGGCACTACAGAAACACTTCGATGA
AATGAAAGACGTTACGATCGCCGATCTTTTTTGCTAAAGATGGTGATCGTTTTTCTAAGTTCT
CCGCAACCTTCGACGATCAGATGCTGGTGGATTACTCCAAAAACCGCATCACTGAAGAGACG
CTGGCGAAATTACAGGATCTGGCGAAAGAGTGCGATCTGGCGGGCGCGATTAAAGTCGATGTT
CTCTGGCGAGAAGATCAACCGCACTGAAAACCGCGCCGTGCTGCACGTAGCGCTGCGTAACC
GTAGCAATACCCCGATTTTGGTTGATGGCAAAGACGTAATGCCGGAAGTCAACGCGGTGCTG
GAGAAGATGAAAACCTTCTCAGAAGCGATTATTTCCGGTGAGTGGAAAGGTTATACCGGCAA
AGCAATCACTGACGTAGTGAACATCGGGATCGGCGGTTCTGACCTCGGCCCATACATGGTGA
CCGAAGCTCTGCGTCCGTACAAAAACCACCTGAACATGCACTTTGTTTCTAACGTCGATGGG
ACTCACATCGCGGAAGTGCTGAAAAAAGTAAACCCGGAACACGCTGTTCTTGGTAGCATC
TAAACCTTCACCACTCAGGAACTATGACCAACGCCCATAGCGCGCGTGACTGGTTCCTGA
AAGCGGCAGGTGATGAGAAGCACGTTGCAAAACACTTTGCGGCGCTTTCACCAATGCCAAA
GCCGTTGGCGAGTTTGGTATTGATACTGCCAACATGTTTCGAGTTCTGGGACTGGGTTGGCGG
CCGTTACTCTTTGTGGTCAGCGATTGGCCTGTCGATTGTTCTCTCCATCGGCTTTGATAACT
TCGTTGAACTGCTTTCCGGCGCACACGCGATGGACAAGCATTCTCCACCACGCTGCCGAG
AAAAACCTGCCTGTACTGCTGGCGCTGATTGGCATCTGGTACAACAATTTCTTTGGTGCGGA
AACTGAAGCGATTCTGCCGTATGACCAGTATATGCACCGTTTTCGCGGCGTACTTCCAGCAGG
GCAATATGGAGTCCAACGGTAAGTATGTTGACCGTAACGGTAACGTTGTGGATTACCAGACT
GGCCCCGATTATCTGGGGTGAACCAGGCACCTAACGGTCAGCACGCGTTCTACCAGCTGATCCA
CCAGGGAACCAAAATGGTACCGTGCGATTTCATCGCTCCGGCTATCACCATAACCCGCTCT
CTGATCATCACCAGAACTGCTGTCTAACTTCTTCGCCCAGACCGAAGCGTGCGGCTTTGGT
AAATCCCGCGAAGTGGTTGAGCAGGAATATCGTGATCAGGGTAAAGATCCGGCAACGCTTGA
CTACGTGGTGCCGTTCAAAGTATTGGAAGGTAACCGCCCGACCAACTCCATCCTGCTGCGTG
AAATCACTCCGTTTCAGCCTGGGTGCGTTGATTGCGCTGTATGAGCACAAAATCTTTACTCAG
GGCGTGATCCTGAACATCTTCACCTTCGACCAGTGGGGCGTGGAAGTGGGTAAACAGCTGGC

GAACCGTATTCTGCCAGAGCTGAAAGATGATAAAGAACTGGAAGTGCTGTTTCAGGGTCCGA
TCAGCAGCCACGATAGCTCGACCAATGGTCTGATTAACCGCTATAAAGCGTGGCGCGGTtaa

SEQ ID NO:13

ATGAAAAACATCAATCCAACGCAGACCGCTGCCTGGCAGGCACTACAGAAACACTTCGATGA
 AATGAAAGACGTTACGATCGCCGATCTTTTTTGCTAAAGATGGTGATCGTTTTTCTAAGTTCT
 CCGCAACCTTCGACGATCAGATGCTGGTGGATTACTCCAAAAACCGCATCACTGAAGAGACG
 CTGGCGAAATTACAGGATCTGGCGAAAGAGTGCGATCTGGCGGGCGCGATTAAAGTCGATGTT
 CTCTGGCGAGAAGATCAACCGCACTGAAAACCGCGCCGTGCTGCACGTAGCGCTGCGTAACC
 GTAGCAATACCCCGATTTTGGTTGATGGCAAAGACGTAATGCCGGAAGTCAACGCGGTGCTG
 GAGAAGATGAAAACCTTCTCAGAAGCGATTATTTCCGGTGAGTGGAAGGTTATACCGGCAA
 AGCAATCACTGACGTAGTGAACATCGGGATCGGCGGTTCTGACCTCGGCCCATAACATGGTGA
 CCGAAGCTCTGCGTCCGTACAAAAACCACCTGAACATGCACTTTGTTTCTAACGTCGATGGG
 ACTCACATCGCGGAAGTGCTGAAAAAAGTAAACCCGGAAACCACGCTGTTCTTGGTAGCATC
 TAAAACCTTCACCACTCAGGAACTATGACCAACGCCCATAGCGCGCGTGACTGGTTCCTGA
 AAGCGGCAGGTGATGAGAAGCACGTTGCAAAACACTTTGCGGCGCTTTCCACCAATGCCAAA
 GCCGTTGGCGAGTTTGGTATTGATACTGCCAACATGTTTCGAGTTCTGGGACTGGGTGGCGG
 CCGTTACTCTTTGTGGTCAGCGATTGGCCTGTCGATTGTTCTCTCCATCGGCTTTGATAACT
 TCGTTGAACTGCTTTCCGGCGCACACGCGATGGACAAGCATTTCTCCACCACGCCTGCCGAG
 AAAACCTGCCTGTACTGCTGGCGCTGATTGGCATCTGGTACAACAATTTCTTTGGTGCGGA
 AACTGAAGCGATTCTGCCGTATGACCAGTATATGCACCGTTTCGCGGCGTACTTCCAGCAGG
 GCAATATGGAGTCCAACGTAAGTATGTTGACCGTAACGGTAACGTTGTGGATTACCAGACT
 GGCCCGATTATCTGGGGTGAACCAGGCACTAACGGTCAGCACGCGTTCTACCAGCTGATCCA
 CCAGGGAACCAAAATGGTACCGTGCGATTTCATCGCTCCGGCTATCACCCATAACCCGCTCT
 CTGATCATCACCAGAACTGCTGTCTAACTTCTTCGCCCAGACCGAAGCGCTGGCGTTTGGT
 AAATCCCGCGAAGTGGTTGAGCAGGAATATCGTGATCAGGGTAAAGATCCGGCAACGCTTGA
 CTACGTGGTGCCGTTCAAAGTATTCGAAGGTAACCGCCCGACCAACTCCATCCTGCTGCGTG
 AAATCACTCCGTTTCAGCCTGGGTGCGTTGATTGCGCTGTATGAGCACAAAATCTTTACTCAG
 GGCGTGATCCTGAACATCTTCACCTTCGACCAGTGGGGCGTGGAAGTGGGTAAACAGCTGGC
 GAACCGTATTTCTGCCAGAGCTGAAAGATGATAAAGAAATCCTGGAAGTGCTGTTTCAGGGTC
 CGAGCAGCCACGATAGCTCGACCAATGGTCTGATTAACCGCTATAAAGCGTGCGCGGTTaa

5

SEQ ID NO:14

ATGAAAAACATCAATCCAACGCAGACCGCTGCCTGGCAGGCACTACAGAAACACTTCGATGA
 AATGAAAGACGTTACGATCGCCGATCTTTTTTGCTAAAGATGGTGATCGTTTTTCTAAGTTCT
 CCGCAACCTTCGACGATCAGATGCTGGTGGATTACTCCAAAAACCGCATCACTGAAGAGACG
 CTGGCGAAATTACAGGATCTGGCGAAAGAGTGCGATCTGGCGGGCGCGATTAAAGTCGATGTT
 CTCTGGCGAGAAGATCAACCGCACTGAAAACCGCGCCGTGCTGCACGTAGCGCTGCGTAACC
 GTAGCAATACCCCGATTTTGGTTGATGGCAAAGACGTAATGCCGGAAGTCAACGCGGTGCTG
 GAGAAGATGAAAACCTTCTCAGAAGCGATTATTTCCGGTGAGTGGAAGGTTATACCGGCAA
 AGCAATCACTGACGTAGTGAACATCGGGATCGGCGGTTCTGACCTCGGCCCATAACATGGTGA
 CCGAAGCTCTGCGTCCGTACAAAAACCACCTGAACATGCACTTTGTTTCTAACGTCGATGGG
 ACTCACATCGCGGAAGTGCTGAAAAAAGTAAACCCGGAAACCACGCTGTTCTTGGTAGCATC
 TAAAACCTTCACCACTCAGGAACTATGACCAACGCCCATAGCGCGCGTGACTGGTTCCTGA
 AAGCGGCAGGTGATGAGAAGCACGTTGCAAAACACTTTGCGGCGCTTTCCACCAATGCCAAA
 GCCGTTGGCGAGTTTGGTATTGATACTGCCAACATGTTTCGAGTTCTGGGACTGGGTGGCGG
 CCGTTACTCTTTGTGGTCAGCGATTGGCCTGTCGATTGTTCTCTCCATCGGCTTTGATAACT
 TCGTTGAACTGCTTTCCGGCGCACACGCGATGGACAAGCATTTCTCCACCACGCCTGCCGAG
 AAAACCTGCCTGTACTGCTGGCGCTGATTGGCATCTGGTACAACAATTTCTTTGGTGCGGA

AACTGAAGCGATTCTGCCGTATGACCAGTATATGCACCGTTTCGCGGCGTACTTCCAGCAGG
GCAATATGGAGTCCAACGGTAAGTATGTTGACCGTAACGGTAACGTTGTGGATTACCAGACT
GGCCCGATTATCTGGGGTGAACCAGGCACTAACGGTCAGCACGCGTTCTACCAGCTGATCCA
CCAGGGAACCAAAATGGTACCGTGCGATTTTCATCGCTCCGGCTATCACCCATAACCCGCTCT
CTGATCATCACCAGAACTGCTGTCTAACTTCTTCGCCCAGACCGAAGCGCTGGCGTTTGGT
AAATCCCGCGAAGTGGTTGAGCAGGAATATCGTGATCAGGGTAAAGATCCGGCAACGCTTGA
CTACGTGGTGCCGTTCAAAGTATTCTGAAGGTAACCGCCCGACCAACTCCATCCTGCTGCGTG
AAATCACTCCGTTTCAGCCTGGGTGCGTTGATTGCGCTGTATGAGCACAAAATCTTTACTCAG
GGCGTGATCCTGAACATCTTCACCTTCGACCAGTGGGGCGTGGAAGTGGGTAAACAGCTGGC
GAACCGTATTCTGCCAGAGCTGAAAGATGATAAAGAAATCAGCCTGGAAGTGCTGTTTCAGG
GTCCGAGCCACGATAGCTCGACCAATGGTCTGATTAACCGCTATAAAGCGTGGCGCGGTtaa

SEQ ID NO:15

ATGAAAAACATCAATCCAACGCAGACCGCTGCCTGGCAGGCACTACAGAAACACTTCGATGA
AATGAAAGACGTTACGATCGCCGATCTTTTTGCTAAAGATGGTGATCGTTTTTCTAAGTTCT
CCGCAACCTTCGACGATCAGATGCTGGTGGATTACTCCAAAAACCGCATCACTGAAGAGACG
CTGGCGAAATTACAGGATCTGGCGAAAGAGTGCGATCTGGCGGGCGCGATTAAAGTCGATGTT
CTCTGGCGAGAAGATCAACCGCACTGAAAACCGCGCCGTGCTGCACGTAGCGCTGCGTAACC
GTAGCAATACCCCGATTTTGGTTGATGGCAAAGACGTAATGCCGGAAGTCAACGCGGTGCTG
GAGAAGATGAAAACCTTCTCAGAAGCGATTATTTCCGGTGAGTGGAAGGTTATACCGGCAA
AGCAATCACTGACGTAGTGAACATCGGGATCGGCGGTTCTGACCTCGGCCCATACATGGTGA
CCGAAGCTCTGCGTCCGTACAAAAACCACCTGAACATGCACCTTTGTTTCTAACGTCGATGGG
ACTCACATCGCGGAAGTGCTGAAAAAAGTAAACCCGGAAACCACGCTGTTCTTGGTAGCATC
TAAAACCTTCACCACTCAGGAACTATGACCAACGCCCATAGCGCGCGTGACTGGTTCCTGA
AAGCGGCAGGTGATGAGAAGCACGTTGCAAAACACTTTGCGGCGCTTTCACCAATGCCAAA
GCCGTTGGCGAGTTTGGTATTGATACTGCCAACATGTTTCGAGTTCTGGGACTGGGTGGCGG
CCGTTACTCTTTGTGGTCAGCGATTGGCCTGTCGATTGTTCTCTCCATCGGCTTTGATAACT
TCGTTGAACTGCTTTCCGGCGCACACGCGATGGACAAGCATTTCTCCACCACGCTGCCGAG
AAAAACCTGCCTGTACTGCTGGCGCTGATTGGCATCTGGTACAACAATTTCTTTGGTGCGGA
AACTGAAGCGATTCTGCCGTATGACCAGTATATGCACCGTTTCGCGGCGTACTTCCAGCAGG
GCAATATGGAGTCCAACGGTAAGTATGTTGACCGTAACGGTAACGTTGTGGATTACCAGACT
GGCCCGATTATCTGGGGTGAACCAGGCACTAACGGTCAGCACGCGTTCTACCAGCTGATCCA
CCAGGGAACCAAAATGGTACCGTGCGATTTTCATCGCTCCGGCTATCACCCATAACCCGCTCT
CTGATCATCACCAGAACTGCTGTCTAACTTCTTCGCCCAGACCGAAGCGCTGGCGTTTGGT
AAATCCCGCGAAGTGGTTGAGCAGGAATATCGTGATCAGGGTAAAGATCCGGCAACGCTTGA
CTACGTGGTGCCGTTCAAAGTATTCTGAAGGTAACCGCCCGACCAACTCCATCCTGCTGCGTG
AAATCACTCCGTTTCAGCCTGGGTGCGTTGATTGCGCTGTATGAGCACAAAATCTTTACTCAG
GGCGTGATCCTGAACATCTTCACCTTCGACCAGTGGGGCGTGGAAGTGGGTAAACAGCTGGC
GAACCGTATTCTGCCAGAGCTGAAAGATGATAAAGAAATCAGCAGCCTGGAAGTGCTGTTTC
AGGGTCCGCACGATAGCTCGACCAATGGTCTGATTAACCGCTATAAAGCGTGGCGCGGTtaa

SEQ ID NO:16

ATGAAAAACATCAATCCAACGCAGACCGCTGCCTGGCAGGCACTACAGAAACACTTCGATGA
AATGAAAGACGTTACGATCGCCGATCTTTTTGCTAAAGATGGTGATCGTTTTTCTAAGTTCT
CCGCAACCTTCGACGATCAGATGCTGGTGGATTACTCCAAAAACCGCATCACTGAAGAGACG
CTGGCGAAAATTACAGGATCTGGCGAAAGAGTGCGATCTGGCGGGCGCGATTAAAGTCGATGTT
CTCTGGCGAGAAGATCAACCGCACTGAAAACCGCGCCGTGCTGCACGTAGCGCTGCGTAACC
GTAGCAATACCCGATTTTGGTTGATGGCAAAGACGTAATGCCGGAAGTCAACGCGGTGCTG
GAGAAGATGAAAACCTTCTCAGAAGCGATTATTTCCGGTGAGTGGAAGGTTATACCGGCAA

AGCAATCACTGACGTAGTGAACATCGGGATCGGCGGTTCTGACCTCGGCCCATACATGGTGA
CCGAAGCTCTGCGTCCGTACAAAAACCACCTGAACATGCACTTTGTTTCTAACGTCGATGGG
ACTCACATCGCGGAAGTGCTGAAAAAAGTAAACCCGGAAACCACGCTGTTCTTGGTAGCATC
TAAAACCTTCACCACTCAGGAACTATGACCAACGCCCATAGCGCGCGTGACTGGTTCCTGA
AAGCGGCAGGTGATGAGAAGCACGTTGCAAAACACTTTGCGGCGCTTTCCACCAATGCCAAA
GCCGTTGGCGAGTTTGGTATTGATACTGCCAACATGTTTCGAGTTCTGGGACTGGGTTGGCGG
CCGTTACTCTTTGTGGTCAGCGATTGGCCTGTCGATTGTTCTCTCCATCGGCTTTGATAACT
TCGTTGAACGCTTTCCGGCGCACACGCGATGGACAAGCATTTCTCCACCACGCCTGCCGAG
AAAAACCTGCCTGTACTGCTGGCGCTGATTGGCATCTGGTACAACAATTTCTTTGGTGCGGA
AACTGAAGCGATTCTGCCGTATGACCAGTATATGCACCGTTTCGCGGCGTACTTCCAGCAGG
GCAATATGGAGTCCAACGGTAAGTATGTTGACCGTAACGGTAACGTTGTGGATTACCAGACT
GGCCCGATTATCTGGGGTGAACCAGGCACTAACGGTCAGCACGCGTTCTACCAGCTGATCCA
CCAGGGAACCAAAATGGTACCGTGCGATTTTCATCGCTCCGGCTATCACCCATAACCCGCTCT
CTGATCATCACCAGAACTGCTGTCTAACTTCTTCGCCCAGACCGAAGCGCTGGCGTTTGGT
AAATCCCGCGAAGTGTTGAGCAGGAATATCGTGATCAGGGTAAAGATCCGGCAACGCTTGA
CTACGTGGTGCCGTTCAAAGTATTCGAAGGTAACCGCCCGACCAACTCCATCCTGCTGCGTG
AAATCACTCCGTTTCAGCCTGGGTGCGTTGATTGCGCTGTATGAGCACAAAATCTTTACTCAG
GGCGTGATCCTGAACATCTTCACCTTCGACCAGTGGGGCGTGGAACCTGGGTAAACAGCTGGC
GAACCGTATTCTGCCAGAGCTGAAAGATGATAAAGAAATCAGCAGCCACGATAGCTCGACCA
ATGGTCTGATTAACCGCTATAAACTGGAAGTGCTGTTTCAGGGTCCGGCGTGCGCGGTTaa

5

SEQ ID NO:17

MKNINPTQTAAWQALQKHFDKMDVTIADLFAKDGDGRFSKFSATFDDQMLVDYSKNRITEET
LAKLQDLAKECDLAGAIKSMFSGEKINRTENRAVLHVALRNRNTPILVDGKDVMPEVNAVL
EKMKTFSIAIISGEWKGYTGKAITDVVNIGIGGSDLGPYMVTEALRPYKNHLNMFVSNVDG
THIAEVLKKVNPETTLFLVASKTFTTQETMTNAHSARDWFLKAAGDEKHVAKHFAALSTNAK
AVGEFGIDTANMFEFWDWVGGRYSLWSAIGLSIVLSIGFDNFVELLSGAHAMDKHFSTTPAE
KNLPVLLALIGIWYNNFFGAETEAILPYDQYMHRFAAYFQQGNMESNGKYVDRNGNVVDYQT
GPIIWGEPGTNGQHAFYQLIHQGTGMVPCDFIAPAIITHNPLSDHHQKLLSNFFAQTEALAFG
KSREVVEQEYRDQKDPATLDYVVPFKVFEGNRPTNSILLREITPFSLGALIALYEHKIFTQ
GVILNIFTFDQWGVELGKQLANRILPELKDDKEISSHDSSTNGLINRYKAWRG

SEQ ID NO:18

MKNINPTQTAAWQALQKHFDKMDVTIADLFAKDGDGRFSKFSATFDDQMLVDYSKNRITEET
LAKLQDLAKECDLAGAIKSMFSGEIKINRTENRAVLHVALRNRSNTPLEVLFGQPIILVDGKDV
MPEVNAVLEKMKTFSEAIISGEWKGYTGKAITDVVNIGIGGSDLGPYMVTEALRPYKNHLNM
HFVSNVDGTHIAEVLKKVNPETTLFLVASKTFTTQETMTNAHSARDWFLKAAGDEKHAVAKHF
AALSTNAKAVGEFGIDTANMFEFWDWVGGRYSLWSAIGLSIVLSIGFDNFVELLSGAHAMDK
HFSTTPAEKNLPVLLALIGIWYNNFFGAETEAILPYDQYMHRFAAYFQQGNMESNGKYVDRN
GNVVDYQTGPPIIWGEPGTNGQHAFYQLIHQGTKMVPCDFIAPAITHNPLSDHHQKLLSNFFA
QTEALAFGKSREVVEQEYRDQGKDPATLDYVVPFKVFEGNRPTNSILLREITPFSLGALIAL
YEHKIFTQGVILNIFTFDQWGVLELKGQLANRILPELKDDKEISSHDSSTNGLINRYKAWRG

5 SEQ ID NO:19

MKNINPTQTAAWQALQKHFDKMDVTIADLFAKDGDGRFSKFSATFDDQMLVDYSKNRITEET
LAKLQDLAKECDLAGAIKSMFSGEIKINRTENRAVLHVALRNRSNTPILEVLFGQPIILVDGKDV
MPEVNAVLEKMKTFSEAIISGEWKGYTGKAITDVVNIGIGGSDLGPYMVTEALRPYKNHLNM
HFVSNVDGTHIAEVLKKVNPETTLFLVASKTFTTQETMTNAHSARDWFLKAAGDEKHAVAKHF

AALSTNAKAVGEFGIDTANMFEFWDWVGGRYSLWSAIGLSIVLSIGFDNFVELLSGAHAMDK
HFSTTPAEKNLPVLLALIGIWYNNFFGAETEAILPYDQYMHRFAAYFQQGNMESNGKYVDRN
GNVVDYQTGPPIIWGEPGTNGQHAFYQLIHQGTKMVPCDFIAPAITHNPLSDHHQKLLSNFFA
QTEALAFGKSREVVEQEYRDQGKDPATLDYVVPFKVFEGNRPTNSILLREITPFSLGALIAL
YEHKIFTQGVILNIFTFDQWGVLELKGQLANRILPELKDDKEISSHDSSTNGLINRYKAWRG

10 SEQ ID NO:20

MKNINPTQTAAWQALQKHFDKMDVTIADLFAKDGDGRFSKFSATFDDQMLVDYSKNRITEET
LAKLQDLAKECDLAGAIKSMFSGEIKINRTENRAVLHVALRNRSNTPILEVLFGQPIILVDGKDV
PEVNAVLEKMKTFSEAIISGEWKGYTGKAITDVVNIGIGGSDLGPYMVTEALRPYKNHLNM
FVSNVDGTHIAEVLKKVNPETTLFLVASKTFTTQETMTNAHSARDWFLKAAGDEKHAVAKHFA
ALSTNAKAVGEFGIDTANMFEFWDWVGGRYSLWSAIGLSIVLSIGFDNFVELLSGAHAMDKH
FSTTPAEKNLPVLLALIGIWYNNFFGAETEAILPYDQYMHRFAAYFQQGNMESNGKYVDRNG
NVVDYQTGPPIIWGEPGTNGQHAFYQLIHQGTKMVPCDFIAPAITHNPLSDHHQKLLSNFFAQ
TEALAFGKSREVVEQEYRDQGKDPATLDYVVPFKVFEGNRPTNSILLREITPFSLGALIALY
EHKIFTQGVILNIFTFDQWGVLELKGQLANRILPELKDDKEISSHDSSTNGLINRYKAWRG

SEQ ID NO:21

MKNINPTQTAAWQALQKHFDKMDVTIADLFAKDGDGRFSKFSATFDDQMLVDYSKNRITEET
LAKLQDLAKECDLAGAIKSMFSGEIKINRTENRAVLHVALRNRSNTPILVDGKDVMPVNAVLE
EKMKTTFSEAIISGELEVLFGQPWKGYTGKAITDVVNIGIGGSDLGPYMVTEALRPYKNHLNM
HFVSNVDGTHIAEVLKKVNPETTLFLVASKTFTTQETMTNAHSARDWFLKAAGDEKHAVAKHF
AALSTNAKAVGEFGIDTANMFEFWDWVGGRYSLWSAIGLSIVLSIGFDNFVELLSGAHAMDK
HFSTTPAEKNLPVLLALIGIWYNNFFGAETEAILPYDQYMHRFAAYFQQGNMESNGKYVDRN
GNVVDYQTGPPIIWGEPGTNGQHAFYQLIHQGTKMVPCDFIAPAITHNPLSDHHQKLLSNFFA
QTEALAFGKSREVVEQEYRDQGKDPATLDYVVPFKVFEGNRPTNSILLREITPFSLGALIAL
YEHKIFTQGVILNIFTFDQWGVLELKGQLANRILPELKDDKEISSHDSSTNGLINRYKAWRG

15

SEQ ID NO:22

MKNINPTQTAAWQALQKHFDKMDVTIADLFAKDGDRFSKFSATFDDQMLVDYSKNRITEET
LAKLQDLAKECDLAGAIKSMFSGEKNRTENRAVLHVALRNRSTPILVDGKDVMPVNAVL
EKMKTFSIAIISGEWKGYTGKAITDVVNIGIGGSDLGPYMTALRPNHNLNMFVSNVDG
THIAEVLKKVNPETTLFLVASKTFTTQETMTNAHSARDWFLKAAGDEKHVAKHFAALSTNAK
AVGEFGIDTANMFEFWDWVGGRYSLWSAIGLSIVLSIGFDNFVELLSGAHAMDKHFSTTPAE
KNLPVLLALIGIWYNNFFGAETEAILPYDQYMRFAAYFQQGNMESNGKYVDRNGNVVDYQT
GPIIWGEPGTNGQHAFYQLIHQGTKMVPCDFIAPAITHLEVLFQGNPLSDHHQKLLSNFFA
QTEALAFGKSREVVEQEYRDQGKDPATLDYVVPFKVFEGNRPTNSILLREITPFSLGALIAL
YEHKIFTQGVILNIFTFDQWGVELGKQLANRILPELKDDKEISSHDSSTNGLINRYKAWRG

5 SEQ ID NO:23

MKNINPTQTAAWQALQKHFDKMDVTIADLFAKDGDRFSKFSATFDDQMLVDYSKNRITEET
LAKLQDLAKECDLAGAIKSMFSGEKNRTENRAVLHVALRNRSTPILVDGKDVMPVNAVL
EKMKTFSIAIISGEWKGYTGKAITDVVNIGIGGSDLGPYMTALRPNHNLNMFVSNVDG
THIAEVLKKVNPETTLFLVASKTFTTQETMTNAHSARDWFLKAAGDEKHVAKHFAALSTNAK
AVGEFGIDTANMFEFWDWVGGRYSLWSAIGLSIVLSIGFDNFVELLSGAHAMDKHFSTTPAE
KNLPVLLALIGIWYNNFFGAETEAILPYDQYMRFAAYFQQGNMESNGKYVDRNGNVVDYQT
GPIIWGEPGTNGQHAFYQLIHQGTKMVPCDFIAPAITHNPLSDHHQKLLSNFFAQTEALAFG

KSREVVEQEYRDQGKDPATLDYVVPFKVFEGNRPTNSILLREITPFSLGALIALYEHKIFTQ
GVILNIFTFDQWGVELGKQLANRILPELEVLFQGNPKDDKEISSHDSSTNGLINRYKAWRG

SEQ ID NO:24

10

MKNINPTQTAAWQALQKHFDKMDVTIADLFAKDGDRFSKFSATFDDQMLVDYSKNRITEET
LAKLQDLAKECDLAGAIKSMFSGEKNRTENRAVLHVALRNRSTPILVDGKDVMPVNAVL
EKMKTFSIAIISGEWKGYTGKAITDVVNIGIGGSDLGPYMTALRPNHNLNMFVSNVDG
THIAEVLKKVNPETTLFLVASKTFTTQETMTNAHSARDWFLKAAGDEKHVAKHFAALSTNAK
AVGEFGIDTANMFEFWDWVGGRYSLWSAIGLSIVLSIGFDNFVELLSGAHAMDKHFSTTPAE
KNLPVLLALIGIWYNNFFGAETEAILPYDQYMRFAAYFQQGNMESNGKYVDRNGNVVDYQT
GPIIWGEPGTNGQHAFYQLIHQGTKMVPCDFIAPAITHNPLSDHHQKLLSNFFAQTEALAFG
KSREVVEQEYRDQGKDPATLDYVVPFKVFEGNRPTNSILLREITPFSLGALIALYEHKIFTQ
GVILNIFTFDQWGVELGKQLANRILPELKLEVLFQGNPKDDKEISSHDSSTNGLINRYKAWRG

SEQ ID NO:25

MKNINPTQTAAWQALQKHFDKMDVTIADLFAKDGDRFSKFSATFDDQMLVDYSKNRITEET
LAKLQDLAKECDLAGAIKSMFSGEKNRTENRAVLHVALRNRSTPILVDGKDVMPVNAVL
EKMKTFSIAIISGEWKGYTGKAITDVVNIGIGGSDLGPYMTALRPNHNLNMFVSNVDG
THIAEVLKKVNPETTLFLVASKTFTTQETMTNAHSARDWFLKAAGDEKHVAKHFAALSTNAK
AVGEFGIDTANMFEFWDWVGGRYSLWSAIGLSIVLSIGFDNFVELLSGAHAMDKHFSTTPAE
KNLPVLLALIGIWYNNFFGAETEAILPYDQYMRFAAYFQQGNMESNGKYVDRNGNVVDYQT
GPIIWGEPGTNGQHAFYQLIHQGTKMVPCDFIAPAITHNPLSDHHQKLLSNFFAQTEALAFG
KSREVVEQEYRDQGKDPATLDYVVPFKVFEGNRPTNSILLREITPFSLGALIALYEHKIFTQ
GVILNIFTFDQWGVELGKQLANRILPELKDLEVLFQGNPKDDKEISSHDSSTNGLINRYKAWRG

15

SEQ ID NO:26

MKNINPTQTAAWQALQKHFDKMDVTIADLFAKDGDRFSKFSATFDDQMLVDYSKNRITEET
LAKLQDLAKECDLAGAIKSMFSGEKINRTENRAVLHVALRNRSTPIILVDGKDVMPEVNAVL
EKMKTFFSEAIISGEWKGYTGKAITDVVNIGIGGSDLGPMVTEALRPYKNHLMHFVSNVDG
THIAEVLKKVNPETTLFLVASKTFTTQETMTNAHSARDWFLKAAGDEKHAVKHFAALSTNAK
AVGEFGIDTANMFEFWDWVGGRYSLWSAIGLSIVLSIGFDNFVELLSGAHAMDKHFSTTPAE
KNLPVLLALIGIWIYNNFFGAETAILPYDQYMHRFAAYFQQGNMESNGKYVDRNGNVVDYQT
GPIIWGEPTNGQHAFYQLIHQGTKMVPCDFIAPAITHNPLSDHHQKLLSNFFAQTEALAFG
KSREVEQEYRDQGKDPATLDYVVPFKVFEGNRPTNSILLREITPFSLGALIALYEHKIFTQ
GVILNIFTFDQWGVDELGKQLANRILPELKDDLEVLFGQGPKEISSHDSSTNGLINRYKAWRG

5 SEQ ID NO:27

MKNINPTQTAAWQALQKHFDKMDVTIADLFAKDGDRFSKFSATFDDQMLVDYSKNRITEET
LAKLQDLAKECDLAGAIKSMFSGEKINRTENRAVLHVALRNRSTPIILVDGKDVMPEVNAVL
EKMKTFFSEAIISGEWKGYTGKAITDVVNIGIGGSDLGPMVTEALRPYKNHLMHFVSNVDG
THIAEVLKKVNPETTLFLVASKTFTTQETMTNAHSARDWFLKAAGDEKHAVKHFAALSTNAK
AVGEFGIDTANMFEFWDWVGGRYSLWSAIGLSIVLSIGFDNFVELLSGAHAMDKHFSTTPAE
KNLPVLLALIGIWIYNNFFGAETAILPYDQYMHRFAAYFQQGNMESNGKYVDRNGNVVDYQT
GPIIWGEPTNGQHAFYQLIHQGTKMVPCDFIAPAITHNPLSDHHQKLLSNFFAQTEALAFG
KSREVEQEYRDQGKDPATLDYVVPFKVFEGNRPTNSILLREITPFSLGALIALYEHKIFTQ
GVILNIFTFDQWGVDELGKQLANRILPELKDDKLEVLFGQGPKEISSHDSSTNGLINRYKAWRG

10 SEQ ID NO:28

MKNINPTQTAAWQALQKHFDKMDVTIADLFAKDGDRFSKFSATFDDQMLVDYSKNRITEET
LAKLQDLAKECDLAGAIKSMFSGEKINRTENRAVLHVALRNRSTPIILVDGKDVMPEVNAVL
EKMKTFFSEAIISGEWKGYTGKAITDVVNIGIGGSDLGPMVTEALRPYKNHLMHFVSNVDG
THIAEVLKKVNPETTLFLVASKTFTTQETMTNAHSARDWFLKAAGDEKHAVKHFAALSTNAK
AVGEFGIDTANMFEFWDWVGGRYSLWSAIGLSIVLSIGFDNFVELLSGAHAMDKHFSTTPAE
KNLPVLLALIGIWIYNNFFGAETAILPYDQYMHRFAAYFQQGNMESNGKYVDRNGNVVDYQT
GPIIWGEPTNGQHAFYQLIHQGTKMVPCDFIAPAITHNPLSDHHQKLLSNFFAQTEALAFG
KSREVEQEYRDQGKDPATLDYVVPFKVFEGNRPTNSILLREITPFSLGALIALYEHKIFTQ
GVILNIFTFDQWGVDELGKQLANRILPELKDDKELEVLFGQGPKEISSHDSSTNGLINRYKAWRG

SEQ ID NO:29

15 MKNINPTQTAAWQALQKHFDKMDVTIADLFAKDGDRFSKFSATFDDQMLVDYSKNRITEET
LAKLQDLAKECDLAGAIKSMFSGEKINRTENRAVLHVALRNRSTPIILVDGKDVMPEVNAVL
EKMKTFFSEAIISGEWKGYTGKAITDVVNIGIGGSDLGPMVTEALRPYKNHLMHFVSNVDG
THIAEVLKKVNPETTLFLVASKTFTTQETMTNAHSARDWFLKAAGDEKHAVKHFAALSTNAK
AVGEFGIDTANMFEFWDWVGGRYSLWSAIGLSIVLSIGFDNFVELLSGAHAMDKHFSTTPAE
KNLPVLLALIGIWIYNNFFGAETAILPYDQYMHRFAAYFQQGNMESNGKYVDRNGNVVDYQT
GPIIWGEPTNGQHAFYQLIHQGTKMVPCDFIAPAITHNPLSDHHQKLLSNFFAQTEALAFG
KSREVEQEYRDQGKDPATLDYVVPFKVFEGNRPTNSILLREITPFSLGALIALYEHKIFTQ
GVILNIFTFDQWGVDELGKQLANRILPELKDDKEILEVLFGQGPSSHDSSTNGLINRYKAWRG

SEQ ID NO:30

MKNINPTQTAAWQALQKHFDKMDVTIADLFAKDGDRFSKFSATFDDQMLVDYSKNRITEET
LAKLQDLAKECDLAGAIKSMFSGEIKINRTENRAVLHVALRNRSNTPILVDGKDVMPPEVNAVL
EKMKTFFSEAIISGEWKGYTGKAITDVVNIGIGGSDLGPYMVTEALRPYKNHLMHFVSNVDG
THIAEVLKKVNPETTLFLVASKTFTTQETMTNAHSARDWFLKAAGDEKHAVAKHFAALSTNAK
AVGEFGIDTANMFEFWDWVGGRYSLWSAIGLSIVLSIGFDNFVELLSGAHAMDKHFSTTPAE
KNLPVLLALIGIWYNNFFGAETAILPYDQYMHRFAAYFQQGNMESNGKYVDRNGNVVDYQT
GPIIWGEPTNGQHAFYQLIHQGTGMVPCDFIAPAITHNPLSDHHQKLLSNFFAQTEALAFG
KSREVEQEYRDQGKDPATLDYVVPFKVFEGNRPTNSILLREITPFSLGALIALYEHKIFTQ
GVILNIFTFDQWGVELGKQLANRILPELKDDKEISLEVLFQGPSHDSSTNGLINRYKAWRG

5 SEQ ID NO:31

MKNINPTQTAAWQALQKHFDKMDVTIADLFAKDGDRFSKFSATFDDQMLVDYSKNRITEET
LAKLQDLAKECDLAGAIKSMFSGEIKINRTENRAVLHVALRNRSNTPILVDGKDVMPPEVNAVL
EKMKTFFSEAIISGEWKGYTGKAITDVVNIGIGGSDLGPYMVTEALRPYKNHLMHFVSNVDG
THIAEVLKKVNPETTLFLVASKTFTTQETMTNAHSARDWFLKAAGDEKHAVAKHFAALSTNAK
AVGEFGIDTANMFEFWDWVGGRYSLWSAIGLSIVLSIGFDNFVELLSGAHAMDKHFSTTPAE
KNLPVLLALIGIWYNNFFGAETAILPYDQYMHRFAAYFQQGNMESNGKYVDRNGNVVDYQT
GPIIWGEPTNGQHAFYQLIHQGTGMVPCDFIAPAITHNPLSDHHQKLLSNFFAQTEALAFG
KSREVEQEYRDQGKDPATLDYVVPFKVFEGNRPTNSILLREITPFSLGALIALYEHKIFTQ
GVILNIFTFDQWGVELGKQLANRILPELKDDKEISLEVLFQGPHDSSTNGLINRYKAWRG

SEQ ID NO:32

10

MKNINPTQTAAWQALQKHFDKMDVTIADLFAKDGDRFSKFSATFDDQMLVDYSKNRITEET
LAKLQDLAKECDLAGAIKSMFSGEIKINRTENRAVLHVALRNRSNTPILVDGKDVMPPEVNAVL
EKMKTFFSEAIISGEWKGYTGKAITDVVNIGIGGSDLGPYMVTEALRPYKNHLMHFVSNVDG
THIAEVLKKVNPETTLFLVASKTFTTQETMTNAHSARDWFLKAAGDEKHAVAKHFAALSTNAK
AVGEFGIDTANMFEFWDWVGGRYSLWSAIGLSIVLSIGFDNFVELLSGAHAMDKHFSTTPAE
KNLPVLLALIGIWYNNFFGAETAILPYDQYMHRFAAYFQQGNMESNGKYVDRNGNVVDYQT
GPIIWGEPTNGQHAFYQLIHQGTGMVPCDFIAPAITHNPLSDHHQKLLSNFFAQTEALAFG
KSREVEQEYRDQGKDPATLDYVVPFKVFEGNRPTNSILLREITPFSLGALIALYEHKIFTQ
GVILNIFTFDQWGVELGKQLANRILPELKDDKEISSHDSSTNGLINRYKLEVLFQGPAWRG

SEQ ID NO:33

ATGGGCCCAGAAGAAGAATTCGGCATGAGCCTGATCAAGCATAACTCTTGCGTCATTACCAC
GGAGAATGGTAAGTTCACGGGCTTGGGCGTTTATGACCGTTTCGTCTGGTTCCGACCCACG
CTGACCCGGGTAAAGAAATCCAGGTTGACGGTATCACGACCAAAGTGATTGATAGCTATGAT
CTCTATAATAAGAACGGCATCAAGCTGGAATCACGGTGCTGAACTGGACCGTAATGAAAA
GTTTCGTGATATCCGTCGCTATATTCGAATAACGAGGATGACTACCCAAATGCAATCTGG
CGCTGCTGGCAAATCAGCCGGAACCGACGATCATCAACGTGGGTGACGTGGTGAGCTATGGC
AATATCCTGCTGAGCGGTAACCAGACCGCGCGTATGCTGAAGTATTCTATCCGACGAAAAG
CGGCTATTGCGGCGGCGTGCTCTATAAGATTGGTCAAGTCCTGGGCATCCACGTCGGCGGTA
ATGGCCGCGATGGTTTCAGCGCGATGCTGCTGCGTAGCTATTTACCCGACGTCCAGTGATAA

15

SEQ ID NO:34

MGPEEEFGMSLIKHNSCVITTENGKFTGLGVYDRFVVVPTHADPGKEIQVDGITTVIDSYI
LYNKNGIKLEITVLKLDNRNEKFRDIRRYIPNNEDDYPNCNLALLANQPEPTIINVGDVVSYSY
NILLSGNQATARMLKYSYPTKSGYCGGVLYKIGQVLGIHVGGNGRDFGSAMLLRSYFTDVQ

5 SEQ ID NO:35

ATGAAAAAACGGCAATTGCGATAGCGGTTGCGCTAGCTGGTTTTGCCACGGTGGCGCAGGC
TGGCCCAGAAGAAGAATTCGGCATGAGCCTGATCAAGCATAACTCTTGCGTCATTACCACGG
AGAATGGTAAGTTCACGGGCTTGGGCGTTTATGACCGTTTCGTCGTGGTTCCGACCCACGCT
GACCCGGGTAAAGAAATCCAGGTTGACGGTATCACGACCAAAGTGATTGATAGCTATGATCT
CTATAATAAGAACGGCATCAAGCTGGAATCACGGTGCTGAAACTGGACCGTAATGAAAAGT
TTCGTGATATCCGTGCTATATTCCGAATAACGAGGATGACTACCCAAATTGCAATCTGGCG
CTGCTGGCAAATCAGCCGGAACCGACGATCATCAACGTGGGTGACGTGGTGAGCTATGGCAA
TATCCTGCTGAGCGGTAACCAGACCGCGCGTATGCTGAAGTATTCCTATCCGACGAAAAGCG
GCTATTGCGGCGGCGTGCTCTATAAGATTGGTCAAGTCCTGGGCATCCACGTCGGCGGTAAT
GGCCGCGATGGTTTCAGCGCGATGCTGCTGCGTAGCTATTTACCGACGTCCAGTGATAA

SEQ ID NO:36

10

MKKTAIAIAVALAGFATVAQAGPEEEFGMSLIKHNSCVITTENGKFTGLGVYDRFVVVPTHAD
DPGKEIQVDGITTVIDSYDLYNKNGIKLEITVLKLDNRNEKFRDIRRYIPNNEDDYPNCNLAL
LLANQPEPTIINVGDVVSYSYGNILLSGNQATARMLKYSYPTKSGYCGGVLYKIGQVLGIHVGGN
GRDFGSAMLLRSYFTDVQ

SEQ ID NO:37

CTGGAAGTGCTGTTTCAGGGTCCG

15

SEQ ID NO:38

LEVLFGGP

SEQ ID NO:39

EVLFGGP

20

SEQ ID NO:40

VLFQGP

SEQ ID NO:41

LFQGP

25

SEQ ID NO:42

LEVLFGG

30

SEQ ID NO:43

LEVLFG

SEQ ID NO:44

LEVL

35

SEQ ID NO:45

EVLFGG

SEQ ID NO:46

MKKTAIAIAVALAGFATVAQA

40

SEQ ID NO:47

GTGTCCCGTATTATTATGCTGATCCCTACCGGAACCAGCGTCGGTCTGACCAGCGTCAGCCT
TGGCGTGATCCGTGCAATGGAACGCAAAGGCGTTCGTCTGAGCGTTTTCAAACCTATCGCTC
AGCCGCGTACCGGTGGCGATGCGCCCGATCAGACTACGACTATCGTGCGTGCGAACTCTTCC
ACCACGACGGCCGCTGAACCGCTGAAAATGAGCTACGTTGAAGGTCTGCTTTCCAGCAATCA
GAAAGATGTGCTGATGGAAGAGATCGTCGCAAACTACCACGCTAACACCAAAGACGCTGAAG
TCGTTCTGGTTGAAGGTCTGGTCCCGACACGTAAGCACCAGTTTGCCAGTCTCTGAACTAC
GAAATCGCTAAAACGCTGAATGCGGAAATCGTCTTCGTTATGTCTCAGGGCACTGACACCCC
GGAACAGCTGAAAGAGCGTATCGAACTGACCCGCAACAGCTTCGGCGGTGCCAAAAACACCA
ACATCACCGGCGTTATCGTTAAACAACTGAACGCACCGGTTGATGAACAGGGTTCGTACTCGC
CCGGATCTGTCCGAGATTTTCGACGACTCTTCCAAAGCTAAAGTAAACAATGTTGATCCGGC
GAAGCTGCAAGAATCCAGCCCGCTGCCGTTCTCGGCGCTGTGCCGTGGAGCTTTGACCTGA
TCGCGACTCGTGCGATCGATATGGCTCGCCACCTGAATGCGACCATCATCAACGAAGGCGAC

ATCAATACTCGCCGCGTTAAATCCGTCACCTTCTGCGCACGCAGCATTCGCGACATGCTGGA
GCACTTCCGTGCCGTTCTCTGCTGGTGACTTCCGCAGACCGTCCTGACGTGCTGGTGCCG
CTTGCCCTGGCAGCCATGAACGGCGTAGAAATCGGTGCCCTGCTGCTGACTGGCGGTTACGAA
ATGGACGCGCGCATTTCTAAACTGTGCGAACGTGCTTTCGCTACCGGCCTGCCGGTATTTAT
GGTGAACACCAACACCTGGCAGACCTCTCTGAGCCTGCAGAGCTTCAACCTGGAAGTTCCGG
TTGACGATCACGAACGTATCGAGAAAGTTTACGAATACGTTGCTAACTACATCAACGCTGAC
TGGATCGAATCTCTGACTGCCACTTCTGAGCGCAGCCGTCGTCTGTCTCCGCCTGCGTTCCG
TTATCAGCTGACTGAACTTGCGCGCAAAGCGGGCAAACGTATCGTACTGCCGGAAGGTGACG
AACC GCGTACCGTTAAAGCAGCCGCTATCTGTGCTGAACGTGGTATCGCAACTTGCGTACTG
CTGGGTAATCCGGCAGAGATCAACCGTGTTGCAGCGTCTCAGGGTGTAAGTGGGTGCAGG
GATTGAAATCGTTGATCCAGAAGTGGTTCGCGAAAGCTATGTTGGTCTGCTGGTTCGAACTGC
GTAAGAACAAAGGCATGACCGAAACCGTTGCCCGCGAACAGCTGGAAGACAACGTGGTGCTC
GGTACGCTGATGCTGGAACAGGATGAAGTTGATGGTCTGGTTTCCGGTGCTGTTACACTAC
CGCAAACACCATCCGTCCGCCGCTGCAGCTGATCAAACTGCACCGGGCAGCTCCCTGGTAT
CTTCCGTGTTCTTCATGCTGCTGCCGGAACAGGTTTACGTTTACGGTGACTGTGCGATCAAC
CCGGATCCGACCGCTGAACAGCTGGCAGAAATCGCGATTTCAGTCCGCTGATTCCGCTGCGGC
CTTCGGTATCGAACCGCGCGTTGCTATGCTCTCCTACTCCACCGGTACTTCTGGTGCAGGTA
GCGACGTAGAAAAAGTTTCGCGAAGCAACTCGTCTGGCGCAGGAAAAACGTCCTGACCTGATG
ATCGACGGTCCGCTGCAGTACGACGCTGCGGTAATGGCTGACGTTGCGAAATCCAAAGCGCC
GAACTCTCCGGTTGCAGGTGCGGCTACCGTGTTTCATCTTCCCGGATCTGAACACCGGTAACA
CCACCTACAAAGCGGTACAGCGTTCTGCCGACCTGATCTCCATCGGGCCGATGCTGCAGGGT
ATGCGCAAGCCGGTTAACGACCTGTCCCGTGGCGCACTGGTTGACGATATCGTCTACACCAT
CGCGCTGACTGCGATTTCAGTCTGCACAGCAGCAGTAA

SEQ ID NO:48

VSRIIMLIPTGTSVGLTSVSLGVIRAMERKGVRLSVFKPIAQPRTGGDAPDQTTTIVRANSS
TTTAAEPLKMSYVEGLSSNQKDVLMEEIVANYHANTKDAEVVLVEGLVPTRKHQFAQSLNY
EIAKTLNAEIVFVMSQGTDTPEQLKERIELTRNSFGGAKNTNITGVIVNKLNAPVDEQGRTR
PDLSEIFDDSSKAKVNNVDPAKLQESSPLPVLGAVPWSFDLIATRAIDMARHLNATIINEGD
INTRRVKSVTFCARSIPHMLEHFRAGSLLVTSADRPDVLVAACLAAMNGVEIGALLLTGGYE
MDARISKLCEAFATGLPVFMVNTNTWQTSLSLSQSFNLEVPVDDHERIEKVQVEYVANYINAD
WIESLTATSESRRLSPPAFRYQLTELARKAGKRIVLPEGDEPRTVKAAAICAERGIATCVL
LGNPAEINRVAASQGVELGAGIEIVDPEVVRESYVGRVLVELRKNKGMTETVAREQLEDNVVL
GTLMLEQDEVDGLVSGAVHTTANTIRPPLQLIKTAPGSSLVSSVFFMLLPEQVYVYGDCAIN
PDPTAEQLAEIAIQSADSAAFGIEPRVAMLSYSTGTSGAGSDVEKVVREATRLAQEKRPDL
IDGFLQYDAVMADVAKSKAPNSPVAGRATVFIFPDLNTGNTTYKAVQRSADLISIGPMLQG
MRKPVNDLSRGALVDDIVYTIALTAIQSAQQQ

SEQ ID NO:49

GTGTCCCGTATTATTATGCTGATCCCTACCGGAACCAGCGTCGGTCTGACCAGCGTCAGCCT
TGGCGTGATCCGTGCAATGGAACGCAAAGGCGTTCGTCTGAGCGTTTTCAAACCTATCGCTC
AGCCGCGTACCGGTGGCGATGCGCCCGATCAGACTACGACTATCGTGCGTGCGAACTCTTCC
ACCACGACGGCCGCTGAACCGCTGAAAATGAGCTACGTTGAAGGTCTGCTTTCCAGCAATCA
GAAAGATGTGCTGATGGAAGAGATCGTCGCAAACCTACCACGCTAACACCAAAGACGCTGAAG
TCGTTCTGGTTGAAGGTCTGGTCCCGACACGTAAGCACCAGTTTGCCAGTCTCTGAACTAC
GAAATCGCTAAAACGCTGAATGCGGAAATCGTCTTCGTTATGTCTCAGGGCACTGACACCCC
GGAACAGCTGAAAGAGCGTATCGAACTGACCCGCAACAGCTTCGGCGGTGCCAAAAACACCA
ACATCACCGGCGTTATCGTTAACTGAACGCACCGGTTGATGAACAGGGTCGTA CT CGC
CCGGATCTGTCCGAGATTTTCGACGACTCTTCCAAAGCTAAAGTAAACAATGTTGATCCGGC

GAAGCTGCAAGAATCCAGCCCGCTGCCGGTTCTCGGCGCTGTGCCGTGGAGCTTTGACCTGA
TCGCGACTCGTGCGATCGATATGGCTCGCCACCTGAATGCGACCATCATCAACGAAGCGCAC
ATCAATACTCGCCGCGTTAAATCCGTCACCTTTCTGCGCACGCAGCATTCGCGACATGCTGGA
GCACTTCCGTGCCGGTTCTCTGCTGGTGACTTCCGCAGACCGTCCTGACGTGCTGGTGGCCG
CTTGCCCTGGCAGCCATGAACGGCGTAGAAATCGGTGCCCTGCTGCTGACTGGCGGTTACGAA
ATGGACGCGCGCATTTCTAAACTGTGCGAACGTGCTTTTCGCTACCGGCCTGCCGGTATTTAT
GGTGAACACCAACACCTGGCAGACCTCTCTGAGCCTGCAGAGCTTCAACCTGGAAGTCCGG
TTGACGATCACGAACGTATCGAGAAAGTTCAGGAATACGTTGCTAACTACATCAACGCTGAC
TGGATCGAATCTCTGACTGCCACTTCTCTGGAAGTGCTGTTTCAGGGTCCGGAGCGCAGCCG
TCGCTGTCTCCGCCTGCGTTCCGTTATCAGCTGACTGAACTTGCGCGCAAAGCGGGCAAAC
GTATCGTACTGCCGGAAGGTGACGAACCGCGTACCGTTAAAGCAGCCGCTATCTGTGCTGAA
CGTGGTATCGCAACTTGCGTACTGCTGGGTAATCCGGCAGAGATCAACCGTGTTCAGCGTC
TCAGGGTGTAGAACTGGGTGCAGGGATTGAAATCGTTGATCCAGAAAGTGGTTCGCGAAAGCT
ATGTTGGTCGTCTGGTCAACTGCGTAAGAACAAAGGCATGACCGAAACCGTTGCCCGCGAA
CAGCTGGAAGACAACGTGGTGCTCGGTACGCTGATGCTGGAACAGGATGAAGTTGATGGTCT
GGTTTCCGGTGCTGTTACACTACCGCAAACACCATCCGTCCGCCGCTGCAGCTGATCAAAA
CTGCACCGGGCAGCTCCCTGGTATCTTCCGTGTTCTTCATGCTGCTGCCGGAACAGGTTTAC
GTTTACGGTGACTGTGCGATCAACCCGGATCCGACCGCTGAACAGCTGGCAGAAATCGCGAT
TCAGTCCGCTGATTCCGCTGCGGCCCTTCGGTATCGAACCGCGCGTTGCTATGCTCTCTACT
CCACCGGTACTTCTGGTGCAGGTAGCGACGTAGAAAAAGTTCGCGAAGCAACTCGTCTGGCG
CAGGAAAAACGTCTGACCTGATGATCGACGGTCCGCTGCAGTACGACGCTGCGGTAATGGC
TGACGTTGCGAAATCCAAAGCGCCGAACTCTCCGGTTGCAGGTGCGGCTACCGTGTTTCATCT
TCCCGGATCTGAACACCGGTAACACCACCTACAAAGCGGTACAGCGTTCTGCCGACCTGATC
TCCATCGGGCCGATGCTGCAGGGTATGCGCAAGCCGGTTAACGACCTGTCCCGTGGCGCACT
GGTTGACGATATCGTCTACCATCGCGCTGACTGCGATTTCAGTCTGCACAGCAGCAGTAA

5

SEQ ID NO:50

VSRIIMLIPTGTSVGLTSVSLGVIRAMERKGVRLSVFKPIAQPRITGGDAPDQTTTIVRANSS
TTTAAEPLKMSYVEGLLSNQKDVLMEEIVANYHANTKDAEVVLVEGLVPTRKHQFAQSLNY
EIAKTLNAEIVFVMSQGTDTPEQLKERIELTRNSFGGAKNTNITGVIVNKLNAPVDEQGRTR
PDLSEIFDDSSKAKVNNVDPKQLQESSPLPVLGAVPWSFDLIATRAIDMARHLNATIINEGD
INTRRVKSVTFCARSIPHMLEHFRAAGSLVTSADRPDVLVAACLAAMNGVEIGALLLTGGYE
MDARISKLCERAFATGLPVFMVNTNTWQTSLSLQSFNLEVPVDDHERIEKVQEYVANYINAD
WIESLTATSLEVLFOGPERSRRLSPPAFRYQLTELARKAGKRIVLPEGDEPRTVKAAAI CAE
RGIATCVLLGNPAEINRVAASQGVELGAGIEIVDPEVVRESYVGRVLVELRKNKGMTETVARE
QLEDNVVLGTLMLEQDEVDGLVSGAVHTTANTIRPPLQLIKTAPGSSLVSSVFFMMLLPEQVY
VYGDCAINPDPTAEQLAEIAIQSADSAAAFGIEPRVAMLSYSTGTSAGSDVEKVVREATRLA
QEKRPDL MIDGPLQYDAVMADVAKSKAPNSPVAGRATVFI FPD LNTGN TTYKAVQRSADLI
SIGPMLQGMKPVNDLSRGALVDDIVYTI ALTAIQSAQQQ

SEQ ID NO:51

GTGTCCCGTATTATTATGCTGATCCCTACCGGAACCAGCGTCGGTCTGACCAGCGTCAGCCT
TGGCGTGATCCGTGCAATGGAACGCAAAGGCGTTCGTCTGAGCGTTTTCAAACCTATCGCTC
AGCCGCGTACCGGTGGCGATGCGCCCGATCAGACTACGACTATCGTGCGTGCGAACTCTTCC
ACCACGACGGCCGCTGAACCGCTGAAAATGAGCTACGTTGAAGGTCTGCTTTCCAGCAATCA
GAAAGATGTGCTGATGGAAGAGATCGTCGCAAACTACCACGCTAACACCAAAGACGCTGAAG
TCGTTCTGGTTGAAGGTCTGGTCCCGACACGTAAGCACCAGTTTGCCAGTCTCTGAACTAC
GAAATCGCTAAAACGCTGAATGCGGAAATCGTCTTCGTTATGTCTCAGGGCACTGACACCCC
GGAACAGCTGAAAGAGCGTATCGAACTGACCCGCAACAGCTTCGGCGGTGCCAAAAACACCA

ACATCACCGGCGTTATCGTTAACAACTGAACGCACCGGTTGATGAACAGGGTCGTA CT CGC
CCGGATCTGTCCGAGATTTTCGACGACTCTTCCAAAGCTAAAGTAAACAATGTTGATCCGGC
GAAGCTGCAAGAATCCAGCCCGCTGCCGGTTCTCGGCGCTGTGCCGTGGAGCTTTGACCTGA
TCGCGACTCGTGCGATCGATATGGCTCGCCACCTGAATGCGACCATCATCAACGAAGGCGAC
ATCAATACTCGCCGCGTTAAATCCGTCACTTTCTGCGCACGCAGCATTCCGCACATGCTGGA
GCACTTCCGTGCCGGTTCTCTGCTGGTGACTTCCGCAGACCGTCCTGACGTGCTGGTGGCCG
CTTGCCCTGGCAGCCATGAACGGCGTAGAAATCGGTGCCCTGCTGCTGACTGGCGGTTACGAA
ATGGACGCGCGCATTTCTAAACTGTGCGAACGTGCTTTCGCTACCGGCCTGCCGGTATTTAT
GGTGAACACCAACACCTGGCAGACCTCTCTGAGCCTGCAGAGCTTCAACCTGGAAGTTCCGG
TTGACGATCACGAACGTATCGAGAAAGTTCAGGAATACGTTGCTAACTACATCAACGCTGAC
TGGATCGAATCTCTGACTGCCACTTCTGAGCTGGAAGTGCTGTTTCAGGGTCCGCGCAGCCG
TCGTCTGTCTCCGCCTGCGTTCCGTTATCAGCTGACTGAACTTGCGCGCAAAGCGGGCAAAC
GTATCGTACTGCCGGAAGGTGACGAACCGCGTACCGTTAAAGCAGCCGCTATCTGTGCTGAA
CGTGGTATCGCAACTTGCGTACTGCTGGGTAATCCGGCAGAGATCAACCGTGTTGCAGCGTC
TCAGGGTG TAGAACTGGGTGCAGGGATTGAAATCGTTGATCCAGAAGTGGTTCGCGAAAGCT
ATGTTGGTCGTCTGGTCGAACTGCGTAAGAACAAAGGCATGACCGAAACCGTTGCCCGCGAA
CAGCTGGAAGACAACGTGGTGCTCGGTACGCTGATGCTGGAACAGGATGAAGTTGATGGTCT
GGTTTCCGGTGCTGTTTCACTACTACCGCAAACACCATCCGTCCGCCGCTGCAGCTGATCAAAA
CTGCACCGGGCAGCTCCCTGGTATCTTCCGTGTTCTTCATGCTGCTGCCGGAACAGGTTTAC
GTTTACGGTGACTGTGCGATCAACCCGGATCCGACCGCTGAACAGCTGGCAGAAATCGCGAT
TCAGTCCGCTGATTCCGCTGCGGCCTTCGGTATCGAACCGCGCGTTGCTATGCTCTCCTACT
CCACCGGTACTTCTGGTGCAGGTAGCGACGTAGAAAAAGTTCGCGAAGCAACTCGTCTGGCG
CAGGAAAAACGTCCTGACCTGATGATCGACGGTCCGCTGCAGTACGACGCTGCGGTAATGGC
TGACGTTGCGAAATCCAAAGCGCCGAACTCTCCGGTTGCAGGTGCGCTACCGTGTTTCATCT
TCCCGGATCTGAACACCGGTAACACCACCTACAAAGCGGTACAGCGTTCTGCCGACCTGATC
TCCATCGGGCCGATGCTGCAGGGTATGCGCAAGCCGGTTAACGACCTGTCCCGTGGCGCACT
GTTTGACGATATCGTCTACACCATCGCGCTGACTGCGATTGAGTCTGCACAGCAGCAGTAA

5 SEQ ID NO:52

VSRIIMLIPTGTSVGLTSVSLGVIRAMERKGVRLSVFKPIAQPRTG DAPDQTTTIVRANSS
TTTAAEPLKMSYVEGLLSSNQKDVLMEEIVANYHANTKDAEVVLVEGLVPTRKHQFAQSLNY
EIAKTLNAEIVFVMSQGTDTPEQLKERIELTRNSFGAKNTNITGVIVNKL NAPVDEQGRTR
PDLSEIFDDSSKAKVNNVDPAKLQESSPLPVLGAVPWSFDLIATRAIDMARHLNATI INEGD
INTRRVKSVTF CARSIPHMLEHFRA G SLLVTSADRPDLVLAACLAAMNGVEIGALLLTGGYE
MDARISKLCERAFATGLPVFMVNTNTWQTSLSLSQSFNLEVPVDDHERIEKVQEYVANYINAD
WIESLTATSELEVLFQGPRSRRLSPPAFRYQLTELARKAGKRIVLPEGDEPRTVKA A ICAE
RGIATCVLLGNPAEINRVAASQGV ELGAGIEIVDPEVVRESYVGRLEVELRKNKGMTETVARE
QLEDNVVLGTLML EQDEVDGLVSGAVHTTANTIRPPLQLIKTAPGSSLVSSVFFMLLPEQVY
VYGDCAINPDPTAEQLAEIAIQSADSAAAFGIEPRVAMLSYSTGTSGAGSDVEKVREATRLA
QEKRPDL MIDGPLQYDAAVMADVAKSKAPNSPVAGRATVFIFPDLNTGNTTYKAVQRSADLI
SIGPMLQGMRKPVNDLSRGALVDDIVYTIALTAIQSAQQQ

SEQ ID NO:53

GTGTCCCGTATTATTATGCTGATCCCTACCGGAACCAGCGTCGGTCTGACCAGCGTCAGCCT
TGGCGTGATCCGTGCAATGGAACGCAAAGGCGTTCGTCTGAGCGTTTTCAAACCTATCGCTC
AGCCGCGTACCGGTGGCGATGCGCCCGATCAGACTACGACTATCGTGCGTGCGAACTCTTCC
ACCACGACGGCCGCTGAACCGCTGAAAATGAGCTACGTTGAAGGTCTGCTTTCCAGCAATCA
GAAAGATGTGCTGATGGAAGAGATCGTCGCAAACCTACCACGCTAACACCAAAGACGCTGAAG
TCGTTCTGGTTGAAGGTCTGGTCCCGACACGTAAGCACCAGTTTGCCCAGTCTCTGAACTAC

GAAATCGCTAAAACGCTGAATGCGGAAATCGTCTTCGTTATGTCTCAGGGCACTGACACCCC
GGAACAGCTGAAAGAGCGTATCGAACTGACCCGCAACAGCTTCGGCGGTGCCAAAAACACCA
ACATCACCGGCGTTATCGTTAACAACTGAACGCACCGGTTGATGAACAGGGTCGTACTCGC
CCGGATCTGTCCGAGATTTTCGACGACTCTTCCAAAGCTAAAGTAAACAATGTTGATCCGGC
GAAGCTGCAAGAATCCAGCCCGCTGCCGGTTCTCGGCGCTGTGCCGTGGAGCTTTGACCTGA
TCGCGACTCGTGCGATCGATATGGCTCGCCACCTGAATGCGACCATCATCAACGAAGGCGAC
ATCAATACTCGCCGCGTTAAATCCGTCACCTTTCTGCGCACGCAGCATTCGCGACATGCTGGA
GCACTTCCGTCGCCGGTTCTCTGCTGGTGACTTCCGCGAGCCGTCCTGACGTGCTGGTGCCCG
CTTGCCCTGGCAGCCATGAACGGCGTAGAAATCGGTGCCCTGCTGCTGACTGGCGGTTACGAA
ATGGACGCGCGCATTTCTAAACTGTGCGAACGTGCTTTCGCTACCGGCCCTGCCGGTATTTAT
GGTGAACACCAACACCTGGCAGACCTCTCTGAGCCTGCAGAGCTTCAACCTGGAAGTTCGG
TTGACGATCACGAACGTATCGAGAAAGTTCAGGAATACGTTGCTAACTACATCAACGCTGAC
TGGATCGAATCTCTGACTGCCACTTCTGAGCGCAGCCGTCGTCTGGAAGTGCTGTTTCAGGG
TCCGTCTCCGCTGCGTTCCGTTATCAGCTGACTGAACTTGCGCGCAAAGCGGGCAAACGTA
TCGTACTGCCGGAAGGTGACGAACCGCGTACCGTTAAAGCAGCCGCTATCTGTGCTGAACGT
GGTATCGCAACTTGCGTACTGCTGGGTAATCCGGCAGAGATCAACCGTGTGTCAGCGTCTCA
GGGTGTAGAACTGGGTGCAGGGATTGAAATCGTTGATCCAGAAGTGGTTCGCGAAAGCTATG
TTGGTCGTCTGGTCGAACTGCGTAAGAACAAAGGCATGACCGAAACCGTTGCCCGCGAACAG
CTGGAAGACAACGTGGTGCTCGGTACGCTGATGCTGGAACAGGATGAAGTTGATGGTCTGGT
TTCCGGTGCTGTTACACTACCGCAAACACCATCCGTCCGCCGCTGCAGCTGATCAAACTG
CACCGGGCAGCTCCCTGGTATCTTCCGTGTTCTTCATGCTGCTGCCGGAACAGGTTTACGTT
TACGGTGACTGTGCGATCAACCCGGATCCGACCGCTGAACAGCTGGCAGAAATCGCGATTCA
GTCCGCTGATTCCGCTGCGGCCTTCGGTATCGAACCGCGCGTTGCTATGCTCTCCTACTCCA
CCGGTACTTCTGGTGCAGGTAGCGACGTAGAAAAAGTTCGCGAAGCAACTCGTCTGGCGCAG
GAAAAACGTCCTGACCTGATGATCGACGGTCCGCTGCAGTACGACGCTGCGGTAATGGCTGA
CGTTGCGAAATCCAAAGCGCCGAACCTCTCCGGTTGCAGGTGCGGCTACCGTGTTTCATCTTCC
CGGATCTGAACACCGGTAACACCACCTACAAAGCGGTACAGCGTTCTGCCGACCTGATCTCC
ATCGGGCCGATGCTGCAGGGTATGCGCAAGCCGGTTAACGACCTGTCCCGTGCGGCACTGGT
TGACGATATCGTCTACACCATCGCGCTGACTGCGATTTCAGTCTGCACAGCAGCAGTAA

5

SEQ ID NO:54

VSRIIMLIPTGTSVGLTSVSLGVIRAMERKGVRLSVFKPIAQPRTGGDAPDQTTTIVRANSS
TTTAAEPLKMSYVEGLLSSNQKDVLMEEIVANYHANTKDAEVVLVEGLVPTRKHQFAQSLNY
EIAKTLNAEIVFVMSQGTDTPEQLKERIELTRNSFGGAKNTNITGVIVNKLNPVDEQGRTR
PDLSEIFDDSSKAKVNNVDPAKLQESSPLPVLGAVPWSFDLIATRAIDMARHLNATIINEGD
INTRRVKSVTFCARSIPHMLEHFRAGSLVTSADRPDVLVAACLAAMNGVEIGALLLTGGYE
MDARISKLCERAFATGLPVFMVNTNTWQTSLSLQSFNLEVPVDDHERIEKVQVEYVANYINAD
WIESLTATSESRRLLEVLFQGPSPPAFRYQLTELARKAGKRIVLPEGDEPRTVKAACAER
GIATCVLLGNPAEINRVAASQGVELGAGIEIVDPEVVRESYVGRLEVELRKNKGMTETVAREQ
LEDNVVLGTLMLEQDEVDGLVSGAVHTTANTIRPPLQLIKTAPGSSLVSSVFFMLLPEQVYV
YGDCAINPDPTAEQLAEIAIQSADSAAAFGIEPRVAMLSYSTGTSGAGSDVEKVVREATRLAQ
EKRPDL MIDGPLQYDAAVMADVAKSKAPNSPVAGRATVFIFPDLNTGNTTYKAVQRSADLIS
IGPMLQGMRKPVNDLSRGALVDDIVYTIALTAIQSAQQQ

10

SEQ ID NO:55

GTGTCCCGTATTATTATGCTGATCCCTACCGGAACCAGCGTCGGTCTGACCAGCGTCAGCCT
TGGCGTGATCCGTGCAATGGAACGCAAAGGCGTTCGTCTGAGCGTTTTCAAACCTATCGCTC
AGCCGCGTACCGGTGGCGATGCGCCCGATCAGACTACGACTATCGTGCGTGCGAACTCTTCC
ACCACGACGGCCGCTGAACCGCTGAAAATGAGCTACGTTGAAGGTCTGCTTTCCAGCAATCA

GAAAGATGTGCTGATGGAAGAGATCGTCGCAAACTACCACGCTAACACCAAAGACGCTGAAG
TCGTTCTGGTTGAAGGTCTGGTCCCGACACGTAAGCACCAGTTTGCCAGTCTCTGAACTAC
GAAATCGCTAAAACGCTGAATGCGGAAATCGTCTTCGTTATGTCTCAGGGCACTGACACCCC
GGAACAGCTGAAAGAGCGTATCGAACTGACCCGCAACAGCTTCGGCGGTGCCAAAAACACCA
ACATCACCGGCGTTATCGTTAAACAACTGAACGCACCGGTTGATGAACAGGGTTCGTACTCGC
CCGGATCTGTCCGAGATTTTCGACGACTCTTCCAAAGCTAAAGTAAACAATGTTGATCCGGC
GAAGCTGCAAGAATCCAGCCCGCTGCCGGTTCTCGGCGCTGTGCCGTGGAGCTTTGACCTGA
TCGCGACTCGTGCGATCGATATGGCTCGCCACCTGAATGCGACCATCATCAACGAAGGCGAC
ATCAATACTCGCCGCGTTAAATCCGTCACTTTCTGCGCACGCAGCATTCCGCACATGCTGGA
GCACTTCCGTGCCGGTTCTCTGCTGGTGACTTCCGCAGACCGTCTGACGTGCTGGTGGCCG
CTTGCCCTGGCAGCCATGAACGGCGTAGAAATCGGTGCCCTGCTGCTGACTGGCGGTTACGAA
ATGGACGCGCGCATTTCTAAACTGTGCGAACGTGCTTTCGCTACCGGCCTGCCGGTATTTAT
GGTGAACACCAACACCTGGCAGACCTCTCTGAGCCTGCAGAGCTTCAACCTGGAAGTTCCGG
TTGACGATCACGAACGTATCGAGAAAGTTTCAAGGAATACGTTGCTAACTACATCAACGCTGAC
TGGATCGAATCTCTGACTGCCACTTCTGAGCGCAGCCGTCGTCTGTCTCCGCCTGCGTTCCG
TTATCAGCTGACTGAACTTGCGCGCAAAGCGGGCAAACGTATCGTACTGGAAGTGCTGTTTC
AGGGTCCGGAAGGTGACGAACCGCTACCGTTAAAGCAGCCGCTATCTGTGCTGAACGTGGT
ATCGCAACTTGCGTACTGCTGGGTAATCCGGCAGAGATCAACCGTGTTGCAGCGTCTCAGGG
TGTAAGAACTGGGTGCAGGGATTGAAATCGTTGATCCAGAAGTGGTTTCGCGAAAGCTATGTTG
GTCGTCTGGTGAAGTGCAGTAAGAACAAAGGCATGACCGAAACCGTTGCCCGCGAACAGCTG
GAAGACAACGTGGTGCTCGGTACGCTGATGCTGGAACAGGATGAAGTTGATGGTCTGGTTTC
CGGTGCTGTTTCACTACCGCAAACACCATCCGTCCGCCGCTGCAGCTGATCAAACTGCAC
CGGGCAGCTCCCTGGTATCTTCCGTGTTCTTCATGCTGCTGCCGGAACAGGTTTACGTTTAC
GGTACTGTGCGATCAACCCGGATCCGACCGCTGAACAGCTGGCAGAAATCGCGATTACAGTC
CGCTGATTCCGCTGCCGCCCTTCGGTATCGAACCGCGCGTTGCTATGCTCTCCTACTCCACCG
GTACTTCTGGTGCAGGTAGCGACGTAGAAAAAGTTTCGCGAAGCAACTCGTCTGGCGCAGGAA
AAACGTCTGACCTGATGATCGACGGTCCGCTGCAGTACGACGCTGCCGTAATGGCTGACGT
TGCGAAATCCAAAGCGCCGAACTCTCCGGTTGCAGGTGCGCGTACCGTGTTTATCTTCCCGG
ATCTGAACACCGGTAACACCACCTACAAAGCGGTACAGCGTTCTGCCGACCTGATCTCCATC
GGGCCGATGCTGCAGGTATGCGCAAGCCGGTTAACGACCTGTCCCGTGGCGCACTGGTTGA
CGATATCGTCTACACCATCGCGCTGACTGCGATTGAGTCTGCACAGCAGCAGTAA

5

SEQ ID NO:56

VSRIIMLIPTGTSVGLTSVSLGVIRAMERKGVRLSVFKPIAQPRTGGDAPDQTTTIVRANSS
TTTAAEPLKMSYVEGLLSSNQKDVLMEEIVANYHANTKDAEVVLVEGLVPTRKHQFAQSLNY
EIAKTLNAEIVFVMSQGTDTPEQLKERIELTRNSFGGAKNTNITGVIIVNKLNPVDEQGRTR
PDLSEIFDDSSKAKVNNVDPAKLQESSPLPVLGAVPWSFDLIATRAIDMARHLNATIINEGD
INTRRVKSVTFCARSIPHMLEHFRAGSLVTSADRPDVLVAACLAAMNGVEIGALLLTGGYE
MDARISKLCEAFATGLPVFMVNTNTWQTSLSLQSFNLEVPVDDHERIEKVQEYVANYINAD
WIESLTATSESRRLSPPAFRYQLTELARKAGKRIVLEVLFGPEGDEPRTVKAAICAERG
IATCVLLGNPAEINRVAASQGVELGAGIEIVDPEVVRESYVGRLEVELRKNKGMTETVAREQL
EDNVVLGTLMLEQDEVDGLVSGAVHTTANTIRPPLQLIKTAPGSSLVSSVFFMLLPEQVYVY
GDCAINPDPTAEQLAEIAIQSADSAAAFGIEPRVAMLSYSTGTSGAGSDVEKVREATRLAQE
KRPDLMIDGPLQYDAAVMADVAKSKAPNSPVAGRATVFIFFDLNTGNTTYKAVQRSADLISI
GPMLQGMKRPVNDLSRGALVDDIVYTIALTAIQSAQQQ

SEQ ID NO:57

ATGTCCTCACGTAAAGAGCTTGCCAATGCTATTCGTGCGCTGAGCATGGACGCAGTACAGAA
AGCCAAATCCGGTCACCCGGGTGCCCCCTATGGGTATGGCTGACATTGCCGAAGTCCTGTGGC

GTGATTTCTGAAACACAACCCGCGAGAATCCGTCCTGGGCTGACCGTGACCGCTTCGTGCTG
TCCAACGGCCACGGCTCCATGCTGATCTACAGCCTGCTGCACCTCACCGGTTACGATCTGCC
GATGGAAGAACTGAAAACTTCCGTGAGCTGCACTCTAAACTCCGGGTCACCCGGAAGTGG
GTTACACCGCTGGTGTGGAAACCACCGGTCCGCTGGGTGAGGGTATTGCCAACGCAGTC
GGTATGGCGATTGCAGAAAAACGCTGGCGGCGCAGTTTAACCGTCCGGGCCACGACATTGT
CGACCACTACACCTACGCCCTTCATGGGCGACGGCTGCATGATGGAAGGCATCTCCACGAAG
TTTGCTCTCTGGCGGGTACGCTGAAGCTGGGTAAACTGATTGCATTCTACGATGACAACGGT
ATTTCTATCGATGGTCACGTTGAAGGCTGGTTCACCGACGACACCGCAATGCGTTTCGAAGC
TTACGGCTGGCACGTTATTCGCGACATCGACGGTCATGACGCGGCATCTATCAAACGCGCAG
TAGAAGAAGCGCGCGCAGTGACTGACAAACCTTCCCTGCTGATGTGCAAACCATCATCGGT
TTCGGTTCCCCGAACAAAGCCGGTACCCACGACTCCCACGGTGCGCCGCTGGGCGACGCTGA
AATTGCCCTGACCCGCGAACAACCTGGGCTGGAAATATGCGCCGTTGAAATCCCGTCTGAAA
TCTATGCTCAGTGGGATGCGAAAGAAGCAGGCCAGGCGAAAGAATCCGCATGGAACGAGAAA
TTCGCTGCTTACGCGAAAGCTTATCCGCGAGGAAGCCGCTGAATTTACCCGCCGTATGAAAGG
CGAAATGCCGTCTGACTTCGACGCTAAAGCGAAAGAGTTCATCGCTAAACTGCAGGCTAATC
CGGCGAAAATCGCCAGCCGTAAAGCGTCTCAGAATGCTATCGAAGCGTTCGGTCCGCTGTTG
CCGGAATTCCCTCGGCGGTTCTGCTGACCTGGCGCCGTCTAACCTGACCTGTGGTCTGGTTC
TAAAGCAATCAACGAAGATGCTGCGGGTAACTACATCCACTACGGTGTTCGCGAGTTCGGTA
TGACCGCGATTGCTAACGGTATCTCCCTGCACGGTGGCTTCCCTGCCGTACACCTCCACCTTC
CTGATGTTTCGTGGAATACGCACGTAACGCCGTACGTATGGCTGCGCTGATGAAACAGCGTCA
GGTATGGTTTACACCCACGACTCCATCGGTCTGGGCGAAGACGGCCCGACTCACCAGCCGG
TTGAGCAGGTCGCTTCTCTGCGCGTAACCCCGAACATGTCTACATGGCGTCCGTGTGACCAG
GTTGAATCCGCGGTGCGGTGGAATACGGTGTTGAGCGTCAGGACGGCCCGACCGCACTGAT
CCTCTCCCGTCAGAACCTGGCGCAGCAGGAACGAACCTGAAGAGCAACTGGCAAACATCGCGC
GCGGTGGTTATGTGCTGAAAGACTGCGCCGGTCAGCCGGAACCTGATTTTCATCGCTACCGGT
TCAGAAAGTTGAACTGGCTGTTGCTGCCTACGAAAACTGACTGCCGAAGGCGTGAAAGCGCG
CGTGGTGTCCATGCCGTCTACCGACGCATTTGACAAGCAGGATGCTGCTTACCGTGAATCCG
TACTGCCGAAAGCGGTTACTGCACGCGTTGCTGTAGAAGCGGGTATTGCTGACTACTGGTAC
AAGTATGTTGGCCTGAACGGTGCTATCGTCGGTATGACCACCTTCGGTGAATCTGCTCCGGC
AGAGCTGCTGTTTGAAGAGTTCGGCTTCACTGTTGATAACGTTGTTGCGAAAGCAAAAGAAC
TGCTGTAA

SEQ ID NO:58

ATGTCCTCACGTAAAGAGCTTGCCAATGCTATTCGTGCGCTGAGCATGGACGCGAGTACAGAA
AGCCAAATCCGGTCACCCGGGTGCCCCCTATGGGTATGGCTGACATTGCCGAAGTCCTGTGGC
GTGATTTCTGAAACACAACCCGCGAGAATCCGTCCTGGGCTGACCGTGACCGCTTCGTGCTG
TCCAACGGCCACGGCTCCATGCTGATCTACAGCCTGCTGCACCTCACC GGTTACGATCTGCC
GATGGAAGAACTGAAAACTTCCGTCAGCTGCACTCTAAAACTCCGGGTACCCGGAAGTGG
GTTACACCGCTGGTGTGGAACCACCACCGGTCCGCTGGGTGAGGGTATTGCCAACGCGAGTC
GGTATGGCGATTGCAGAAAAACGCTGGCGGCGCAGTTTAACCGTCCGGGCCACGACATTGT
CGACCACTACACCTACGCCCTTCATGGGCGACGGCTGCATGATGGAAGGCATCTCCACGAAG
TTTGCTCTCTGGCGGGTACGCTGAAGCTGGGTAACTGATTGCATTCTACGATGACAACGGT
ATTTCTATCGATGGTCACGTTGAAGGCTGGTTCACCGACGACACCGCAATGCGTTTCTGAAGC
TTACGGCTGGCACGTTATTCGCGACATCGACGGTCATGACGCGGCATCTATCAAACGCGCAG
TAGAAGAAGCGCGCGCAGTGA CTGACAAACCTTCCCTGCTGATGTGCAAACCATCATCGGT
TTCGGTTCCCCGAACAAAGCCGGTACCCACGACTCCCACGGTGCGCCGCTGGGCGACGCTGA
AATTGCCCTGACCCGCGAACA ACTGGGCTGGAAATATGCGCCGTTGAAATCCCGTCTGAAA
TCTATGCTCAGTGGGATGCGAAAGAAGCAGGCCAGGCGAAAGAATCCGCATGGAACGAGAAA
TTCGCTGCTTACGCGAAAGCTTATCCGCGAGGAAGCCGCTGAATTTACCCGCCGTATGAAAGG
CGAAATGCCGTCTGACTTCGACGCTAAAGCGAAAGAGTTCATCGCTAAACTGCAGGCTAATC

CGGCGAAAATCGCCAGCCGTAAAGCGTCTCAGAATGCTATCGAAGCGTTCGGTCCGCTGTTG
CCGGAATTCTCGGCGGTTCTGCTGACCTGGCGCCGTCTAACCTGACCCTGTGGTCTGGTTC
TAAAGCAATCAACGAAGATGCTGCGGGTAACTACATCCACTACGGTGTTTCGCGAGTTCGGTA
TGACCGCGATTGCTAACGGTATCTCCCTGCACGGTGGCTTCCCTGCCGTACACCTCCACCTTC
CTGATGTTCTGTGGAATACGCACGTAACGCCGTACGTATGGCTGCGCTGATGAAACAGCGTCA
GGTGATGGTTTACACCCACGACTCCATCGGTCTGGGCGAAGACGGCCCGACTCACCAGCCGG
TTGAGCAGGTCGCTTCTCTGCGCGTAACCCCGAACATGTCTACATGGCGTCCGTGTGACCAG
GTTGAATCCGCGGTGCGGTGGAAATACGGTGTTGAGCGTCAGGACGGCCCCGACCGCACTGAT
CCTCTCCCGTCAGAACCTGGCGCAGCAGGAACGA ACTGAAGAGCAACTGGCAAACATCGCGC
GCGGTGGTTATGTGCTGAAAGACTGCGCCGGTCAGCCGGA ACTGATTTTCATCGCTACCGGT
TCAGAAGTTGAACTGGCTGTTGCTGCCTACGAAAACTGACTGCCGAAGGCGTGAAAGCGCG
CGTGGTGTCCATGCCGTCTACCGACGCATTTGACAAGCAGGATGCTGCTTACCGTGAATCCG
TACTGCCGAAAGCGGTTACTGCACGCGTTGCTGTAGAAGCGGGTATTGCTGACTACTGGTAC
AAGTATGTTGGCCTGAACGGTGCTATCGTCGGTATGACCACCTTCCTGGAAGTGCTGTTTCA
GGGTCCGGGTGAATCTGCTCCGGCAGAGCTGCTGTTTGAAGAGTTCGGCTTCACTGTTGATA
ACGTTGTTGCGAAAGCAAAAGAACTGCTGTAA

SEQ ID NO:59

ATGTCCTCACGTAAAGAGCTTGCCAATGCTATTCGTGCGCTGAGCATGGACGCAGTACAGAA
 AGCCAAATCCGGTCACCCGGGTGCCCCATATGGGTATGGCTGACATTGCCGAAGTCCTGTGGC
 GTGATTTCTCTGAAACACAACCCGCAGAATCCGTCCTGGGCTGACCGTGACCGCTTCGTGCTG
 TCCAACGGCCACGGCTCCATGCTGATCTACAGCCTGCTGCACCTCACCGGTTACGATCTGCC
 GATGGAAGAACTGAAAACTTCCGTGAGCTGCACTCTAAAACTCCGGGTACCCGGAAGTGG
 GTTACACCGCTGGTGTGGAAACCACCACCGGTCCGCTGGGTGAGGGTATTGCCAACGCAGTC
 GGTATGGCGATTGCAGAAAAACGCTGGCGGCGCAGTTTAAACCGTCCGGGGCCACGACATTGT
 CGACCACTACACCTACGCCTTCATGGGCGACGGCTGCATGATGGAAGGCATCTCCACGAAG
 TTTGCTCTCTGGCGGGTACGCTGAAGCTGGGTAACTGATTGCATTCTACGATGACAACGGT
 ATTTCTATCGATGGTCACGTTGAAGGCTGGTTACCGACGACACCGCAATGCGTTTTCGAAGC
 TTACGGCTGGCACGTTATTTCGCGACATCGACGGTCATGACGCGGCATCTATCAAACGCGCAG
 TAGAAGAAGCGCGCGCAGTGACTGACAAACCTTCCCTGCTGATGTGCAAACCATCATCGGT
 TTCGGTTCCCCGAACAAAGCCGGTACCCACGACTCCCACGGTGCGCCGCTGGGCGACGCTGA
 AATTGCCCTGACCCGCGAACAACCTGGGCTGGAAATATGCGCCGTTGAAATCCCGTCTGAAA
 TCTATGCTCAGTGGGATGCGAAAGAAGCAGGCCAGGCGAAAGAATCCGCATGGAACGAGAAA
 TTCGCTGCTTACGCGAAAGCTTATCCGCAGGAAGCCGCTGAATTTACCCGCCGTATGAAAGG
 CGAAATGCCGTCTGACTTCGACGCTAAAGCGAAAGAGTTCATCGCTAAACTGCAGGCTAATC
 CGGCGAAAAATCGCCAGCCGTAAAGCGTCTCAGAAATGCTATCGAAGCGTTCCGGTCCGCTGTTG
 CCGGAATTCCCTCGGCGGTTCTGCTGACCTGGCGCCGTCTAACCTGACCCTGTGGTCTGGTTC
 TAAAGCAATCAACGAAGATGCTGCGGGTAACTACATCCACTACGGTGTTCGCGAGTTCGGTA
 TGACCGCGATTGCTAACGGTATCTCCCTGCACGGTGGCTTCCTGCCGTACACCTCCACCTTC
 CTGATGTTTCGTGGAATACGCACGTAACGCCGTACGTATGGCTGCGCTGATGAAACAGCGTCA
 GGTGATGGTTTACACCCACGACTCCATCGGTCTGGGCGAAGACGGCCCGACTCACCAGCCGG
 TTGAGCAGGTCGCTTCTCTGCGCGTAACCCCGAACATGTCTACATGGCGTCCGTGTGACCAG
 GTTGAATCCGCGGTGCGGTGGAATACGGTGTTGAGCGTCAGGACGGCCCGACCGCACTGAT
 CCTCTCCCGTCAGAACCTGGCGCAGCAGGAACGAACCTGAAGAGCAACTGGCAAACATCGCGC
 GCGGTGGTTATGTGCTGAAAGACTGCGCCGGTCAGCCGGAACCTGATTTTTCATCGCTACCGGT
 TCAGAAGTTGAACTGGCTGTTGCTGCCTACGAAAACTGACTGCCGAAGGCGTGAAAGCGCG
 CGTGGTGTCCATGCCGTCTACCGACGCATTTGACAAGCAGGATGCTGCTTACCGTGAATCCG
 TACTGCCGAAAGCGGTTACTGCACGCGTTGCTGTAGAAGCGGGTATTGCTGACTACTGGTAC
 AAGTATGTTGGCCTGAACGGTGCTATCGTCGGTATGACCACCTTCGGTCTGGAAGTGCTGTT

TCAGGGTCCGGAATCTGCTCCGGCAGAGCTGCTGTTTGAAGAGTTCGGCTTCACTGTTGATA
 ACGTTGTTGCGAAAGCAAAAGAACTGCTGTAA

SEQ ID NO:60

ATGTCCTCACGTAAAGAGCTTGCCAATGCTATTCGTGCGCTGAGCATGGACGCAGTACAGAA
AGCCAAATCCGGTCACCCGGGTGCCCCATATGGGTATGGCTGACATTGCCGAAGTCCTGTGGC
GTGATTTCTCTGAAACACAACCCGCAGAATCCGTCCTGGGCTGACCGTGACCGCTTCGTGCTG
TCCAACGGCCACGGCTCCATGCTGATCTACAGCCTGCTGCACCTCACCGGTTACGATCTGCC
GATGGAAGAACTGAAAACTTCCGTCAGCTGCACTCTAAAACTCCGGGTACCCGGAAGTGG
GTTACACCGCTGGTGTGGAAACCACCACCGGTCCGCTGGGTGAGGGTATTGCCAACGCAGTC
GGTATGGCGATTGCAGAAAAACGCTGGCGGCGCAGTTTAACCGTCCGGGGCCACGACATTGT
CGACCACTACACCTACGCCTTCATGGGCGACGGCTGCATGATGGAAGGCATCTCCACGAAG
TTTGCTCTCTGGCGGGTACGCTGAAGCTGGGTAAACTGATTGCATTCTACGATGACAACGGT
ATTTCTATCGATGGTCACGTTGAAGGCTGGTTACCCGACGACACCGCAATGCGTTTTCGAAGC
TTACGGCTGGCACGTTATTCGCGACATCGACGGTCATGACGCGGCATCTATCAAACGCGCAG
TAGAAGAAGCGCGCGCAGTGACTGACAAACCTTCCCTGCTGATGTGCAAAACCATCATCGGT
TTCGGTTCCCCGAACAAAGCCGGTACCCACGACTCCCACGGTGCGCCGCTGGGCGACGCTGA
AATTGCCCTGACCCGCGAACAACCTGGGCTGGAAATATGCGCCGTTGAAATCCCGTCTGAAA
TCTATGCTCAGTGGGATGCGAAAGAAGCAGGCCAGGCGAAAGAATCCGCATGGAACGAGAAA
TTCGCTGCTTACGCGAAAGCTTATCCGCAGGAAGCCGCTGAATTTACCCGCCGTATGAAAGG
CGAAATGCCGTCTGACTTCGACGCTAAAGCGAAAGAGTTCATCGCTAAACTGCAGGCTAATC
CGGCGAAAAATCGCCAGCCGTAAAGCGTCTCAGAATGCTATCGAAGCGTTCCGTCGCTGTTG
CCGGAATTCCCTCGGCGGTTCTGCTGACCTGGCGCCGTCTAACCTGACCCTGTGGTCTGGTTC
TAAAGCAATCAACGAAGATGCTGCGGGTAACTACATCCACTACGGTGTTCGCGAGTTCGGTA
TGACCGCGATTGCTAACGGTATCTCCCTGCACGGTGGCTTCCTGCCGTACACCTCCACCTTC
CTGATGTTCTGTGGAATACGCACGTAACGCCGTACGTATGGCTGCGCTGATGAAACAGCGTCA
GGTGATGGTTTACACCCACGACTCCATCGGTCTGGGCGAAGACGGCCCGACTCACCAGCCGG
TTGAGCAGGTCGCTTCTCTGCGCGTAACCCCGAACATGTCTACATGGCGTCCGTGTGACCAG
GTTGAATCCGCGGTCGCGTGGAATACGGTGTTGAGCGTCAGGACGGCCCGACCGCACTGAT
CCTCTCCCGTCAGAACCTGGCGCAGCAGGAACGAACCTGAAGAGCAACTGGCAAACATCGCGC
GCGGTGGTTATGTGCTGAAAGACTGCGCCGGTCAGCCGGAACCTGATTTTTCATCGCTACCGGT
TCAGAAGTTGAACTGGCTGTTGCTGCCTACGAAAACTGACTGCCGAAGGCGTGAAAGCGCG
CGTGGTGTCCATGCCGTCTACCGACGCATTTGACAAGCAGGATGCTGCTTACCGTGAATCCG
TACTGCCGAAAGCGGTTACTGCACGCGTTGCTGTAGAAGCGGGTATTGCTGACTACTGGTAC
AAGTATGTTGGCCTGAACGGTGCTATCGTCGGTATGACCACCTTCGGTGAACCTGGAAGTGCT
GTTTCAGGGTCCGTCTGCTCCGGCAGAGCTGCTGTTTGAAGAGTTCGGCTTCACTGTTGATA
ACGTTGTTGCGAAAGCAAAAGAACTGCTGTAA

5

SEQ ID NO:61

ATGTCCTCACGTAAAGAGCTTGCCAATGCTATTCGTGCGCTGAGCATGGACGCAGTACAGAA
AGCCAAATCCGGTCACCCGGGTGCCCCATATGGGTATGGCTGACATTGCCGAAGTCCTGTGGC
GTGATTTCTCTGAAACACAACCCGCAGAATCCGTCCTGGGCTGACCGTGACCGCTTCGTGCTG
TCCAACGGCCACGGCTCCATGCTGATCTACAGCCTGCTGCACCTCACCGGTTACGATCTGCC
GATGGAAGAACTGAAAACTTCCGTCAGCTGCACTCTAAAACTCCGGGTACCCGGAAGTGG
GTTACACCGCTGGTGTGGAAACCACCACCGGTCCGCTGGGTGAGGGTATTGCCAACGCAGTC
GGTATGGCGATTGCAGAAAAACGCTGGCGGCGCAGTTTAACCGTCCGGGGCCACGACATTGT
CGACCACTACACCTACGCCTTCATGGGCGACGGCTGCATGATGGAAGGCATCTCCACGAAG
TTTGCTCTCTGGCGGGTACGCTGAAGCTGGGTAAACTGATTGCATTCTACGATGACAACGGT
ATTTCTATCGATGGTCACGTTGAAGGCTGGTTACCCGACGACACCGCAATGCGTTTTCGAAGC

TTACGGCTGGCACGTTATTTCGCGACATCGACGGTCATGACGCGGCATCTATCAAACGCGCAG
TAGAAGAAGCGCGCGCAGTGACTGACAAACCTTCCCTGCTGATGTGCAAACCATCATCGGT
TTCGGTTCCCCGAACAAAGCCGGTACCCACGACTCCCACGGTGCGCCGCTGGGCGACGCTGA
AATTGCCCTGACCCGCGAACAACTGGGCTGGAAATATGCGCCGTTTCGAAATCCCGTCTGAAA
TCTATGCTCAGTGGGATGCGAAAGAAGCAGGCCAGGCGAAAGAATCCGCATGGAACGAGAAA
TTCGCTGCTTACGCGAAAGCTTATCCGCAGGAAGCCGCTGAATTTACCCGCCGTATGAAAGG
CGAAATGCCGTCTGACTTCGACGCTAAAGCGAAAGAGTTCATCGCTAAACTGCAGGCTAATC
CGGCGAAAAATCGCCAGCCGTAAAGCGTCTCAGAATGCTATCGAAGCGTTCGGTCCGCTGTTG
CCGGAATTCCCTCGGCGGTTCTGCTGACCTGGCGCCGTCTAACCTGACCCTGTGGTCTGGTTC
TAAAGCAATCAACGAAGATGCTGCGGGTAACTACATCCACTACGGTGTTTCGCGAGTTCGGTA
TGACCGCGATTGCTAACGGTATCTCCCTGCACGGTGGCTTCCCTGCCGTACACCTCCACCTTC
CTGATGTTTCGTGGAATACGCACGTAACGCCGTACGTATGGCTGCGCTGATGAAACAGCGTCA
GGTGATGGTTTACACCCACGACTCCATCGGTCTGGGCGAAGACGGCCCGACTCACCAGCCGG
TTGAGCAGGTCGCTTCTCTGCGCGTAACCCCGAACATGTCTACATGGCGTCCGTGTGACCAG
GTTGAATCCGCGGTCGCGTGGAATAACGGTGTTGAGCGTCAGGACGGCCCGACCGCACTGAT
CCTCTCCCGTCAGAACCTGGCGCAGCAGGAACGAACCTGAAGAGCAACTGGCAAACATCGCGC
GCGGTGGTTATGTGCTGAAAGACTGCGCCGGTCAGCCGGAACCTGATTTTCATCGCTACCGGT
TCAGAAGTTGAACTGGCTGTTGCTGCCTACGAAAACTGACTGCCGAAGGCGTGAAAGCGCG
CGTGGTGTCCATGCCGTCTACCGACGCATTTGACAAGCAGGATGCTGCTTACCGTGAATCCG
TACTGCCGAAAGCGGTTACTGCACGCGTTGCTGTAGAAGCGGGTATTGCTGACTACTGGTAC
AAGTATGTTGGCCTGAACGGTGCTATCGTCGGTATGACCACCTTCGGTGAATCTCTGGAAGT
GCTGTTTCAGGGTCCGGCTCCGGCAGAGCTGCTGTTTGAAGAGTTCGGCTTCACTGTTGATA
ACGTTGTTGCGAAAGCAAAAGAACTGCTGTAA

SEQ ID NO:62

ATGTCCTCACGTAAAGAGCTTGCCAATGCTATTCGTGCGCTGAGCATGGACGCAGTACAGAA
AGCCAAATCCGGTCACCCGGGTGCCCCTATGGGTATGGCTGACATTGCCGAAGTCCTGTGGC
GTGATTTCTGAAACACAACCCGCAGAATCCGTCCTGGGCTGACCGTGACCGCTTCGTGCTG
TCCAACGGCCACGGCTCCATGCTGATCTACAGCCTGCTGCACCTCACCGGTTACGATCTGCC
GATGGAAGAACTGAAAACTTCCGTCAGCTGCACTCTAAACTCCGGGTACCCGGAAGTGG
GTTACACCGCTGGTGTGGAAACCACCACCGGTCCGCTGGGTGAGGGTATTGCCAACGCAGTC
GGTATGGCGATTGCAGAAAAACGCTGGCGGCGCAGTTTAACCGTCCGGGCCACGACATTGT
CGACCACTACACCTACGCCTTCATGGGCGACGGCTGCATGATGGAAGGCATCTCCACGAAG
TTTGCTCTCTGGCGGGTACGCTGAAGCTGGGTAAACTGATTGCATTCTACGATGACAACGGT
ATTTCTATCGATGGTCACGTTGAAGGCTGGTTCACCGACGACACCGCAATGCGTTTCAAGC
TTACGGCTGGCACGTTATTTCGCGACATCGACGGTCATGACGCGGCATCTATCAAACGCGCAG
TAGAAGAAGCGCGCGCAGTGACTGACAAACCTTCCCTGCTGATGTGCAAACCATCATCGGT
TTCGGTTCCCCGAACAAAGCCGGTACCCACGACTCCCACGGTGCGCCGCTGGGCGACGCTGA
AATTGCCCTGACCCGCGAACAACCTGGGCTGGAAATATGCGCCGTTGAAATCCCGTCTGAAA
TCTATGCTCAGTGGGATGCGAAAGAAGCAGGCCAGGCGAAAGAATCCGCATGGAACGAGAAA
TTCGCTGCTTACGCGAAAGCTTATCCGCAGGAAGCCGCTGAATTTACCCGCCGTATGAAAGG
CGAAATGCCGTCTGACTTCGACGCTAAAGCGAAAGAGTTCATCGCTAAACTGCAGGCTAATC
CGGCGAAAAATCGCCAGCCGTAAAGCGTCTCAGAAATGCTATCGAAGCGTTCGGTCCGCTGTTG
CCGGAATTCCCTCGGCGGGTCTGCTGACCTGGCGCCGTCTAACCTGACCCTGTGGTCTGGTTC
TAAAGCAATCAACGAAGATGCTGCGGGTAACTACATCCACTACGGTGTTTCGCGAGTTCGGTA
TGACCGCGATTGCTAACGGTATCTCCCTGCACGGTGGCTTCCTGCCGTACACCTCCACCTTC
CTGATGTTTCGTGGAATACGCACGTAACGCCGTACGTATGGCTGCGCTGATGAAACAGCGTCA
GGTGATGGTTTTACACCCACGACTCCATCGGTCTGGGCGAAGACGGCCCGACTCACCAGCCGG
TTGAGCAGGTCGCTTCTCTGCGCGTAACCCCGAACATGTCTACATGGCGTCCGTGTGACCAG
GTTGAATCCGCGGTGCGGTGGAAATACGGTGTTGAGCGTCAGGACGGCCCGACCGCACTGAT

CCTCTCCCGTCAGAACCTGGCGCAGCAGGAACGAACTGAAGAGCAACTGGCAAACATCGCGC
GCGGTGGTTATGTGCTGAAAGACTGCGCCGGTCAGCCGGAAGTATTTTCATCGCTACCGGT
TCAGAAGTTGAACTGGCTGTTGCTGCCTACGAAAACTGACTGCCGAAGGCGTGAAAGCGCG
CGTGGTGTCCATGCCGTCTACCGACGCATTTGACAAGCAGGATGCTGCTTACCGTGAATCCG
TACTGCCGAAAGCGGTTACTGCACGCGTTGCTGTAGAAGCGGGTATTGCTGACTACTGGTAC
AAGTATGTTGGCCTGAACGGTGCTATCGTCCGTATGACCACCTTCGGTGAATCTGCTCCGCT
GGAAGTGCTGTTTCAGGGTCCGGCAGAGCTGCTGTTTGAAGAGTTCGGCTTCACTGTTGATA
ACGTTGTTGCGAAAGCAAAAGAACTGCTGTAA

5 SEQ ID NO:63

MSSRKELANAI RALSMDAVQKAKSGHPGAPMGMA DIAEVLWRDFLKHNPQNPSWADRDRFVL
SNHGSM LIYSLHLTG YDLPMEELKNFRQLH SKTPGHPEVG YTAGVETTTG PLGQGIANAV
GMAIAEKT LAAQFNRP GHDIVDHYTYAFMGD GCMMEG ISHEVCSLAGTLKLGLKLI AFYDDNG
ISIDGHVEGWFTDDTAMRFEAYGWHVIRIDIDGHDAASIKRAVEEARAVTDKPSLLMCKTIIG
FGSPNKAGTHDSHGAPLGDAEIALTREQLGWKYAPFEIPSEIYAQWDAKEAGQAKESAWNEK
FAAYAKAYPQEAAEFTRRMKGEMPSDFDAKAKEFI AKLQANPAKIASRKASQNAIEAFGPLL
PEFLGGSADLAPSNLTLWSGSKAINEDAAGNYIHYGVREFGMTAIANGISLHGGLPYTSTF
LMFVEYARNAVRMAALMKQRQVMVYTHDSIGLGEDGPTHQPVEQIASLRVTPNMSTWRPCDQ
VESAVAWKYGVERQDGPTALILSRQNL AQQERTEEQ LANIARGGYVLKDCAGQPELIFIATG
SEVELAVAA YEKLTAEGVKARVVSMPSTDAFDKQDAAYRESVLPKAVTARVAVEAGIADYWY
KYVGLNGAIVGMTTFGESAPAEQLFEEFGFTVDNVVAKAKALL

SEQ ID NO:64

MSSRKELANAI RALSMDAVQKAKSGHPGAPMGMADIAEVLWRDFLKHNPQNPSWADRDRFVL
 SNHGGSMLIYSLHLTLGYDLPMEELKNFRQLHSKTPGHPEVG YTAGVETTTGPLGQGIANAV
 GMAIAEKT LAAQFNRP GHDIVDHYTYAFMGDGCMMEGISHEVCSLAGTLKLGLKLI AFYDDNG
 ISIDGHVEGWFTDDTAMRFEAYGWHVIRDIDGHDAASIKRAVEE ARAVTDKPSLLMCKTIIG
 FGSPNKAGTHDSHGAPLGDAEIALTREQLGWKYAPFEIPSEIYAQWDAKEAGQAKESAWNEK
 FAAYAKAYPQEAAEFTRRMKGEMP SDFDAKAKEFI AKLQANPAKIASRKASQNAIEAFGPLL
 PEFLGGSADLAPSNLTLWSGSKAINEDAAGNYIHYGVREFGMTAIANGISLHGGFLPYTSTF
 LMFVEYARNAVRMAALMKQRQVMVYTHDSIGLGEDGPTHQPVEQVASLRVTPNMSTWRPCDQ
 VESAVAWKYGVERQDGPTALILSRQNL AQQERTEEQLANIARGGYVLKDCAGQPELIFIATG
 SEVELAVAAYEKLTAEGVKARVVSMPSTDAFDKQDAAYRESVLPKAVTARVAVEAGIADYWY
 KYVGLNGAIVGMTTFLEVL FQGPGESAPAE LLFE EFGFTVDNVVAKAKELL

5

SEQ ID NO:65

MSSRKELANAI RALSMDAVQKAKSGHPGAPMGMADIAEVLWRDFLKHNPQNPSWADRDRFVL
 SNHGGSMLIYSLHLTLGYDLPMEELKNFRQLHSKTPGHPEVG YTAGVETTTGPLGQGIANAV
 GMAIAEKT LAAQFNRP GHDIVDHYTYAFMGDGCMMEGISHEVCSLAGTLKLGLKLI AFYDDNG
 ISIDGHVEGWFTDDTAMRFEAYGWHVIRDIDGHDAASIKRAVEE ARAVTDKPSLLMCKTIIG
 FGSPNKAGTHDSHGAPLGDAEIALTREQLGWKYAPFEIPSEIYAQWDAKEAGQAKESAWNEK
 FAAYAKAYPQEAAEFTRRMKGEMP SDFDAKAKEFI AKLQANPAKIASRKASQNAIEAFGPLL
 PEFLGGSADLAPSNLTLWSGSKAINEDAAGNYIHYGVREFGMTAIANGISLHGGFLPYTSTF
 LMFVEYARNAVRMAALMKQRQVMVYTHDSIGLGEDGPTHQPVEQVASLRVTPNMSTWRPCDQ
 VESAVAWKYGVERQDGPTALILSRQNL AQQERTEEQLANIARGGYVLKDCAGQPELIFIATG
 SEVELAVAAYEKLTAEGVKARVVSMPSTDAFDKQDAAYRESVLPKAVTARVAVEAGIADYWY
 KYVGLNGAIVGMTTFGLEVL FQGPESAPAE LLFE EFGFTVDNVVAKAKELL

10

SEQ ID NO:66

MSSRKELANAI RALSMDAVQKAKSGHPGAPMGMADIAEVLWRDFLKHNPQNPSWADRDRFVL
 SNHGGSMLIYSLHLTLGYDLPMEELKNFRQLHSKTPGHPEVG YTAGVETTTGPLGQGIANAV
 GMAIAEKT LAAQFNRP GHDIVDHYTYAFMGDGCMMEGISHEVCSLAGTLKLGLKLI AFYDDNG
 ISIDGHVEGWFTDDTAMRFEAYGWHVIRDIDGHDAASIKRAVEE ARAVTDKPSLLMCKTIIG
 FGSPNKAGTHDSHGAPLGDAEIALTREQLGWKYAPFEIPSEIYAQWDAKEAGQAKESAWNEK
 FAAYAKAYPQEAAEFTRRMKGEMP SDFDAKAKEFI AKLQANPAKIASRKASQNAIEAFGPLL
 PEFLGGSADLAPSNLTLWSGSKAINEDAAGNYIHYGVREFGMTAIANGISLHGGFLPYTSTF
 LMFVEYARNAVRMAALMKQRQVMVYTHDSIGLGEDGPTHQPVEQVASLRVTPNMSTWRPCDQ
 VESAVAWKYGVERQDGPTALILSRQNL AQQERTEEQLANIARGGYVLKDCAGQPELIFIATG
 SEVELAVAAYEKLTAEGVKARVVSMPSTDAFDKQDAAYRESVLPKAVTARVAVEAGIADYWY
 KYVGLNGAIVGMTTFGELEVL FQGPSAPAE LLFE EFGFTVDNVVAKAKELL

SEQ ID NO:67

MSSRKELANAI RALSMDAVQKAKSGHPGAPMG MADI AEVLWRDFLKHNPQNPSWADRDRFVL
SNGHGSMLIYSL LHLTG YDLPMEELKNFRQLH SKTPGHPEVGYTAGVETTTGPLGQG IANAV
GMAIAEKT LAAQFNRP GHDIVDHYTYAFMGDGCMMEGISHEVC SLAGTLKLGLKLI AFYDDNG
ISIDGHVEGWFTDDTAMRFEAYGWHVIRDI DGHDAASIKRAVEE EARA VTDKPSLLMCKTIIG
FGSPNKAGTHDSHGAPLGDAEIALTREQLGWKYAPFEIPSEIYAQWDAKEAGQAKESAWNEK
FAAYAKAYPQEAAEFTRRMKGEMPSDFDAKAKEFI AKLQANPAKIASRKASQNAIEAFGPLL
PEFLGGSADLAPSNLTLWSGSKAINEDAAGNYIHYGVREFGMTAIANGISLHGGFLPYTSTF
LMFVEYARNAVRMAALMKQRQVMVYTHDSIGLGEDGPTHQPVEQVASLRVTPNMSTWRPCDQ
VESAVAWKYGVERQDGPTALILSRQNL AQQERTEEQ LANIARGGYVLKDCAGQPELIFIATG
SEVELAVAAYEKLTAEGVKARVVSMPSTDAFDKQDAAYRESVLPKAVTARVAVEAGIADYWY
KYVGLNGAIVGMTTFGESLEVL FQGPAPAE LLFE EFGFTVDNVVAKAKELL

5 SEQ ID NO:68

MSSRKELANAI RALSMDAVQKAKSGHPGAPMG MADI AEVLWRDFLKHNPQNPSWADRDRFVL
SNGHGSMLIYSL LHLTG YDLPMEELKNFRQLH SKTPGHPEVGYTAGVETTTGPLGQG IANAV
GMAIAEKT LAAQFNRP GHDIVDHYTYAFMGDGCMMEGISHEVC SLAGTLKLGLKLI AFYDDNG
ISIDGHVEGWFTDDTAMRFEAYGWHVIRDI DGHDAASIKRAVEE EARA VTDKPSLLMCKTIIG
FGSPNKAGTHDSHGAPLGDAEIALTREQLGWKYAPFEIPSEIYAQWDAKEAGQAKESAWNEK
FAAYAKAYPQEAAEFTRRMKGEMPSDFDAKAKEFI AKLQANPAKIASRKASQNAIEAFGPLL
PEFLGGSADLAPSNLTLWSGSKAINEDAAGNYIHYGVREFGMTAIANGISLHGGFLPYTSTF
LMFVEYARNAVRMAALMKQRQVMVYTHDSIGLGEDGPTHQPVEQVASLRVTPNMSTWRPCDQ
VESAVAWKYGVERQDGPTALILSRQNL AQQERTEEQ LANIARGGYVLKDCAGQPELIFIATG
SEVELAVAAYEKLTAEGVKARVVSMPSTDAFDKQDAAYRESVLPKAVTARVAVEAGIADYWY
KYVGLNGAIVGMTTFGESAPLEVL FQGP AE LLFE EFGFTVDNVVAKAKELL

10 SEQ ID NO:69
CTGGAAGTGCTGTTTCAGGGTCCG

SEQ ID NO:70
MKNINPTQTAA

15 SEQ ID NO:71
ATGAAAAACATCAATCCAACGCAGACCGCTGCC

SEQ ID NO:72
MLEVL FQGPKNINPTQTAA

20 SEQ ID NO:73
ATGCTGGAAGTGCTGTTTCAGGGTCCGAAAAACATCAATCCAACGCAGACCGCTGCC

25 SEQ ID NO:74
MKLEVL FQGPININPTQTAA

SEQ ID NO:75
ATGAAACTGGAAGTGCTGTTTCAGGGTCCGAACATCAATCCAACGCAGACCGCTGCC

30 SEQ ID NO:76
MKNLEVL FQGPININPTQTAA

SEQ ID NO:77
ATGAAAAACCTGGAAGTGCTGTTTCAGGGTCCGATCAATCCAACGCAGACCGCTGCC

35 SEQ ID NO:78
MLEVL FQGPA

SEQ ID NO:79
ATGCTGGAAGTGCTGTTTCAGGGTCCGGCTGCC

5 SEQ ID NO:80
MKLEVLFGGPA

SEQ ID NO:81
ATGAAACTGGAAGTGCTGTTTCAGGGTCCGGCC

10 SEQ ID NO:82
MKNLEVLFGGP

15 SEQ ID NO:83
ATGAAAAACCTGGAAGTGCTGTTTCAGGGTCCG

SEQ ID NO:84
MKNINLEVLFGGPTQTAA

20 SEQ ID NO:85
ATGAAAAACATCAATCTGGAAGTGCTGTTTCAGGGTCCAACGCAGACCGCTGCC

SEQ ID NO:86
TAAWQALEVLFGGPQKH

25 SEQ ID NO:87
ACCGCTGCCTGGCAGGCACTAGAAGTGCTGTTTCAGGGTCCGCAGAAACAC

LISTADO DE SECUENCIAS

30 <110> GreenLight Biosciences, Inc.
Blake, William J
Cunningham, Drew S

35 <120> PROTEÍNAS GENOMODIFICADAS CON UN SITIO DE ESCISIÓN DE PROTEASA

<130> G0830.70013WO00

40 <140> Todavía sin asignar
<141> 05/08/2014

<150> US 61/987.518
<151> 02/05/2014

45 <150> US 61/862.363
<151> 05/08/2013

<160> 87

50 <170> PatentIn versión 3.5

<210> 1
<211> 1650
<212> ADN

55 <213> Secuencia artificial

<220>
<223> Oligonucleótido sintético

60 <400> 1

ES 2 704 697 T3

atgaaaaaca tcaatccaac gcagaccgct gcctggcagg cactacagaa acacttcgat	60
gaaatgaaag acgttacgat cgccgatctt tttgctaaag acggcgatcg tttttctaag	120
ttctccgcaa ccttcgacga tcagatgctg gtggattact ccaaaaaccg catcactgaa	180
gagacgctgg cgaaattaca ggatctggcg aaagagtgcg atctggcggg cgcgattaag	240
tcgatgttct ctggcgagaa gatcaaccgc actgaaaacc gcgccgtgct gcacgtagcg	300
ctgcgtaacc gtagcaatac cccgattttg gttgatggca aagacgtaat gccggaagtc	360
aacgcggtgc tggagaagat gaaaaccttc tcagaagcga ttatttccgg tgagtggaaa	420
ggttataccg gcaaagcaat cactgacgta gtgaacatcg ggatcggcgg ttctgacctc	480
ggcccataca tggtgaccga agctctgcgt ccgtacaaaa accacctgaa catgcacttt	540
gtttctaacg tcgatgggac tcacatcgcg gaagtgcgtg aaaaagtaaa cccggaacc	600
acgctgttct tggtagcatc taaaaccttc accactcagg aaactatgac caacgcccat	660
agcgcgcgtg actggttcct gaaagcggca ggtgatgaaa aacacgttg caaaacacttt	720
gcggcgcttt ccaccaatgc caaagccgtt ggcgagtttg gtattgatac tgccaacatg	780
ttcgagttct gggactgggt tggcggccgt tactctttgt ggtcagcgat tggcctgtcg	840
attgttctct ccatcggtt tgataacttc gttgaactgc tttccggcgc acacgcgatg	900
gacaagcatt tctccaccac gcctgccgag aaaaacctgc ctgtactgct ggcgctgatt	960
ggcatctggt acaacaattt ctttggtgcg gaaactgaag cgattctgcc gtatgaccag	1020
tatatgcacc gtttcgcggc gtacttccag cagggcaata tggagtccaa cggttaagtat	1080
gttgaccgta acggtaacgt tgtggattac cagactggcc cgattatctg gggatgaacca	1140
ggcactaacg gtcagcacgc gttctaccag ctgatccacc aggaaccaa aatggtaccg	1200
tgcgatttca tcgctccggc tatcaccat aaccgcgtct ctgatcatca ccagaaactg	1260
ctgtctaact tcttcgcca gaccgaagcg ctggcgtttg gtaaattccc cgaagtgggt	1320
gagcaggaat atcgtgatca gggtaaagat ccggcaacgc ttgactacgt ggtgccgttc	1380
aaagtattcg aaggtaaccg cccgaccaac tccatcctgc tgcgtgaaat cactccgttc	1440
agcctgggtg cgttgattgc gctgtatgag cacaaaatct ttactcagg cgtgatcctg	1500
aacatcttca ccttcgacca gtggggcggt gaactgggta aacagctggc gaaccgtatt	1560
ctgccagagc tgaaagatga taaagaaatc agcagccacg atagctcgac caatggctctg	1620
attaaccgct ataaagcgtg gcgcggttaa	1650

<210> 2
 <211> 1674
 <212> ADN
 <213> Secuencia artificial

<220>

<223>Oligonucleótido sintético

<400> 2

5

atgaaaaaca	tcaatccaac	gcagaccgct	gcctggcagg	cactacagaa	acacttcgat	60
gaaatgaaag	acgttacgat	cgccgatctt	tttgctaaag	atggtgatcg	tttttctaag	120
ttctccgcaa	ccttcgacga	tcagatgctg	gtggattact	ccaaaaaccg	catcactgaa	180
gagacgctgg	cgaaattaca	ggatctggcg	aaagagtgcg	atctggcggg	cgcgattaag	240
tcgatgttct	ctggcgagaa	gatcaaccgc	actgaaaacc	gcgccgtgct	gcacgtagcg	300
ctgcgtaacc	gtagcaatac	cccgttgga	gtgctgtttc	aggggtccgat	tttggttgat	360
ggcaaagacg	taatgccgga	agtcaacgcg	gtgctggaga	agatgaaaac	cttctcagaa	420
gcgattatth	ccggtgagt	gaaaggttat	accggcaaag	caatcactga	cgtagtgaac	480
atcgggatcg	gcggttctga	cctcggccca	tacatggtga	ccgaagctct	gcgtccgtac	540
aaaaaccacc	tgaacatgca	ctttgtttct	aacgtcgtatg	ggactcacat	cgcggaagt	600
ctgaaaaaag	taaaccgga	aaccacgctg	ttcttggtag	catctaaaac	cttcaccact	660
caggaaacta	tgaccaacgc	ccatagcgcg	cgtgactggt	tcctgaaagc	ggcaggtgat	720
gagaagcacg	ttgcaaaaca	ctttgcggcg	ctttccacca	atgccaaagc	cggtggcgag	780
tttggtattg	atactgccaa	catgttcgag	ttctgggact	gggttggcgg	ccgttactct	840
ttgtggtcag	cgattggcct	gtcgattgtt	ctctccatcg	gctttgataa	cttcgttgaa	900
ctgctttccg	gcgcacacgc	gatggacaag	catttctcca	ccacgcctgc	cgagaaaaac	960
ctgcctgtac	tgctggcgct	gattggcatc	tggtacaaca	atttctttgg	tgcggaact	1020
gaagcgattc	tgccgtatga	ccagtatatg	caccgtttcg	cggcgactct	ccagcagggc	1080
aatatggagt	ccaacggtaa	gtatgttgac	cgtaacggta	acgttggtga	ttaccagact	1140
ggcccgatta	tctgggggtga	accaggcact	aacggtcagc	acgcgttcta	ccagctgac	1200
caccagggaa	ccaaaatggt	accgtgcgat	ttcatcgctc	cggctatcac	ccataaccgc	1260
ctctctgac	atcaccagaa	actgctgtct	aacttcttcg	cccagaccga	agcgctggcg	1320
tttggtaaat	cccgcgaagt	ggttgagcag	gaatatcggt	atcagggtaa	agatccggca	1380
acgcttgact	acgtggtgcc	gttcaaagta	ttcgaaggta	accgcccgc	caactccatc	1440
ctgctgcgtg	aaatcactcc	gttcagcctg	ggtgcgttga	ttgcgctgta	tgagcacaaa	1500
atctttactc	agggcggtgat	cctgaacatc	ttcaccttcg	accagtgggg	cgtggaactg	1560
ggtaaacagc	tggcgaaccg	tattctgcca	gagctgaaag	atgataaaga	aatcagcagc	1620
cacgatagct	cgaccaatgg	tctgattaac	cgctataaag	cgtggcgcg	ttaa	1674

	<210> 3
	<211> 1674
	<212> ADN
5	<213> Secuencia artificial
	<220>
	<223> Oligonucleótido sintético
10	<400> 3

ES 2 704 697 T3

```

atgaaaaaca tcaatccaac gcagaccgct gcctggcagg cactacagaa acacttcgat      60
gaaatgaaag acgttacgat cgccgatctt tttgctaaag atggtgatcg tttttctaag      120
ttctccgcaa ccttcgacga tcagatgctg gtggattact ccaaaaaccg catcactgaa      180
gagacgctgg cgaaattaca ggatctggcg aaagagtgcg atctggcggg cgcgattaag      240
tcgatgttct ctggcgagaa gatcaaccgc actgaaaacc gcgccgtgct gcacgtagcg      300
ctgcgttaacc gtagcaatac cccgattctg gaagtgcgtg ttcaggggtcc gttggttgat      360
ggcaaagacg taatgccgga agtcaacgcg gtgctggaga agatgaaaac cttctcagaa      420
gcgattatth cgggtgagtg gaaaggttat accggcaaag caatcactga cgtagtgaac      480
atcgggatcg gcggttctga cctcggccca tacatggtga ccgaagctct gcgtccgtac      540
aaaaaccacc tgaacatgca ctttgtttct aacgtcgatg ggactcacat cgcggaagtg      600
ctgaaaaaag taaaccgga aaccacgctg ttcttggtag catctaaaac cttcaccact      660
caggaaacta tgaccaacgc ccatagcgcg cgtgactggt tcctgaaagc ggcaggtgat      720
gagaagcacg ttgcaaaaca ctttgcgggc ctttccacca atgccaaagc cgttggcgag      780

tttggtattg atactgccaa catgttcgag ttctgggact gggttggcgg ccgttactct      840
ttgtggtcag cgattggcct gtcgattgtt ctctccatcg gctttgataa cttcgttgaa      900
ctgctttccg gcgcacacgc gatggacaag catttctcca ccacgcctgc cgagaaaaac      960
ctgcctgtac tgctggcgct gattggcatc tggtaacaaca atttctttgg tgcggaaact     1020
gaagcgattc tgccgtatga ccagtatatg caccgtttcg cggcgtactt ccagcagggc     1080
aatatggagt ccaacggtaa gtatgttgac cgtaacggtg acgttgtgga ttaccagact     1140
ggccccgatta tctgggggtg accaggcact aacggtcagc acgcgttcta ccagctgac     1200
caccagggaa ccaaaatggt accgtgcgat ttcacgctc cggctatcac ccataaaccg     1260
ctctctgac atcaccagaa actgctgtct aacttcttcg ccagaccga agcgtggcg     1320
tttggtaaat cccgcgaagt ggttgagcag gaatatcgtg atcagggtaa agatccggca     1380
acgcttgact acgtggtgcc gttcaaagta ttcgaaggta accgcccgc caactccatc     1440
ctgctgcgtg aaatcactcc gttcagcctg ggtgcgttga ttgcgctgta tgagcacaaa     1500
atctttactc agggcgatg cctgaacatc ttcaccttcg accagtgggg cgtggaactg     1560
ggtaaacagc tggcgaaccg tattctgccg gagctgaaag atgataaaga aatcagcagc     1620
cacgatagct cgaccaatgg tctgattaac cgctataaag cgtggcgcgg ttaa      1674

```

<210> 4
 <211> 1671
 <212> ADN
 <213> Secuencia artificial

5

<220>
 <223> Oligonucleótido sintético

ES 2 704 697 T3

<400> 4

```

atgaaaaaca tcaatccaac gcagaccgct gcctggcagg cactacagaa acacttcgat      60
gaaatgaaag acgttacgat cgccgatctt tttgctaaag atggtgatcg tttttctaag      120
ttctccgcaa ccttcgacga tcagatgctg gtggattact ccaaaaaccg catcactgaa      180
gagacgctgg cgaaattaca ggatctggcg aaagagtgcg atctggcggg cgcgattaag      240
tcgatgttct ctggcgagaa gatcaaccgc actgaaaacc gcgccgtgct gcacgtagcg      300
ctgcgtaacc gtagcaatac cccgattttg gaagtgcgtt ttcagggtcc ggttgatggc      360
aaagacgtaa tgccggaagt caacgcggtg ctggagaaga tgaaaacctt ctcagaagcg      420
attatctccg gtgagtggaa aggttatacc ggcaaaagaa tcaactgacgt agtgaacatc      480
gggatcggcg gttctgacct cgcccatac atggtgaccg aagctctgcg tccgtacaaa      540
aaccacctga acatgcactt tgtttctaac gtcgatggga ctcacatcgc ggaagtgcgt      600
aaaaaagtaa acccggaaac cacgctgttc ttggtagcat ctaaaacctt caccactcag      660

gaaactatga ccaacgcca tagcgcgct gactgggtcc tgaaagcggc aggtgatgag      720
aagcacgttg caaaacactt tgccgctt tccaccaatg ccaaagccgt tggcgagttt      780
ggtattgata ctgccaacat gttcgagttc tgggactggg ttggcggccg ttactctttg      840
tggtcagcga ttggcctgtc gattgttctc tccatcggct ttgataactt cgttgaactg      900
ctttccggcg cacacgcgat ggacaagcat ttctccacca cgcctgccga gaaaaacctg      960
cctgtactgc tggcgctgat tggcatctgg tacaacaatt tctttggtgc ggaaactgaa     1020
gcgattctgc cgtatgacca gtatatgcac cgtttcgcgg cgtacttcca gcagggaat     1080
atggagtcca acggttaagta tgttgaccgt aacggtaacg ttgtggatta ccagactggc     1140
ccgattatct ggggtgaacc aggcactaac ggtcagcacg cgttctacca gctgatccac     1200
cagggaacca aaatggtacc gtgcgatttc atcgctccgg ctatcaccca taacccgctc     1260
tctgatcatc accagaaact gctgtctaac ttcttcgccc agaccgaagc gctggcgttt     1320
ggtaaatccc gcgaagtggg tgagcaggaa tatcgtgatc agggtaaaga tccggcaacg     1380
cttgactacg tgggtgccgtt caaagtattc gaaggtaacg gcccgaccaa ctccatcctg     1440
ctgcgtgaaa tcaactccgtt cagcctgggt gcgttgattg cgctgtatga gcacaaaatc     1500
tttactcagg gcgtgatcct gaacatcttc accttcgacc agtggggcgt ggaactgggt     1560
aaacagctgg cgaaccgtat tctgccagag ctgaaagatg ataaagaaat cagcagccac     1620
gatagctcga ccaatggtct gattaaccgc tataaagcgt ggcgcggtta a             1671

```

<210> 5

<211> 1674

<212> ADN

<213> Secuencia artificial

<220>

<223>Oligonucleótido sintético

ES 2 704 697 T3

<400> 5

atgaaaaaca tcaatccaac gcagaccgct gcctggcagg cactacagaa acacttcgat	60
gaaatgaaag acgttacgat cgccgatctt tttgctaaag atggtgatcg tttttctaag	120
ttctccgcaa ccttcgacga tcagatgctg gtggattact ccaaaaaccg catcactgaa	180
gagacgctgg cgaaattaca ggatctggcg aaagagtgcg atctggcggg cgcgattaag	240
tcgatgttct ctggcgagaa gatcaaccgc actgaaaacc gcgccgtgct gcacgtagcg	300
ctgctgaacc gtagcaatac cccgattttg gttgatggca aagacgtaat gccggaagtc	360
aacgcggtgc tggagaagat gaaaaccttc tcagaagcga ttatttccgg tgagctggaa	420
gtgctgtttc aggggtccgtg gaaagggttat accggcaaag caatcactga cgtagtgaac	480
atcgggatcg gcggttctga cctcggccca tacatggtga ccgaagctct gcgtccgtac	540
aaaaaccacc tgaacatgca ctttgtttct aacgtcgatg ggactcacat cgcggaagtg	600
ctgaaaaaag taaacccgga aaccacgctg ttcttggtag catctaaaac cttcaccact	660
caggaaacta tgaccaacgc ccatagcgcg cgtgactggg tcctgaaagc ggcaggtgat	720
gagaagcacg ttgcaaaaca ctttgcgggc ctttccacca atgccaaagc cgttggcgag	780
tttggtattg atactgccaa catgttcgag ttctgggact gggttggcgg ccgttactct	840
ttgtggtcag cgattggcct gtcgattgtt ctctccatcg gctttgataa cttcgttgaa	900
ctgctttccg gcgcacacgc gatggacaag catttctcca ccacgcctgc cgagaaaaac	960
ctgcctgtac tgctggcgct gattggcatc tggtaacaaca atttctttgg tgcggaaact	1020
gaagcgattc tgccgtatga ccagtatatg caccgtttcg cggcgactct ccagcagggc	1080
aatatggagt ccaacggtaa gtatgttgac cgtaacggta acgttgtgga ttaccagact	1140
ggcccgatta tctggggtga accaggcact aacggtcagc acgcgttcta ccagctgac	1200
caccagggaa ccaaaatggg accgtgcgat ttcacgctc cggctatcac ccataaaccg	1260
ctctctgac atcaccagaa actgctgtct aacttcttcg ccagaccga agcgtggcg	1320
tttggtaaat cccgcgaagt ggttgagcag gaatatcgtg atcagggtaa agatccggca	1380
acgcttgact acgtggtgcc gttcaaagta ttcgaaggta accgcccgc caactccatc	1440
ctgctgcgtg aaatcactcc gttcagcctg ggtgcgttga ttgcgctgta tgagcacaaa	1500
atctttactc agggcgatgat cctgaacatc ttcaccttcg accagtgggg cgaggaaactg	1560
ggtaaacagc tggcgaaccg tattctgccg gagctgaaag atgataaaga aatcagcagc	1620
cacgatagct cgaccaatgg tctgattaac cgctataaag cgtggcgcgg ttaa	1674

5

<210> 6

<211> 1674

<212> ADN

<213> Secuencia artificial

10

<220>

<223> Oligonucleótido sintético

<400> 6

atgaaaaaca tcaatccaac gcagaccgct gcctggcagg cactacagaa acacttcgat	60
gaaatgaaag acgttacgat cgccgatctt tttgctaaag atggtgatcg tttttctaag	120
ttctccgcaa ccttcgacga tcagatgctg gtggattact ccaaaaaccg catcactgaa	180
gagacgctgg cgaaattaca ggatctggcg aaagagtgcg atctggcggg cgcgattaag	240
tcgatgttct ctggcgagaa gatcaaccgc actgaaaacc gcgccgtgct gcacgtagcg	300
ctgcgtaacc gtagcaatac cccgattttg gttgatggca aagacgtaat gccggaagtc	360
aacgcggtgc tggagaagat gaaaaccttc tcagaagcga ttatttccgg tgagtggaaa	420
ggttataccg gcaaagcaat cactgacgta gtgaacatcg ggatcggcgg ttctgacctc	480
ggcccataca tggtagccga agctctgcgt ccgtacaaaa accacctgaa catgcacttt	540
gtttctaacg tcgatgggac tcacatcgcg gaagtgctga aaaaagtaaa cccggaaaacc	600
acgctgttct tggtagcatc taaaaccttc accactcagg aaactatgac caacgcccac	660
agcgcgcgtg actggttcct gaaagcggca ggtgatgaga agcacgttgc aaaacacttt	720
gcggcgcttt ccaccaatgc caaagccgtt ggcgagtttg gtattgatac tgccaacatg	780
ttcgagttct gggactgggt tggcgccgt tactctttgt ggtcagcgat tggcctgtcg	840
attgttctct ccatcggtt tgataacttc gttgaactgc tttccggcgc acacgcgatg	900
gacaagcatt tctccaccac gcctgccgag aaaaacctgc ctgtactgct ggcgctgatt	960
ggcatctggt acaacaattt ctttggtgcg gaaactgaag cgattctgcc gtatgaccag	1020
tatatgcacc gtttcgcggc gtacttccag cagggaataa tggagtccaa cggttaagtat	1080
gttgaccgta acggtaacgt tgtggattac cagactggcc cgattatctg ggggtgaacca	1140
ggcactaacg gtcagcacgc gttctaccag ctgatccacc agggaaccaa aatggtaccg	1200
tgcgatttca tcgctccggc tatcaccat ctggaagtgc tgtttcaggg tccgaacccg	1260
ctctctgatc atcaccagaa actgctgtct aacttcttcg ccagaccga agcgtggcg	1320
tttggtaaat cccgcgaagt ggttgagcag gaatatcgtg atcagggtaa agatccggca	1380
acgcttgact acgtggtgcc gttcaaagta ttogaaggta accgcccgc caactccatc	1440
ctgctgcgtg aaatcactcc gttcagcctg ggtgcgttga ttgcgctgta tgagcacaaa	1500
atctttactc agggcgtgat cctgaacatc ttcaccttcg accagtgggg cgtggaactg	1560
ggtaaacagc tggcgaaccg tattctgcca gagctgaaag atgataaaga aatcagcagc	1620
cacgatagct cgaccaatgg tctgattaac cgctataaag cgtggcgcgg ttaa	1674

5
10

<210> 7
<211> 1671
<212> ADN
<213> Secuencia artificial

<220>
<223> Oligonucleótido sintético

<400> 7

ES 2 704 697 T3

atgaaaaaca tcaatccaac gcagaccgct gcctggcagg cactacagaa acacttcgat	60
gaaatgaaag acgttacgat cgccgatctt tttgctaaag atgggtgatcg tttttctaag	120
ttctccgcaa ccttcgacga tcagatgctg gtggattact ccaaaaaccg catcactgaa	180
gagacgctgg cgaaattaca ggatctggcg aaagagtgcg atctggcggg cgcgattaag	240
tcgatgttct ctggcgagaa gatcaaccgc actgaaaacc gcgccgtgct gcacgtagcg	300
ctgcgtaacc gtagcaatac cccgattttg gttgatggca aagacgtaat gccggaagtc	360
aacgcggtgc tggagaagat gaaaaccttc tcagaagcga ttatttccgg tgagtggaaa	420
ggttataccg gcaaagcaat cactgacgta gtgaacatcg ggatcggcgg ttctgacctc	480
ggcccatata tggtgaccga agctctgcgt ccgtacaaaa accacctgaa catgcacttt	540
gtttctaacg tcgatgggac tcacatcgcg gaagtgcga aaaaagtaaa cccggaacc	600
acgctgttct tggtagcatc taaaaccttc accactcagg aaactatgac caacgcccat	660
agcgcgcgtg actggttcct gaaagcggca ggtgatgaga agcacgttgc aaaacacttt	720
gcggcgcttt ccaccaatgc caaagccgtt ggcgagtttg gtattgatac tgccaacatg	780
ttcagattct gggactgggt tggcggccgt tactctttgt ggtcagcgat tggcctgtcg	840
attgttctct ccacgcgctt tgataacttc gttgaactgc tttccggcgc acacgcgatg	900
gacaagcatt tctccaccac gcctgccgag aaaaacctgc ctgtactgct ggcgctgatt	960
ggcatctggt acaacaattt ctttggtgcg gaaactgaag cgattctgcc gtatgaccag	1020
tatatgcacc gtttcgcggc gtacttccag cagggcaata tggagtccaa cggttaagtat	1080
gttgaccgta acggtaacgt tgtggattac cagactggcc cgattatctg gggtaacca	1140
ggcactaacg gtcagcacgc gttctaccag ctgatccacc agggaaccaa aatggtaccg	1200
tgcgatttca tcgctccggc tatcaccat aaccgcctct ctgatcatca ccagaaactg	1260
ctgtctaact tcttcgcca gaccgaagcg ctggcgtttg gtaaattccg cgaagtgggt	1320
gagcaggaat atcgtgatca gggtaaagat ccggcaacgc ttgactacgt ggtgccgttc	1380
aaagtattcg aaggtaaccg cccgaccaac tccatcctgc tgcgtgaaat cactccgttc	1440
agcctgggtg cgttgattgc gctgtatgag cacaaaatct ttactcaggg cgtgatcctg	1500
aacatcttca ccttcgacca gtggggcggtg gaactgggta aacagctggc gaaccgtatt	1560
ctgccagagc tggaaagtgc gtttcagggt ccgaaagatg ataaagaaat cagcagccac	1620
gatagctcga ccaatgggtct gattaaccgc tataaagcgt ggcgcgggta a	1671

<210> 8

<211> 1674

<212> ADN

<213> Secuencia artificial

<220>

<223>Oligonucleótido sintético

<400> 8

5
atgaaaaaca tcaatccaac gcagaccgct gcctggcagg cactacagaa acacttcgat 60
gaaatgaaag acgttacgat cgccgatctt tttgctaaag atggatgatcg tttttctaag 120
ttctccgcaa ccttcgacga tcagatgctg gtggattact ccaaaaaccg catcactgaa 180
gagacgctgg cgaaattaca ggatctggcg aaagagtgcg atctggcggg cgcgattaag 240
tcgatgttct ctggcgagaa gatcaaccgc actgaaaacc gcgccgtgct gcacgtagcg 300
ctgcgtaacc gtagcaatac cccgattttg gttgatggca aagacgtaat gccggaagtc 360
aacgcggtgc tggagaagat gaaaaccttc tcagaagcga ttatttccgg tgagtggaaa 420
ggttataaccg gcaaagcaat cactgacgta gtgaacatcg ggatcggcgg ttctgacctc 480
ggcccataca tggtgaccga agctctgcgt ccgtacaaaa accacctgaa catgcacttt 540
gtttctaacg tcgatgggac tcacatcgcg gaagtgctga aaaaagtaaa cccggaacc 600
acgctgttct tggtagcatc taaaaccttc accactcagg aactatgac caacgcccat 660
agcgcgcgtg actggttctt gaaagcggca ggtgatgaga agcacgttgc aaaacacttt 720
gcggcgcttt ccaccaatgc caaagccgtt ggcgagtttg gtattgatac tgccaacatg 780
ttcgagttct gggactgggt tggcggccgt tactctttgt ggtcagcgat tggcctgtcg 840
attgttctct ccatcggctt tgataacttc gttgaactgc tttccggcgc acacgcgatg 900
gacaagcatt tctccaccac gcctgccgag aaaaacctgc ctgtactgct ggcgctgatt 960
ggcatctggt acaacaattt ctttggtgcg gaaactgaag cgattctgcc gtatgaccag 1020
tatatgcacc gtttcgcggc gtacttccag cagggcaata tggagtccaa cggttaagtat 1080
gttgaccgta acggtaacgt tgtggattac cagactggcc cgattatctg gggatgaacca 1140
ggcactaacg gtcagcacgc gttctaccag ctgatccacc agggaaacaa aatggtaccg 1200
tgcgatttca tcgctccggc tatcaccat aaccgctct ctgatcatca ccagaaactg 1260
ctgtctaact tcttcgcca gaccgaagcg ctggcgtttg gtaaatcccg cgaagtgggt 1320
gagcaggaat atcgtgatca gggtaaagat ccggcaacgc ttgactacgt ggtgccgttc 1380
aaagtattcg aaggtaaccg cccgaccaac tccatcctgc tgcgtgaaat cactccgttc 1440
agcctgggtg cggttgattgc gctgtatgag cacaaaatct ttactcaggg cgtgatcctg 1500
aacatcttca ccttcgacca gtggggcggt gaactgggta aacagctggc gaaccgtatt 1560
ctgccagagc tgaaactgga agtgctgttt cagggctccg atgataaaga aatcagcagc 1620
cacgatagct cgaccaatgg tctgattaac cgctataaag cgtggcgcggt ttaa 1674

	<210> 9
	<211> 1674
	<212> ADN
5	<213> Secuencia artificial
	<220>
	<223> Oligonucleótido sintético
10	<400> 9

ES 2 704 697 T3

atgaaaaaca tcaatccaac gcagaccgct gcctggcagg cactacagaa acacttcgat	60
gaaatgaaag acgttacgat cgccgatctt tttgctaaag atggtgatcg tttttctaag	120
ttctccgcaa ccttcgacga tcagatgctg gtggattact caaaaaacg catcactgaa	180
gagacgctgg cgaaattaca ggatctggcg aaagagtgcg atctggcggg cgcgattaag	240
tcgatgttct ctggcgagaa gatcaaccgc actgaaaacc gcgccgtgct gcacgtagcg	300
ctgcgtaacc gtagcaatac cccgattttg gttgatggca aagacgtaat gccggaagtc	360
aacgcggtgc tggagaagat gaaaaccttc tcagaagcga ttatttcggtg tgagtggaaa	420
ggttataccg gcaaagcaat cactgacgta gtgaacatcg ggatcggcgg ttctgacctc	480
ggcccatata tggtgaccga agctctgcgt ccgtacaaaa accacctgaa catgcacttt	540
gtttctaacg tcgatgggac tcacatcgcg gaagtgcgtg aaaaagtaaa cccggaaacc	600
acgctgttct tggtagcatc taaaaccttc accactcagg aaactatgac caacgcccat	660
agcgcgcgtg actgggtcct gaaagcggca ggtgatgaga agcacgttgc aaaacacttt	720
gcggcgcttt ccaccaatgc caaagccgtt ggcgagtttg gtattgatac tgccaacatg	780
ttcgagttct gggactgggt tggcggccgt tactctttgt ggtcagcgat tggcctgtcg	840
attgttctct ccatcggctt tgataacttc gttgaactgc tttccggcgc acacgcgatg	900
gacaagcatt tctccaccac gcctgccgag aaaaacctgc ctgtactgct ggcgctgatt	960
ggcatctggt acaacaattt ctttggtgcg gaaactgaag cgattctgcc gtatgaccag	1020
tatatgcacc gtttcgcggc gtacttccag cagggcaata tggagtccaa cggtaagtat	1080
gttgaccgta acggtaacgt tgtggattac cagactggcc cgattatctg gggatgaacca	1140
ggcactaacg gtcagcacgc gttctaccag ctgatccacc aggaaccaa aatggtagcg	1200
tgcgatttca tcgctccggc tatcaccat aaccgctct ctgatcatca ccagaaactg	1260
ctgtctaact tcttcgcca gaccgaagcg ctggcgtttg gtaaataccg cgaagtgggt	1320
gagcaggaat atcgtgatca gggtaaagat ccggcaacgc ttgactacgt ggtgccgttc	1380
aaagtattcg aaggtaacg cccgaccaac tccatcctgc tgcgtgaaat cactccgttc	1440
agcctgggtg cgttgattgc gctgtatgag cacaaaatct ttactcagg cgtgatcctg	1500
aacatcttca ccttcgacca gtggggcgtg gaactgggta aacagctggc gaaccgtatt	1560
ctgccagagc tgaaagatct ggaagtgcgt tttcagggtc cggataaaga aatcagcagc	1620
cacgatagct cgaccaatgg tctgattaac cgctataaag cgtggcgcgg ttaa	1674

<210> 10

<211> 1674

<212> ADN

<213> Secuencia artificial

<220>

<223>Oligonucleótido sintético

<400> 10

5

atgaaaaaca tcaatccaac gcagaccgct gcctggcagg cactacagaa acacttcgat	60
gaaatgaaag acgttacgat cgccgatctt tttgctaaag atggtgatcg tttttctaag	120
ttctccgcaa ccttcgacga tcagatgctg gtggattact ccaaaaaccg catcactgaa	180
gagacgctgg cgaaattaca ggatctggcg aaagagtgcg atctggcggg cgcgattaag	240
tcgatgttct ctggcgagaa gatcaaccgc actgaaaacc gcgccgtgct gcacgtagcg	300
ctgcgtaacc gtagcaatac cccgattttg gttgatggca aagacgtaat gccggaagtc	360
aacgcggtgc tggagaagat gaaaaccttc tcagaagcga ttatttccgg tgagtggaaa	420
ggttataccg gcaaagcaat cactgacgta gtgaacatcg ggatcggcgg ttctgacctc	480
ggcccatata tggtagaccga agctctgctg ccgtacaaaa accacctgaa catgcacttt	540
gtttctaacg tcgatgggac tcacatcgcg gaagtgcgta aaaaagtaaa cccggaaacc	600
acgctgttct tggtagcatc taaaaccttc accactcagg aaactatgac caacgcccatt	660
agcgcgcgctg actggttcct gaaagcggca ggtgatgaga agcacgttgc aaaacacttt	720
gcggcgcttt ccaccaatgc caaagccgtt ggcgagtttg gtattgatac tgccaacatg	780
ttcgagttct gggactgggt tggcggccgt tactctttgt ggtcagcgat tggcctgtcg	840
attgttctct ccacggtctt tgataacttc gttgaactgc tttccggcgc acacgcgatg	900
gacaagcatt tctccaccac gcctgccgag aaaaacctgc ctgtactgct ggcgctgatt	960
ggcatctggt acaacaattt ctttggcgcg gaaactgaag cgattctgcc gtatgaccag	1020
tatatgcacc gtttcgcggc gtacttccag cagggcaata tggagtccaa cggttaagtat	1080
gttgaccgta acggtaacgt tgtggattac cagactggcc cgattatctg gggatgaacca	1140
ggcactaacg gtcagcacgc gttctaccag ctgatccacc aggaaccaa aatggtaccg	1200
tgcgatttca tcgctccggc tatcaccat aaccgcctct ctgatcatca ccagaaactg	1260
ctgtctaact tcttcgcca gaccgaagcg ctggcgtttg gtaaatcccg cgaagtgggt	1320
gagcaggaat atcgtgatca gggtaaagat ccggcaacgc ttgactacgt ggtgccgttc	1380
aaagtattcg aaggtaaccg cccgaccaac tccatcctgc tgcgtgaaat cactccgttc	1440
agcctgggtg cgttgattgc gctgtatgag caaaaatct ttactcaggg cgtgatcctg	1500
aacatcttca ccttcgacca gtggggcgctg gaactgggta aacagctggc gaaccgtatt	1560
ctgccagagc tgaaagatga tctggaagtg ctgtttcagg gtccgaaaga aatcagcagc	1620
cacgatagct cgaccaatgg tctgattaac cgctataaag cgtggcgcgg ttaa	1674

ES 2 704 697 T3

<210> 11
<211> 1674
<212> ADN
<213> Secuencia artificial
5
<220>
<223> Oligonucleótido sintético
10
<400> 11
atgaaaaaca tcaatccaac gcagaccgct gcctggcagg cactacagaa acacttcgat 60

gaaatgaaag acgttacgat cgccgatcctt tttgctaaag atggtgatcg tttttctaag 120
 ttctccgcaa ccttcgacga tcagatgctg gtggattact ccaaaaaccg catcactgaa 180
 gagacgctgg cgaaattaca ggatctggcg aaagagtgcg atctggcggg cgcgattaag 240
 tcgatgttct ctggcgagaa gatcaaccgc actgaaaacc gcgcgctgct gcacgtagcg 300
 ctgcgtaacc gtagcaatac cccgattttg gttgatggca aagacgtaat gccggaagtc 360
 aacgcggtgc tggagaagat gaaaaccttc tcagaagcga ttatttcggg tgagtggaaa 420
 gggtataaccg gcaaagcaat cactgacgta gtgaacatcg ggatcggcgg ttctgacctc 480
 ggcccataca tggtgaccga agctctgcgt ccgtacaaaa accacctgaa catgcacttt 540
 gtttctaacg tcgatgggac tcacatcgcg gaagtgctga aaaaagtaaa cccggaaacc 600
 acgctgttct tggtagcatc taaaaccttc accactcagg aaactatgac caacgcccac 660
 agcgcgctg actggttcct gaaagcggca ggtgatgaga agcacgttgc aaaacacttt 720
 gcggcgcttt ccaccaatgc caaagccgtt ggcgagtgtg gtattgatac tgccaacatg 780
 ttcgagttct gggactgggt tggcgccgt tactctttgt ggtcagcgat tggcctgtcg 840
 attgttctct ccacggtt tgataacttc gttgaactgc tttccggcgc acacgcgatg 900
 gacaagcatt tctccaccac gcctgccgag aaaaacctgc ctgtactgct ggcgctgatt 960
 ggcatctggt acaacaattt ctttggcgcg gaaactgaag cgattctgcc gtatgaccag 1020
 tatatgcacc gtttcgcggc gtacttccag cagggcaata tggagtccaa cggttaagtat 1080
 gttgaccgta acggtaacgt tgtggattac cagactggcc cgattatctg ggggtgaacca 1140
 ggactaaccg gtcagcacgc gttctaccag ctgatccacc agggaaacaa aatggtaccg 1200
 tgcgatttca tcgctccggc tatcaccat aaccgctct ctgatcatca ccagaaactg 1260
 ctgtctaact tcttcgcca gaccgaagcg ctggcgtttg gtaaatcccg cgaagtgggt 1320
 gagcaggaat atcgtgatca gggtaaagat ccggcaacgc ttgactacgt ggtgccgttc 1380
 aaagtattcg aaggtaaccg cccgaccaac tccatcctgc tgcgtgaaat cactccgttc 1440
 agcctgggtg cgttgattgc gctgtatgag caaaaatct ttactcaggg cgtgatcctg 1500
 aacatcttca ccttcgacca gtggggcggtg gaactgggta aacagctggc gaaccgtatt 1560
 ctgccagagc tgaaagatga taaactggaa gtgctgtttc agggtcggga aatcagcagc 1620
 cacgatagct cgaccaatgg tctgattaac cgctataaag cgtggcgcg ttaa 1674

<210> 12
 <211> 1674
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Oligonucleótido sintético

<400> 12

atgaaaaaca tcaatccaac gcagaccgct gcctggcagg cactacagaa acacttcgat	60
gaaatgaaag acgttacgat cgccgatctt tttgctaaag atggtgatcg tttttctaag	120
ttctccgcaa ccttcgacga tcagatgctg gtggattact ccaaaaaccg catcactgaa	180
gagacgctgg cgaaattaca ggatctggcg aaagagtgcg atctggcggg cgcgattaag	240
tcgatgttct ctggcgagaa gatcaaccgc actgaaaacc gcgccgtgct gcacgtagcg	300
ctgcgtaacc gtagcaatac cccgattttg gttgatggca aagacgtaat gccggaagtc	360
aacgcggtgc tggagaagat gaaaaccttc tcagaagcga ttatttccgg tgagtggaaa	420
ggttataacc gcaaagcaat cactgacgta gtgaacatcg ggatcggcgg ttctgacctc	480
ggcccataca tggtagaccga agctctgcgt ccgtacaaaa accacctgaa catgcacttt	540
gtttctaacg tcgatgggac tcacatcgcg gaagtgctga aaaaagtaaa cccggaacc	600
acgctgttct tggtagcatc taaaaccttc accactcagg aaactatgac caacgcccac	660
agcgcgcgtg actggttcct gaaagcggca ggtgatgaga agcacgttgc aaaacacttt	720
gcggcgcttt ccaccaatgc caaagccgtt ggcgagtttg gtattgatac tgccaacatg	780
ttcgagttct gggactgggt tggcggccgt tactctttgt ggtcagcgat tggcctgtcg	840
attgttctct ccacggtt tgataacttc gttgaactgc tttccggcgc acacgcgatg	900
gacaagcatt tctccaccac gcctgccgag aaaaacctgc ctgtactgct ggcgctgatt	960
ggcatctggt acaacaattt ctttgggtgcg gaaactgaag cgattctgcc gtatgaccag	1020
tatatgcacc gtttcgcggc gtacttccag cagggcaata tggagtccaa cggttaagtat	1080
gttgaccgta acggtaacgt tgtggattac cagactggcc cgattatctg gggatgaacca	1140
ggcactaacg gtcagcacgc gttctaccag ctgatccacc agggaaccaa aatggtaccg	1200
tgcgatttca tcgctccggc tatcaccat aaccgctct ctgatcatca ccagaaactg	1260
ctgtctaact tcttcgcca gaccgaagcg ctggcgtttg gtaaatcccg cgaagtgggt	1320
gagcaggaat atcgtgatca gggtaaagat ccggcaacgc ttgactacgt ggtgccgttc	1380
aaagtattcg aaggtaaccg cccgaccaac tccatcctgc tgcgtgaaat cactccgttc	1440
agcctgggtg cgttgattgc gctgtatgag cacaaaatct ttactcaggg cgtgatcctg	1500
aacatcttca ccttcgacca gtggggcgtg gaactgggta aacagctggc gaaccgtatt	1560
ctgccagagc tgaaagatga taaagaactg gaagtgctgt ttcagggtcc gatcagcagc	1620
cacgatagct cgaccaatgg tctgattaac cgctataaag cgtggcgcgg ttaa	1674

ES 2 704 697 T3

<210> 13
<211> 1674
<212> ADN
<213> Secuencia artificial

5

<220>
<223> Oligonucleótido sintético

<400> 13

10

```

atgaaaaaca tcaatccaac gcagaccgct gcctggcagc cactacagaa acacttcgat      60
gaaatgaaag acgttacgat cgccgatctt ttgtctaaag atggatgatcg tttttctaag      120
ttctccgcaa ccttcgacga tcagatgctg gtggattact ccaaaaaccg catcactgaa      180
gagacgctgg cgaaattaca ggatctggcg aaagagtgcg atctggcggg cgcgattaag      240
tcgatgttct ctggcgagaa gatcaaccgc actgaaaacc gcgccgtgct gcacgtagcg      300
ctgcgtaacc gtagcaatac cccgattttg gttgatggca aagacgtaat gccggaagtc      360
aacgcggtgc tggagaagat gaaaaccttc tcagaagcga ttatttccgg tgagtggaaa      420
ggttataccg gcaaagcaat cactgacgta gtgaacatcg ggatcggcgg ttctgacctc      480
ggcccataca tggtgaccga agctctgctg ccgtacaaaa accacctgaa catgcacttt      540
gtttctaacg tcgatgggac tcacatcgcg gaagtgctga aaaaagtaaa cccggaacc      600
acgctgttct tggtagcatc taaaaccttc accactcagg aaactatgac caacgcccac      660
agcgcgcgtg actggttctt gaaagcggca ggtgatgaga agcacgtgc aaaacacttt      720
gcggcgcttt ccaccaatgc caaagccgtt ggcgagtttg gtattgatac tgccaacatg      780
ttcgagttct gggactgggt tggcgccgtt tactctttgt ggtcagcgat tggcctgtcg      840
attgttctct ccatcggtt tgataacttc gttgaactgc tttccggcgc acacgcgatg      900
gacaagcatt tctccaccac gcctgccgag aaaaacctgc ctgtactgct ggcgctgatt      960
ggcatctggt acaacaattt ctttggtgcg gaaactgaag cgattctgcc gtatgaccag      1020
tatatgcacc gtttcgcggc gtacttccag cagggcaata tggagtcaa cggtaaagtat      1080
gttgaccgta acggtaacgt tgtggattac cagactggcc cgattatctg gggatgaacca      1140
ggcactaacg gtcagcacgc gttctaccag ctgatccacc aggaaccaa aatggtaccg      1200
tgcgatttca tcgctccggc tatcaccat aaccgctct ctgatcatca ccagaaactg      1260
ctgtctaact tcttcgcca gaccgaagcg ctggcgtttg gtaaatcccg cgaagtgggt      1320
gagcaggaat atcgtgatca ggttaaagat ccggcaacgc ttgactacgt ggtgccgttc      1380
aaagtattcg aaggtaacg cccgaccaac tccatcctgc tgcgtgaaat cactccgttc      1440
agcctgggtg cgttgattgc gctgtatgag caaaaatct ttactcaggg cgtgatcctg      1500
aacatcttca ccttcgacca gtggggcgtg gaactgggta aacagctggc gaaccgtatt      1560
ctgccagagc tgaaagatga taaagaaatc ctggaagtgc tgtttcaggg tccgagcagc      1620
cacgatagct cgaccaatgg tctgattaac cgctataaag cgtggcgcgg ttaa      1674

```

	<210> 14
	<211> 1674
	<212> ADN
5	<213> Secuencia artificial
	<220>
	<223> Oligonucleótido sintético
10	<400> 14

atgaaaaaca tcaatccaac gcagaccgct gcctggcagg cactacagaa acacttcgat	60
gaaatgaaag acgttacgat cgccgatctt ttgctaaag atggtgatcg tttttctaag	120
ttctccgcaa ccttcgacga tcagatgctg gtggattact ccaaaaaccg catcactgaa	180
gagacgctgg cgaaattaca ggatctggcg aaagagtgcg atctggcggg cgcgattaag	240
tcgatgttct ctggcgagaa gatcaaccgc actgaaaacc gcgccgtgct gcacgtagcg	300
ctgcgtaacc gtagcaatac cccgattttg gttgatggca aagacgtaat gccggaagtc	360
aacgcggtgc tggagaagat gaaaaccttc tcagaagcga ttatttccgg tgagtggaaa	420
ggttataaccg gcaaagcaat cactgacgta gtgaacatcg ggatcggcgg ttctgacctc	480
ggcccataca tggtgaccga agctctgctt ccgtacaaaa accacctgaa catgcacttt	540
gtttctaacg tcgatgggac tcacatcgcg gaagtgcgtg aaaaagtaaa cccggaaacc	600
acgctgttct tggtagcatc taaaaccttc accactcagg aaactatgac caacgcccac	660
agcgcgcgtg actggttcct gaaagcggca ggtgatgaga agcacgttgc aaaacacttt	720
gcggcgcttt ccaccaatgc caaagccgtt ggcgagtttg gtattgatac tgccaacatg	780
ttcgagttct gggactgggt tggcgccgtt tactctttgt ggtcagcgat tggcctgtcg	840
attgttctct ccatcggtt tgataacttc gttgaactgc tttccggcgc acacgcgatg	900
gacaagcatt tctccaccac gcctgccgag aaaaacctgc ctgtactgct ggcgtgatt	960
ggcatctggg acaacaattt ctttggtgcg gaaactgaag cgattctgcc gtatgaccag	1020
tatatgcacc gtttcgcggc gtacttccag cagggcaata tggagtccaa cggtaagtat	1080
gttgaccgta acggtaacgt tgtggattac cagactggcc cgattatctg gggagaacca	1140
ggcactaacg gtcagcacgc gttctaccag ctgatccacc agggaaacaa aatggtaccg	1200
tgcgatttca tcgctccggc tatcaccat aaccgcgtct ctgatcatca ccagaaactg	1260
ctgtctaact tcttcgcca gaccgaagcg ctggcgtttg gtaaataccg cgaagtgggt	1320
gagcaggaat atcgtgatca gggtaaagat ccggcaacgc ttgactacgt ggtgccgttc	1380
aaagtattcg aaggtaaccg cccgaccaac tccatcctgc tgcgtgaaat cactccgttc	1440
agcctgggtg cgttgattgc gctgtatgag cacaaaatct ttactcaggg cgtgatcctg	1500
aacatcttca ccttcgacca gtggggcggt gaactgggtg aacagctggc gaaccgtatt	1560
ctgccagagc tgaaagatga taaagaaatc agcctggaag tgctgtttca gggccgagc	1620
cacgatagct cgaccaatgg tctgattaac cgctataaag cgtggcgcggt ttaa	1674

<210> 15
 <211> 1674
 <212> ADN
 <213> Secuencia artificial

ES 2 704 697 T3

<220>

<223>Oligonucleótido sintético

<400> 15

5

```

atgaaaaaca tcaatccaac gcagaccgct gcctggcagg cactacagaa acacttcgat      60
gaaatgaaag acgttacgat cgccgatctt tttgctaaag atggatgatcg tttttctaag      120
ttctccgcaa ccttcgacga tcagatgctg gtggattact ccaaaaaccg catcactgaa      180
gagacgctgg cgaaattaca ggatctggcg aaagagtgcg atctggcggg cgcgattaag      240
tcgatgttct ctggcgagaa gatcaaccgc actgaaaacc gcgccgtgct gcacgtagcg      300
ctgcgtaacc gtagcaatac cccgattttg gttgatggca aagacgtaat gccggaagtc      360
aacgcggtgc tggagaagat gaaaaccttc tcagaagcga ttatttcggg tgagtggaaa      420
ggttataccg gcaaagcaat cactgacgta gtgaacatcg ggatcgggcg ttctgacctc      480
ggcccataca tggtgaccga agctctgcgt ccgtacaaaa accacctgaa catgcacttt      540
gtttctaacg tcgatgggac tcacatcgcg gaagtgctga aaaaagtaaa cccggaaacc      600
acgctgttct tggtagcatc taaaaccttc accactcagg aaactatgac caacgcccat      660
agcgcgctg actggttcct gaaagcggca ggtgatgaga agcacgttgc aaaacacttt      720
gcggcgcttt ccaccaatgc caaagccgtt ggcgagtgtg gtattgatac tgccaacatg      780
ttcgagtctt gggactgggt tggcgccgt tactctttgt ggtcagcgat tggcctgtcg      840
attgttctct ccatcggtt tgataacttc gttgaactgc tttccggcgc acacgcgatg      900
gacaagcatt tctccaccac gcctgccgag aaaaacctgc ctgtactgct ggcgctgatt      960
ggcatctggt acaacaatth ctttggtgcg gaaactgaag cgattctgcc gtatgaccag      1020
tatatgcacc gtttcgcggc gtacttccag cagggcaata tggagtcaa cggtaatgat      1080
gttgaccgta acggtaacgt tgtggattac cagactggcc cgattatctg ggtgaacca      1140
ggcactaacg gtcagcacgc gttctaccag ctgatccacc aggaaccaaa aatggatacc      1200
tgcgatttca tcgctccggc tatcaccat aaccgctct ctgatcatca ccagaaactg      1260
ctgtctaact tcttcgcca gaccgaagcg ctggcggttg gtaaatcccg cgaagtgggt      1320
gagcaggaat atcgtgatca gggtaaagat ccggcaacgc ttgactacgt ggtgccgttc      1380
aaagtattcg aaggtaacgg cccgaccaac tccatcctgc tgcgtgaaat cactccgttc      1440
agcctgggtg cggtgattgc gctgtatgag cacaaaatct ttactcaggg cgtgatcctg      1500
aacatcttca ccttcgacca gtggggcgct gaactgggta aacagctggc gaaccgtatt      1560

ctgccagagc tgaaagatga taaagaaatc agcagcctgg aagtgtgtt tcagggtccg      1620
cacgatagct cgaccaatgg tctgattaac cgctataaag cgtggcgcgg ttaa      1674

```

10

<210> 16

<211> 1674

<212> ADN

<213> Secuencia artificial

15

<220>

<223>Oligonucleótido sintético

<400> 16

5

atgaaaaaca tcaatccaac gcagaccgct gcctggcagg cactacagaa acacttcgat	60
gaaatgaaag acgttacgat cgccgatctt tttgctaaag atggtgatcg tttttctaag	120
ttctccgcaa ccttcgacga tcagatgctg gtggattact ccaaaaaccg catcactgaa	180
gagacgctgg cgaaattaca ggatctggcg aaagagtgcg atctggcggg cgcgattaag	240
tcgatgttct ctggcgagaa gatcaaccgc actgaaaacc gcgccgtgct gcacgtagcg	300
ctgcgtaacc gtagcaatac cccgattttg gttgatggca aagacgtaat gccggaagtc	360
aacgcggtgc tggagaagat gaaaaccttc tcagaagcga ttatttccgg tgagtggaaa	420
ggttataccg gcaaagcaat cactgacgta gtgaacatcg ggatcggcgg ttctgacctc	480
ggcccataca tggtagaccga agctctgctg ccgtacaaaa accacctgaa catgcacttt	540
gtttctaacg tcgatgggac tcacatcgcg gaagtgctga aaaaagtaaa cccggaacc	600
acgctgttct tggtagcatc taaaaccttc accactcagg aaactatgac caacgcccac	660
agcgcgctg actggttcct gaaagcggca ggtgatgaga agcacgttgc aaaacacttt	720
gcggcgcttt ccaccaatgc caaagccgtt ggcgagtttg gtattgatac tgccaacatg	780
ttcgagttct gggactgggt tggcgccgt tactctttgt ggtcagcgat tggcctgtcg	840
attgttctct ccacggtt tgataacttc gttgaactgc tttccggcgc acacgcgatg	900
gacaagcatt tctccaccac gcctgccgag aaaaacctgc ctgtactgct ggcgtgatt	960
ggcatctggt acaacaattt ctttgggtgc gaaactgaag cgattctgcc gtatgaccag	1020
tatatgcacc gtttcgcggc gtacttccag cagggcaata tggagtccaa cgtaagtat	1080
gttgaccgta acggtaacgt tgtggattac cagactggcc cgattatctg ggtgaacca	1140
ggcactaacg gtcagcacgc gttctaccag ctgatccacc agggaaacaa aatggtaccg	1200
tgcgatttca tcgctccggc tatcacccat aaccgctct ctgatcatca ccagaaactg	1260
ctgtctaact tcttcgcca gaccgaagcg ctggcgtttg gtaaatcccg cgaagtgggt	1320
gagcaggaat atcgtgatca gggtaaagat ccggcaacgc ttgactacgt ggtgccgttc	1380
aaagtattcg aaggtaaccg cccgaccaac tccatcctgc tgcgtgaaat cactccgttc	1440
agcctgggtg cgttgattgc gctgtatgag caaaaaatct ttactcaggg cgtgatcctg	1500
aacatcttca ctttcgacca gtggggcgtg gaactgggta aacagctggc gaaccgtatt	1560
ctgccagagc tgaaagatga taaagaaatc agcagccacg atagctcgac caatggtctg	1620
attaaccgct ataaactgga agtgctgttt cagggctccg cgtggcgcg ttaa	1674

<210> 17
 <211> 549
 <212> PRT
 <213> Secuencia artificial

5

<220>
 <223> Polipéptido sintético

<400> 17

10

```

Met  Lys  Asn  Ile  Asn  Pro  Thr  Gln  Thr  Ala  Ala  Trp  Gln  Ala  Leu  Gln
1      5      10      15

Lys  His  Phe  Asp  Glu  Met  Lys  Asp  Val  Thr  Ile  Ala  Asp  Leu  Phe  Ala
      20      25      30

Lys  Asp  Gly  Asp  Arg  Phe  Ser  Lys  Phe  Ser  Ala  Thr  Phe  Asp  Asp  Gln
      35      40      45

Met  Leu  Val  Asp  Tyr  Ser  Lys  Asn  Arg  Ile  Thr  Glu  Glu  Thr  Leu  Ala
      50      55      60

Lys  Leu  Gln  Asp  Leu  Ala  Lys  Glu  Cys  Asp  Leu  Ala  Gly  Ala  Ile  Lys
65      70      75      80

Ser  Met  Phe  Ser  Gly  Glu  Lys  Ile  Asn  Arg  Thr  Glu  Asn  Arg  Ala  Val
      85      90      95

Leu  His  Val  Ala  Leu  Arg  Asn  Arg  Ser  Asn  Thr  Pro  Ile  Leu  Val  Asp
      100     105     110

Gly  Lys  Asp  Val  Met  Pro  Glu  Val  Asn  Ala  Val  Leu  Glu  Lys  Met  Lys
      115     120     125

Thr  Phe  Ser  Glu  Ala  Ile  Ile  Ser  Gly  Glu  Trp  Lys  Gly  Tyr  Thr  Gly
      130     135     140

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

Gly  Pro  Tyr  Met  Val  Thr  Glu  Ala  Leu  Arg  Pro  Tyr  Lys  Asn  His  Leu
      165     170     175

Asn  Met  His  Phe  Val  Ser  Asn  Val  Asp  Gly  Thr  His  Ile  Ala  Glu  Val
    
```

ES 2 704 697 T3

180					185					190					
Leu	Lys	Lys	Val	Asn	Pro	Glu	Thr	Thr	Leu	Phe	Leu	Val	Ala	Ser	Lys
		195					200					205			
Thr	Phe	Thr	Thr	Gln	Glu	Thr	Met	Thr	Asn	Ala	His	Ser	Ala	Arg	Asp
	210					215					220				
Trp	Phe	Leu	Lys	Ala	Ala	Gly	Asp	Glu	Lys	His	Val	Ala	Lys	His	Phe
225					230					235					240
Ala	Ala	Leu	Ser	Thr	Asn	Ala	Lys	Ala	Val	Gly	Glu	Phe	Gly	Ile	Asp
				245					250					255	
Thr	Ala	Asn	Met	Phe	Glu	Phe	Trp	Asp	Trp	Val	Gly	Gly	Arg	Tyr	Ser
			260					265					270		
Leu	Trp	Ser	Ala	Ile	Gly	Leu	Ser	Ile	Val	Leu	Ser	Ile	Gly	Phe	Asp
		275					280					285			
Asn	Phe	Val	Glu	Leu	Leu	Ser	Gly	Ala	His	Ala	Met	Asp	Lys	His	Phe
	290					295					300				
Ser	Thr	Thr	Pro	Ala	Glu	Lys	Asn	Leu	Pro	Val	Leu	Leu	Ala	Leu	Ile
305					310					315					320
Gly	Ile	Trp	Tyr	Asn	Asn	Phe	Phe	Gly	Ala	Glu	Thr	Glu	Ala	Ile	Leu
				325					330					335	
Pro	Tyr	Asp	Gln	Tyr	Met	His	Arg	Phe	Ala	Ala	Tyr	Phe	Gln	Gln	Gly
			340					345					350		
Asn	Met	Glu	Ser	Asn	Gly	Lys	Tyr	Val	Asp	Arg	Asn	Gly	Asn	Val	Val
		355					360					365			
Asp	Tyr	Gln	Thr	Gly	Pro	Ile	Ile	Trp	Gly	Glu	Pro	Gly	Thr	Asn	Gly
	370					375					380				
Gln	His	Ala	Phe	Tyr	Gln	Leu	Ile	His	Gln	Gly	Thr	Lys	Met	Val	Pro
385					390					395					400
Cys	Asp	Phe	Ile	Ala	Pro	Ala	Ile	Thr	His	Asn	Pro	Leu	Ser	Asp	His
				405					410					415	
His	Gln	Lys	Leu	Leu	Ser	Asn	Phe	Phe	Ala	Gln	Thr	Glu	Ala	Leu	Ala
			420					425					430		

ES 2 704 697 T3

Phe Gly Lys Ser Arg Glu Val Val Glu Gln Glu Tyr Arg Asp Gln Gly
435 440 445

Lys Asp Pro Ala Thr Leu Asp Tyr Val Val Pro Phe Lys Val Phe Glu
450 455 460

Gly Asn Arg Pro Thr Asn Ser Ile Leu Leu Arg Glu Ile Thr Pro Phe
465 470 475 480

Ser Leu Gly Ala Leu Ile Ala Leu Tyr Glu His Lys Ile Phe Thr Gln
485 490 495

Gly Val Ile Leu Asn Ile Phe Thr Phe Asp Gln Trp Gly Val Glu Leu
500 505 510

Gly Lys Gln Leu Ala Asn Arg Ile Leu Pro Glu Leu Lys Asp Asp Lys
515 520 525

Glu Ile Ser Ser His Asp Ser Ser Thr Asn Gly Leu Ile Asn Arg Tyr
530 535 540

Lys Ala Trp Arg Gly
545

<210> 18
<211> 557
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido sintético

<400> 18

Met Lys Asn Ile Asn Pro Thr Gln Thr Ala Ala Trp Gln Ala Leu Gln
1 5 10 15

Lys His Phe Asp Glu Met Lys Asp Val Thr Ile Ala Asp Leu Phe Ala
20 25 30

Lys Asp Gly Asp Arg Phe Ser Lys Phe Ser Ala Thr Phe Asp Asp Gln
35 40 45

Met Leu Val Asp Tyr Ser Lys Asn Arg Ile Thr Glu Glu Thr Leu Ala
50 55 60

Lys Leu Gln Asp Leu Ala Lys Glu Cys Asp Leu Ala Gly Ala Ile Lys
65 70 75 80

Ser Met Phe Ser Gly Glu Lys Ile Asn Arg Thr Glu Asn Arg Ala Val

ES 2 704 697 T3

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

ES 2 704 697 T3

Gly Ala Glu Thr Glu Ala Ile Leu Pro Tyr Asp Gln Tyr Met His Arg
340 345 350

Phe Ala Ala Tyr Phe Gln Gln Gly Asn Met Glu Ser Asn Gly Lys Tyr
355 360 365

Val Asp Arg Asn Gly Asn Val Val Asp Tyr Gln Thr Gly Pro Ile Ile
370 375 380

Trp Gly Glu Pro Gly Thr Asn Gly Gln His Ala Phe Tyr Gln Leu Ile
385 390 395 400

His Gln Gly Thr Lys Met Val Pro Cys Asp Phe Ile Ala Pro Ala Ile
405 410 415

Thr His Asn Pro Leu Ser Asp His His Gln Lys Leu Leu Ser Asn Phe
420 425 430

Phe Ala Gln Thr Glu Ala Leu Ala Phe Gly Lys Ser Arg Glu Val Val
435 440 445

Glu Gln Glu Tyr Arg Asp Gln Gly Lys Asp Pro Ala Thr Leu Asp Tyr
450 455 460

Val Val Pro Phe Lys Val Phe Glu Gly Asn Arg Pro Thr Asn Ser Ile
465 470 475 480

Leu Leu Arg Glu Ile Thr Pro Phe Ser Leu Gly Ala Leu Ile Ala Leu
485 490 495

Tyr Glu His Lys Ile Phe Thr Gln Gly Val Ile Leu Asn Ile Phe Thr
500 505 510

Phe Asp Gln Trp Gly Val Glu Leu Gly Lys Gln Leu Ala Asn Arg Ile
515 520 525

Leu Pro Glu Leu Lys Asp Asp Lys Glu Ile Ser Ser His Asp Ser Ser
530 535 540

Thr Asn Gly Leu Ile Asn Arg Tyr Lys Ala Trp Arg Gly
545 550 555

<210> 19

<211> 557

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido sintético

ES 2 704 697 T3

<400> 19

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

ES 2 704 697 T3

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

ES 2 704 697 T3

Tyr Glu His Lys Ile Phe Thr Gln Gly Val Ile Leu Asn Ile Phe Thr
500 505 510

Phe Asp Gln Trp Gly Val Glu Leu Gly Lys Gln Leu Ala Asn Arg Ile
515 520 525

Leu Pro Glu Leu Lys Asp Asp Lys Glu Ile Ser Ser His Asp Ser Ser
530 535 540

Thr Asn Gly Leu Ile Asn Arg Tyr Lys Ala Trp Arg Gly
545 550 555

<210> 20

<211> 556

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido sintético

<400> 20

Met Lys Asn Ile Asn Pro Thr Gln Thr Ala Ala Trp Gln Ala Leu Gln
1 5 10 15

Lys His Phe Asp Glu Met Lys Asp Val Thr Ile Ala Asp Leu Phe Ala
20 25 30

Lys Asp Gly Asp Arg Phe Ser Lys Phe Ser Ala Thr Phe Asp Asp Gln
35 40 45

Met Leu Val Asp Tyr Ser Lys Asn Arg Ile Thr Glu Glu Thr Leu Ala
50 55 60

Lys Leu Gln Asp Leu Ala Lys Glu Cys Asp Leu Ala Gly Ala Ile Lys
65 70 75 80

Ser Met Phe Ser Gly Glu Lys Ile Asn Arg Thr Glu Asn Arg Ala Val
85 90 95

Leu His Val Ala Leu Arg Asn Arg Ser Asn Thr Pro Ile Leu Glu Val
100 105 110

Leu Phe Gln Gly Pro Val Asp Gly Lys Asp Val Met Pro Glu Val Asn
115 120 125

Ala Val Leu Glu Lys Met Lys Thr Phe Ser Glu Ala Ile Ile Ser Gly
130 135 140

ES 2 704 697 T3

Glu	Trp	Lys	Gly	Tyr	Thr	Gly	Lys	Ala	Ile	Thr	Asp	Val	Val	Asn	Ile	145	150	155	160
Gly	Ile	Gly	Gly	Ser	Asp	Leu	Gly	Pro	Tyr	Met	Val	Thr	Glu	Ala	Leu	165	170	175	
Arg	Pro	Tyr	Lys	Asn	His	Leu	Asn	Met	His	Phe	Val	Ser	Asn	Val	Asp	180	185	190	
Gly	Thr	His	Ile	Ala	Glu	Val	Leu	Lys	Lys	Val	Asn	Pro	Glu	Thr	Thr	195	200	205	
Leu	Phe	Leu	Val	Ala	Ser	Lys	Thr	Phe	Thr	Thr	Gln	Glu	Thr	Met	Thr	210	215	220	
Asn	Ala	His	Ser	Ala	Arg	Asp	Trp	Phe	Leu	Lys	Ala	Ala	Gly	Asp	Glu	225	230	235	240
Lys	His	Val	Ala	Lys	His	Phe	Ala	Ala	Leu	Ser	Thr	Asn	Ala	Lys	Ala	245	250	255	
Val	Gly	Glu	Phe	Gly	Ile	Asp	Thr	Ala	Asn	Met	Phe	Glu	Phe	Trp	Asp	260	265	270	
Trp	Val	Gly	Gly	Arg	Tyr	Ser	Leu	Trp	Ser	Ala	Ile	Gly	Leu	Ser	Ile	275	280	285	
Val	Leu	Ser	Ile	Gly	Phe	Asp	Asn	Phe	Val	Glu	Leu	Leu	Ser	Gly	Ala	290	295	300	
His	Ala	Met	Asp	Lys	His	Phe	Ser	Thr	Thr	Pro	Ala	Glu	Lys	Asn	Leu	305	310	315	320
Pro	Val	Leu	Leu	Ala	Leu	Ile	Gly	Ile	Trp	Tyr	Asn	Asn	Phe	Phe	Gly	325	330	335	
Ala	Glu	Thr	Glu	Ala	Ile	Leu	Pro	Tyr	Asp	Gln	Tyr	Met	His	Arg	Phe	340	345	350	
Ala	Ala	Tyr	Phe	Gln	Gln	Gly	Asn	Met	Glu	Ser	Asn	Gly	Lys	Tyr	Val	355	360	365	
Asp	Arg	Asn	Gly	Asn	Val	Val	Asp	Tyr	Gln	Thr	Gly	Pro	Ile	Ile	Trp	370	375	380	
Gly	Glu	Pro	Gly	Thr	Asn	Gly	Gln	His	Ala	Phe	Tyr	Gln	Leu	Ile	His	385	390	395	400

ES 2 704 697 T3

Gln Gly Thr Lys Met Val Pro Cys Asp Phe Ile Ala Pro Ala Ile Thr
405 410 415

His Asn Pro Leu Ser Asp His His Gln Lys Leu Leu Ser Asn Phe Phe
420 425 430

Ala Gln Thr Glu Ala Leu Ala Phe Gly Lys Ser Arg Glu Val Val Glu
435 440 445

Gln Glu Tyr Arg Asp Gln Gly Lys Asp Pro Ala Thr Leu Asp Tyr Val
450 455 460

Val Pro Phe Lys Val Phe Glu Gly Asn Arg Pro Thr Asn Ser Ile Leu
465 470 475 480

Leu Arg Glu Ile Thr Pro Phe Ser Leu Gly Ala Leu Ile Ala Leu Tyr
485 490 495

Glu His Lys Ile Phe Thr Gln Gly Val Ile Leu Asn Ile Phe Thr Phe
500 505 510

Asp Gln Trp Gly Val Glu Leu Gly Lys Gln Leu Ala Asn Arg Ile Leu
515 520 525

Pro Glu Leu Lys Asp Asp Lys Glu Ile Ser Ser His Asp Ser Ser Thr
530 535 540

Asn Gly Leu Ile Asn Arg Tyr Lys Ala Trp Arg Gly
545 550 555

<210> 21

<211> 557

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido sintético

<400> 21

Met Lys Asn Ile Asn Pro Thr Gln Thr Ala Ala Trp Gln Ala Leu Gln
1 5 10 15

Lys His Phe Asp Glu Met Lys Asp Val Thr Ile Ala Asp Leu Phe Ala
20 25 30

Lys Asp Gly Asp Arg Phe Ser Lys Phe Ser Ala Thr Phe Asp Asp Gln
35 40 45

ES 2 704 697 T3

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

ES 2 704 697 T3

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

ES 2 704 697 T3

<210> 22
 <211> 557
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Polipéptido sintético

<400> 22

```

Met Lys Asn Ile Asn Pro Thr Gln Thr Ala Ala Trp Gln Ala Leu Gln
1           5           10           15

Lys His Phe Asp Glu Met Lys Asp Val Thr Ile Ala Asp Leu Phe Ala
          20           25           30

Lys Asp Gly Asp Arg Phe Ser Lys Phe Ser Ala Thr Phe Asp Asp Gln
          35           40           45

Met Leu Val Asp Tyr Ser Lys Asn Arg Ile Thr Glu Glu Thr Leu Ala
50           55           60

Lys Leu Gln Asp Leu Ala Lys Glu Cys Asp Leu Ala Gly Ala Ile Lys
65           70           75           80

Ser Met Phe Ser Gly Glu Lys Ile Asn Arg Thr Glu Asn Arg Ala Val
          85           90           95

Leu His Val Ala Leu Arg Asn Arg Ser Asn Thr Pro Ile Leu Val Asp
          100          105          110

Gly Lys Asp Val Met Pro Glu Val Asn Ala Val Leu Glu Lys Met Lys
          115          120          125

Thr Phe Ser Glu Ala Ile Ile Ser Gly Glu Trp Lys Gly Tyr Thr Gly
130           135           140

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

Gly Pro Tyr Met Val Thr Glu Ala Leu Arg Pro Tyr Lys Asn His Leu
          165          170          175

Asn Met His Phe Val Ser Asn Val Asp Gly Thr His Ile Ala Glu Val
          180          185          190

Leu Lys Lys Val Asn Pro Glu Thr Thr Leu Phe Leu Val Ala Ser Lys
195           200           205
  
```

ES 2 704 697 T3

Thr	Phe	Thr	Thr	Gln	Glu	Thr	Met	Thr	Asn	Ala	His	Ser	Ala	Arg	Asp	210	215	220
Trp	Phe	Leu	Lys	Ala	Ala	Gly	Asp	Glu	Lys	His	Val	Ala	Lys	His	Phe	225	230	235 240
Ala	Ala	Leu	Ser	Thr	Asn	Ala	Lys	Ala	Val	Gly	Glu	Phe	Gly	Ile	Asp	245	250	255
Thr	Ala	Asn	Met	Phe	Glu	Phe	Trp	Asp	Trp	Val	Gly	Gly	Arg	Tyr	Ser	260	265	270
Leu	Trp	Ser	Ala	Ile	Gly	Leu	Ser	Ile	Val	Leu	Ser	Ile	Gly	Phe	Asp	275	280	285
Asn	Phe	Val	Glu	Leu	Leu	Ser	Gly	Ala	His	Ala	Met	Asp	Lys	His	Phe	290	295	300
Ser	Thr	Thr	Pro	Ala	Glu	Lys	Asn	Leu	Pro	Val	Leu	Leu	Ala	Leu	Ile	305	310	315 320
Gly	Ile	Trp	Tyr	Asn	Asn	Phe	Phe	Gly	Ala	Glu	Thr	Glu	Ala	Ile	Leu	325	330	335
Pro	Tyr	Asp	Gln	Tyr	Met	His	Arg	Phe	Ala	Ala	Tyr	Phe	Gln	Gln	Gly	340	345	350
Asn	Met	Glu	Ser	Asn	Gly	Lys	Tyr	Val	Asp	Arg	Asn	Gly	Asn	Val	Val	355	360	365
Asp	Tyr	Gln	Thr	Gly	Pro	Ile	Ile	Trp	Gly	Glu	Pro	Gly	Thr	Asn	Gly	370	375	380
Gln	His	Ala	Phe	Tyr	Gln	Leu	Ile	His	Gln	Gly	Thr	Lys	Met	Val	Pro	385	390	395 400
Cys	Asp	Phe	Ile	Ala	Pro	Ala	Ile	Thr	His	Leu	Glu	Val	Leu	Phe	Gln	405	410	415
Gly	Pro	Asn	Pro	Leu	Ser	Asp	His	His	Gln	Lys	Leu	Leu	Ser	Asn	Phe	420	425	430
Phe	Ala	Gln	Thr	Glu	Ala	Leu	Ala	Phe	Gly	Lys	Ser	Arg	Glu	Val	Val	435	440	445
Glu	Gln	Glu	Tyr	Arg	Asp	Gln	Gly	Lys	Asp	Pro	Ala	Thr	Leu	Asp	Tyr	450	455	460

ES 2 704 697 T3

Val Val Pro Phe Lys Val Phe Glu Gly Asn Arg Pro Thr Asn Ser Ile
465 470 475 480

Leu Leu Arg Glu Ile Thr Pro Phe Ser Leu Gly Ala Leu Ile Ala Leu
485 490 495

Tyr Glu His Lys Ile Phe Thr Gln Gly Val Ile Leu Asn Ile Phe Thr
500 505 510

Phe Asp Gln Trp Gly Val Glu Leu Gly Lys Gln Leu Ala Asn Arg Ile
515 520 525

Leu Pro Glu Leu Lys Asp Asp Lys Glu Ile Ser Ser His Asp Ser Ser
530 535 540

Thr Asn Gly Leu Ile Asn Arg Tyr Lys Ala Trp Arg Gly
545 550 555

<210> 23

<211> 556

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido sintético

<400> 23

Met Lys Asn Ile Asn Pro Thr Gln Thr Ala Ala Trp Gln Ala Leu Gln
1 5 10 15

Lys His Phe Asp Glu Met Lys Asp Val Thr Ile Ala Asp Leu Phe Ala
20 25 30

Lys Asp Gly Asp Arg Phe Ser Lys Phe Ser Ala Thr Phe Asp Asp Gln
35 40 45

Met Leu Val Asp Tyr Ser Lys Asn Arg Ile Thr Glu Glu Thr Leu Ala
50 55 60

Lys Leu Gln Asp Leu Ala Lys Glu Cys Asp Leu Ala Gly Ala Ile Lys
65 70 75 80

Ser Met Phe Ser Gly Glu Lys Ile Asn Arg Thr Glu Asn Arg Ala Val
85 90 95

Leu His Val Ala Leu Arg Asn Arg Ser Asn Thr Pro Ile Leu Val Asp
100 105 110

ES 2 704 697 T3

Gly Lys Asp Val Met Pro Glu Val Asn Ala Val Leu Glu Lys Met Lys
 115 120 125
 Thr Phe Ser Glu Ala Ile Ile Ser Gly Glu Trp Lys Gly Tyr Thr Gly
 130 135 140
 Lys Ala Ile Thr Asp Val Val Asn Ile Gly Ile Gly Gly Ser Asp Leu
 145 150 155 160
 Gly Pro Tyr Met Val Thr Glu Ala Leu Arg Pro Tyr Lys Asn His Leu
 165 170 175
 Asn Met His Phe Val Ser Asn Val Asp Gly Thr His Ile Ala Glu Val
 180 185 190
 Leu Lys Lys Val Asn Pro Glu Thr Thr Leu Phe Leu Val Ala Ser Lys
 195 200 205
 Thr Phe Thr Thr Gln Glu Thr Met Thr Asn Ala His Ser Ala Arg Asp
 210 215 220
 Trp Phe Leu Lys Ala Ala Gly Asp Glu Lys His Val Ala Lys His Phe
 225 230 235 240
 Ala Ala Leu Ser Thr Asn Ala Lys Ala Val Gly Glu Phe Gly Ile Asp
 245 250 255
 Thr Ala Asn Met Phe Glu Phe Trp Asp Trp Val Gly Gly Arg Tyr Ser
 260 265 270
 Leu Trp Ser Ala Ile Gly Leu Ser Ile Val Leu Ser Ile Gly Phe Asp
 275 280 285
 Asn Phe Val Glu Leu Leu Ser Gly Ala His Ala Met Asp Lys His Phe
 290 295 300
 Ser Thr Thr Pro Ala Glu Lys Asn Leu Pro Val Leu Leu Ala Leu Ile
 305 310 315 320
 Gly Ile Trp Tyr Asn Asn Phe Phe Gly Ala Glu Thr Glu Ala Ile Leu
 325 330 335
 Pro Tyr Asp Gln Tyr Met His Arg Phe Ala Ala Tyr Phe Gln Gln Gly
 340 345 350
 Asn Met Glu Ser Asn Gly Lys Tyr Val Asp Arg Asn Gly Asn Val Val
 355 360 365

ES 2 704 697 T3

Asp Tyr Gln Thr Gly Pro Ile Ile Trp Gly Glu Pro Gly Thr Asn Gly
370 375 380

Gln His Ala Phe Tyr Gln Leu Ile His Gln Gly Thr Lys Met Val Pro
385 390 395 400

Cys Asp Phe Ile Ala Pro Ala Ile Thr His Asn Pro Leu Ser Asp His
405 410 415

His Gln Lys Leu Leu Ser Asn Phe Phe Ala Gln Thr Glu Ala Leu Ala
420 425 430

Phe Gly Lys Ser Arg Glu Val Val Glu Gln Glu Tyr Arg Asp Gln Gly
435 440 445

Lys Asp Pro Ala Thr Leu Asp Tyr Val Val Pro Phe Lys Val Phe Glu
450 455 460

Gly Asn Arg Pro Thr Asn Ser Ile Leu Leu Arg Glu Ile Thr Pro Phe
465 470 475 480

Ser Leu Gly Ala Leu Ile Ala Leu Tyr Glu His Lys Ile Phe Thr Gln
485 490 495

Gly Val Ile Leu Asn Ile Phe Thr Phe Asp Gln Trp Gly Val Glu Leu
500 505 510

Gly Lys Gln Leu Ala Asn Arg Ile Leu Pro Glu Leu Glu Val Leu Phe
515 520 525

Gln Gly Pro Lys Asp Asp Lys Glu Ile Ser Ser His Asp Ser Ser Thr
530 535 540

Asn Gly Leu Ile Asn Arg Tyr Lys Ala Trp Arg Gly
545 550 555

<210> 24

<211> 557

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido sintético

<400> 24

Met Lys Asn Ile Asn Pro Thr Gln Thr Ala Ala Trp Gln Ala Leu Gln
1 5 10 15

ES 2 704 697 T3

Lys His Phe Asp Glu Met Lys Asp Val Thr Ile Ala Asp Leu Phe Ala
 20 25 30

Lys Asp Gly Asp Arg Phe Ser Lys Phe Ser Ala Thr Phe Asp Asp Gln
 35 40 45

Met Leu Val Asp Tyr Ser Lys Asn Arg Ile Thr Glu Glu Thr Leu Ala
 50 55 60

Lys Leu Gln Asp Leu Ala Lys Glu Cys Asp Leu Ala Gly Ala Ile Lys
 65 70 75 80

Ser Met Phe Ser Gly Glu Lys Ile Asn Arg Thr Glu Asn Arg Ala Val
 85 90 95

Leu His Val Ala Leu Arg Asn Arg Ser Asn Thr Pro Ile Leu Val Asp
 100 105 110

Gly Lys Asp Val Met Pro Glu Val Asn Ala Val Leu Glu Lys Met Lys
 115 120 125

Thr Phe Ser Glu Ala Ile Ile Ser Gly Glu Trp Lys Gly Tyr Thr Gly
 130 135 140

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

Gly Pro Tyr Met Val Thr Glu Ala Leu Arg Pro Tyr Lys Asn His Leu
 165 170 175

Asn Met His Phe Val Ser Asn Val Asp Gly Thr His Ile Ala Glu Val
 180 185 190

Leu Lys Lys Val Asn Pro Glu Thr Thr Leu Phe Leu Val Ala Ser Lys
 195 200 205

Thr Phe Thr Thr Gln Glu Thr Met Thr Asn Ala His Ser Ala Arg Asp
 210 215 220

Trp Phe Leu Lys Ala Ala Gly Asp Glu Lys His Val Ala Lys His Phe
 225 230 235 240

Ala Ala Leu Ser Thr Asn Ala Lys Ala Val Gly Glu Phe Gly Ile Asp
 245 250 255

Thr Ala Asn Met Phe Glu Phe Trp Asp Trp Val Gly Gly Arg Tyr Ser
 260 265 270

ES 2 704 697 T3

Leu Trp Ser Ala Ile Gly Leu Ser Ile Val Leu Ser Ile Gly Phe Asp
 275 280 285
 Asn Phe Val Glu Leu Leu Ser Gly Ala His Ala Met Asp Lys His Phe
 290 295 300
 Ser Thr Thr Pro Ala Glu Lys Asn Leu Pro Val Leu Leu Ala Leu Ile
 305 310 315 320
 Gly Ile Trp Tyr Asn Asn Phe Phe Gly Ala Glu Thr Glu Ala Ile Leu
 325 330 335
 Pro Tyr Asp Gln Tyr Met His Arg Phe Ala Ala Tyr Phe Gln Gln Gly
 340 345 350
 Asn Met Glu Ser Asn Gly Lys Tyr Val Asp Arg Asn Gly Asn Val Val
 355 360 365
 Asp Tyr Gln Thr Gly Pro Ile Ile Trp Gly Glu Pro Gly Thr Asn Gly
 370 375 380
 Gln His Ala Phe Tyr Gln Leu Ile His Gln Gly Thr Lys Met Val Pro
 385 390 395 400
 Cys Asp Phe Ile Ala Pro Ala Ile Thr His Asn Pro Leu Ser Asp His
 405 410 415
 His Gln Lys Leu Leu Ser Asn Phe Phe Ala Gln Thr Glu Ala Leu Ala
 420 425 430
 Phe Gly Lys Ser Arg Glu Val Val Glu Gln Glu Tyr Arg Asp Gln Gly
 435 440 445
 Lys Asp Pro Ala Thr Leu Asp Tyr Val Val Pro Phe Lys Val Phe Glu
 450 455 460
 Gly Asn Arg Pro Thr Asn Ser Ile Leu Leu Arg Glu Ile Thr Pro Phe
 465 470 475 480
 Ser Leu Gly Ala Leu Ile Ala Leu Tyr Glu His Lys Ile Phe Thr Gln
 485 490 495
 Gly Val Ile Leu Asn Ile Phe Thr Phe Asp Gln Trp Gly Val Glu Leu
 500 505 510
 Gly Lys Gln Leu Ala Asn Arg Ile Leu Pro Glu Leu Lys Leu Glu Val

ES 2 704 697 T3

515

520

525

Leu Phe Gln Gly Pro Asp Asp Lys Glu Ile Ser Ser His Asp Ser Ser
530 535 540

Thr Asn Gly Leu Ile Asn Arg Tyr Lys Ala Trp Arg Gly
545 550 555

<210> 25

<211> 557

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido sintético

<400> 25

Met Lys Asn Ile Asn Pro Thr Gln Thr Ala Ala Trp Gln Ala Leu Gln
1 5 10 15

Lys His Phe Asp Glu Met Lys Asp Val Thr Ile Ala Asp Leu Phe Ala
20 25 30

Lys Asp Gly Asp Arg Phe Ser Lys Phe Ser Ala Thr Phe Asp Asp Gln
35 40 45

Met Leu Val Asp Tyr Ser Lys Asn Arg Ile Thr Glu Glu Thr Leu Ala
50 55 60

Lys Leu Gln Asp Leu Ala Lys Glu Cys Asp Leu Ala Gly Ala Ile Lys
65 70 75 80

Ser Met Phe Ser Gly Glu Lys Ile Asn Arg Thr Glu Asn Arg Ala Val
85 90 95

Leu His Val Ala Leu Arg Asn Arg Ser Asn Thr Pro Ile Leu Val Asp
100 105 110

Gly Lys Asp Val Met Pro Glu Val Asn Ala Val Leu Glu Lys Met Lys
115 120 125

Thr Phe Ser Glu Ala Ile Ile Ser Gly Glu Trp Lys Gly Tyr Thr Gly
130 135 140

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

Gly Pro Tyr Met Val Thr Glu Ala Leu Arg Pro Tyr Lys Asn His Leu
165 170 175

ES 2 704 697 T3

Asn	Met	His	Phe	Val	Ser	Asn	Val	Asp	Gly	Thr	His	Ile	Ala	Glu	Val	180	185	190
Leu	Lys	Lys	Val	Asn	Pro	Glu	Thr	Thr	Leu	Phe	Leu	Val	Ala	Ser	Lys	195	200	205
Thr	Phe	Thr	Thr	Gln	Glu	Thr	Met	Thr	Asn	Ala	His	Ser	Ala	Arg	Asp	210	215	220
Trp	Phe	Leu	Lys	Ala	Ala	Gly	Asp	Glu	Lys	His	Val	Ala	Lys	His	Phe	225	230	235
Ala	Ala	Leu	Ser	Thr	Asn	Ala	Lys	Ala	Val	Gly	Glu	Phe	Gly	Ile	Asp	245	250	255
Thr	Ala	Asn	Met	Phe	Glu	Phe	Trp	Asp	Trp	Val	Gly	Gly	Arg	Tyr	Ser	260	265	270
Leu	Trp	Ser	Ala	Ile	Gly	Leu	Ser	Ile	Val	Leu	Ser	Ile	Gly	Phe	Asp	275	280	285
Asn	Phe	Val	Glu	Leu	Leu	Ser	Gly	Ala	His	Ala	Met	Asp	Lys	His	Phe	290	295	300
Ser	Thr	Thr	Pro	Ala	Glu	Lys	Asn	Leu	Pro	Val	Leu	Leu	Ala	Leu	Ile	305	310	315
Gly	Ile	Trp	Tyr	Asn	Asn	Phe	Phe	Gly	Ala	Glu	Thr	Glu	Ala	Ile	Leu	325	330	335
Pro	Tyr	Asp	Gln	Tyr	Met	His	Arg	Phe	Ala	Ala	Tyr	Phe	Gln	Gln	Gly	340	345	350
Asn	Met	Glu	Ser	Asn	Gly	Lys	Tyr	Val	Asp	Arg	Asn	Gly	Asn	Val	Val	355	360	365
Asp	Tyr	Gln	Thr	Gly	Pro	Ile	Ile	Trp	Gly	Glu	Pro	Gly	Thr	Asn	Gly	370	375	380
Gln	His	Ala	Phe	Tyr	Gln	Leu	Ile	His	Gln	Gly	Thr	Lys	Met	Val	Pro	385	390	395
Cys	Asp	Phe	Ile	Ala	Pro	Ala	Ile	Thr	His	Asn	Pro	Leu	Ser	Asp	His	405	410	415
His	Gln	Lys	Leu	Leu	Ser	Asn	Phe	Phe	Ala	Gln	Thr	Glu	Ala	Leu	Ala			

ES 2 704 697 T3

420

425

430

Phe Gly Lys Ser Arg Glu Val Val Glu Gln Glu Tyr Arg Asp Gln Gly
435 440 445

Lys Asp Pro Ala Thr Leu Asp Tyr Val Val Pro Phe Lys Val Phe Glu
450 455 460

Gly Asn Arg Pro Thr Asn Ser Ile Leu Leu Arg Glu Ile Thr Pro Phe
465 470 475 480

Ser Leu Gly Ala Leu Ile Ala Leu Tyr Glu His Lys Ile Phe Thr Gln
485 490 495

Gly Val Ile Leu Asn Ile Phe Thr Phe Asp Gln Trp Gly Val Glu Leu
500 505 510

Gly Lys Gln Leu Ala Asn Arg Ile Leu Pro Glu Leu Lys Asp Leu Glu
515 520 525

Val Leu Phe Gln Gly Pro Asp Lys Glu Ile Ser Ser His Asp Ser Ser
530 535 540

Thr Asn Gly Leu Ile Asn Arg Tyr Lys Ala Trp Arg Gly
545 550 555

<210> 26

<211> 557

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido sintético

<400> 26

Met Lys Asn Ile Asn Pro Thr Gln Thr Ala Ala Trp Gln Ala Leu Gln
1 5 10 15

Lys His Phe Asp Glu Met Lys Asp Val Thr Ile Ala Asp Leu Phe Ala
20 25 30

Lys Asp Gly Asp Arg Phe Ser Lys Phe Ser Ala Thr Phe Asp Asp Gln
35 40 45

Met Leu Val Asp Tyr Ser Lys Asn Arg Ile Thr Glu Glu Thr Leu Ala
50 55 60

Lys Leu Gln Asp Leu Ala Lys Glu Cys Asp Leu Ala Gly Ala Ile Lys
65 70 75 80

ES 2 704 697 T3

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

ES 2 704 697 T3

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

<210> 27
 <211> 557
 <212> PRT
 <213> Secuencia artificial

 <220>
 <223> Polipéptido sintético

ES 2 704 697 T3

<400> 27

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

ES 2 704 697 T3

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

ES 2 704 697 T3

Ser Leu Gly Ala Leu Ile Ala Leu Tyr Glu His Lys Ile Phe Thr Gln
485 490 495

Gly Val Ile Leu Asn Ile Phe Thr Phe Asp Gln Trp Gly Val Glu Leu
500 505 510

Gly Lys Gln Leu Ala Asn Arg Ile Leu Pro Glu Leu Lys Asp Asp Lys
515 520 525

Leu Glu Val Leu Phe Gln Gly Pro Glu Ile Ser Ser His Asp Ser Ser
530 535 540

Thr Asn Gly Leu Ile Asn Arg Tyr Lys Ala Trp Arg Gly
545 550 555

<210> 28

<211> 557

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido sintético

<400> 28

Met Lys Asn Ile Asn Pro Thr Gln Thr Ala Ala Trp Gln Ala Leu Gln
1 5 10 15

Lys His Phe Asp Glu Met Lys Asp Val Thr Ile Ala Asp Leu Phe Ala
20 25 30

Lys Asp Gly Asp Arg Phe Ser Lys Phe Ser Ala Thr Phe Asp Asp Gln
35 40 45

Met Leu Val Asp Tyr Ser Lys Asn Arg Ile Thr Glu Glu Thr Leu Ala
50 55 60

Lys Leu Gln Asp Leu Ala Lys Glu Cys Asp Leu Ala Gly Ala Ile Lys
65 70 75 80

Ser Met Phe Ser Gly Glu Lys Ile Asn Arg Thr Glu Asn Arg Ala Val
85 90 95

Leu His Val Ala Leu Arg Asn Arg Ser Asn Thr Pro Ile Leu Val Asp
100 105 110

Gly Lys Asp Val Met Pro Glu Val Asn Ala Val Leu Glu Lys Met Lys
115 120 125

Thr Phe Ser Glu Ala Ile Ile Ser Gly Glu Trp Lys Gly Tyr Thr Gly

ES 2 704 697 T3

130						135						140			
Lys	Ala	Ile	Thr	Asp	Val	Val	Asn	Ile	Gly	Ile	Gly	Gly	Ser	Asp	Leu
145					150					155					160
Gly	Pro	Tyr	Met	Val	Thr	Glu	Ala	Leu	Arg	Pro	Tyr	Lys	Asn	His	Leu
				165					170					175	
Asn	Met	His	Phe	Val	Ser	Asn	Val	Asp	Gly	Thr	His	Ile	Ala	Glu	Val
			180					185					190		
Leu	Lys	Lys	Val	Asn	Pro	Glu	Thr	Thr	Leu	Phe	Leu	Val	Ala	Ser	Lys
		195					200					205			
Thr	Phe	Thr	Thr	Gln	Glu	Thr	Met	Thr	Asn	Ala	His	Ser	Ala	Arg	Asp
	210					215					220				
Trp	Phe	Leu	Lys	Ala	Ala	Gly	Asp	Glu	Lys	His	Val	Ala	Lys	His	Phe
225					230					235					240
Ala	Ala	Leu	Ser	Thr	Asn	Ala	Lys	Ala	Val	Gly	Glu	Phe	Gly	Ile	Asp
				245					250					255	
Thr	Ala	Asn	Met	Phe	Glu	Phe	Trp	Asp	Trp	Val	Gly	Gly	Arg	Tyr	Ser
			260					265					270		
Leu	Trp	Ser	Ala	Ile	Gly	Leu	Ser	Ile	Val	Leu	Ser	Ile	Gly	Phe	Asp
		275					280					285			
Asn	Phe	Val	Glu	Leu	Leu	Ser	Gly	Ala	His	Ala	Met	Asp	Lys	His	Phe
	290					295					300				
Ser	Thr	Thr	Pro	Ala	Glu	Lys	Asn	Leu	Pro	Val	Leu	Leu	Ala	Leu	Ile
305					310					315					320
Gly	Ile	Trp	Tyr	Asn	Asn	Phe	Phe	Gly	Ala	Glu	Thr	Glu	Ala	Ile	Leu
				325					330					335	
Pro	Tyr	Asp	Gln	Tyr	Met	His	Arg	Phe	Ala	Ala	Tyr	Phe	Gln	Gln	Gly
			340					345					350		
Asn	Met	Glu	Ser	Asn	Gly	Lys	Tyr	Val	Asp	Arg	Asn	Gly	Asn	Val	Val
		355					360					365			
Asp	Tyr	Gln	Thr	Gly	Pro	Ile	Ile	Trp	Gly	Glu	Pro	Gly	Thr	Asn	Gly
	370					375					380				

ES 2 704 697 T3

Gln His Ala Phe Tyr Gln Leu Ile His Gln Gly Thr Lys Met Val Pro
385 390 395 400

Cys Asp Phe Ile Ala Pro Ala Ile Thr His Asn Pro Leu Ser Asp His
405 410 415

His Gln Lys Leu Leu Ser Asn Phe Phe Ala Gln Thr Glu Ala Leu Ala
420 425 430

Phe Gly Lys Ser Arg Glu Val Val Glu Gln Glu Tyr Arg Asp Gln Gly
435 440 445

Lys Asp Pro Ala Thr Leu Asp Tyr Val Val Pro Phe Lys Val Phe Glu
450 455 460

Gly Asn Arg Pro Thr Asn Ser Ile Leu Leu Arg Glu Ile Thr Pro Phe
465 470 475 480

Ser Leu Gly Ala Leu Ile Ala Leu Tyr Glu His Lys Ile Phe Thr Gln
485 490 495

Gly Val Ile Leu Asn Ile Phe Thr Phe Asp Gln Trp Gly Val Glu Leu
500 505 510

Gly Lys Gln Leu Ala Asn Arg Ile Leu Pro Glu Leu Lys Asp Asp Lys
515 520 525

Glu Leu Glu Val Leu Phe Gln Gly Pro Ile Ser Ser His Asp Ser Ser
530 535 540

Thr Asn Gly Leu Ile Asn Arg Tyr Lys Ala Trp Arg Gly
545 550 555

<210> 29

<211> 557

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido sintético

<400> 29

ES 2 704 697 T3

Met Lys Asn Ile Asn Pro Thr Gln Thr Ala Ala Trp Gln Ala Leu Gln
 1 5 10 15

Lys His Phe Asp Glu Met Lys Asp Val Thr Ile Ala Asp Leu Phe Ala
 20 25 30

Lys Asp Gly Asp Arg Phe Ser Lys Phe Ser Ala Thr Phe Asp Asp Gln

ES 2 704 697 T3

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

ES 2 704 697 T3

Asn	Phe	Val	Glu	Leu	Leu	Ser	Gly	Ala	His	Ala	Met	Asp	Lys	His	Phe	290	295	300
Ser	Thr	Thr	Pro	Ala	Glu	Lys	Asn	Leu	Pro	Val	Leu	Leu	Ala	Leu	Ile	305	310	315
Gly	Ile	Trp	Tyr	Asn	Asn	Phe	Phe	Gly	Ala	Glu	Thr	Glu	Ala	Ile	Leu	325	330	335
Pro	Tyr	Asp	Gln	Tyr	Met	His	Arg	Phe	Ala	Ala	Tyr	Phe	Gln	Gln	Gly	340	345	350
Asn	Met	Glu	Ser	Asn	Gly	Lys	Tyr	Val	Asp	Arg	Asn	Gly	Asn	Val	Val	355	360	365
Asp	Tyr	Gln	Thr	Gly	Pro	Ile	Ile	Trp	Gly	Glu	Pro	Gly	Thr	Asn	Gly	370	375	380
Gln	His	Ala	Phe	Tyr	Gln	Leu	Ile	His	Gln	Gly	Thr	Lys	Met	Val	Pro	385	390	395
Cys	Asp	Phe	Ile	Ala	Pro	Ala	Ile	Thr	His	Asn	Pro	Leu	Ser	Asp	His	405	410	415
His	Gln	Lys	Leu	Leu	Ser	Asn	Phe	Phe	Ala	Gln	Thr	Glu	Ala	Leu	Ala	420	425	430
Phe	Gly	Lys	Ser	Arg	Glu	Val	Val	Glu	Gln	Glu	Tyr	Arg	Asp	Gln	Gly	435	440	445
Lys	Asp	Pro	Ala	Thr	Leu	Asp	Tyr	Val	Val	Pro	Phe	Lys	Val	Phe	Glu	450	455	460
Gly	Asn	Arg	Pro	Thr	Asn	Ser	Ile	Leu	Leu	Arg	Glu	Ile	Thr	Pro	Phe	465	470	475
Ser	Leu	Gly	Ala	Leu	Ile	Ala	Leu	Tyr	Glu	His	Lys	Ile	Phe	Thr	Gln	485	490	495
Gly	Val	Ile	Leu	Asn	Ile	Phe	Thr	Phe	Asp	Gln	Trp	Gly	Val	Glu	Leu	500	505	510
Gly	Lys	Gln	Leu	Ala	Asn	Arg	Ile	Leu	Pro	Glu	Leu	Lys	Asp	Asp	Lys	515	520	525
Glu	Ile	Leu	Glu	Val	Leu	Phe	Gln	Gly	Pro	Ser	Ser	His	Asp	Ser	Ser	530	535	540

ES 2 704 697 T3

Thr Asn Gly Leu Ile Asn Arg Tyr Lys Ala Trp Arg Gly
545 550 555

<210> 30
<211> 557
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido sintético

<400> 30

Met Lys Asn Ile Asn Pro Thr Gln Thr Ala Ala Trp Gln Ala Leu Gln
1 5 10 15

Lys His Phe Asp Glu Met Lys Asp Val Thr Ile Ala Asp Leu Phe Ala
20 25 30

Lys Asp Gly Asp Arg Phe Ser Lys Phe Ser Ala Thr Phe Asp Asp Gln
35 40 45

Met Leu Val Asp Tyr Ser Lys Asn Arg Ile Thr Glu Glu Thr Leu Ala
50 55 60

Lys Leu Gln Asp Leu Ala Lys Glu Cys Asp Leu Ala Gly Ala Ile Lys
65 70 75 80

Ser Met Phe Ser Gly Glu Lys Ile Asn Arg Thr Glu Asn Arg Ala Val
85 90 95

Leu His Val Ala Leu Arg Asn Arg Ser Asn Thr Pro Ile Leu Val Asp
100 105 110

Gly Lys Asp Val Met Pro Glu Val Asn Ala Val Leu Glu Lys Met Lys
115 120 125

Thr Phe Ser Glu Ala Ile Ile Ser Gly Glu Trp Lys Gly Tyr Thr Gly
130 135 140

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

Gly Pro Tyr Met Val Thr Glu Ala Leu Arg Pro Tyr Lys Asn His Leu
165 170 175

Asn Met His Phe Val Ser Asn Val Asp Gly Thr His Ile Ala Glu Val
180 185 190

ES 2 704 697 T3

Leu Lys Lys Val Asn Pro Glu Thr Thr Leu Phe Leu Val Ala Ser Lys
 195 200 205
 Thr Phe Thr Thr Gln Glu Thr Met Thr Asn Ala His Ser Ala Arg Asp
 210 215 220
 Trp Phe Leu Lys Ala Ala Gly Asp Glu Lys His Val Ala Lys His Phe
 225 230 235 240
 Ala Ala Leu Ser Thr Asn Ala Lys Ala Val Gly Glu Phe Gly Ile Asp
 245 250 255
 Thr Ala Asn Met Phe Glu Phe Trp Asp Trp Val Gly Gly Arg Tyr Ser
 260 265 270
 Leu Trp Ser Ala Ile Gly Leu Ser Ile Val Leu Ser Ile Gly Phe Asp
 275 280 285
 Asn Phe Val Glu Leu Leu Ser Gly Ala His Ala Met Asp Lys His Phe
 290 295 300
 Ser Thr Thr Pro Ala Glu Lys Asn Leu Pro Val Leu Leu Ala Leu Ile
 305 310 315 320
 Gly Ile Trp Tyr Asn Asn Phe Phe Gly Ala Glu Thr Glu Ala Ile Leu
 325 330 335
 Pro Tyr Asp Gln Tyr Met His Arg Phe Ala Ala Tyr Phe Gln Gln Gly
 340 345 350
 Asn Met Glu Ser Asn Gly Lys Tyr Val Asp Arg Asn Gly Asn Val Val
 355 360 365
 Asp Tyr Gln Thr Gly Pro Ile Ile Trp Gly Glu Pro Gly Thr Asn Gly
 370 375 380
 Gln His Ala Phe Tyr Gln Leu Ile His Gln Gly Thr Lys Met Val Pro
 385 390 395 400
 Cys Asp Phe Ile Ala Pro Ala Ile Thr His Asn Pro Leu Ser Asp His
 405 410 415
 His Gln Lys Leu Leu Ser Asn Phe Phe Ala Gln Thr Glu Ala Leu Ala
 420 425 430
 Phe Gly Lys Ser Arg Glu Val Val Glu Gln Glu Tyr Arg Asp Gln Gly
 435 440 445

ES 2 704 697 T3

Lys Asp Pro Ala Thr Leu Asp Tyr Val Val Pro Phe Lys Val Phe Glu
450 455 460

Gly Asn Arg Pro Thr Asn Ser Ile Leu Leu Arg Glu Ile Thr Pro Phe
465 470 475 480

Ser Leu Gly Ala Leu Ile Ala Leu Tyr Glu His Lys Ile Phe Thr Gln
485 490 495

Gly Val Ile Leu Asn Ile Phe Thr Phe Asp Gln Trp Gly Val Glu Leu
500 505 510

Gly Lys Gln Leu Ala Asn Arg Ile Leu Pro Glu Leu Lys Asp Asp Lys
515 520 525

Glu Ile Ser Leu Glu Val Leu Phe Gln Gly Pro Ser His Asp Ser Ser
530 535 540

Thr Asn Gly Leu Ile Asn Arg Tyr Lys Ala Trp Arg Gly
545 550 555

<210> 31

<211> 557

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido sintético

<400> 31

Met Lys Asn Ile Asn Pro Thr Gln Thr Ala Ala Trp Gln Ala Leu Gln
1 5 10 15

Lys His Phe Asp Glu Met Lys Asp Val Thr Ile Ala Asp Leu Phe Ala
20 25 30

Lys Asp Gly Asp Arg Phe Ser Lys Phe Ser Ala Thr Phe Asp Asp Gln
35 40 45

Met Leu Val Asp Tyr Ser Lys Asn Arg Ile Thr Glu Glu Thr Leu Ala
50 55 60

Lys Leu Gln Asp Leu Ala Lys Glu Cys Asp Leu Ala Gly Ala Ile Lys
65 70 75 80

Ser Met Phe Ser Gly Glu Lys Ile Asn Arg Thr Glu Asn Arg Ala Val
85 90 95

ES 2 704 697 T3

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

Asn Met Glu Ser Asn Gly Lys Tyr Val Asp Arg Asn Gly Asn Val Val
 355 360 365
 Asp Tyr Gln Thr Gly Pro Ile Ile Trp Gly Glu Pro Gly Thr Asn Gly
 370 375 380
 Gln His Ala Phe Tyr Gln Leu Ile His Gln Gly Thr Lys Met Val Pro
 385 390 395 400
 Cys Asp Phe Ile Ala Pro Ala Ile Thr His Asn Pro Leu Ser Asp His
 405 410 415
 His Gln Lys Leu Leu Ser Asn Phe Phe Ala Gln Thr Glu Ala Leu Ala
 420 425 430
 Phe Gly Lys Ser Arg Glu Val Val Glu Gln Glu Tyr Arg Asp Gln Gly
 435 440 445
 Lys Asp Pro Ala Thr Leu Asp Tyr Val Val Pro Phe Lys Val Phe Glu
 450 455 460
 Gly Asn Arg Pro Thr Asn Ser Ile Leu Leu Arg Glu Ile Thr Pro Phe
 465 470 475 480
 Ser Leu Gly Ala Leu Ile Ala Leu Tyr Glu His Lys Ile Phe Thr Gln
 485 490 495
 Gly Val Ile Leu Asn Ile Phe Thr Phe Asp Gln Trp Gly Val Glu Leu
 500 505 510
 Gly Lys Gln Leu Ala Asn Arg Ile Leu Pro Glu Leu Lys Asp Asp Lys
 515 520 525
 Glu Ile Ser Ser Leu Glu Val Leu Phe Gln Gly Pro His Asp Ser Ser
 530 535 540
 Thr Asn Gly Leu Ile Asn Arg Tyr Lys Ala Trp Arg Gly
 545 550 555

<210> 32
 <211> 557
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Polipéptido sintético

<400> 32

ES 2 704 697 T3

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

ES 2 704 697 T3

Thr Ala Asn Met Phe Glu Phe Trp Asp Trp Val Gly Gly Arg Tyr Ser
 260 265 270
 Leu Trp Ser Ala Ile Gly Leu Ser Ile Val Leu Ser Ile Gly Phe Asp
 275 280 285
 Asn Phe Val Glu Leu Leu Ser Gly Ala His Ala Met Asp Lys His Phe
 290 295 300
 Ser Thr Thr Pro Ala Glu Lys Asn Leu Pro Val Leu Leu Ala Leu Ile
 305 310 315 320
 Gly Ile Trp Tyr Asn Asn Phe Phe Gly Ala Glu Thr Glu Ala Ile Leu
 325 330 335
 Pro Tyr Asp Gln Tyr Met His Arg Phe Ala Ala Tyr Phe Gln Gln Gly
 340 345 350
 Asn Met Glu Ser Asn Gly Lys Tyr Val Asp Arg Asn Gly Asn Val Val
 355 360 365
 Asp Tyr Gln Thr Gly Pro Ile Ile Trp Gly Glu Pro Gly Thr Asn Gly
 370 375 380
 Gln His Ala Phe Tyr Gln Leu Ile His Gln Gly Thr Lys Met Val Pro
 385 390 395 400
 Cys Asp Phe Ile Ala Pro Ala Ile Thr His Asn Pro Leu Ser Asp His
 405 410 415
 His Gln Lys Leu Leu Ser Asn Phe Phe Ala Gln Thr Glu Ala Leu Ala
 420 425 430
 Phe Gly Lys Ser Arg Glu Val Val Glu Gln Glu Tyr Arg Asp Gln Gly
 435 440 445
 Lys Asp Pro Ala Thr Leu Asp Tyr Val Val Pro Phe Lys Val Phe Glu
 450 455 460
 Gly Asn Arg Pro Thr Asn Ser Ile Leu Leu Arg Glu Ile Thr Pro Phe
 465 470 475 480
 Ser Leu Gly Ala Leu Ile Ala Leu Tyr Glu His Lys Ile Phe Thr Gln
 485 490 495
 Gly Val Ile Leu Asn Ile Phe Thr Phe Asp Gln Trp Gly Val Glu Leu
 500 505 510

ES 2 704 697 T3

Gly Lys Gln Leu Ala Asn Arg Ile Leu Pro Glu Leu Lys Asp Asp Lys
515 520 525

Glu Ile Ser Ser His Asp Ser Ser Thr Asn Gly Leu Ile Asn Arg Tyr
530 535 540

Lys Leu Glu Val Leu Phe Gln Gly Pro Ala Trp Arg Gly
545 550 555

5 <210> 33
<211> 558
<212> ADN
<213> Secuencia artificial

10 <220>
<223> Oligonucleótido sintético
<400> 33

```
atgggcccag aagaagaatt cggcatgagc ctgatcaagc ataactcttg cgtcattacc      60
acggagaatg gtaagttcac gggcttgggc gtttatgacc gtttcgtcgt ggttccgacc      120
cacgctgacc cgggtaaaga aatccagggt gacggtatca cgaccaaagt gattgatagc      180
tatgatctct ataataagaa cggcatcaag ctggaaatca cggtgctgaa actggaccgt      240
aatgaaaagt ttcgtgatat ccgtcgctat attccgaata acgaggatga ctacccaaat      300
tgcaatctgg cgctgctggc aaatcagccg gaaccgacga tcatcaacgt gggtgacgtg      360
gtgagctatg gcaatatcct gctgagcggc aaccagaccg cgcgtatgct gaagtattcc      420
tatccgacga aaagcggcta ttgcggcggc gtgctctata agattgggtca agtcctgggc      480
atccacgtcg gcggtaatgg ccgcgatggc ttcagcgcga tgctgctgcg tagctatttc      540
accgacgtcc agtgataa                                     558
```

15 <210> 34
<211> 184
<212> PRT
<213> Secuencia artificial

20 <220>
<223> Polipéptido sintético
<400> 34

ES 2 704 697 T3

Met Gly Pro Glu Glu Glu Phe Gly Met Ser Leu Ile Lys His Asn Ser
1 5 10 15

Cys Val Ile Thr Thr Glu Asn Gly Lys Phe Thr Gly Leu Gly Val Tyr
20 25 30

Asp Arg Phe Val Val Val Pro Thr His Ala Asp Pro Gly Lys Glu Ile
35 40 45

Gln Val Asp Gly Ile Thr Thr Lys Val Ile Asp Ser Tyr Asp Leu Tyr
50 55 60

Asn Lys Asn Gly Ile Lys Leu Glu Ile Thr Val Leu Lys Leu Asp Arg
65 70 75 80

Asn Glu Lys Phe Arg Asp Ile Arg Arg Tyr Ile Pro Asn Asn Glu Asp
85 90 95

Asp Tyr Pro Asn Cys Asn Leu Ala Leu Leu Ala Asn Gln Pro Glu Pro
100 105 110

Thr Ile Ile Asn Val Gly Asp Val Val Ser Tyr Gly Asn Ile Leu Leu
115 120 125

Ser Gly Asn Gln Thr Ala Arg Met Leu Lys Tyr Ser Tyr Pro Thr Lys
130 135 140

Ser Gly Tyr Cys Gly Gly Val Leu Tyr Lys Ile Gly Gln Val Leu Gly
145 150 155 160

Ile His Val Gly Gly Asn Gly Arg Asp Gly Phe Ser Ala Met Leu Leu
165 170 175

Arg Ser Tyr Phe Thr Asp Val Gln
180

<210> 35
<211> 618
<212> ADN
<213> Secuencia artificial

<220>
<223> Oligonucleótido sintético

<400> 35

ES 2 704 697 T3

```

atgaaaaaaa cggcaattgc gatagcgggtt gcgctagctg gttttgccac ggtggcgag      60
gctggcccag aagaagaatt cggcatgagc ctgatcaagc ataactcttg cgtcattacc      120
acggagaaatg gtaagttcac gggcttgggc gtttatgacc gtttcgctgt ggttccgacc      180
cacgctgacc cgggtaaaga aatccagggtt gacggtatca cgaccaaagt gattgatagc      240
tatgatctct ataataagaa cggcatcaag ctggaaatca cggtgctgaa actggaccgt      300
aatgaaaagt ttcgtgatat ccgtcgctat attccgaata acgaggatga ctacccaaat      360
tgcaatctgg cgctgctggc aaatcagccg gaaccgacga tcatcaacgt gggtgacgtg      420
gtgagctatg gcaatatcct gctgagcggg aaccagaccg cgcgtatgct gaagtattcc      480

tatccgacga aaagcggcta ttgcggcggc gtgctctata agattgggtca agtcctgggc      540
atccacgtcg gcggtaatgg ccgcgatggg ttcagcgcga tgctgctgcg tagctatttc      600
accgacgtcc agtgataa                                     618

```

<210> 36
 <211> 204
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Polipéptido sintético

<400> 36

ES 2 704 697 T3

```

Met Lys Lys Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly Phe Ala
1      5      10      15

Thr Val Ala Gln Ala Gly Pro Glu Glu Glu Phe Gly Met Ser Leu Ile
      20      25      30

Lys His Asn Ser Cys Val Ile Thr Thr Glu Asn Gly Lys Phe Thr Gly
      35      40      45

Leu Gly Val Tyr Asp Arg Phe Val Val Val Pro Thr His Ala Asp Pro
      50      55      60

Gly Lys Glu Ile Gln Val Asp Gly Ile Thr Thr Lys Val Ile Asp Ser
      65      70      75      80

Tyr Asp Leu Tyr Asn Lys Asn Gly Ile Lys Leu Glu Ile Thr Val Leu
      85      90      95

Lys Leu Asp Arg Asn Glu Lys Phe Arg Asp Ile Arg Arg Tyr Ile Pro
      100      105      110

Asn Asn Glu Asp Asp Tyr Pro Asn Cys Asn Leu Ala Leu Leu Ala Asn
      115      120      125

Gln Pro Glu Pro Thr Ile Ile Asn Val Gly Asp Val Val Ser Tyr Gly
      130      135      140

Asn Ile Leu Leu Ser Gly Asn Gln Thr Ala Arg Met Leu Lys Tyr Ser
      145      150      155      160

Tyr Pro Thr Lys Ser Gly Tyr Cys Gly Gly Val Leu Tyr Lys Ile Gly
      165      170      175

Gln Val Leu Gly Ile His Val Gly Gly Asn Gly Arg Asp Gly Phe Ser

```

180

185

190

Ala Met Leu Leu Arg Ser Tyr Phe Thr Asp Val Gln 195 200

5 <210> 37
 <211> 24
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Oligonucleótido sintético

ES 2 704 697 T3

	<400> 37 ctggaagtgc tgttcaggg tccg	24
5	<210> 38 <211> 8 <212> PRT <213> Secuencia artificial	
10	<220> <223> Polipéptido sintético	
	<400> 38	
15		Leu Glu Val Leu Phe Gln Gly Pro 1 5
20	<210> 39 <211> 7 <212> PRT <213> Secuencia artificial	
	<220> <223> Polipéptido sintético	
25	<400> 39	
		Glu Val Leu Phe Gln Gly Pro 1 5
30	<210> 40 <211> 6 <212> PRT <213> Secuencia artificial	
35	<220> <223> Polipéptido sintético	
	<400> 40	
		Val Leu Phe Gln Gly Pro 1 5
40	<210> 41 <211> 5 <212> PRT <213> Secuencia artificial	
45	<220> <223> Polipéptido sintético	
	<400> 41	
50		Leu Phe Gln Gly Pro 1 5
55	<210> 42 <211> 7 <212> PRT <213> Secuencia artificial	
	<220> <223> Polipéptido sintético	
60		

	<400> 42	
		Leu Glu Val Leu Phe Gln Gly
		1 5
5	<210> 43 <211> 6 <212> PRT <213> Secuencia artificial	
10	<220> <223> Polipéptido sintético	
	<400> 43	
		Leu Glu Val Leu Phe Gln
		1 5
15	<210> 44 <211> 5 <212> PRT <213> Secuencia artificial	
20	<220> <223> Polipéptido sintético	
	<400> 44	
25		Leu Glu Val Leu Phe
		1 5
30	<210> 45 <211> 6 <212> PRT <213> Secuencia artificial	
35	<220> <223> Polipéptido sintético	
	<400> 45	
		Glu Val Leu Phe Gln Gly
		1 5
40	<210> 46 <211> 21 <212> PRT <213> Secuencia artificial	
45	<220> <223> Polipéptido sintético	
	<400> 46	
		Met Lys Lys Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly Phe Ala
		1 5 10 15
50		Thr Val Ala Gln Ala
		20

<210> 47
 <211> 2145
 <212> ADN
 <213> Secuencia artificial

5

<220>
 <223> Oligonucleótido sintético

<400> 47

10

gtgtcccgtgta ttattatgct gatccctacc ggaaccagcg tcggtctgac cagcgtcagc	60
cttggcgtga tccgtgcaat ggaacgcaaa ggcgttcgtc tgagcgtttt caaacctatc	120
gctcagccgc gtaccggtgg cgatgcgccc gatcagacta cgactatcgt gcgtgcgaac	180
tcttccacca cgacggccgc tgaaccgctg aaaatgagct acgttgaagg tctgctttcc	240
agcaatcaga aagatgtgct gatggaagag atcgtcgcaa actaccacgc taacaccaaaa	300
gacgctgaag tcgttctggt tgaaggctctg gtcccgcacac gtaagcacca gtttgcccag	360
tctctgaact acgaaatcgc taaaacgctg aatgcggaaa tcgtcttcgt tatgtctcag	420
ggcactgaca ccccggaaca gctgaaagag cgtatcgaaac tgacccgcaa cagcttcggc	480
ggtgccaaaa acaccaacat caccggcgctt atcgttaaca aactgaacgc accggttgat	540
gaacaggggtc gtactcgcgc ggatctgtcc gagattttcg acgactcttc caaagctaaa	600
gtaaacaatg ttgatccggc gaagctgcaa gaatccagcc cgctgccggc tctcggcgct	660
gtgccgtgga gctttgacct gatcgcgact cgtgcgatcg atatggctcg ccacctgaat	720
gcgaccatca tcaacgaagg cgacatcaat actcgccgcg ttaaataccgt cactttctgc	780
gcacgcagca ttccgcacat gctggagcac ttccgtgccg gttctctgct ggtgacttcc	840
gcagaccgtc ctgacgtgct ggtggccgct tgcctggcag ccatgaacgg cgtagaaatc	900
ggtgccctgc tgctgactgg cggttacgaa atggacgcgc gcatttctaa actgtgcgaa	960

cgtgcttttcg ctaccggcct gccggtatatt atggtgaaca ccaacacctg gcagacctct 1020
 ctgagcctgc agagcttcaa cctggaagtt ccggttgacg atcacgaacg tatcgagaaa 1080
 gttcaggaat acgttgctaa ctacatcaac gctgactgga tcgaatctct gactgccact 1140
 tctgagcgca gccgtcgtct gtctccgcct gcgttccgtt atcagctgac tgaacttgcg 1200
 cgcaaagcgg gcaaacgtat cgtactgccg gaaggtgacg aaccgcgtac cgtaaagca 1260
 gccgctatct gtgctgaacg tggatcgca acttgctgac tgctgggtaa tccggcagag 1320
 atcaaccgtg ttgcagcgtc tcagggtgta gaactgggtg cagggtattga aatcgttgat 1380
 ccagaagtgg ttcgcgaaaag ctatgttggc cgtctggtcg aactgcgtaa gaacaaaggc 1440
 atgaccgaaa ccgttgcccg cgaacagctg gaagacaacg tgggtgctcg tacgtgatg 1500
 ctggaacagg atgaagttga tggctctggt tccggtgctg ttcacactac cgcaaacacc 1560
 atccgtccgc cgctgcagct gatcaaaact gcaccgggca gctccctggc atcttccgtg 1620
 ttcttcatgc tgctgccgga acagggttac gtttacgggt actgtgcgat caaccggat 1680
 ccgaccgctg aacagctggc agaaatcgcg attcagtcgc ctgattccgc tgcggccttc 1740
 ggtatcgaac cgcgcgttgc tatgctctcc tactccaccg gtacttctgg tgcaggtagc 1800
 gacgtagaaa aagttcgcga agcaactcgt ctggcgcagg aaaaacgtcc tgacctgatg 1860
 atcgacggtc cgctgcagta cgacgctgcg gtaatggctg acgttgcgaa atccaaagcg 1920
 ccgaactctc cggttgacg tgcgcgtacc gtgttcatct tcccggatct gaacaccggt 1980
 aacaccacct acaaagcggc acagcgttct gccgacctga tctccatcgg gccgatgctg 2040
 cagggtatgc gcaagccggt taacgacctg tcccgtggcg cactggttga cgatatcgtc 2100
 tacaccatcg cgctgactgc gattcagtct gcacagcagc agtaa 2145

<210> 48

<211> 714

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido sintético

<400> 48

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

ES 2 704 697 T3

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

ES 2 704 697 T3

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

ES 2 704 697 T3

545		550		555		560									
Pro	Thr	Ala	Glu	Gln	Leu	Ala	Glu	Ile	Ala	Ile	Gln	Ser	Ala	Asp	Ser
				565					570					575	
Ala	Ala	Ala	Phe	Gly	Ile	Glu	Pro	Arg	Val	Ala	Met	Leu	Ser	Tyr	Ser
			580					585					590		
Thr	Gly	Thr	Ser	Gly	Ala	Gly	Ser	Asp	Val	Glu	Lys	Val	Arg	Glu	Ala
		595					600					605			
Thr	Arg	Leu	Ala	Gln	Glu	Lys	Arg	Pro	Asp	Leu	Met	Ile	Asp	Gly	Pro
	610					615					620				
Leu	Gln	Tyr	Asp	Ala	Ala	Val	Met	Ala	Asp	Val	Ala	Lys	Ser	Lys	Ala
625					630					635					640
Pro	Asn	Ser	Pro	Val	Ala	Gly	Arg	Ala	Thr	Val	Phe	Ile	Phe	Pro	Asp
				645					650					655	
Leu	Asn	Thr	Gly	Asn	Thr	Thr	Tyr	Lys	Ala	Val	Gln	Arg	Ser	Ala	Asp
			660					665					670		
Leu	Ile	Ser	Ile	Gly	Pro	Met	Leu	Gln	Gly	Met	Arg	Lys	Pro	Val	Asn
		675					680					685			
Asp	Leu	Ser	Arg	Gly	Ala	Leu	Val	Asp	Asp	Ile	Val	Tyr	Thr	Ile	Ala
	690					695					700				
Leu	Thr	Ala	Ile	Gln	Ser	Ala	Gln	Gln	Gln						
705					710										

<210> 49
 <211> 2169
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Oligonucleótido sintético

<400> 49

ES 2 704 697 T3

gtgtcccgta ttattatgct gatccctacc ggaaccagcg tcggtctgac cagcgtcagc	60
cttggcgtga tccgtgcaat ggaacgcaaa ggcgttcgtc tgagcgtttt caaacctatc	120
gctcagccgc gtaccggtgg cgatgcgccc gatcagacta cgactatcgt gcgtgcgaac	180
tcttccacca cgacggccgc tgaaccgctg aaaatgagct acgttgaagg tctgctttcc	240
agcaatcaga aagatgtgct gatggaagag atcgtcgcaa actaccacgc taacaccaaa	300
gacgctgaag tcgttctggt tgaaggctctg gtcccgacac gtaagcacca gtttgcccag	360

tctctgaact	acgaaatcgc	taaaacgctg	aatgcggaaa	tcgtcttcgt	tatgtctcag	420
ggcactgaca	ccccggaaca	gctgaaagag	cgtatcgaac	tgacccgcaa	cagcttcggc	480
ggtgccaaaa	acaccaacat	caccggcggt	atcgttaaca	aactgaacgc	accggttgat	540
gaacagggtc	gtactcgccc	ggatctgtcc	gagattttcg	acgactcttc	caaagctaaa	600
gtaaacaatg	ttgatccggc	gaagctgcaa	gaatccagcc	cgctgccggg	tctcggcgct	660
gtgccgtgga	gctttgacct	gatcgcgact	cgtgcgatcg	atatggctcg	ccacctgaat	720
gcgaccatca	tcaacgaagg	cgacatcaat	actcgccgcg	ttaaatccgt	cactttctgc	780
gcacgcagca	ttccgcacat	gctggagcac	ttccgtgccg	gttctctgct	ggtgacttcc	840
gcagaccgtc	ctgacgtgct	ggtggccgct	tgccctggcag	ccatgaacgg	cgtagaaatc	900
ggtgccctgc	tgctgactgg	cggttacgaa	atggacgcgc	gcattttctaa	actgtgcgaa	960
cgtgctttcg	ctaccggcct	gccggtat	atggtgaaca	ccaacacctg	gcagacctct	1020
ctgagcctgc	agagcttcaa	cctggaagtt	ccggttgacg	atcacgaacg	tatcgagaaa	1080
gttcaggaat	acgttgctaa	ctacatcaac	gctgactgga	tcgaatctct	gactgccact	1140
tctctggaag	tgctgtttca	gggtccggag	cgcagccgtc	gtctgtctcc	gcctgcgttc	1200
cgttatcagc	tgactgaact	tgccgcgcaa	gcgggcaaac	gtatcgtact	gccggaagg	1260
gacgaaccgc	gtaccgttaa	agcagccgct	atctgtgctg	aacgtgggat	cgcaacttgc	1320
gtactgctgg	gtaatccggc	agagatcaac	cgtgttgacg	cgtctcaggg	tgtagaactg	1380
ggtgcagggg	ttgaaatcgt	tgatccagaa	gtggttcgcg	aaagctatgt	tggtcgtctg	1440
gtcgaactgc	gtaagaacaa	aggcatgacc	gaaaccgttg	cccgcgaaca	gctggaagac	1500
aacgtgggtc	tcgggtacgt	gatgctggaa	caggatgaag	ttgatgggtc	ggtttccggt	1560
gctgttcaca	ctaccgcaaa	caccatccgt	ccgccgctgc	agctgatcaa	aactgcaccg	1620
ggcagctccc	tggtatcttc	cgtgttcttc	atgctgctgc	cggaacagg	ttacgtttac	1680
ggtgactgtg	cgatcaaccc	ggatccgacc	gctgaacagc	tggcagaaat	cgcgattcag	1740
tccgctgatt	ccgctgcggc	cttcgggtatc	gaaccgcgcg	ttgctatgct	ctcctactcc	1800
accggtactt	ctggtgcagg	tagcgacgta	gaaaaagtgc	gcgaagcaac	tcgtctggcg	1860
caggaaaaac	gtcctgacct	gatgatcgac	ggtccgctgc	agtacgacgc	tgccgtaatg	1920
gctgacgttg	cgaaatccaa	agcgcgcaac	tctccggttg	caggtcgcgc	taccgtgttc	1980
atcttcccgg	atctgaacac	cggtaacacc	acctacaaag	cggtaacagc	ttctgccgac	2040
ctgatctcca	tcggggccgat	gctgcagggg	atgcgcaagc	cggttaacga	cctgtcccgt	2100
ggcgactgg	ttgacgat	cgtctacacc	atcgcgctga	ctgcgattca	gtctgcacag	2160
cagcagtaa						2169

<210> 50
 <211> 722
 <212> PRT
 <213> Secuencia artificial

5

<220>
 <223> Polipéptido sintético

10

<400> 50

Val Ser Arg Ile Ile Met Leu Ile Pro Thr Gly Thr Ser Val Gly Leu
 1 5 10 15

Thr Ser Val Ser Leu Gly Val Ile Arg Ala Met Glu Arg Lys Gly Val
 20 25 30

Arg Leu Ser Val Phe Lys Pro Ile Ala Gln Pro Arg Thr Gly Gly Asp
 35 40 45

Ala Pro Asp Gln Thr Thr Thr Ile Val Arg Ala Asn Ser Ser Thr Thr
 50 55 60

Thr Ala Ala Glu Pro Leu Lys Met Ser Tyr Val Glu Gly Leu Leu Ser
 65 70 75 80

Ser Asn Gln Lys Asp Val Leu Met Glu Glu Ile Val Ala Asn Tyr His
 85 90 95

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

Thr Arg Lys His Gln Phe Ala Gln Ser Leu Asn Tyr Glu Ile Ala Lys
 115 120 125

Thr Leu Asn Ala Glu Ile Val Phe Val Met Ser Gln Gly Thr Asp Thr
 130 135 140

Pro Glu Gln Leu Lys Glu Arg Ile Glu Leu Thr Arg Asn Ser Phe Gly
 145 150 155 160

Gly Ala Lys Asn Thr Asn Ile Thr Gly Val Ile Val Asn Lys Leu Asn
 165 170 175

Ala Pro Val Asp Glu Gln Gly Arg Thr Arg Pro Asp Leu Ser Glu Ile
 180 185 190

Phe Asp Asp Ser Ser Lys Ala Lys Val Asn Asn Val Asp Pro Ala Lys
 195 200 205

ES 2 704 697 T3

Leu Gln Glu Ser Ser Pro Leu Pro Val Leu Gly Ala Val Pro Trp Ser
 210 215 220

Phe Asp Leu Ile Ala Thr Arg Ala Ile Asp Met Ala Arg His Leu Asn
 225 230 235 240

Ala Thr Ile Ile Asn Glu Gly Asp Ile Asn Thr Arg Arg Val Lys Ser
 245 250 255

Val Thr Phe Cys Ala Arg Ser Ile Pro His Met Leu Glu His Phe Arg
 260 265 270

Ala Gly Ser Leu Leu Val Thr Ser Ala Asp Arg Pro Asp Val Leu Val
 275 280 285

Ala Ala Cys Leu Ala Ala Met Asn Gly Val Glu Ile Gly Ala Leu Leu
 290 295 300

Leu Thr Gly Gly Tyr Glu Met Asp Ala Arg Ile Ser Lys Leu Cys Glu
 305 310 315 320

Arg Ala Phe Ala Thr Gly Leu Pro Val Phe Met Val Asn Thr Asn Thr
 325 330 335

Trp Gln Thr Ser Leu Ser Leu Gln Ser Phe Asn Leu Glu Val Pro Val
 340 345 350

Asp Asp His Glu Arg Ile Glu Lys Val Gln Glu Tyr Val Ala Asn Tyr
 355 360 365

Ile Asn Ala Asp Trp Ile Glu Ser Leu Thr Ala Thr Ser Leu Glu Val
 370 375 380

Leu Phe Gln Gly Pro Glu Arg Ser Arg Arg Leu Ser Pro Pro Ala Phe
 385 390 395 400

Arg Tyr Gln Leu Thr Glu Leu Ala Arg Lys Ala Gly Lys Arg Ile Val
 405 410 415

Leu Pro Glu Gly Asp Glu Pro Arg Thr Val Lys Ala Ala Ala Ile Cys
 420 425 430

Ala Glu Arg Gly Ile Ala Thr Cys Val Leu Leu Gly Asn Pro Ala Glu
 435 440 445

Ile Asn Arg Val Ala Ala Ser Gln Gly Val Glu Leu Gly Ala Gly Ile
 450 455 460

ES 2 704 697 T3

Glu 465	Ile	Val	Asp	Pro	Glu 470	Val	Val	Arg	Glu	Ser 475	Tyr	Val	Gly	Arg	Leu 480
Val	Glu	Leu	Arg	Lys 485	Asn	Lys	Gly	Met	Thr 490	Glu	Thr	Val	Ala	Arg 495	Glu
Gln	Leu	Glu	Asp 500	Asn	Val	Val	Leu	Gly 505	Thr	Leu	Met	Leu	Glu 510	Gln	Asp
Glu	Val	Asp 515	Gly	Leu	Val	Ser	Gly 520	Ala	Val	His	Thr	Thr 525	Ala	Asn	Thr
Ile	Arg 530	Pro	Pro	Leu	Gln	Leu 535	Ile	Lys	Thr	Ala	Pro 540	Gly	Ser	Ser	Leu
Val 545	Ser	Ser	Val	Phe	Phe 550	Met	Leu	Leu	Pro	Glu 555	Gln	Val	Tyr	Val	Tyr 560
Gly	Asp	Cys	Ala	Ile 565	Asn	Pro	Asp	Pro	Thr 570	Ala	Glu	Gln	Leu	Ala 575	Glu
Ile	Ala	Ile	Gln 580	Ser	Ala	Asp	Ser	Ala 585	Ala	Ala	Phe	Gly	Ile 590	Glu	Pro
Arg	Val	Ala	Met 595	Leu	Ser	Tyr	Ser 600	Thr	Gly	Thr	Ser	Gly 605	Ala	Gly	Ser
Asp 610	Val	Glu	Lys	Val	Arg	Glu 615	Ala	Thr	Arg	Leu	Ala 620	Gln	Glu	Lys	Arg
Pro 625	Asp	Leu	Met	Ile	Asp 630	Gly	Pro	Leu	Gln	Tyr 635	Asp	Ala	Ala	Val	Met 640
Ala	Asp	Val	Ala	Lys 645	Ser	Lys	Ala	Pro	Asn 650	Ser	Pro	Val	Ala	Gly 655	Arg
Ala	Thr	Val	Phe 660	Ile	Phe	Pro	Asp	Leu 665	Asn	Thr	Gly	Asn 670	Thr	Thr	Tyr
Lys	Ala	Val	Gln 675	Arg	Ser	Ala	Asp 680	Leu	Ile	Ser	Ile	Gly 685	Pro	Met	Leu
Gln 690	Gly	Met	Arg	Lys	Pro	Val	Asn 695	Asp	Leu	Ser	Arg 700	Gly	Ala	Leu	Val
Asp 705	Asp	Ile	Val	Tyr	Thr 710	Ile	Ala	Leu	Thr	Ala 715	Ile	Gln	Ser	Ala	Gln 720

Gln Gln

<210> 51
<211> 2169
<212> ADN
<213> Secuencia artificial

<220>
<223> Oligonucleótido sintético

<400> 51

```

gtgtcccgtta ttattatgct gatccctacc ggaaccagcg tcggtctgac cagcgtcagc      60
cttggcgtga tccgtgcaat ggaacgcaaa ggcgttcgtc tgagcgtttt caaacctatc      120
gctcagccgc gtaccggtgg cgatgcgccc gatcagacta cgactatcgt gcgtagcgaac      180
tcttccacca cgacggccgc tgaaccgctg aaaatgagct acgttgaagg tctgctttcc      240
agcaatcaga aagatgtgct gatggaagag atcgtcgcaa actaccacgc taacacccaaa      300
gacgctgaag tcgttctggt tgaaggctctg gtcccgacac gtaagcacca gtttgcccag      360
tctctgaact acgaaatcgc taaaacgctg aatgcggaaa tcgtcttcgt tatgtctcag      420
ggcactgaca ccccggaaca gctgaaagag cgtatcgaac tgaccgcgaa cagcttcggc      480
ggtgccaaaa acaccaacat caccggcggt atcgttaaca aactgaacgc accggttgat      540
gaacagggtc gtactcgccc ggatctgtcc gagattttcg acgactcttc caaagctaaa      600
gtaaacaaatg ttgatccggc gaagctgcaa gaatccagcc cgctgccggg tctcggcgct      660
gtgccgtgga gctttgacct gatcgcgact cgtgcgatcg atatggctcg ccacctgaat      720
gcgaccatca tcaacgaagg cgacatcaat actcgccgcg ttaaaccggt cactttctgc      780
gcacgcagca ttccgcacat gctggagcac ttccgtgccg gttctctgct ggtgacttcc      840
gcagaccgtc ctgacgtgct ggtggccgct tgcctggcag ccatgaacgg cgtagaaatc      900
ggtgccctgc tgctgactgg cggttacgaa atggacgcgc gcatttctaa actgtgcgaa      960
cgtgctttcg ctaccggcct gccggtatct atggtgaaca ccaacacctg gcagacctct     1020
ctgagcctgc agagcttcaa cctggaagtt ccggttgacg atcacgaacg tatcgagaaa     1080
gttcaggaat acgttgctaa ctacatcaac gctgactgga tcgaatctct gactgccact     1140
tctgagctgg aagtgtgtt tcagggtccg cgcagccgtc gtctgtctcc gcctgcgttc     1200
cgttatcagc tgactgaact tgcgcgcaaa gcgggcaaac gtatcgact gccggaaggt     1260
gacgaaccgc gtaccgttaa agcagccgct atctgtgctg aacgtggtat cgcaacttgc     1320
gtactgctgg gtaatccggc agagatcaac cgtgttgacg cgtctcaggg tgtagaactg     1380
ggtgcaggga ttgaaatcgt tgatccagaa gtggttcgcg aaagctatgt tggtcgtctg     1440

```

ES 2 704 697 T3

```

gtcgaactgc gtaagaacaa aggcattgacc gaaaccgttg cccgcgaaca gctggaagac      1500
aacgtggtgc tcggtacgct gatgctggaa caggatgaag ttgatggtct ggtttccggt      1560
gctgttcaca ctaccgcaaa caccatccgt ccgccgctgc agctgatcaa aactgcaccg      1620
ggcagctccc tggatatctt cgtgttcttc atgctgctgc cggaacaggt ttacgtttac      1680
ggtgactgtg cgatcaaccc ggatccgacc gctgaacagc tggcagaaat cgcgattcag      1740
tccgctgatt ccgctgcggc cttcggtatc gaaccgcgcg ttgctatgct ctcctactcc      1800
accggtactt ctggtgcagg tagcgacgta gaaaaagtgc gcgaagcaac tcgtctggcg      1860
caggaaaaac gtcctgacct gatgatcgac ggtccgctgc agtacgacgc tgcggtaatg      1920
gctgacgttg cgaaatccaa agcgccgaac tctccggttg caggtcgcgc taccgtgttc      1980
atcttcccg atctgaacac cggtaacacc acctacaaag cggtacagcg ttctgccgac      2040
ctgatctcca tcgggccgat gctgcagggt atgcgcaagc cggttaacga cctgtcccgt      2100
ggcgactgg ttgacgatat cgtctacacc atcgcgctga ctgcgattca gtctgcacag      2160
cagcagtaa                                     2169

```

<210> 52

<211> 722

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Polipéptido sintético

<400> 52

ES 2 704 697 T3

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

ES 2 704 697 T3

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

ES 2 704 697 T3

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

ES 2 704 697 T3

Asp Val Glu Lys Val Arg Glu Ala Thr Arg Leu Ala Gln Glu Lys Arg
610 615 620

Pro Asp Leu Met Ile Asp Gly Pro Leu Gln Tyr Asp Ala Ala Val Met
625 630 635 640

Ala Asp Val Ala Lys Ser Lys Ala Pro Asn Ser Pro Val Ala Gly Arg
645 650 655

Ala Thr Val Phe Ile Phe Pro Asp Leu Asn Thr Gly Asn Thr Thr Tyr
660 665 670

Lys Ala Val Gln Arg Ser Ala Asp Leu Ile Ser Ile Gly Pro Met Leu
675 680 685

Gln Gly Met Arg Lys Pro Val Asn Asp Leu Ser Arg Gly Ala Leu Val
690 695 700

Asp Asp Ile Val Tyr Thr Ile Ala Leu Thr Ala Ile Gln Ser Ala Gln
705 710 715 720

Gln Gln

<210> 53
<211> 2166
<212> ADN
<213> Secuencia artificial

<220>
<223>Oligonucleótido sintético

<400> 53

```
gtgtcccgtta ttattatgct gatccctacc ggaaccagcg tcggtctgac cagcgtcagc      60
cttggcgtga tccgtgcaat ggaacgcaaa ggcgttcgtc tgagcgtttt caaacctatc      120
gctcagccgc gtaccggtgg cgatgcgcc gatcagacta cgactatcgt gcgtgcgaac      180
tcttccacca cgacggccgc tgaaccgctg aaaatgagct acgttgaagg tctgctttcc      240
agcaatcaga aagatgtgct gatggaagag atcgtcgcaa actaccacgc taacaccaaa      300
gacgctgaag tcgttctggt tgaaggctcgt gtcccgacac gtaagcacca gtttgcccag      360
tctctgaact acgaaatcgc taaaacgctg aatgcggaaa tcgtcttcgt tatgtctcag      420
ggcactgaca ccccggaaca gctgaaagag cgtatcgaac tgacccgcaa cagcttcggc      480
ggtgccaaaa acaccaacat caccggcggt atcgtaaca aactgaacgc accggttgat      540
gaacagggtc gtactcgccc ggatctgtcc gagattttcg acgactcttc caaagctaaa      600
gtaaacaatg ttgatccggc gaagctgcaa gaatccagcc cgctgccggt tctcggcgct      660
```

ES 2 704 697 T3

gtgccgtgga gctttgacct gatcgcgact cgtgcgatcg atatggctcg ccacctgaat	720
gcgaccatca tcaacgaagg cgacatcaat actcgccgcg ttaaaccctg cactttctgc	780
gcacgcagca ttccgcacat gctggagcac ttccgtgccg gttctctgct ggtgacttcc	840
gcagaccgtc ctgacgtgct ggtggccgct tgcctggcag ccatgaacgg cgtagaaatc	900
ggtgccctgc tgctgactgg cggttacgaa atggacgcgc gcatttctaa actgtgcgaa	960
cgtgctttcg ctaccggcct gccggtatct atggtgaaca ccaacacctg gcagacctct	1020
ctgagcctgc agagcttcaa cctggaagtt ccggttgacg atcacgaacg tatcgagaaa	1080
gttcaggaat acgttgctaa ctacatcaac gctgactgga tcgaatctct gactgccact	1140
tctgagcgca gccgtcgtct ggaagtgtct tttcagggtc cgtctccgcc tgcgttccgt	1200
tatcagctga ctgaacttgc gcgcaaagcg ggcaaactga tcgtactgcc ggaaggtgac	1260
gaaccgcgta ccgttaaagc agccgctatc tgtgctgaac gtggtatcgc aacttgcgta	1320
ctgctgggta atccggcaga gatcaaccgt gttgcagcgt ctccagggtgt agaactgggt	1380
gcagggattg aaatcgttga tccagaagtg gttcgcgaaa gctatgttgg tctctggtc	1440
gaactgcgta agaacaaagg catgaccgaa accgttgccc gcgaacagct ggaagacaac	1500
gtggtgctcg gtacgtgat gctggaacag gatgaagttg atggtctggt ttccggtgct	1560
gttcacacta ccgcaaacac catccgtccg ccgctgcagc tgatcaaaac tgcaccgggc	1620
agctccctgg tatcttccgt gttcttcatg ctgctgccgg aacaggttta cgtttacggt	1680
gactgtgcga tcaaccggga tccgaccgct gaacagctgg cagaaatcgc gattcagtc	1740
gctgattccg ctgcggcctt cggtatcgaa ccgcgcgttg ctatgctctc ctactccacc	1800
ggtacttctg gtgcaggtag cgacgtagaa aaagttcgcg aagcaactcg tctggcgag	1860
gaaaaacgtc ctgacctgat gatcgacggc ccgctgcagt acgacgctgc ggtaatggct	1920
gacgttgca aatccaaagc gccgaactct ccggttgagc gtcgcgctac cgtgttcac	1980
ttcccggatc tgaacaccgg taacaccacc taaaagcgg tacagcgttc tgccgacctg	2040
atctccatcg ggccgatgct gcaggggatg cgcaagccgg ttaacgacct gtcccgtggc	2100
gcactgggtg acgatatcgt ctacaccatc gcgctgactg cgattcagtc tgcacagcag	2160
cagtaa	2166

<210> 54
 <211> 721
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Polipéptido sintético

<400> 54

ES 2 704 697 T3

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

ES 2 704 697 T3

Val	Thr	Phe	Cys	Ala	Arg	Ser	Ile	Pro	His	Met	Leu	Glu	His	Phe	Arg	260	265	270
Ala	Gly	Ser	Leu	Leu	Val	Thr	Ser	Ala	Asp	Arg	Pro	Asp	Val	Leu	Val	275	280	285
Ala	Ala	Cys	Leu	Ala	Ala	Met	Asn	Gly	Val	Glu	Ile	Gly	Ala	Leu	Leu	290	295	300
Leu	Thr	Gly	Gly	Tyr	Glu	Met	Asp	Ala	Arg	Ile	Ser	Lys	Leu	Cys	Glu	305	310	315
Arg	Ala	Phe	Ala	Thr	Gly	Leu	Pro	Val	Phe	Met	Val	Asn	Thr	Asn	Thr	325	330	335
Trp	Gln	Thr	Ser	Leu	Ser	Leu	Gln	Ser	Phe	Asn	Leu	Glu	Val	Pro	Val	340	345	350
Asp	Asp	His	Glu	Arg	Ile	Glu	Lys	Val	Gln	Glu	Tyr	Val	Ala	Asn	Tyr	355	360	365
Ile	Asn	Ala	Asp	Trp	Ile	Glu	Ser	Leu	Thr	Ala	Thr	Ser	Glu	Arg	Ser	370	375	380
Arg	Arg	Leu	Glu	Val	Leu	Phe	Gln	Gly	Pro	Ser	Pro	Pro	Ala	Phe	Arg	385	390	395
Tyr	Gln	Leu	Thr	Glu	Leu	Ala	Arg	Lys	Ala	Gly	Lys	Arg	Ile	Val	Leu	405	410	415
Pro	Glu	Gly	Asp	Glu	Pro	Arg	Thr	Val	Lys	Ala	Ala	Ala	Ile	Cys	Ala	420	425	430
Glu	Arg	Gly	Ile	Ala	Thr	Cys	Val	Leu	Leu	Gly	Asn	Pro	Ala	Glu	Ile	435	440	445
Asn	Arg	Val	Ala	Ala	Ser	Gln	Gly	Val	Glu	Leu	Gly	Ala	Gly	Ile	Glu	450	455	460
Ile	Val	Asp	Pro	Glu	Val	Val	Arg	Glu	Ser	Tyr	Val	Gly	Arg	Leu	Val	465	470	475
Glu	Leu	Arg	Lys	Asn	Lys	Gly	Met	Thr	Glu	Thr	Val	Ala	Arg	Glu	Gln	485	490	495
Leu	Glu	Asp	Asn	Val	Val	Leu	Gly	Thr	Leu	Met	Leu	Glu	Gln	Asp	Glu	500	505	510

ES 2 704 697 T3

Val Asp Gly Leu Val Ser Gly Ala Val His Thr Thr Ala Asn Thr Ile
515 520 525

Arg Pro Pro Leu Gln Leu Ile Lys Thr Ala Pro Gly Ser Ser Leu Val
530 535 540

Ser Ser Val Phe Phe Met Leu Leu Pro Glu Gln Val Tyr Val Tyr Gly
545 550 555 560

Asp Cys Ala Ile Asn Pro Asp Pro Thr Ala Glu Gln Leu Ala Glu Ile
565 570 575

Ala Ile Gln Ser Ala Asp Ser Ala Ala Ala Phe Gly Ile Glu Pro Arg
580 585 590

Val Ala Met Leu Ser Tyr Ser Thr Gly Thr Ser Gly Ala Gly Ser Asp
595 600 605

Val Glu Lys Val Arg Glu Ala Thr Arg Leu Ala Gln Glu Lys Arg Pro
610 615 620

Asp Leu Met Ile Asp Gly Pro Leu Gln Tyr Asp Ala Ala Val Met Ala
625 630 635 640

Asp Val Ala Lys Ser Lys Ala Pro Asn Ser Pro Val Ala Gly Arg Ala
645 650 655

Thr Val Phe Ile Phe Pro Asp Leu Asn Thr Gly Asn Thr Thr Tyr Lys
660 665 670

Ala Val Gln Arg Ser Ala Asp Leu Ile Ser Ile Gly Pro Met Leu Gln
675 680 685

Gly Met Arg Lys Pro Val Asn Asp Leu Ser Arg Gly Ala Leu Val Asp
690 695 700

Asp Ile Val Tyr Thr Ile Ala Leu Thr Ala Ile Gln Ser Ala Gln Gln
705 710 715 720

Gln

<210> 55
<211> 2163
<212> ADN
<213> Secuencia artificial

<220>
<223> Oligonucleótido sintético

<400> 55

gtgtcccgtgta ttattatgct gatccctacc ggaaccagcg tcggtctgac cagcgtcagc	60
cttgggcgtga tccgtgcaat ggaacgcaaa ggcgttcgctc tgagcgtttt caaacctatc	120
gctcagccgc gtaccggtgg cgatgcgccc gatcagacta cgactatcgt gcgtgcgaac	180
tcttccacca cgacggccgc tgaaccgctg aaaatgagct acgttgaagg tctgctttcc	240
agcaatcaga aagatgtgct gatggaagag atcgtcgcaa actaccacgc taacaccaaa	300
gacgtgaag tcgttctggt tgaaggctctg gtcccgcacac gtaagcacca gtttgccag	360
tctctgaact acgaaatcgc taaaacgctg aatgcggaaa tcgtcttcgt tatgtctcag	420
ggcactgaca ccccggaaca gctgaaagag cgtatcgaaac tgaccgcgcaa cagcttcggc	480
ggtgccaaaa acaccaacat caccggcggtt atcgttaaca aactgaacgc accggttgat	540
gaacagggtc gtactcgcgc ggatctgtcc gagattttcg acgactcttc caaagctaaa	600
gtaaacaaatg ttgatccggc gaagctgcaa gaatccagcc cgctgccggt tctcggcgct	660
gtgccgtgga gctttgacct gatcgcgact cgtgcgatcg atatggctcg ccacctgaat	720
gcgaccatca tcaacgaagg cgacatcaat actcgccgcg ttaaaccggt cactttctgc	780
gcacgcagca ttccgcacat gctggagcac ttccgtgccg gttctctgct ggtgacttcc	840
gcagaccgtc ctgacgtgct ggtggccgct tgcctggcag ccatgaacgg cgtagaaatc	900
ggtgccctgc tgctgactgg cggttacgaa atggacgcgc gcattttctaa actgtgcgaa	960
cgtgctttcg ctaccggcct gccggtattt atggtgaaca ccaacacctg gcagacctct	1020
ctgagcctgc agagcttcaa cctggaagtt ccggttgacg atcacgaacg tatcgagaaa	1080
gttcaggaat acgttgctaa ctacatcaac gctgactgga tcgaatctct gactgccact	1140
tctgagcgca gccgtcgtct gtctccgcct gcgttccggt atcagctgac tgaacttgcg	1200
cgcaaagcgg gcaaacgtat cgtactggaa gtgctgtttc aggggccgga aggtgacgaa	1260
ccgcgtaccg ttaaagcagc cgctatctgt gctgaacgtg gtatcgcaac ttgcgtactg	1320
ctgggtaatc cggcagagat caaccgtgtt gcagcgtctc aggggtgtaga actgggtgca	1380
gggattgaaa tcgttgatcc agaagtgggt cgcgaaagct atgttggtcg tctggtcgaa	1440
ctgcgtaaga acaaaggcat gaccgaaacc gttgcccgcg aacagctgga agacaacgtg	1500
gtgctcggta cgctgatgct ggaacaggat gaagttgatg gtctggtttc cggtgctgtt	1560
cacactaccg caaacaccat ccgtccgccg ctgcagctga tcaaaactgc accgggcagc	1620
tccctggtat cttccgtgtt cttcatgctg ctgccggaac aggtttacgt ttacggtgac	1680
tgtgcgatca acccggatcc gaccgctgaa cagctggcag aaatcgcgat tcagtcgct	1740
gattccgctg cggccttcgg tatcgaaccg cgcgttgcta tgctctccta ctccaccggt	1800

ES 2 704 697 T3

acttctggtg caggtagcga cgtagaaaaa gttcgcgaag caactcgtct ggcgcaggaa 1860
aaacgtcctg acctgatgat cgacgggtccg ctgcagtacg acgctgcggt aatggctgac 1920
gttgcgaaat ccaaagcgcc gaactctccg gttgcaggtc gcgctaccgt gttcatcttc 1980
ccggatctga acaccggtaa caccacctac aaagcggtag agcgttctgc cgacctgac 2040
tccatcgggc cgatgctgca gggtagcgcg aagccggta acgacctgtc ccgtggcgca 2100
ctgggtgacg atatcgtcta caccatcgcg ctgactgcga ttcagtctgc acagcagcag 2160
taa 2163

<210> 56
<211> 720
<212> PRT
<213> Secuencia artificial
<220>
<223> Polipéptido sintético
<400> 56

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

ES 2 704 697 T3

145		150		155		160									
Gly	Ala	Lys	Asn	Thr	Asn	Ile	Thr	Gly	Val	Ile	Val	Asn	Lys	Leu	Asn
			165					170						175	
Ala	Pro	Val	Asp	Glu	Gln	Gly	Arg	Thr	Arg	Pro	Asp	Leu	Ser	Glu	Ile
			180					185					190		
Phe	Asp	Asp	Ser	Ser	Lys	Ala	Lys	Val	Asn	Asn	Val	Asp	Pro	Ala	Lys
		195					200					205			
Leu	Gln	Glu	Ser	Ser	Pro	Leu	Pro	Val	Leu	Gly	Ala	Val	Pro	Trp	Ser
	210					215					220				
Phe	Asp	Leu	Ile	Ala	Thr	Arg	Ala	Ile	Asp	Met	Ala	Arg	His	Leu	Asn
225					230					235					240
Ala	Thr	Ile	Ile	Asn	Glu	Gly	Asp	Ile	Asn	Thr	Arg	Arg	Val	Lys	Ser
				245					250					255	
Val	Thr	Phe	Cys	Ala	Arg	Ser	Ile	Pro	His	Met	Leu	Glu	His	Phe	Arg
			260					265					270		
Ala	Gly	Ser	Leu	Leu	Val	Thr	Ser	Ala	Asp	Arg	Pro	Asp	Val	Leu	Val
		275					280					285			
Ala	Ala	Cys	Leu	Ala	Ala	Met	Asn	Gly	Val	Glu	Ile	Gly	Ala	Leu	Leu
	290					295					300				
Leu	Thr	Gly	Gly	Tyr	Glu	Met	Asp	Ala	Arg	Ile	Ser	Lys	Leu	Cys	Glu
305					310					315					320
Arg	Ala	Phe	Ala	Thr	Gly	Leu	Pro	Val	Phe	Met	Val	Asn	Thr	Asn	Thr
				325					330					335	
Trp	Gln	Thr	Ser	Leu	Ser	Leu	Gln	Ser	Phe	Asn	Leu	Glu	Val	Pro	Val
			340					345					350		
Asp	Asp	His	Glu	Arg	Ile	Glu	Lys	Val	Gln	Glu	Tyr	Val	Ala	Asn	Tyr
		355					360					365			
Ile	Asn	Ala	Asp	Trp	Ile	Glu	Ser	Leu	Thr	Ala	Thr	Ser	Glu	Arg	Ser
	370					375					380				
Arg	Arg	Leu	Ser	Pro	Pro	Ala	Phe	Arg	Tyr	Gln	Leu	Thr	Glu	Leu	Ala
385					390					395					400

Arg	Lys	Ala	Gly	Lys	Arg	Ile	Val	Leu	Glu	Val	Leu	Phe	Gln	Gly	Pro		
				405					410					415			
Glu	Gly	Asp	Glu	Pro	Arg	Thr	Val	Lys	Ala	Ala	Ala	Ile	Cys	Ala	Glu		
			420					425					430				
Arg	Gly	Ile	Ala	Thr	Cys	Val	Leu	Leu	Gly	Asn	Pro	Ala	Glu	Ile	Asn		
		435					440					445					
Arg	Val	Ala	Ala	Ser	Gln	Gly	Val	Glu	Leu	Gly	Ala	Gly	Ile	Glu	Ile		
	450					455					460						
Val	Asp	Pro	Glu	Val	Val	Arg	Glu	Ser	Tyr	Val	Gly	Arg	Leu	Val	Glu		
465					470				475						480		
Leu	Arg	Lys	Asn	Lys	Gly	Met	Thr	Glu	Thr	Val	Ala	Arg	Glu	Gln	Leu		
			485					490						495			
Glu	Asp	Asn	Val	Val	Leu	Gly	Thr	Leu	Met	Leu	Glu	Gln	Asp	Glu	Val		
			500					505					510				
Asp	Gly	Leu	Val	Ser	Gly	Ala	Val	His	Thr	Thr	Ala	Asn	Thr	Ile	Arg		
		515					520					525					
Pro	Pro	Leu	Gln	Leu	Ile	Lys	Thr	Ala	Pro	Gly	Ser	Ser	Leu	Val	Ser		
		530				535					540						
Ser	Val	Phe	Phe	Met	Leu	Leu	Pro	Glu	Gln	Val	Tyr	Val	Tyr	Gly	Asp		
545					550				555						560		
Cys	Ala	Ile	Asn	Pro	Asp	Pro	Thr	Ala	Glu	Gln	Leu	Ala	Glu	Ile	Ala		
				565					570					575			
Ile	Gln	Ser	Ala	Asp	Ser	Ala	Ala	Ala	Phe	Gly	Ile	Glu	Pro	Arg	Val		
			580					585					590				
Ala	Met	Leu	Ser	Tyr	Ser	Thr	Gly	Thr	Ser	Gly	Ala	Gly	Ser	Asp	Val		
		595					600					605					
Glu	Lys	Val	Arg	Glu	Ala	Thr	Arg	Leu	Ala	Gln	Glu	Lys	Arg	Pro	Asp		
	610					615					620						
Leu	Met	Ile	Asp	Gly	Pro	Leu	Gln	Tyr	Asp	Ala	Ala	Val	Met	Ala	Asp		
625					630				635						640		
Val	Ala	Lys	Ser	Lys	Ala	Pro	Asn	Ser	Pro	Val	Ala	Gly	Arg	Ala	Thr		
				645				650						655			

ES 2 704 697 T3

Val Phe Ile Phe Pro Asp Leu Asn Thr Gly Asn Thr Thr Tyr Lys Ala
660 665 670

Val Gln Arg Ser Ala Asp Leu Ile Ser Ile Gly Pro Met Leu Gln Gly
675 680 685

Met Arg Lys Pro Val Asn Asp Leu Ser Arg Gly Ala Leu Val Asp Asp
690 695 700

Ile Val Tyr Thr Ile Ala Leu Thr Ala Ile Gln Ser Ala Gln Gln Gln
705 710 715 720

<210> 57
<211> 1992
<212> ADN
<213> Secuencia artificial

<220>
<223> Oligonucleótido sintético

<400> 57

atgtcctcac	gtaaagagct	tgccaatgct	attcgtgcgc	tgagcatgga	cgagtacag	60
aaagccaaat	ccggtcacc	gggtgcccct	atgggtatgg	ctgacattgc	cgaagtcctg	120
tggcgtgatt	tcctgaaaca	caaccgcag	aatccgtcct	gggctgaccg	tgaccgcttc	180
gtgctgtcca	acggccacgg	ctccatgctg	atctacagcc	tgctgcacct	caccggttac	240
gatctgccga	tggaagaact	gaaaaacttc	cgtcagctgc	actctaaaac	tccgggtcac	300
ccggaagtgg	gttacaccgc	tggtgtggaa	accaccaccg	gtccgctggg	tcagggtatt	360
gccaacgcag	tcggtatggc	gattgcagaa	aaaacgctgg	cggcgcagtt	taaccgtccg	420
ggccacgaca	ttgtcgacca	ctacacctac	gccttcatgg	gcgacggctg	catgatggaa	480
ggcatctccc	acgaagtttg	ctctctggcg	ggtacgctga	agctgggtaa	actgattgca	540
ttctacgatg	acaacggtat	ttctatcgat	ggtcacgttg	aaggctgggt	caccgacgac	600
accgcaatgc	gtttcgaagc	ttacggctgg	cacgttattc	gcgacatcga	cggtcatgac	660
gcggcatcta	tcaaacgcgc	agtagaagaa	gcgcgcgcag	tgactgacaa	accttccctg	720
ctgatgtgca	aaaccatcat	cggtttcggg	tccccgaaca	aagccggtac	ccacgactcc	780
cacggtgcgc	cgctgggcca	cgctgaaatt	gccctgaccc	gcgaacaact	gggctggaaa	840
tatgcgccgt	tcgaaatccc	gtctgaaatc	tatgctcagt	gggatgcgaa	agaagcaggc	900
caggcgaaag	aatccgcatg	gaacgagaaa	ttcgctgctt	acgcgaaagc	ttatccgcag	960
gaagccgctg	aatttaccgc	ccgtatgaaa	ggcgaaatgc	cgtctgactt	cgacgctaaa	1020
gcgaaagagt	tcatcgctaa	actgcaggct	aatccggcca	aaatcgccag	ccgtaaagcg	1080
tctcagaatg	ctatcgaagc	gttcgggtccg	ctgttgccgg	aattcctcgg	cggttctgct	1140

ES 2 704 697 T3

gacctggcgc cgtctaacct gaccctgtgg tctggttcta aagcaatcaa cgaagatgct 1200
 gcgggtaact acatccacta cgggtgttcgc gagttcggta tgaccgcgat tgctaacggt 1260
 atctccctgc acggtggcctt cctgccgtac acctccacct tcctgatggt cgtggaatac 1320
 gcacgtaacg ccgtacgtat ggctgcgctg atgaaacagc gtcaggtgat ggtttacacc 1380
 cactgactcca tcggtctggg cgaagacggc ccgactcacc agccggttga gcaggtcgct 1440
 tctctgcgcg taaccccgaa catgtctaca tggcgtccgt gtgaccaggt tgaatccgcg 1500
 gtgcgctgga aatacgggtg tgagcgtcag gacggcccg cgcactgat cctctcccg 1560
 cagaacctgg cgcagcagga acgaactgaa gagcaactgg caaacatcgc gcgcggtggt 1620
 tatgtgctga aagactgcgc cggtcagccg gaactgattt tcatcgctac cggttcagaa 1680
 gttgaactgg ctgttgctgc ctacgaaaaa ctgactgccg aaggcgtgaa agcgcgcgtg 1740
 gtgtccatgc cgtctaccga cgcatttgac aagcaggatg ctgcttaccg tgaatccgta 1800
 ctgccgaaag cggttactgc acgcgttgct gtagaagcgg gtattgctga ctactggtac 1860
 aagtatgttg gcctgaacgg tgctatcgtc ggtatgacca ccttcggtga atctgctccg 1920
 gcagagctgc tgtttgaaga gttcggcttc actgttgata acgttggtgc gaaagcaaaa 1980
 gaactgctgt aa 1992

<210> 58
 <211> 2016
 <212> ADN
 <213> Secuencia artificial

<220>
 <223>Oligonucleótido sintético

<400> 58

atgtcctcac gtaaagagct tgccaatgct attcgtgcgc tgagcatgga cgcagtacag 60
 aaagccaaat ccggtcaccc gggtgcccct atgggtatgg ctgacattgc cgaagtccctg 120
 tggcgtgatt tcctgaaaca caaccgcag aatccgtcct gggctgaccg tgaccgcttc 180
 gtgctgtcca acggccacgg ctccatgctg atctacagcc tgctgcacct caccggttac 240
 gatctgccga tggaagaact gaaaaacttc cgtcagctgc actctaaaac tccgggtcac 300
 ccggaagtgg gttacaccgc tgggtgtggaa accaccaccg gtccgctggg tcagggtatt 360
 gccaacgcag tcggtatggc gattgcagaa aaaacgctgg cggcgcagtt taaccgtccg 420
 ggccacgaca ttgtcgacca ctacacctac gccttcatgg gcgacggctg catgatggaa 480
 ggcattctcc acgaagtttg ctctctggcg ggtacgctga agctgggtaa actgattgca 540
 ttctacgatg acaacgggtat ttctatcgat ggtcacgttg aaggctggtt caccgacgac 600
 accgcaatgc gtttcgaagc ttacggctgg cacgttattc gcgacatcga cggtcatgac 660

ES 2 704 697 T3

gcggcatcta tcaaacgcgc agtagaagaa gcgcgcgcag tgactgacaa accttccctg 720
ctgatgtgca aaaccatcat cggtttcggg tccccgaaca aagccggtac ccacgactcc 780
cacggtgcgc cgctgggcga cgctgaaatt gccctgaccc gcgaacaact gggctggaaa 840
tatgcgcgct tcgaaatccc gtctgaaatc tatgctcagt gggatgcgaa agaagcaggc 900
caggcgaaag aatccgcatg gaacgagaaa ttcgctgctt acgcgaaagc ttatccgcag 960
gaagccgctg aatttaccgc cggtatgaaa ggcgaaatgc cgtctgactt cgacgctaaa 1020
gcgaaagagt tcatcgctaa actgcaggct aatccggcga aaatcgccag ccgtaaagcg 1080
tctcagaatg ctatcgaagc gttcgggtccg ctgttgccgg aattcctcgg cggttctgct 1140
gacctggcgc cgtctaacct gacctgtgg tctggttcta aagcaatcaa cgaagatgct 1200
gcgggtaact acatccacta cgggtgttcgc gagttcggta tgaccgcgat tgctaacggt 1260
atctccctgc acggtggctt cctgccgtac acctccacct tcctgatgtt cgtggaatac 1320
gcacgtaacg ccgtacgtat ggctgcgctg atgaaacagc gtcagggtgat ggtttacacc 1380
cacgactcca tcggtctggg cgaagacggc ccgactcacc agccggttga gcaggtcgct 1440
tctctgcgcg taaccccgaa catgtctaca tggcgtccgt gtgaccaggt tgaatccgcg 1500
gtcgcgtgga aatacgggtg tgagcgtcag gacggcccg cgcactgat cctctcccg 1560
cagaacctgg cgcagcagga acgaactgaa gagcaactgg caaacatcgc gcgcggtggt 1620
tatgtgctga aagactgcgc cggtcagccg gaactgattt tcatcgctac cggttcagaa 1680
gttgaactgg ctgttgctgc ctacgaaaaa ctgactgccg aaggcgtgaa agcgcgcgtg 1740
gtgtccatgc cgtctaccga cgcatttgac aagcaggatg ctgcttaccg tgaatccgta 1800
ctgccgaaag cggttactgc acgcgttgct gtagaagcgg gtattgctga ctactggtac 1860
aagtatgttg gcctgaacgg tgctatcgtc ggtatgacca ccttcctgga agtgctgttt 1920
cagggtccgg gtgaatctgc tccggcagag ctgctgtttg aagagttcgg cttcactggt 1980
gataacgttg ttgcgaaagc aaaagaactg ctgtaa 2016

<210> 59
<211> 2016
<212> ADN
<213> Secuencia artificial

<220>
<223>Oligonucleótido sintético

<400> 59

atgtcctcac gtaaagagct tgccaatgct attcgtgcgc tgagcatgga cgcagtacag 60
aaagccaaat ccggtcaccg ggggtgccct atgggtatgg ctgacattgc cgaagtcctg 120
tggcgtgatt tcctgaaaca caaccgcgag aatccgtcct gggctgaccg tgaccgcttc 180
gtgctgtcca acggccacgg ctccatgctg atctacagcc tgctgcacct caccgggttac 240

gatctgccga	tggaagaact	gaaaaacttc	cgtcagctgc	actctaaaac	tccgggtcac	300
ccggaagtgg	gttacaccgc	tggtgtggaa	accaccaccg	gtccgctggg	tcagggtatt	360
gccaacgcag	tcggtatggc	gattgcagaa	aaaacgctgg	cggcgcagtt	taaccgtccg	420
ggccacgaca	ttgtcgacca	ctacacctac	gccttcatgg	gcgacggctg	catgatggaa	480
ggcatctccc	acgaagtttg	ctctctggcg	ggtacgctga	agctgggtaa	actgattgca	540
ttctacgatg	acaacgggat	ttctatcgat	ggtcacgttg	aaggctgggt	caccgacgac	600
accgcaatgc	gtttcgaagc	ttacggctgg	cacgttattc	gcgacatcga	cggtcatgac	660
gcggcatcta	tcaaacgcgc	agtagaagaa	gcgcgcgcag	tgactgacaa	accttccctg	720
ctgatgtgca	aaaccatcat	cggtttcggg	tccccgaaca	aagccggtac	ccacgactcc	780
cacggtgcgc	cgctgggcga	cgctgaaatt	gccctgaccc	gcgaacaact	gggctggaaa	840
tatgcgccgt	tcgaaatccc	gtctgaaatc	tatgctcagt	gggatgcgaa	agaagcaggc	900
caggcgaaag	aatccgcatg	gaacgagaaa	ttcgctgctt	acgcgaaagc	ttatccgcag	960
gaagccgctg	aattttacccg	ccgtatgaaa	ggcgaaatgc	cgtctgactt	cgacgctaaa	1020
gcgaaagagt	tcatcgctaa	actgcaggct	aatccggcga	aaatcgccag	ccgtaaagcg	1080
tctcagaatg	ctatcgaagc	gttcgggtccg	ctggtgccgg	aattcctcgg	cggttctgct	1140
gacctggcgc	cgtctaacct	gaccctgtgg	tctggttcta	aagcaatcaa	cgaagatgct	1200
gcgggtaact	acatccacta	cggtgttcgc	gagttcggta	tgaccgcgat	tgctaacggg	1260
atctccctgc	acggtggcct	cctgccgtac	acctccacct	tcctgatggt	cgtggaatac	1320
gcacgtaacg	ccgtacgtat	ggctgcgctg	atgaaacagc	gtcagggtgat	ggtttacacc	1380
cacgactcca	tcggtctggg	cgaagacggc	ccgactcacc	agccggttga	gcaggtcgct	1440
tctctgcgcg	taaccccgaa	catgtctaca	tggcgtccgt	gtgaccagggt	tgaatccgcg	1500
gtcgcgtgga	aatacgggtg	tgagcgtcag	gacggcccca	ccgcactgat	cctctcccgt	1560
cagaacctgg	cgcagcagga	acgaactgaa	gagcaactgg	caaacatcgc	gcgcgggtgg	1620
tatgtgctga	aagactgcgc	cggtcagccg	gaactgattt	tcatcgctac	cggttcagaa	1680
gttgaactgg	ctgttgctgc	ctacgaaaaa	ctgactgccg	aaggcgtgaa	agcgcgcgtg	1740
gtgtccatgc	cgtctaccga	cgcatttgac	aagcagggat	ctgcttaccg	tgaatccgta	1800
ctgccgaaag	cggttactgc	acgcgttgct	gtagaagcgg	gtattgctga	ctactggtac	1860
aagtatgttg	gcctgaacgg	tgctatcgtc	ggtatgacca	ccttcgggtc	ggaagtgctg	1920
tttcagggtc	cggaatctgc	tccggcagag	ctgctgtttg	aagagttcgg	cttccactgtt	1980
gataacgttg	ttgcgaaagc	aaaagaactg	ctgtaa			2016

<210> 60
 <211> 2016
 <212> ADN
 <213> Secuencia artificial

<220>

<223>Oligonucleótido sintético

<400> 60

5

atgtcctcac	gtaaagagct	tgccaatgct	attcgtgcgc	tgagcatgga	cgagtacag	60
aaagccaaat	ccggtcaccc	gggtgcccct	atgggtatgg	ctgacattgc	cgaagtcctg	120
tggcgtgatt	tcctgaaaca	caacccgcag	aatccgtcct	gggctgaccg	tgaccgcttc	180
gtgctgtcca	acggccacgg	ctccatgctg	atctacagcc	tgctgcacct	caccggttac	240
gatctgccga	tggaagaact	gaaaaacttc	cgtcagctgc	actctaaaac	tccgggtcac	300
ccggaagtgg	gttacaccgc	tggtgtggaa	accaccaccg	gtccgctggg	tcagggtatt	360
gccaacgcag	tcggtatggc	gattgcagaa	aaaacgctgg	cggcgcagtt	taaccgtccg	420
ggccacgaca	ttgtcgacca	ctacacctac	gccttcatgg	gcgacggctg	catgatggaa	480
ggcatctccc	acgaagtttg	ctctctggcg	ggtagcgtga	agctgggtaa	actgattgca	540
ttctacgatg	acaacggtat	ttctatcgat	ggtcacgttg	aaggctgggt	caccgacgac	600
accgcaatgc	gtttcgaagc	ttacggctgg	cacgttattc	gcgacatcga	cggtcatgac	660
gcggcatcta	tcaaacgcgc	agtagaagaa	gcgcgcgcag	tgactgacaa	accttccctg	720
ctgatgtgca	aaaccatcat	cggtttcggt	tccccgaaca	aagccggtac	ccacgactcc	780
cacggtgcgc	cgctgggcga	cgctgaaatt	gccctgacct	gcgaacaact	gggctggaaa	840
tatgcgccgt	tcgaaatccc	gtctgaaatc	tatgctcagt	gggatgcgaa	agaagcaggc	900
caggcgaaaag	aatccgcgatg	gaacgagaaa	ttcgttgctt	acgcgaaagc	ttatccgcag	960
gaagccgctg	aatttaccgg	ccgtatgaaa	ggcgaaatgc	cgtctgactt	cgacgctaaa	1020
gcgaaagagt	tcatcgctaa	actgcaggct	aatccggcga	aaatcgccag	ccgtaaagcg	1080
tctcagaatg	ctatcgaaagc	gttcggtccg	ctgttgccgg	aattcctcgg	cggttctgct	1140
gacctggcgc	cgtctaacct	gacctgtgg	tctggttcta	aagcaatcaa	cgaagatgct	1200
gcgggtaact	acatccacta	cggtgttcgc	gagttcggta	tgaccgcgat	tgctaaccgt	1260
atctccctgc	acggtggctt	cctgccgtac	acctccacct	tcctgatgtt	cgtggaatac	1320
gcacgtaacg	ccgtacgtat	ggctgcgctg	atgaaacagc	gtcaggtgat	ggtttacacc	1380
cacgactcca	tcggtctggg	cgaagacggc	ccgactcacc	agccggttga	gcaggtcgct	1440
tctctgcgcg	taacccccgaa	catgtctaca	tggcgtccgt	gtgaccagg	tgaatccgcg	1500
gtcgcgtgga	aatacgggtg	tgagcgtcag	gacggcccca	ccgcactgat	cctctcccgt	1560
cagaacctgg	cgcagcagga	acgaactgaa	gagcaactgg	caaacatcgc	gcgcgggtgg	1620
tatgtgctga	aagactgcgc	cggtcagccg	gaactgattt	tcatcgctac	cggttcagaa	1680

ES 2 704 697 T3

gttgaactgg ctgttgctgc ctacgaaaaa ctgactgccg aaggcgtgaa agcgcgcgtg	1740
gtgtccatgc cgtctaccga cgcatttgac aagcaggatg ctgcttaccg tgaatccgta	1800
ctgccgaaaag cggttactgc acgcgttgct gtagaagcgg gtattgctga ctactggtac	1860
aagtatgttg gcctgaacgg tgctatcgtc ggtatgacca ccttcggtga actggaagtg	1920
ctgtttcagg gtccgtctgc tccggcagag ctgctgtttg aagagttcgg cttcactggt	1980
gataacgttg ttgcgaaagc aaaagaactg ctgtaa	2016

<210> 61
 <211> 2016
 <212> ADN
 <213> Secuencia artificial

<220>
 <223>Oligonucleótido sintético

<400> 61

atgtcctcac gtaaagagct tgccaatgct attcgtgctg tgagcatgga cgcagtacag	60
aaagccaaat ccggtcaccc gggtgcccct atgggtatgg ctgacattgc cgaagtcctg	120
tggcgtgatt tcctgaaaca caaccgcag aatccgtcct gggctgaccg tgaccgcttc	180
gtgctgtcca acggccacgg ctccatgctg atctacagcc tgctgcacct caccggttac	240
gatctgccga tggaagaact gaaaaacttc cgtcagctgc actctaaaac tccgggtcac	300
ccggaagtgg gttacaccgc tgggtgtgaa accaccaccg gtccgctggg tcagggtatt	360
gccaacgcag tcggtatggc gattgcagaa aaaacgctgg cggcgcagtt taaccgtccg	420
ggccacgaca ttgtcgacca ctacacctac gccttcatgg gcgacggctg catgatggaa	480
ggcatctccc acgaagtttg ctctctggcg ggtacgctga agctgggtaa actgattgca	540
ttctacgatg acaacggtat ttctatcgat ggtcacgttg aaggctggtt caccgacgac	600
accgcaatgc gtttcgaagc ttacggctgg cacgttattc gcgacatcga cggcatgac	660
gcggcatcta tcaaacgcgc agtagaagaa gcgcgcgcag tgactgacaa accttccctg	720
ctgatgtgca aaaccatcat cggtttcggt tccccgaaca aagccggtac ccacgactcc	780
cacggtgctc cgctgggcga cgctgaaatt gccctgaccc gcgaacaact gggctggaaa	840
tatgcgccgt tcgaaatccc gtctgaaatc tatgctcagt gggatgcgaa agaagcaggc	900
caggcgaaag aatccgcatg gaacgagaaa ttcgctgctt acgcgaaagc ttatccgcag	960
gaagccgctg aatttaccgg ccgtatgaaa ggcgaaatgc cgtctgactt cgacgctaaa	1020
gcgaaagagt tcatcgctaa actgcaggct aatccggcga aaatcgccag ccgtaaagcg	1080
tctcagaatg ctatcgaagc gttcgggtccg ctgttgccgg aattcctcgg cggttctgct	1140
gacctggcgc cgtctaacct gacctgtgg tctggttcta aagcaatcaa cgaagatgct	1200

ES 2 704 697 T3

gcgggtaact acatccacta cgggtgttcgc gaggttcggta tgaccgcgat tgctaacggt 1260
atctccctgc acggtggctt cctgccgtac acctccacct tcctgatggt cgtggaatac 1320
gcacgtaacg ccgtacgtat ggctgcgctg atgaaacagc gtcaggatgat ggtttacacc 1380
cacgactcca tcggtctggg cgaagacggc ccgactcacc agccggttga gcaggtcgct 1440
tctctgcgcg taaccccgaa catgtctaca tggcgtccgt gtgaccaggt tgaatccgcg 1500
gtcgcgtgga aatacgggtg tgagcgtcag gacggcccga ccgcactgat cctctcccgt 1560
cagaacctgg cgcagcagga acgaactgaa gagcaactgg caaacatcgc gcgcggtggg 1620
tatgtgctga aagactgcgc cggtcagccg gaactgattt tcatcgctac cggttcagaa 1680
gttgaactgg ctgttgctgc ctacgaaaaa ctgactgccg aaggcgtgaa agcgcgcgtg 1740
gtgtccatgc cgtctaccga cgcatttgac aagcaggatg ctgcttaccg tgaatccgta 1800
ctgccgaaag cggttactgc acgcgttgct gtagaagcgg gtattgctga ctactggtag 1860
aagtatgttg gcctgaacgg tgctatcgtc ggtatgacca ccttcgggtga atctctggaa 1920
gtgctgtttc agggtcggc tccggcagag ctgctgtttg aagagttcgg cttcactggt 1980
gataacggtt ttgcgaaagc aaaagaactg ctgtaa 2016

5 <210> 62
<211> 2016
<212> ADN
<213> Secuencia artificial

10 <220>
<223>Oligonucleótido sintético

<400> 62

atgtcctcac gtaaagagct tgccaatgct attcgtgcgc tgagcatgga cgcagtacag 60
aaagccaaat ccggtcaccc gggtgccct atgggtatgg ctgacattgc cgaagtccg 120
tggcgtgatt tcctgaaaca caaccgcag aatccgtcct gggctgaccg tgaccgcttc 180
gtgctgtcca acggccacgg ctccatgctg atctacagcc tgctgcacct caccggttac 240
gatctgccga tggaagaact gaaaaacttc cgtcagctgc actctaaaac tccgggtcac 300
ccggaagtgg gttacaccgc tgggtgtgaa accaccaccg gtccgctggg tcagggtatt 360
gccaacgcag tcggtatggc gattgcagaa aaaacgctgg cggcgcagtt taaccgtccg 420
ggccacgaca ttgtcgacca ctacacctac gccttcatgg gcgacggctg catgatggaa 480
ggcatctccc acgaagtttg ctctctggcg ggtacgctga agctgggtaa actgattgca 540
ttctacgatg acaacgggtat ttctatcgat ggtcacgttg aaggctgggt caccgacgac 600
accgcaatgc gtttcgaagc ttacggctgg caggttattc gcgacatcga cggtcatgac 660
gcggcatcta tcaaacgcgc agtagaagaa gcgcgcgcag tgactgacaa accttccctg 720
ctgatgtgca aaaccatcat cggtttcggg tccccgaaca aagccgttac ccacgactcc 780

cacggtgcg c gctgggga c gctgaaatt gccctgaccc gcgaacaact gggctggaaa 840
 tatgcccgt tcgaaatccc gtctgaaatc tatgctcagt gggatgcgaa agaagcaggc 900
 caggcgaaag aatccgcatg gaacgagaaa ttcgctgctt acgcgaaagc ttatccgcag 960
 gaagccgctg aatttaccgc ccgtatgaaa ggcgaaatgc cgtctgactt cgacgctaaa 1020
 gcgaaagagt tcatcgctaa actgcaggct aatccggcga aaatcgccag ccgtaaagcg 1080
 tctcagaatg ctatcgaagc gtccgggtccg ctgttgccgg aattcctcgg cggttctgct 1140
 gacctggcgc cgtctaacct gaccctgtgg tctggttcta aagcaatcaa cgaagatgct 1200
 gcgggtaact acatccacta cgggtgttcgc gagttcggta tgaccgcgat tgctaacggt 1260
 atctccctgc acggtggctt cctgccgtac acctccacct tcctgatgtt cgtggaatac 1320
 gcacgtaacg ccgtacgtat ggctgcgctg atgaaacagc gtcagggtgat ggtttacacc 1380
 cacgactcca tcgggtctggg cgaagacggc ccgactcacc agccgggtga gcaggctcgt 1440
 tctctgcgcg taaccccgaa catgtctaca tggcgtccgt gtgaccagggt tgaatccgcg 1500
 gtcgcgtgga aatacgggtg tgagcgtcag gacggcccga ccgactgat cctctcccgt 1560
 cagaacctgg cgcagcagga acgaactgaa gagcaactgg caaacatcgc gcgcgggtgt 1620
 tatgtgctga aagactgcgc cggtcagccg gaactgattt tcatcgctac cggttcagaa 1680
 gttgaactgg ctgttgctgc ctacgaaaaa ctgactgccg aaggcgtgaa agcgcgcgtg 1740
 gtgtccatgc cgtctaccga cgcatttgac aagcaggatg ctgcttaccg tgaatccgta 1800
 ctgccgaaag cggttactgc acgcgttgct gtagaagcgg gtattgctga ctactggtac 1860
 aagtatgttg gcctgaacgg tgctatcgtc ggtatgacca ccttcgggtga atctgctccg 1920
 ctggaagtgc tgtttcaggg tccggcagag ctgctgtttg aagagttcgg cttcactgtt 1980
 gataacgttg ttgcgaaagc aaaagaactg ctgtaa 2016

<210> 63
 <211> 663
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Polipéptido sintético

<400> 63

Met Ser Ser Arg Lys Glu Leu Ala Asn Ala Ile Arg Ala Leu Ser Met
 1 5 10 15

Asp Ala Val Gln Lys Ala Lys Ser Gly His Pro Gly Ala Pro Met Gly
 20 25 30

Met Ala Asp Ile Ala Glu Val Leu Trp Arg Asp Phe Leu Lys His Asn

ES 2 704 697 T3

35	40	45
Pro Gln Asn Pro Ser Trp Ala Asp Arg Asp Arg Phe Val Leu Ser Asn		
50	55	60
Gly His Gly Ser Met Leu Ile Tyr Ser Leu Leu His Leu Thr Gly Tyr		
65	70	75
Asp Leu Pro Met Glu Glu Leu Lys Asn Phe Arg Gln Leu His Ser Lys		
	85	90
Thr Pro Gly His Pro Glu Val Gly Tyr Thr Ala Gly Val Glu Thr Thr		
	100	105
Thr Gly Pro Leu Gly Gln Gly Ile Ala Asn Ala Val Gly Met Ala Ile		
	115	120
Ala Glu Lys Thr Leu Ala Ala Gln Phe Asn Arg Pro Gly His Asp Ile		
	130	135
Val Asp His Tyr Thr Tyr Ala Phe Met Gly Asp Gly Cys Met Met Glu		
145	150	155
Gly Ile Ser His Glu Val Cys Ser Leu Ala Gly Thr Leu Lys Leu Gly		
	165	170
Lys Leu Ile Ala Phe Tyr Asp Asp Asn Gly Ile Ser Ile Asp Gly His		
	180	185
Val Glu Gly Trp Phe Thr Asp Asp Thr Ala Met Arg Phe Glu Ala Tyr		
	195	200
Gly Trp His Val Ile Arg Asp Ile Asp Gly His Asp Ala Ala Ser Ile		
	210	215
Lys Arg Ala Val Glu Glu Ala Arg Ala Val Thr Asp Lys Pro Ser Leu		
225	230	235
Leu Met Cys Lys Thr Ile Ile Gly Phe Gly Ser Pro Asn Lys Ala Gly		
	245	250
Thr His Asp Ser His Gly Ala Pro Leu Gly Asp Ala Glu Ile Ala Leu		
	260	265
Thr Arg Glu Gln Leu Gly Trp Lys Tyr Ala Pro Phe Glu Ile Pro Ser		
	275	280
		285

ES 2 704 697 T3

Glu Ile Tyr Ala Gln Trp Asp Ala Lys Glu Ala Gly Gln Ala Lys Glu
 290 295 300
 Ser Ala Trp Asn Glu Lys Phe Ala Ala Tyr Ala Lys Ala Tyr Pro Gln
 305 310 315 320
 Glu Ala Ala Glu Phe Thr Arg Arg Met Lys Gly Glu Met Pro Ser Asp
 325 330 335
 Phe Asp Ala Lys Ala Lys Glu Phe Ile Ala Lys Leu Gln Ala Asn Pro
 340 345 350
 Ala Lys Ile Ala Ser Arg Lys Ala Ser Gln Asn Ala Ile Glu Ala Phe
 355 360 365
 Gly Pro Leu Leu Pro Glu Phe Leu Gly Gly Ser Ala Asp Leu Ala Pro
 370 375 380
 Ser Asn Leu Thr Leu Trp Ser Gly Ser Lys Ala Ile Asn Glu Asp Ala
 385 390 395 400
 Ala Gly Asn Tyr Ile His Tyr Gly Val Arg Glu Phe Gly Met Thr Ala
 405 410 415
 Ile Ala Asn Gly Ile Ser Leu His Gly Gly Phe Leu Pro Tyr Thr Ser
 420 425 430
 Thr Phe Leu Met Phe Val Glu Tyr Ala Arg Asn Ala Val Arg Met Ala
 435 440 445
 Ala Leu Met Lys Gln Arg Gln Val Met Val Tyr Thr His Asp Ser Ile
 450 455 460
 Gly Leu Gly Glu Asp Gly Pro Thr His Gln Pro Val Glu Gln Ile Ala
 465 470 475 480
 Ser Leu Arg Val Thr Pro Asn Met Ser Thr Trp Arg Pro Cys Asp Gln
 485 490 495
 Val Glu Ser Ala Val Ala Trp Lys Tyr Gly Val Glu Arg Gln Asp Gly
 500 505 510
 Pro Thr Ala Leu Ile Leu Ser Arg Gln Asn Leu Ala Gln Gln Glu Arg
 515 520 525
 Thr Glu Glu Gln Leu Ala Asn Ile Ala Arg Gly Gly Tyr Val Leu Lys
 530 535 540

ES 2 704 697 T3

Asp Cys Ala Gly Gln Pro Glu Leu Ile Phe Ile Ala Thr Gly Ser Glu
545 550 555 560

Val Glu Leu Ala Val Ala Ala Tyr Glu Lys Leu Thr Ala Glu Gly Val
565 570 575

Lys Ala Arg Val Val Ser Met Pro Ser Thr Asp Ala Phe Asp Lys Gln
580 585 590

Asp Ala Ala Tyr Arg Glu Ser Val Leu Pro Lys Ala Val Thr Ala Arg
595 600 605

Val Ala Val Glu Ala Gly Ile Ala Asp Tyr Trp Tyr Lys Tyr Val Gly
610 615 620

Leu Asn Gly Ala Ile Val Gly Met Thr Thr Phe Gly Glu Ser Ala Pro
625 630 635 640

Ala Glu Gln Leu Phe Glu Glu Phe Gly Phe Thr Val Asp Asn Val Val
645 650 655

Ala Lys Ala Lys Ala Leu Leu
660

<210> 64
<211> 671
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido sintético

<400> 64

Met Ser Ser Arg Lys Glu Leu Ala Asn Ala Ile Arg Ala Leu Ser Met
1 5 10 15

Asp Ala Val Gln Lys Ala Lys Ser Gly His Pro Gly Ala Pro Met Gly
20 25 30

Met Ala Asp Ile Ala Glu Val Leu Trp Arg Asp Phe Leu Lys His Asn
35 40 45

Pro Gln Asn Pro Ser Trp Ala Asp Arg Asp Arg Phe Val Leu Ser Asn
50 55 60

Gly His Gly Ser Met Leu Ile Tyr Ser Leu Leu His Leu Thr Gly Tyr
65 70 75 80

ES 2 704 697 T3

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

ES 2 704 697 T3

Phe Asp Ala Lys Ala Lys Glu Phe Ile Ala Lys Leu Gln Ala Asn Pro
 340 345 350
 Ala Lys Ile Ala Ser Arg Lys Ala Ser Gln Asn Ala Ile Glu Ala Phe
 355 360 365
 Gly Pro Leu Leu Pro Glu Phe Leu Gly Gly Ser Ala Asp Leu Ala Pro
 370 375 380
 Ser Asn Leu Thr Leu Trp Ser Gly Ser Lys Ala Ile Asn Glu Asp Ala
 385 390 395 400
 Ala Gly Asn Tyr Ile His Tyr Gly Val Arg Glu Phe Gly Met Thr Ala
 405 410 415
 Ile Ala Asn Gly Ile Ser Leu His Gly Gly Phe Leu Pro Tyr Thr Ser
 420 425 430
 Thr Phe Leu Met Phe Val Glu Tyr Ala Arg Asn Ala Val Arg Met Ala
 435 440 445
 Ala Leu Met Lys Gln Arg Gln Val Met Val Tyr Thr His Asp Ser Ile
 450 455 460
 Gly Leu Gly Glu Asp Gly Pro Thr His Gln Pro Val Glu Gln Val Ala
 465 470 475 480
 Ser Leu Arg Val Thr Pro Asn Met Ser Thr Trp Arg Pro Cys Asp Gln
 485 490 495
 Val Glu Ser Ala Val Ala Trp Lys Tyr Gly Val Glu Arg Gln Asp Gly
 500 505 510
 Pro Thr Ala Leu Ile Leu Ser Arg Gln Asn Leu Ala Gln Gln Glu Arg
 515 520 525
 Thr Glu Glu Gln Leu Ala Asn Ile Ala Arg Gly Gly Tyr Val Leu Lys
 530 535 540
 Asp Cys Ala Gly Gln Pro Glu Leu Ile Phe Ile Ala Thr Gly Ser Glu
 545 550 555 560
 Val Glu Leu Ala Val Ala Ala Tyr Glu Lys Leu Thr Ala Glu Gly Val
 565 570 575
 Lys Ala Arg Val Val Ser Met Pro Ser Thr Asp Ala Phe Asp Lys Gln
 580 585 590

ES 2 704 697 T3

Asp Ala Ala Tyr Arg Glu Ser Val Leu Pro Lys Ala Val Thr Ala Arg
595 600 605

Val Ala Val Glu Ala Gly Ile Ala Asp Tyr Trp Tyr Lys Tyr Val Gly
610 615 620

Leu Asn Gly Ala Ile Val Gly Met Thr Thr Phe Leu Glu Val Leu Phe
625 630 635 640

Gln Gly Pro Gly Glu Ser Ala Pro Ala Glu Leu Leu Phe Glu Glu Phe
645 650 655

Gly Phe Thr Val Asp Asn Val Val Ala Lys Ala Lys Glu Leu Leu
660 665 670

<210> 65

<211> 671

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido sintético

<400> 65

Met Ser Ser Arg Lys Glu Leu Ala Asn Ala Ile Arg Ala Leu Ser Met
1 5 10 15

Asp Ala Val Gln Lys Ala Lys Ser Gly His Pro Gly Ala Pro Met Gly
20 25 30

Met Ala Asp Ile Ala Glu Val Leu Trp Arg Asp Phe Leu Lys His Asn
35 40 45

Pro Gln Asn Pro Ser Trp Ala Asp Arg Asp Arg Phe Val Leu Ser Asn
50 55 60

Gly His Gly Ser Met Leu Ile Tyr Ser Leu Leu His Leu Thr Gly Tyr
65 70 75 80

Asp Leu Pro Met Glu Glu Leu Lys Asn Phe Arg Gln Leu His Ser Lys
85 90 95

Thr Pro Gly His Pro Glu Val Gly Tyr Thr Ala Gly Val Glu Thr Thr
100 105 110

Thr Gly Pro Leu Gly Gln Gly Ile Ala Asn Ala Val Gly Met Ala Ile
115 120 125

ES 2 704 697 T3

Ala	Glu	Lys	Thr	Leu	Ala	Ala	Gln	Phe	Asn	Arg	Pro	Gly	His	Asp	Ile	130	135	140
Val	Asp	His	Tyr	Thr	Tyr	Ala	Phe	Met	Gly	Asp	Gly	Cys	Met	Met	Glu	145	150	155
Gly	Ile	Ser	His	Glu	Val	Cys	Ser	Leu	Ala	Gly	Thr	Leu	Lys	Leu	Gly	165	170	175
Lys	Leu	Ile	Ala	Phe	Tyr	Asp	Asp	Asn	Gly	Ile	Ser	Ile	Asp	Gly	His	180	185	190
Val	Glu	Gly	Trp	Phe	Thr	Asp	Asp	Thr	Ala	Met	Arg	Phe	Glu	Ala	Tyr	195	200	205
Gly	Trp	His	Val	Ile	Arg	Asp	Ile	Asp	Gly	His	Asp	Ala	Ala	Ser	Ile	210	215	220
Lys	Arg	Ala	Val	Glu	Glu	Ala	Arg	Ala	Val	Thr	Asp	Lys	Pro	Ser	Leu	225	230	235
Leu	Met	Cys	Lys	Thr	Ile	Ile	Gly	Phe	Gly	Ser	Pro	Asn	Lys	Ala	Gly	245	250	255
Thr	His	Asp	Ser	His	Gly	Ala	Pro	Leu	Gly	Asp	Ala	Glu	Ile	Ala	Leu	260	265	270
Thr	Arg	Glu	Gln	Leu	Gly	Trp	Lys	Tyr	Ala	Pro	Phe	Glu	Ile	Pro	Ser	275	280	285
Glu	Ile	Tyr	Ala	Gln	Trp	Asp	Ala	Lys	Glu	Ala	Gly	Gln	Ala	Lys	Glu	290	295	300
Ser	Ala	Trp	Asn	Glu	Lys	Phe	Ala	Ala	Tyr	Ala	Lys	Ala	Tyr	Pro	Gln	305	310	315
Glu	Ala	Ala	Glu	Phe	Thr	Arg	Arg	Met	Lys	Gly	Glu	Met	Pro	Ser	Asp	325	330	335
Phe	Asp	Ala	Lys	Ala	Lys	Glu	Phe	Ile	Ala	Lys	Leu	Gln	Ala	Asn	Pro	340	345	350
Ala	Lys	Ile	Ala	Ser	Arg	Lys	Ala	Ser	Gln	Asn	Ala	Ile	Glu	Ala	Phe	355	360	365
Gly	Pro	Leu	Leu	Pro	Glu	Phe	Leu	Gly	Gly	Ser	Ala	Asp	Leu	Ala	Pro	370	375	380

ES 2 704 697 T3

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

ES 2 704 697 T3

625 630 635 640

Phe Gln Gly Pro Glu Ser Ala Pro Ala Glu Leu Leu Phe Glu Glu Phe
645 650 655

Gly Phe Thr Val Asp Asn Val Val Ala Lys Ala Lys Glu Leu Leu
660 665 670

<210> 66

<211> 671

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido sintético

<400> 66

Met Ser Ser Arg Lys Glu Leu Ala Asn Ala Ile Arg Ala Leu Ser Met
1 5 10 15

Asp Ala Val Gln Lys Ala Lys Ser Gly His Pro Gly Ala Pro Met Gly
20 25 30

Met Ala Asp Ile Ala Glu Val Leu Trp Arg Asp Phe Leu Lys His Asn
35 40 45

Pro Gln Asn Pro Ser Trp Ala Asp Arg Asp Arg Phe Val Leu Ser Asn
50 55 60

Gly His Gly Ser Met Leu Ile Tyr Ser Leu Leu His Leu Thr Gly Tyr
65 70 75 80

Asp Leu Pro Met Glu Glu Leu Lys Asn Phe Arg Gln Leu His Ser Lys
85 90 95

Thr Pro Gly His Pro Glu Val Gly Tyr Thr Ala Gly Val Glu Thr Thr
100 105 110

Thr Gly Pro Leu Gly Gln Gly Ile Ala Asn Ala Val Gly Met Ala Ile
115 120 125

Ala Glu Lys Thr Leu Ala Ala Gln Phe Asn Arg Pro Gly His Asp Ile
130 135 140

Val Asp His Tyr Thr Tyr Ala Phe Met Gly Asp Gly Cys Met Met Glu
145 150 155 160

Gly Ile Ser His Glu Val Cys Ser Leu Ala Gly Thr Leu Lys Leu Gly
165 170 175

ES 2 704 697 T3

Lys Leu Ile Ala Phe Tyr Asp Asp Asn Gly Ile Ser Ile Asp Gly His
 180 185 190
 Val Glu Gly Trp Phe Thr Asp Asp Thr Ala Met Arg Phe Glu Ala Tyr
 195 200 205
 Gly Trp His Val Ile Arg Asp Ile Asp Gly His Asp Ala Ala Ser Ile
 210 215 220
 Lys Arg Ala Val Glu Glu Ala Arg Ala Val Thr Asp Lys Pro Ser Leu
 225 230 235 240
 Leu Met Cys Lys Thr Ile Ile Gly Phe Gly Ser Pro Asn Lys Ala Gly
 245 250 255
 Thr His Asp Ser His Gly Ala Pro Leu Gly Asp Ala Glu Ile Ala Leu
 260 265 270
 Thr Arg Glu Gln Leu Gly Trp Lys Tyr Ala Pro Phe Glu Ile Pro Ser
 275 280 285
 Glu Ile Tyr Ala Gln Trp Asp Ala Lys Glu Ala Gly Gln Ala Lys Glu
 290 295 300
 Ser Ala Trp Asn Glu Lys Phe Ala Ala Tyr Ala Lys Ala Tyr Pro Gln
 305 310 315 320
 Glu Ala Ala Glu Phe Thr Arg Arg Met Lys Gly Glu Met Pro Ser Asp
 325 330 335
 Phe Asp Ala Lys Ala Lys Glu Phe Ile Ala Lys Leu Gln Ala Asn Pro
 340 345 350
 Ala Lys Ile Ala Ser Arg Lys Ala Ser Gln Asn Ala Ile Glu Ala Phe
 355 360 365
 Gly Pro Leu Leu Pro Glu Phe Leu Gly Gly Ser Ala Asp Leu Ala Pro
 370 375 380
 Ser Asn Leu Thr Leu Trp Ser Gly Ser Lys Ala Ile Asn Glu Asp Ala
 385 390 395 400
 Ala Gly Asn Tyr Ile His Tyr Gly Val Arg Glu Phe Gly Met Thr Ala
 405 410 415
 Ile Ala Asn Gly Ile Ser Leu His Gly Gly Phe Leu Pro Tyr Thr Ser

ES 2 704 697 T3

420								425								430			
Thr	Phe	Leu	Met	Phe	Val	Glu	Tyr	Ala	Arg	Asn	Ala	Val	Arg	Met	Ala				
435						440						445							
Ala	Leu	Met	Lys	Gln	Arg	Gln	Val	Met	Val	Tyr	Thr	His	Asp	Ser	Ile				
450						455						460							
Gly	Leu	Gly	Glu	Asp	Gly	Pro	Thr	His	Gln	Pro	Val	Glu	Gln	Val	Ala				
465						470						475			480				
Ser	Leu	Arg	Val	Thr	Pro	Asn	Met	Ser	Thr	Trp	Arg	Pro	Cys	Asp	Gln				
			485						490						495				
Val	Glu	Ser	Ala	Val	Ala	Trp	Lys	Tyr	Gly	Val	Glu	Arg	Gln	Asp	Gly				
			500						505						510				
Pro	Thr	Ala	Leu	Ile	Leu	Ser	Arg	Gln	Asn	Leu	Ala	Gln	Gln	Glu	Arg				
			515						520						525				
Thr	Glu	Glu	Gln	Leu	Ala	Asn	Ile	Ala	Arg	Gly	Gly	Tyr	Val	Leu	Lys				
530						535						540							
Asp	Cys	Ala	Gly	Gln	Pro	Glu	Leu	Ile	Phe	Ile	Ala	Thr	Gly	Ser	Glu				
545						550						555			560				
Val	Glu	Leu	Ala	Val	Ala	Ala	Tyr	Glu	Lys	Leu	Thr	Ala	Glu	Gly	Val				
			565						570						575				
Lys	Ala	Arg	Val	Val	Ser	Met	Pro	Ser	Thr	Asp	Ala	Phe	Asp	Lys	Gln				
			580						585						590				
Asp	Ala	Ala	Tyr	Arg	Glu	Ser	Val	Leu	Pro	Lys	Ala	Val	Thr	Ala	Arg				
			595						600						605				
Val	Ala	Val	Glu	Ala	Gly	Ile	Ala	Asp	Tyr	Trp	Tyr	Lys	Tyr	Val	Gly				
610						615						620							
Leu	Asn	Gly	Ala	Ile	Val	Gly	Met	Thr	Thr	Phe	Gly	Glu	Leu	Glu	Val				
625						630						635			640				
Leu	Phe	Gln	Gly	Pro	Ser	Ala	Pro	Ala	Glu	Leu	Leu	Phe	Glu	Glu	Phe				
			645						650						655				
Gly	Phe	Thr	Val	Asp	Asn	Val	Val	Ala	Lys	Ala	Lys	Glu	Leu	Leu					
			660						665						670				

<210> 67
 <211> 671
 <212> PRT
 <213> Secuencia artificial

5

<220>
 <223> Polipéptido sintético

10

<400> 67

```

Met Ser Ser Arg Lys Glu Leu Ala Asn Ala Ile Arg Ala Leu Ser Met
 1          5          10          15

Asp Ala Val Gln Lys Ala Lys Ser Gly His Pro Gly Ala Pro Met Gly
      20          25          30

Met Ala Asp Ile Ala Glu Val Leu Trp Arg Asp Phe Leu Lys His Asn
      35          40          45

Pro Gln Asn Pro Ser Trp Ala Asp Arg Asp Arg Phe Val Leu Ser Asn
      50          55          60

Gly His Gly Ser Met Leu Ile Tyr Ser Leu Leu His Leu Thr Gly Tyr
65          70          75          80

Asp Leu Pro Met Glu Glu Leu Lys Asn Phe Arg Gln Leu His Ser Lys
      85          90          95

Thr Pro Gly His Pro Glu Val Gly Tyr Thr Ala Gly Val Glu Thr Thr
      100          105          110

Thr Gly Pro Leu Gly Gln Gly Ile Ala Asn Ala Val Gly Met Ala Ile
      115          120          125

Ala Glu Lys Thr Leu Ala Ala Gln Phe Asn Arg Pro Gly His Asp Ile
      130          135          140

Val Asp His Tyr Thr Tyr Ala Phe Met Gly Asp Gly Cys Met Met Glu
145          150          155          160

Gly Ile Ser His Glu Val Cys Ser Leu Ala Gly Thr Leu Lys Leu Gly
      165          170          175

Lys Leu Ile Ala Phe Tyr Asp Asp Asn Gly Ile Ser Ile Asp Gly His
      180          185          190

Val Glu Gly Trp Phe Thr Asp Asp Thr Ala Met Arg Phe Glu Ala Tyr
      195          200          205

Gly Trp His Val Ile Arg Asp Ile Asp Gly His Asp Ala Ala Ser Ile

```

ES 2 704 697 T3

210						215						220				
Lys	Arg	Ala	Val	Glu	Glu	Ala	Arg	Ala	Val	Thr	Asp	Lys	Pro	Ser	Leu	
225					230					235					240	
Leu	Met	Cys	Lys	Thr	Ile	Ile	Gly	Phe	Gly	Ser	Pro	Asn	Lys	Ala	Gly	
				245					250					255		
Thr	His	Asp	Ser	His	Gly	Ala	Pro	Leu	Gly	Asp	Ala	Glu	Ile	Ala	Leu	
			260					265					270			
Thr	Arg	Glu	Gln	Leu	Gly	Trp	Lys	Tyr	Ala	Pro	Phe	Glu	Ile	Pro	Ser	
		275					280					285				
Glu	Ile	Tyr	Ala	Gln	Trp	Asp	Ala	Lys	Glu	Ala	Gly	Gln	Ala	Lys	Glu	
290						295					300					
Ser	Ala	Trp	Asn	Glu	Lys	Phe	Ala	Ala	Tyr	Ala	Lys	Ala	Tyr	Pro	Gln	
305					310					315					320	
Glu	Ala	Ala	Glu	Phe	Thr	Arg	Arg	Met	Lys	Gly	Glu	Met	Pro	Ser	Asp	
				325					330					335		
Phe	Asp	Ala	Lys	Ala	Lys	Glu	Phe	Ile	Ala	Lys	Leu	Gln	Ala	Asn	Pro	
			340					345					350			
Ala	Lys	Ile	Ala	Ser	Arg	Lys	Ala	Ser	Gln	Asn	Ala	Ile	Glu	Ala	Phe	
		355					360					365				
Gly	Pro	Leu	Leu	Pro	Glu	Phe	Leu	Gly	Gly	Ser	Ala	Asp	Leu	Ala	Pro	
	370					375					380					
Ser	Asn	Leu	Thr	Leu	Trp	Ser	Gly	Ser	Lys	Ala	Ile	Asn	Glu	Asp	Ala	
385					390					395					400	
Ala	Gly	Asn	Tyr	Ile	His	Tyr	Gly	Val	Arg	Glu	Phe	Gly	Met	Thr	Ala	
				405					410					415		
Ile	Ala	Asn	Gly	Ile	Ser	Leu	His	Gly	Gly	Phe	Leu	Pro	Tyr	Thr	Ser	
			420					425					430			
Thr	Phe	Leu	Met	Phe	Val	Glu	Tyr	Ala	Arg	Asn	Ala	Val	Arg	Met	Ala	
		435					440					445				
Ala	Leu	Met	Lys	Gln	Arg	Gln	Val	Met	Val	Tyr	Thr	His	Asp	Ser	Ile	
450						455					460					

Gly Leu Gly Glu Asp Gly Pro Thr His Gln Pro Val Glu Gln Val Ala
465 470 475 480

Ser Leu Arg Val Thr Pro Asn Met Ser Thr Trp Arg Pro Cys Asp Gln
485 490 495

Val Glu Ser Ala Val Ala Trp Lys Tyr Gly Val Glu Arg Gln Asp Gly
500 505 510

Pro Thr Ala Leu Ile Leu Ser Arg Gln Asn Leu Ala Gln Gln Glu Arg
515 520 525

Thr Glu Glu Gln Leu Ala Asn Ile Ala Arg Gly Gly Tyr Val Leu Lys
530 535 540

Asp Cys Ala Gly Gln Pro Glu Leu Ile Phe Ile Ala Thr Gly Ser Glu
545 550 555 560

Val Glu Leu Ala Val Ala Ala Tyr Glu Lys Leu Thr Ala Glu Gly Val
565 570 575

Lys Ala Arg Val Val Ser Met Pro Ser Thr Asp Ala Phe Asp Lys Gln
580 585 590

Asp Ala Ala Tyr Arg Glu Ser Val Leu Pro Lys Ala Val Thr Ala Arg
595 600 605

Val Ala Val Glu Ala Gly Ile Ala Asp Tyr Trp Tyr Lys Tyr Val Gly
610 615 620

Leu Asn Gly Ala Ile Val Gly Met Thr Thr Phe Gly Glu Ser Leu Glu
625 630 635 640

Val Leu Phe Gln Gly Pro Ala Pro Ala Glu Leu Leu Phe Glu Glu Phe
645 650 655

Gly Phe Thr Val Asp Asn Val Val Ala Lys Ala Lys Glu Leu Leu
660 665 670

<210> 68

<211> 671

<212> PRT

<213> Secuencia artificial

<220>

<223> Polipéptido sintético

<400> 68

Met Ser Ser Arg Lys Glu Leu Ala Asn Ala Ile Arg Ala Leu Ser Met

ES 2 704 697 T3

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

ES 2 704 697 T3

Thr	His	Asp	Ser	His	Gly	Ala	Pro	Leu	Gly	Asp	Ala	Glu	Ile	Ala	Leu	260	265	270	
Thr	Arg	Glu	Gln	Leu	Gly	Trp	Lys	Tyr	Ala	Pro	Phe	Glu	Ile	Pro	Ser	275	280	285	
Glu	Ile	Tyr	Ala	Gln	Trp	Asp	Ala	Lys	Glu	Ala	Gly	Gln	Ala	Lys	Glu	290	295	300	
Ser	Ala	Trp	Asn	Glu	Lys	Phe	Ala	Ala	Tyr	Ala	Lys	Ala	Tyr	Pro	Gln	305	310	315	320
Glu	Ala	Ala	Glu	Phe	Thr	Arg	Arg	Met	Lys	Gly	Glu	Met	Pro	Ser	Asp	325	330	335	
Phe	Asp	Ala	Lys	Ala	Lys	Glu	Phe	Ile	Ala	Lys	Leu	Gln	Ala	Asn	Pro	340	345	350	
Ala	Lys	Ile	Ala	Ser	Arg	Lys	Ala	Ser	Gln	Asn	Ala	Ile	Glu	Ala	Phe	355	360	365	
Gly	Pro	Leu	Leu	Pro	Glu	Phe	Leu	Gly	Gly	Ser	Ala	Asp	Leu	Ala	Pro	370	375	380	
Ser	Asn	Leu	Thr	Leu	Trp	Ser	Gly	Ser	Lys	Ala	Ile	Asn	Glu	Asp	Ala	385	390	395	400
Ala	Gly	Asn	Tyr	Ile	His	Tyr	Gly	Val	Arg	Glu	Phe	Gly	Met	Thr	Ala	405	410	415	
Ile	Ala	Asn	Gly	Ile	Ser	Leu	His	Gly	Gly	Phe	Leu	Pro	Tyr	Thr	Ser	420	425	430	
Thr	Phe	Leu	Met	Phe	Val	Glu	Tyr	Ala	Arg	Asn	Ala	Val	Arg	Met	Ala	435	440	445	
Ala	Leu	Met	Lys	Gln	Arg	Gln	Val	Met	Val	Tyr	Thr	His	Asp	Ser	Ile	450	455	460	
Gly	Leu	Gly	Glu	Asp	Gly	Pro	Thr	His	Gln	Pro	Val	Glu	Gln	Val	Ala	465	470	475	480
Ser	Leu	Arg	Val	Thr	Pro	Asn	Met	Ser	Thr	Trp	Arg	Pro	Cys	Asp	Gln	485	490	495	
Val	Glu	Ser	Ala	Val	Ala	Trp	Lys	Tyr	Gly	Val	Glu	Arg	Gln	Asp	Gly	500	505	510	

ES 2 704 697 T3

Pro Thr Ala Leu Ile Leu Ser Arg Gln Asn Leu Ala Gln Gln Glu Arg
515 520 525

Thr Glu Glu Gln Leu Ala Asn Ile Ala Arg Gly Gly Tyr Val Leu Lys
530 535 540

Asp Cys Ala Gly Gln Pro Glu Leu Ile Phe Ile Ala Thr Gly Ser Glu
545 550 555 560

Val Glu Leu Ala Val Ala Ala Tyr Glu Lys Leu Thr Ala Glu Gly Val
565 570 575

Lys Ala Arg Val Val Ser Met Pro Ser Thr Asp Ala Phe Asp Lys Gln
580 585 590

Asp Ala Ala Tyr Arg Glu Ser Val Leu Pro Lys Ala Val Thr Ala Arg
595 600 605

Val Ala Val Glu Ala Gly Ile Ala Asp Tyr Trp Tyr Lys Tyr Val Gly
610 615 620

Leu Asn Gly Ala Ile Val Gly Met Thr Thr Phe Gly Glu Ser Ala Pro
625 630 635 640

Leu Glu Val Leu Phe Gln Gly Pro Ala Glu Leu Leu Phe Glu Glu Phe
645 650 655

Gly Phe Thr Val Asp Asn Val Val Ala Lys Ala Lys Glu Leu Leu
660 665 670

<210> 69
<211> 24
<212> ADN
<213> Secuencia artificial

<220>
<223> Oligonucleótido sintético

<400> 69
ctggaagtgc tgttcaggg tccg 24

<210> 70
<211> 11
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido sintético

<400> 70

Met Lys Asn Ile Asn Pro Thr Gln Thr Ala Ala
1 5 10

5 <210> 71
 <211> 33
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Oligonucleótido sintético
 <400> 71
 10 atgaaaaaca tcaatccaac gcagaccgct gcc 33
 <210> 72
 <211> 19
 <212> PRT
 15 <213> Secuencia artificial
 <220>
 <223> Polipéptido sintético
 20 <400> 72
 Met Leu Glu Val Leu Phe Gln Gly Pro Lys Asn Ile Asn Pro Thr Gln
 1 5 10 15
 Thr Ala Ala
 25 <210> 73
 <211> 57
 <212> ADN
 <213> Secuencia artificial
 <220>
 30 <223> Oligonucleótido sintético
 <400> 73
 atgctggaag tgctgttca gggtcgaaa aacatcaatc caacgcagac cgctgcc 57
 35 <210> 74
 <211> 19
 <212> PRT
 <213> Secuencia artificial
 40 <220>
 <223> Polipéptido sintético
 <400> 74
 Met Lys Leu Glu Val Leu Phe Gln Gly Pro Asn Ile Asn Pro Thr Gln
 1 5 10 15
 Thr Ala Ala
 45 <210> 75
 <211> 57
 <212> ADN
 50 <213> Secuencia artificial
 <220>
 <223> Oligonucleótido sintético
 55 <400> 75
 atgaaactgg aagtgtgtt tcagggtccg aacatcaatc caacgcagac cgctgcc 57

5 <210> 76
 <211> 19
 <212> PRT
 <213> Secuencia artificial

 <220>
 <223> Polipéptido sintético

 10 <400> 76

 Met Lys Asn Leu Glu Val Leu Phe Gln Gly Pro Ile Asn Pro Thr Gln
 1 5 10 15

 Thr Ala Ala

 15 <210> 77
 <211> 57
 <212> ADN
 <213> Secuencia artificial

 20 <220>
 <223> Oligonucleótido sintético

 <400> 77
 atgaaaaacc tggaagtgc tttcagggt ccgatcaatc caacgcagac cgctgcc 57

 25 <210> 78
 <211> 11
 <212> PRT
 <213> Secuencia artificial

 30 <220>
 <223> Polipéptido sintético

 <400> 78

 Met Leu Glu Val Leu Phe Gln Gly Pro Ala Ala
 1 5 10
 35

 40 <210> 79
 <211> 33
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Oligonucleótido sintético

 45 <400> 79
 atgctggaag tgctgttca gggccggct gcc 33

 50 <210> 80
 <211> 11
 <212> PRT
 <213> Secuencia artificial

 <220>
 <223> Polipéptido sintético

 55 <400> 80

 Met Lys Leu Glu Val Leu Phe Gln Gly Pro Ala
 1 5 10

5 <210> 81
 <211> 33
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Oligonucleótido sintético
 10 <400> 81
 atgaaactgg aagtgtgtt tcagggtccg gcc 33
 <210> 82
 <211> 11
 <212> PRT
 15 <213> Secuencia artificial
 <220>
 <223> Polipéptido sintético
 20 <400> 82
 Met Lys Asn Leu Glu Val Leu Phe Gln Gly Pro
 1 5 10
 25 <210> 83
 <211> 33
 <212> ADN
 <213> Secuencia artificial
 <220>
 30 <223> Oligonucleótido sintético
 <400> 83
 atgaaaaacc tgaagtgtt gttcagggt ccg 33
 35 <210> 84
 <211> 18
 <212> PRT
 <213> Secuencia artificial
 40 <220>
 <223> Polipéptido sintético
 <400> 84
 Met Lys Asn Ile Asn Leu Glu Val Leu Phe Gln Gly Pro Thr Gln Thr
 1 5 10 15
 45 Ala Ala
 <210> 85
 <211> 54
 <212> ADN
 50 <213> Secuencia artificial
 <220>
 <223> Oligonucleótido sintético
 55 <400> 85
 atgaaaaaca tcaatctgga agtgtgttt cagggtccaa cgcagaccgc tgcc 54

5 <210> 86
 <211> 17
 <212> PRT
 <213> Secuencia artificial

 <220>
 <223> Polipéptido sintético

 10 <400> 86

 Thr Ala Ala Trp Gln Ala Leu Glu Val Leu Phe Gln Gly Pro Gln Lys
 1 5 10 15

 His

 15 <210> 87
 <211> 51
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Oligonucleótido sintético

 20 <400> 87
 accgctgcct ggcaggcact agaagtgctg ttcagggtc cgcagaaaca c 51

REIVINDICACIONES

1. Una proteína de fosfoglucoisomerasa recombinante que comprende la secuencia de la SEQ ID NO: 17 con una secuencia de reconocimiento de proteasa localizada después del aminoácido 410, 524, 525, 526, 527, 528, 529, 530, 531, 532 o 545 de la secuencia de la SEQ ID NO: 17.
2. La proteína recombinante de la reivindicación 1, en donde la secuencia de reconocimiento de proteasa es reconocida por una proteasa seleccionada del grupo que consiste en alanina carboxipeptidasa, astacina de *Armillaria mellea*, leucil aminopeptidasa bacteriana, procoagulante de cáncer, cathepsina B, clostripaína, alanil aminopeptidasa de citosol, elastasa, endoproteinasa Arg-C, enteroquinasa, gastricsina, gelatinasa, Gly-X carboxipeptidasa, glicil endopeptidasa, proteasa 3C de rinovirus humano, hipodermina C, serina endopeptidasa específica de IgA, leucil aminopeptidasa, leucil endopeptidasa, lysC, pro-X carboxipeptidasa lisosómica, lisil aminopeptidasa, metionil aminopeptidasa, mixobacteria, nardilisina, endopeptidasa E pancreática, picornaína 2A, picornaína 3C, proendopeptidasa, prolil aminopeptidasa, proproteína convertasa I, proproteína convertasa II, ruselisina, sacaropepsina, semenogelasa, activador de T-plasminógeno, trombina, calicreína tisular, virus del grabado del tabaco (TEV), togavirina, triptofanil aminopeptidasa, activador de U-plasminógeno, V8, venombina A, venombina AB y Xaa-pro aminopeptidasa.
3. La proteína recombinante de la reivindicación 2, en donde la secuencia de reconocimiento de proteasa es reconocida por proteasa 3C de rinovirus humano.
4. La proteína recombinante de la reivindicación 3, en donde la secuencia de reconocimiento de proteasa 3C de rinovirus humano de la proteína recombinante comprende la secuencia de la SEQ ID NO: 38.
5. La proteína recombinante de una cualquiera de las reivindicaciones 1-4, en donde la secuencia de reconocimiento de proteasa se localiza después del aminoácido 526 de la secuencia de la SEQ ID NO: 17.
6. La proteína recombinante de la reivindicación 5, en donde la proteína recombinante comprende la secuencia de la SEQ ID NO: 25.
7. Un ácido nucleico que codifica una proteína de fosfoglucoisomerasa recombinante y que comprende la secuencia de la SEQ ID NO: 1 con una secuencia de nucleótidos que codifica una secuencia de reconocimiento de proteasa localizada después de los codones 410, 524, 525, 526, 527, 528, 529, 530, 531, 532 o 545 de la secuencia de la SEQ ID NO: 1.
8. El ácido nucleico de la reivindicación 7, en donde la secuencia de reconocimiento de proteasa es la secuencia de reconocimiento de proteasa reconocida por una proteasa seleccionada del grupo que consiste en alanina carboxipeptidasa, astacina de *Armillaria mellea*, leucil aminopeptidasa bacteriana, procoagulante de cáncer, cathepsina B, clostripaína, alanil aminopeptidasa de citosol, elastasa, endoproteinasa Arg-C, enteroquinasa, gastricsina, gelatinasa, Gly-X carboxipeptidasa, glicil endopeptidasa, proteasa 3C de rinovirus humano, hipodermina C, serina endopeptidasa específica de IgA, leucil aminopeptidasa, leucil endopeptidasa, lysC, pro-X carboxipeptidasa lisosómica, lisil aminopeptidasa, metionil aminopeptidasa, mixobacteria, nardilisina, endopeptidasa E pancreática, picornaína 2A, picornaína 3C, proendopeptidasa, prolil aminopeptidasa, proproteína convertasa I, proproteína convertasa II, ruselisina, sacaropepsina, semenogelasa, activador de T-plasminógeno, trombina, calicreína tisular, virus del grabado del tabaco (TEV), togavirina, triptofanil aminopeptidasa, activador de U-plasminógeno, V8, venombina A, venombina AB y Xaa-pro aminopeptidasa.
9. El ácido nucleico de la reivindicación 8, en donde la secuencia de reconocimiento de proteasa es reconocida por proteasa 3C de rinovirus humano.
10. El ácido nucleico de la reivindicación 9, en donde la secuencia de reconocimiento de proteasa 3C de rinovirus humano del ácido nucleico comprende la secuencia de la SEQ ID NO: 37.
11. El ácido nucleico de la reivindicación 7-10, en donde la secuencia de reconocimiento de proteasa se localiza después del codón 526 de la secuencia de la SEQ ID NO: 1.
12. El ácido nucleico de la reivindicación 11, en donde el ácido nucleico comprende la secuencia de la SEQ ID NO: 9.
13. Un vector que comprende el ácido nucleico de una cualquiera de las reivindicaciones 7-12, opcionalmente en donde el vector es un vector de clonación o un vector de expresión, opcionalmente, en donde el vector es un plásmido, un fósido, un fagémido, un genoma vírico o un cromosoma artificial, opcionalmente en donde el vector es un plásmido.

14. Una célula que comprende el vector de la reivindicación 13, opcionalmente en donde la célula es una célula bacteriana, una célula fúngica, una célula de mamífero o una célula vegetal, opcionalmente en donde la célula es una célula bacteriana, opcionalmente en donde la célula bacteriana es una célula de *Escherichia coli*.

- 5 15. Una célula que expresa la proteína recombinante de una cualquiera de las reivindicaciones 1-6.

FIGURA 1A

Secuencia de reconocimiento de proteasa de RVH									
Aminoácidos:	L	E	V	L	F	Q	G	P	(SEQ ID NO:38)
Nucleótidos:	CTG	GAA	GTC	CTG	TTC	CAG	GGT	CCG	(SEQ ID NO:69)
Pgi	Secuencia de aminoácidos: M K N I N F I Q T A A ... (SEQ ID NO:70)								
pgi	Secuencia de nucleótidos: AAA AAA AAC ATC AAT CCA ACG CAG ACC GCT GCC (SEQ ID NO:71)								
SEQ ID Nos.:									
Pgi-HRV-1001	Secuencia de aminoácidos: M L E V L F Q G P K N I N F I Q T A A 72								
pgi-hrv-1001	Secuencia de nucleótidos: ATG CTG GAA GTC CTG TTC CAG GGT CCG AAA AAC ATC AAT CCA ACG CAG ACC GCT GCC ... 73								
Pgi-HRV-1002	Secuencia de aminoácidos: M K L E V L F Q G P N I N F I Q T A A 74								
pgi-hrv-1002	Secuencia de nucleótidos: ATG AAA CTG GAA GTC CTG TTC CAG GGT CCG AAC ATC AAT CCA ACG CAG ACC GCT GCC ... 75								
Pgi-HRV-1003	Secuencia de aminoácidos: M K N L E V L F Q G P I N F I Q T A A 76								
pgi-hrv-1003	Secuencia de nucleótidos: ATG AAA AAC CTG GAA GTC CTG TTC CAG GGT CCG ATC AAT CCA ACG CAG ACC GCT GCC ... 77								

FIGURA 1B

Secuencia de reconocimiento de proteasa de RVH									
Aminoácidos:	L	E	V	L	F	Q	G	P	(SEQ ID NO:38)
Nucleótidos:	CTG	GAA	GTC	CTG	TTC	CAG	GGT	CCG	(SEQ ID NO:69)
Pgi	Secuencia de aminoácidos: M K N I N F I Q T A A ... (SEQ ID NO:70)								
pgi	Secuencia de nucleótidos: ATG AAA AAC ATC AAT CCA ACG CAG ACC GCT GCC (SEQ ID NO:71)								
SEQ ID Nos.:									
Pgi-HRV-R001	Secuencia de aminoácidos: M L E V L F Q G P A A ... 78								
pgi-hrv-R001	Secuencia de nucleótidos: ATG CTG GAA GTC CTG TTC CAG GGT CCG GCT GCC 79								
Pgi-HRV-R002	Secuencia de aminoácidos: M K L E V L F Q G P A ... 80								
pgi-hrv-R002	Secuencia de nucleótidos: ATG AAA CTG GAA GTC CTG TTC CAG GGT CCG GCC 81								
Pgi-HRV-R003	Secuencia de aminoácidos: M K N L E V L F Q G P ... 82								
pgi-hrv-R003	Secuencia de nucleótidos: ATG AAA AAC CTG GAA GTC CTG TTC CAG GGT CCG 83								

FIGURA 2

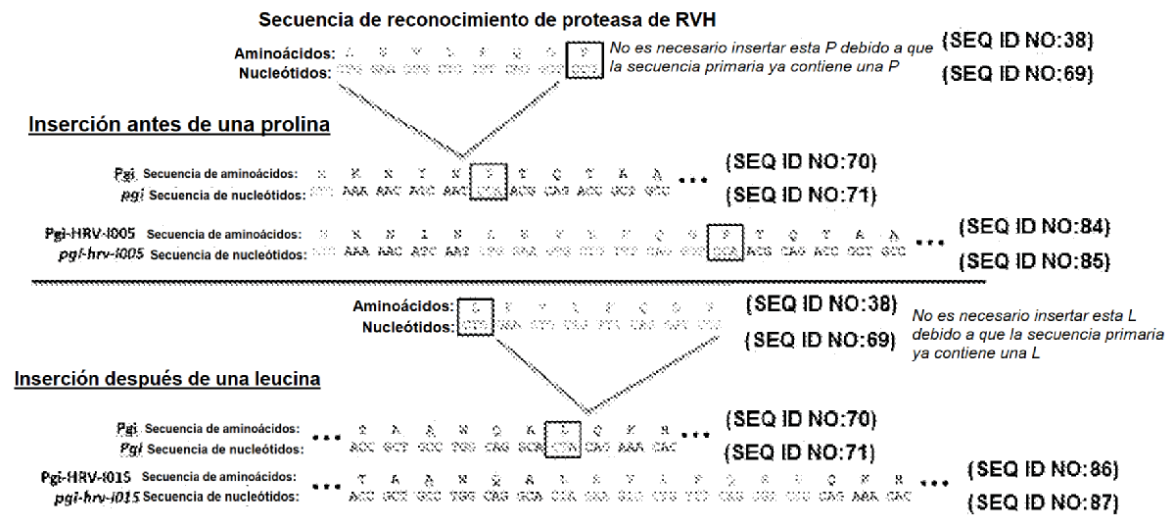


FIGURA 3

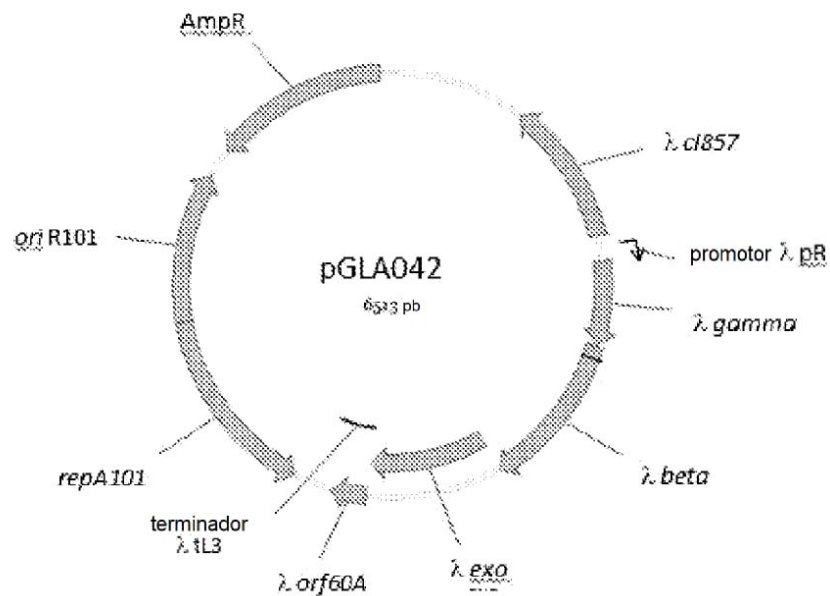


FIGURA 4

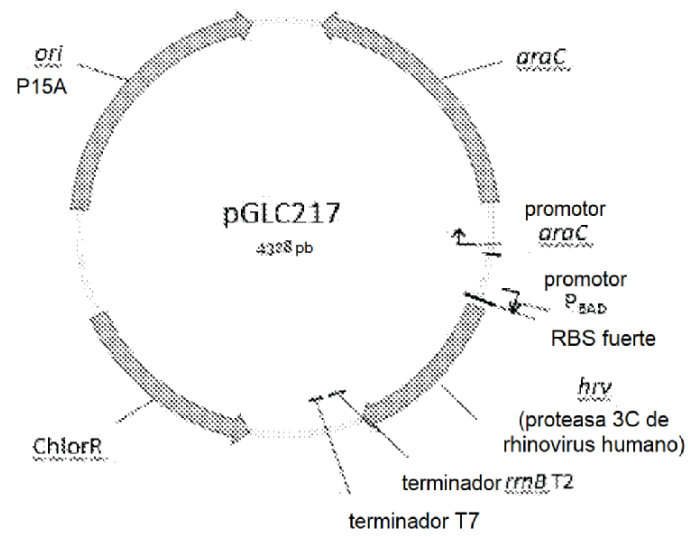


FIGURA 5

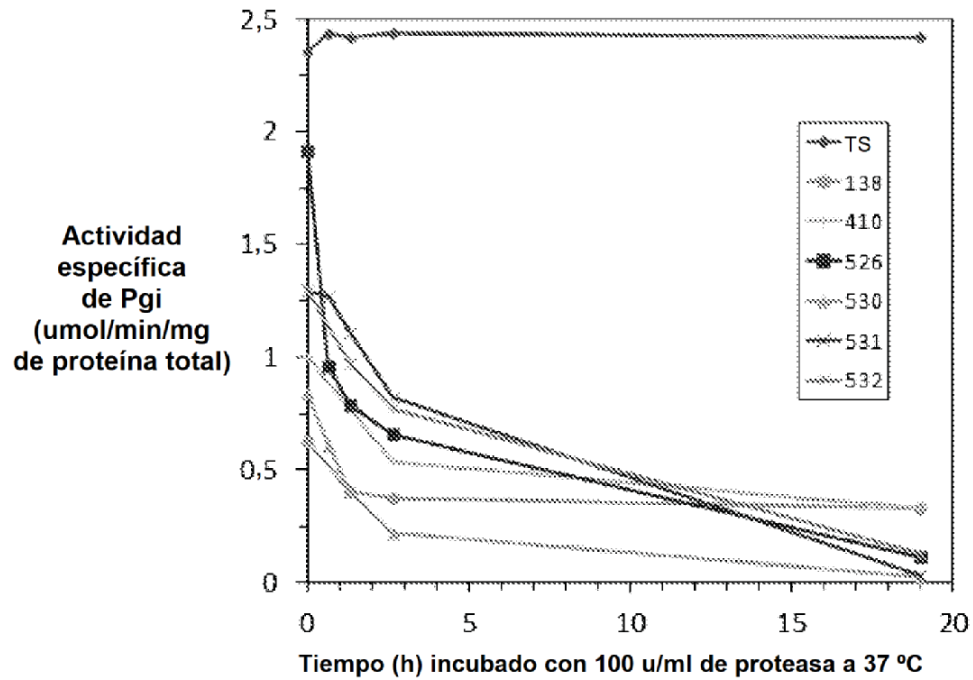


FIGURA 6

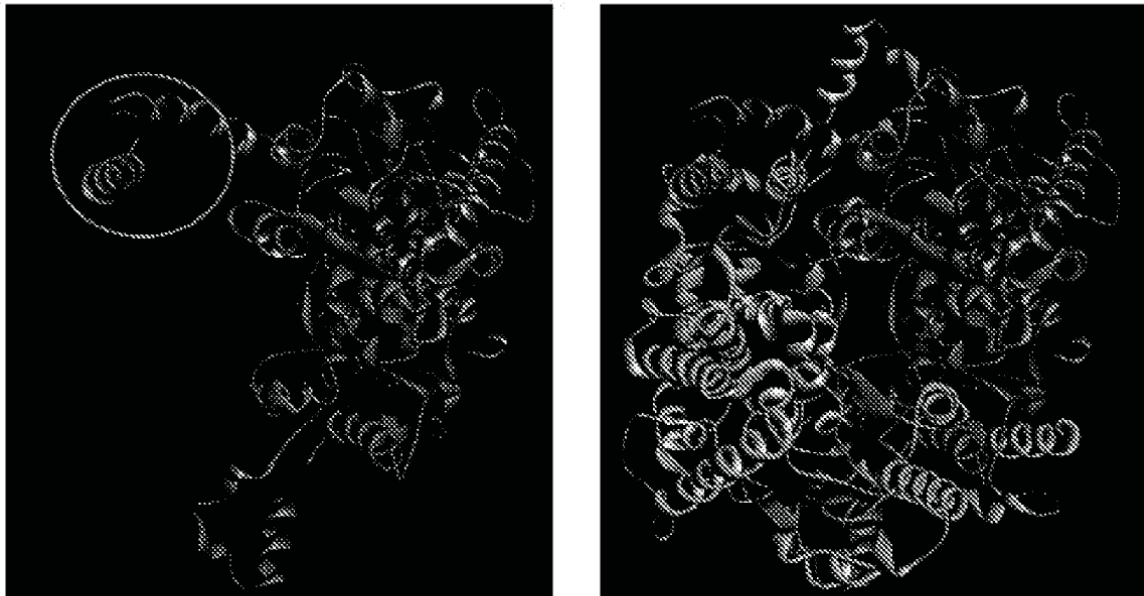


FIGURA 7

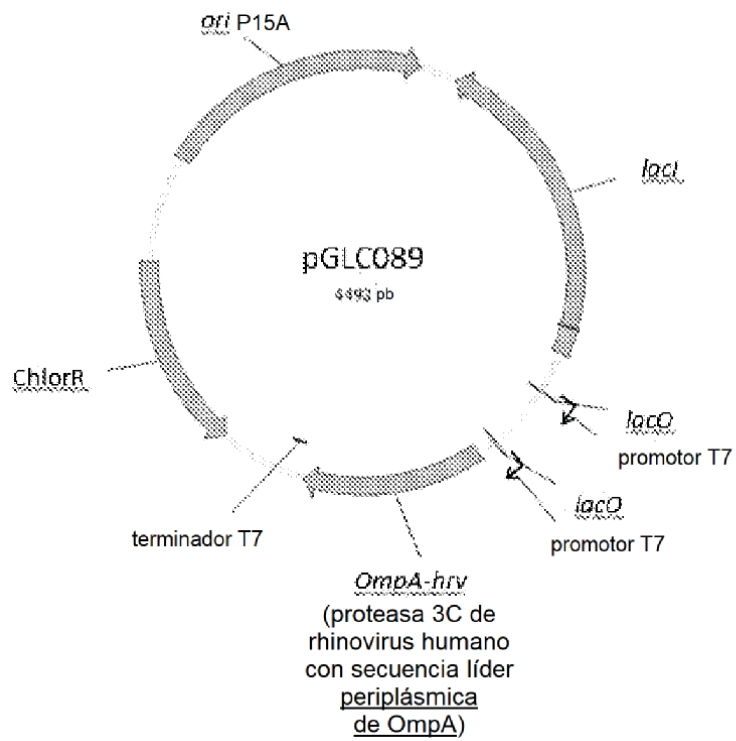


FIGURA 8

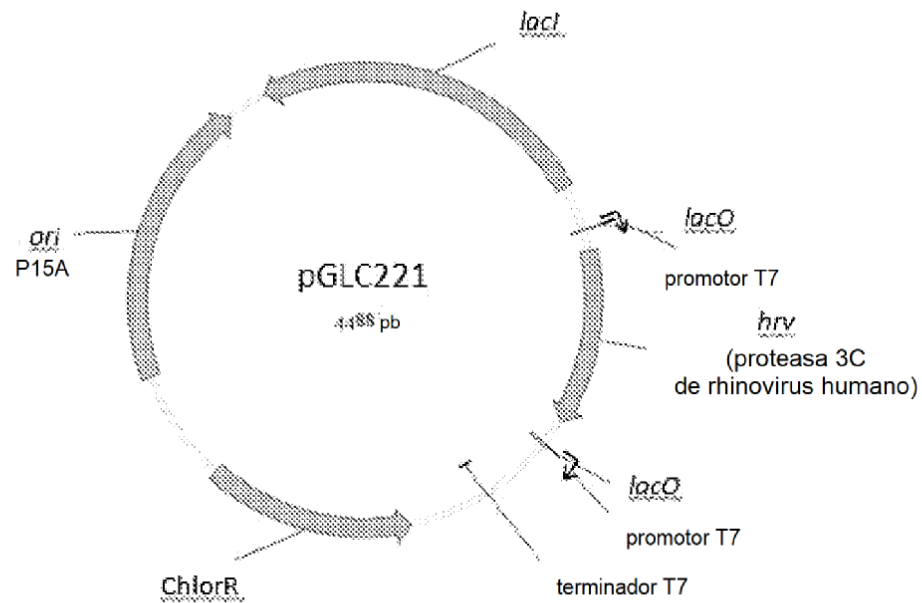


FIGURA 9

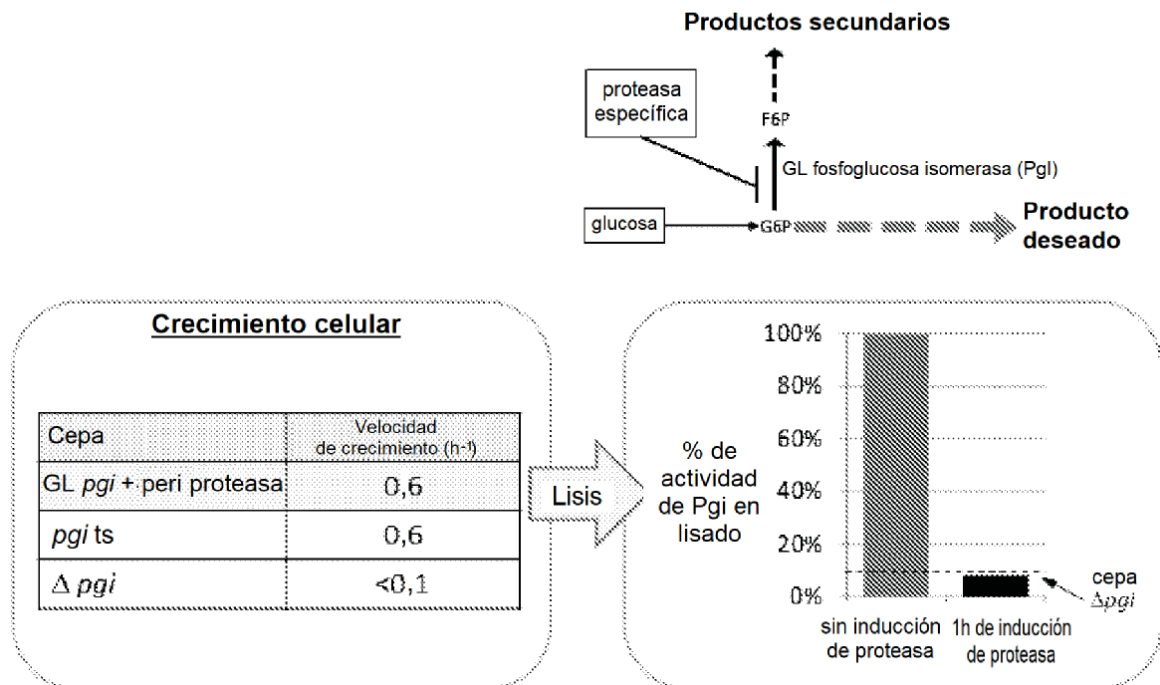


FIGURA 10

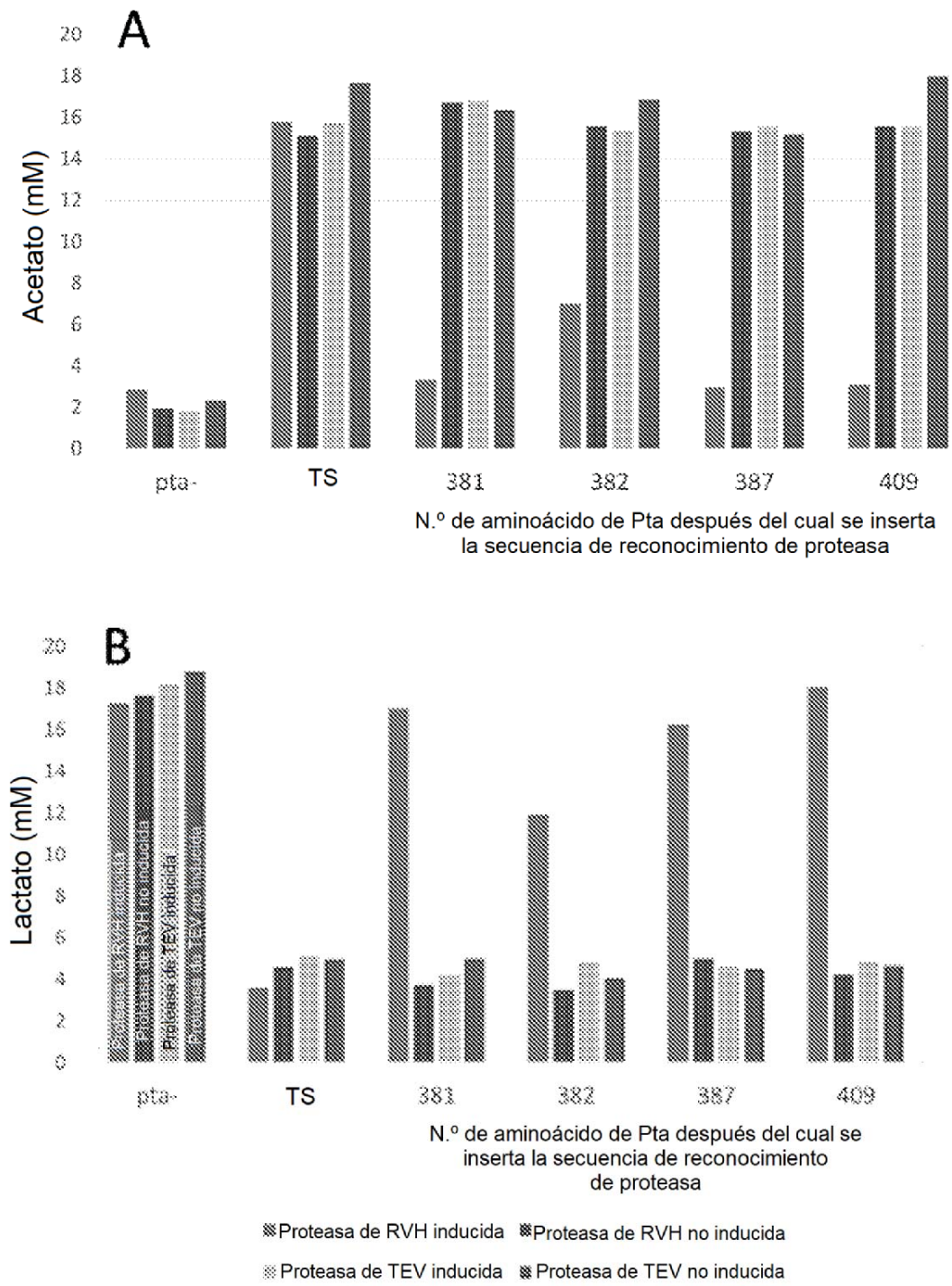


FIGURA 11

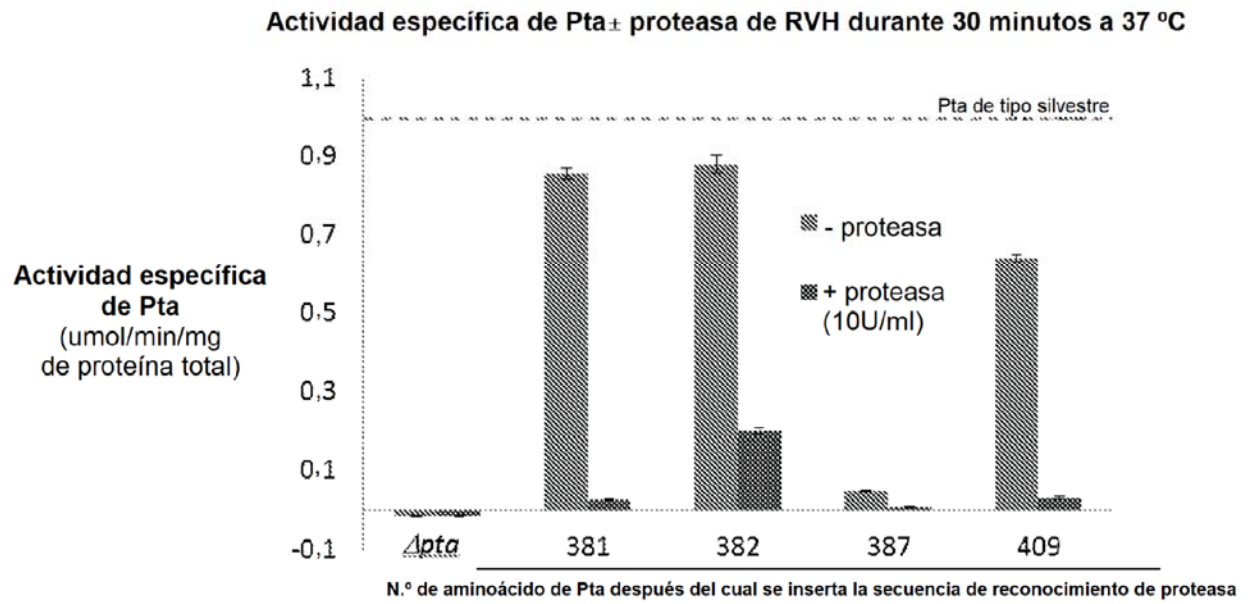


FIGURA 12

