



OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA



11) Número de publicación: 2 457 527

61 Int. Cl.:

C12N 5/10 (2006.01) C07H 21/02 (2006.01) C07H 21/04 (2006.01) C12P 21/02 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Fecha de presentación y número de la solicitud europea: 07.09.2007 E 07837971 (6)
 (97) Fecha y número de publicación de la concesión europea: 08.01.2014 EP 2061878

(54) Título: ARNt Supresor hibrido en células de vertebrados

(30) Prioridad:

08.09.2006 US 843092 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente: 28.04.2014

73) Titular/es:

AMBRX, INC. (100.0%) 10975 NORTH TORREY PINES ROAD, SUITE 100 LA JOLLA CA 92037, US

(72) Inventor/es:

TIAN, FENG; NORMAN, THEA y CHU, STEPHANIE

(74) Agente/Representante:

IZQUIERDO FACES, José

DESCRIPCIÓN

ARNt Supresor hibrido en células de vertebrados

5 CAMPO DE LA INVENCIÓN

10

65

[0001] La invención pertenece al campo de la traducción bioquímica en células de vertebrados. La invención se refiere a métodos y composiciones para producir ARNs de transferencia (ARNts) ortogonales, sintetasas ortogonales y pares de los mismos, en células de vertebrados. La invención también se refiere a composiciones de aminoácidos no naturales, proteínas y métodos de producción de proteínas en células de vertebrados que incluyen aminoácidos no naturales.

ANTECEDENTES DE LA INVENCIÓN

- [0002] El código genético de cada organismo conocido, desde bacterias hasta humanos, codifica los mismos veinte aminoácidos comunes. Las diferentes combinaciones de los mismos veinte aminoácidos naturales de proteínas llevan a cabo todos los procesos complejos de la vida, desde la fotosíntesis hasta la transducción de señal y la respuesta inmune. Para estudiar y modificar la estructura y la función de las proteínas, los científicos han intentado manipular tanto el código genético como la secuencia aminoácida de proteínas. Sin embargo, ha sido complicado eliminar las limitaciones impuestas por el código genético, que limita las proteínas a veinte unidades estructurales estándares genéticamente codificadas, con la rara excepción de la selenocisteína (véase, por ejemplo, A. Bock et al., (1991), Molecular Microbiology 5: 515 20) y la pirrolisina (véase, por ejemplo, G. Srinivasa, et al., (2002), Science 296: 1459 62).
- [0003] Se han realizado algunos progresos para eliminar estas limitaciones, aunque este progreso es limitado, y la capacidad de controlar racionalmente la estructura y la función de proteínas está todavía sin desarrollar. Por ejemplo, algunos químicos han desarrollado métodos y estrategias para sintetizar y manipular las estructuras de moléculas pequeñas (*véase, por ejemplo*, E. J. Corey, & X. M. Cheng, *The logic of chemical synthesis* (Wiley Interscience, New York, 1995)). La síntesis total (*véase, por ejemplo*, B. Merrifield, (1986), Science 232: 341 7 (19686)), y las metodologías semisintéticas (*véanse, por ejemplo*, D. Y. Jackson et al., (1994) Science 266: 243 7; y P. E. Dawson, & S. B. Kent, (2000), Annual Review of Biochemistry 69: 923 60), han hecho posible sintetizar péptidos y proteínas pequeñas, pero estas metodologías tienen utilidad limitada respecto a las proteínas de unos 10 kilo daltons (kDa). Los métodos de mutagénesis, aunque potentes, se limitan a un número restringido de cambios estructurales. En diversos casos, ha sido posible incorporar competitivamente análogos de aminoácidos comunes estructuralmente cercanos mediante proteínas. Véanse, por ejemplo, R. Further, (1998), Protein Science 7: 419 26; K. Kirshenbaum, et al., (2002), ChemBioChem 3: 235 7; y V. Doring et al., (2001), Science 292: 501 4.
- [0004] En un intento de expandir la capacidad de manipular la función y la estructura de proteínas, se han desarrollado métodos *in vitro* que utilizan ARNts ortogonales acilados químicamente que permiten a los aminoácidos no naturales incorporarse selectivamente en respuesta a un codón antisentido, *in vitro* (*véase, por ejemplo*, J. A. Ellman, et al., (1992), Science 255: 197 200). Los aminoácidos con nuevas estructuras y propiedades físicas se incorporaron selectivamente a proteínas para estudiar el plegamiento y la estabilidad proteicos y el reconocimiento y la catálisis biomolecular. Véase, por ejemplo, D. Mendel, et al., (1995), Annual Review of Biophysics and Biomolecular Structure 24: 435 462; y, V. W. Cornish, et al. (Mar. 31, 1995), Angewandte Chemie International Edición en Inglés 34: 621 633. Sin embargo, la naturaleza estoicométrica de este proceso ha limitado severamente la cantidad de proteína que podría generarse.
- [0005] Los aminoácidos no naturales se microinyectaron en células. Por ejemplo, se introdujeron aminoácidos no naturales en el receptor nicotínico de acetilcolina en oocitos de Xenopus (por ejemplo, M. W. Nowak, et al., (1998), In vivo incorporation of unnatural amino acids into ion channels in Xenopus oocyte expression system, Method Enzymol. 293: 504 529) mediante la microinyección de un ARNt disaminoacilado químicamente de Tetrahymena thermophila (por ejemplo, M. E. Saks, et al., (1996), An engineered Tetrahymena tRNAGIn for in vivo incorporation of unnatural amino acids into proteins by nonsense suppression, J. Biol. Chem. 271: 23169 23175), y el ARNm relevante. Esto ha permitido estudios biofísicos detallados del receptor en oocitos mediante la introducción de aminoácidos que incluyen cadenas laterales con propiedades físicas o químicas únicas. Véase, por ejemplo, D. A. Dougherty (2000), Unnatural amino acids as probes of protein structure and function, Curr. Opin. Chem. Biol. 4: 645 652. Desafortunadamente, esta metodología se limita a proteínas en células que pueden microinyectarse y, debido a que el ARNt relevante es acilado químicamente in vitro y no puede reacetilarse, los rendimientos de las proteínas son muy bajos.
 - [0006] Para superar estas limitaciones, se añadieron nuevos componentes a la maquinaria biosintética de la proteína del organismo procariota *Escherichia coli (E. coli)* (por ejemplo, L. Wang, et al., (2001), Science 292: 498 500) que permitieron la codificación genética de aminoácidos no naturales *in vivo*. Se añadieron eficientemente y con alta fidelidad numerosos aminoácidos nuevos con propiedades químicas, físicas o biológicas nuevas, incluyendo marcadores de fotoafinidad y aminoácidos fotoisomerizables, cetoaminoácidos y aminoácidos glicosilados a proteínas de *E. coli* en respuesta al codón ámbar, TAG, utilizando esta metodología. *Véanse, por ejemplo*, J. W.

Chin et al., (2002), Journal of the American Chemical Society 124: 9026 – 9027; J. W. Chin, & P. G. Schultz, (2002), ChemBioChem 11: 1135 – 1137; J. W. Chin, et al., (2002), PNAS United States of America 99: 11020 – 11024; y L. Wang, & P. G. Schultz, (2002), Chem. Comm. 1 – 10. Sin embargo, la maquinaria translacional de procariotas y eucariotas no se conserva muy bien; por lo tanto, los componentes de la maquinaria biosintética añadida a *E. coli* no pueden utilizarse a menudo para incorporar específicamente en el sitio aminoácidos no naturales en proteínas de células de vertebrados. Por ejemplo, el par tirosil - ARNt sintetasa / ARNt de *Methanococcus jannaschii* que se utilizó, en *E. coli* no es ortogonal en las células de vertebrados. Además, la transcripción de ARNt en eucariotas, pero no en procariotas, se lleva a cabo por ARN polimerasa III, y esto supone restricciones en la secuencia primaria de los genes estructurales del ARNt que pueden transcribirse en células de vertebrados. Además, a diferencia de las células procariotas, los ARNts en las células de vertebrados necesitan exportarse desde el núcleo, donde se transcriben, hasta el citoplasma, para funcionar en la traducción. Finalmente, el ribosoma 80S del vertebrado es diferente del ribosoma 70S procariótico. Por lo tanto, es necesario desarrollar componentes mejorados de la maquinaria biosintética para expandir el código genético en vertebrados. La presente invención cubre tanto esta como otras necesidades, como se verá reflejado en el análisis de la presente divulgación.

RESUMEN DE LA INVENCIÓN

5

10

15

20

40

45

50

55

60

65

[0007] La invención proporciona células de vertebrados, en las que la célula de vertebrado no es una célula madre embrionaria humana o una célula humana *in vivo*, con componentes de traducción, por ejemplo, pares de aminoacil – ARNt sintetasas ortogonales (O – RSs) y ARNts ortogonales (O – tRNAs con una secuencia como la descrita en la SEC ID Nº 87) y componentes individuales de los mismos, que se emplean en la maquinaria biosintética de las proteínas de vertebrados para incorporar un aminoácido no natural en una cadena polipeptídica cultivada en una célula de vertebrado.

25 [0008] Las composiciones de la invención incluyen una célula de vertebrado (por ejemplo, una célula de mamífero, una célula aviar, una célula de pez, una célula de reptil, una célula de anfibio, células derivadas de animales no mamíferos, etc.) que comprende una aminoacil - ARNt sintetasa ortogonal (O - RS) (por ejemplo, derivada de un organismo no vertebrado, como Escherichia coli, Bacillus stearothermophilus, etc.), en la que la O - RS aminoacila, preferentemente, al ARNt ortogonal con una secuencia como la descrita en la SEC ID Nº 87 (O - tRNA) con, al 30 menos, un aminoácido no natural en la célula de vertebrado. Opcionalmente, pueden aminoacilarse dos o más O tRNAs en una célula de vertebrado dada. En un aspecto, una O - RS aminoacila una O - tRNA con el aminoácido no natural, por ejemplo, al menos un 40 %, al menos un 45 %, al menos un 50 %, al menos un 60 %, al menos un 75 %, al menos un 80 % o incluso un 90 % o más tan eficazmente como lo hace una O - RS con una secuencia aminoácida, como se describe, por ejemplo, en la SEC ID Nº 86 o 45. En una realización, una O - RS de la 35 divulgación aminoacila el O - tRNA con el aminoácido no natural, por ejemplo, al menos 10 veces, al menos 20 veces, al menos 30 veces, etc. más eficientemente que la O - RS aminoacila el O - tRNA con un aminoácido natural.

[0009] En una realización, la O – RS o una parte de la misma está codificada por una secuencia polinucleótida como se describe en cualquiera de las SECs ID Nº 3 – 35, o una secuencia polinucleótida complementaria de la misma. En otra realización, la O – RS comprende una secuencia aminoácida como la descrita en cualquiera de las SECs ID Nº 36 – 63, y / o la 86, o una variación conservadora de la misma. En otra realización, la O – RS comprende una secuencia aminoácida que es, por ejemplo, idéntica en, al menos, el 90 %, al menos el 95 %, al menos el 98 %, al menos el 99 % o, al menos, el 99,5 % o más, a la de una tirosil aminoacil – ARNt sintetasa natural (TyrRS) y comprende dos o más aminoácidos de los grupos A – E. El grupo A incluye valina, isoleucina, leucina, glicina, serina, alanina o treonina en una posición correspondiente a Tyr37 de una TyrRS de *E. coli.* El grupo B incluye aspartato en una posición correspondiente a Asn126 de una TyrRS de *E. coli.* El grupo D incluye metionina, alanina, valina o tirosina en una posición correspondiente a Phe183 de una TyrRS de *E. coli.* El grupo E incluye serina, metionina, valina, cisteína, treonina o alanina en una posición correspondiente a Leu186 de una TyrRS de *E. coli.*

[0010] En otra realización, la O – RS tiene una o más propiedades enzimáticas aumentadas o mejoradas para los aminoácidos no naturales en comparación a un aminoácido natural. Por ejemplo, las propiedades aumentadas o mejoradas para el aminoácido no natural en comparación a un aminoácido natural incluyen, por ejemplo, un Km más alto, un Km más bajo, un kcat más alto, un kcat más bajo, un kcat / km más bajo, un kcat / km más alto, etc.

[0011] La célula de vertebrado también incluye, opcionalmente, aminoácido (s) no natural (es). La célula de vertebrado también incluye un ARNt ortogonal (O – tRNA) con una secuencia como la descrita en la SEC ID Nº 87, en la que el O – tRNA reconoce un codón selectivo y se aminoacila, preferentemente, con el aminoácido no natural mediante la O – RS. En un aspecto, el O – tRNA regula la incorporación del aminoácido no natural en una proteína en, por ejemplo, al menos el 45 %, al menos el 50 %, al menos el 60 %, al menos el 75 %, al menos el 80 %, al menos el 90 %, al menos el 95 % o el 99 % o la eficacia de un ARNt que comprende o se procesa en una célula procedente de una secuencia polinucleótida como la descrita en la SEC ID Nº 65.

[0012] En otra realización, la célula de vertebrado comprende un ácido nucleico que comprende un polinucleótido

que codifica un polipéptido de interés, en el que el polinucleótido comprende un codón selectivo reconocido por el O – tRNA. En un aspecto, el rendimiento del polipéptido de interés que comprende los aminoácidos no naturales es de, por ejemplo, al menos 2,5 %, al menos 5 %, al menos 10 %, al menos 25 %, al menos 30 %, al menos 40 %, 50 % o más, del obtenido con el polipéptido natural de interés procedente de una célula en la que el polinucleótido carece de codón selectivo. En otro aspecto, la célula produce el polipéptido de interés en ausencia del aminoácido no natural, con un rendimiento que es, por ejemplo, menor del 35 %, menor del 30 %, menor del 20 %, menor del 15 %, menor del 10 %, menor del 5 %, menor del 2,5 %, etc. del rendimiento del polinucleótido en presencia del aminoácido no natural.

5

30

35

40

55

60

- 10 [0013] La invención también proporciona una célula de vertebrado que comprende una aminoacil ARNt sintetasa ortogonal (O RS), un ARNt ortogonal (O tRNA con una secuencia como la descrita en la SE ID Nº 87), un aminoácido no natural y un ácido nucleico que comprende un polinucleótido que codifica un polipéptido de interés. El polinucleótido comprende un codón selectivo reconocido por el O tRNA. Además, la O RS aminoacila, preferentemente, el ARNt ortogonal (O tRNA) con el aminoácido no natural en la célula de vertebrado, y la célula produce el polipéptido de interés en ausencia del aminoácido no natural, con un rendimiento que es, por ejemplo, menor del 30 %, menor del 20 %, menor del 15 %, menor del 10 %, menor del 5 %, menor del 2,5 %, etc. del rendimiento del polipéptido en presencia del aminoácido no natural.
- [0014] Las composiciones que incluyen una célula de vertebrado que comprenden un ARNt ortogonal (O tRNA con una secuencia como la descrita en la SEC ID Nº 87) también son una característica de la invención. Normalmente, el O tRNA regula la incorporación de un aminoácido no natural en una proteína que se codifica por un polinucleótido que comprende un codón selectivo reconocido por el O tRNA *in vivo*. En una realización, el O tRNA regula la incorporación del aminoácido no natural en la proteína con, por ejemplo, al menos el 45 %, al menos el 50 %, al menos el 60 %, al menos el 75 %, al menos el 80 %, al menos el 90 %, al menos el 95 % o incluso el 99 % o más de la eficacia de un ARNt que comprende o se procesa en una célula procedente de una secuencia polinucleótida como la descrita en la SEC ID Nº 65.
 - [0015] En un aspecto de la presente divulgación, el O tRNA se modifica post transcripcionalmente. La divulgación también proporciona un ácido nucleico que codifica un O tRNA en una célula de vertebrado, o un polinucleótido complementario del mismo. En una realización, el ácido nucleico comprende una caja A y una caja B.
 - **[0016]** La invención también proporciona métodos para producir, en una célula de vertebrado, al menos, una proteína que comprende, al menos, un aminoácido no natural (así como proteínas producidas por dichos métodos). Los métodos incluyen, por ejemplo, el cultivo en un medio apropiado de una célula de vertebrado que comprende un ácido nucleico que incluye, al menos, un codón selectivo y codifica la proteína. La célula de vertebrado también comprende un ARNt ortogonal con una secuencia como la descrita en la SEC ID Nº 87 u 88 (O tRNA) que funciona en la célula y reconoce el codón selectivo y una aminoacil ANRt sintetasa ortogonal (O RS) que aminoacila, preferentemente, el O tRNA con el aminoácido no natural, y el medio comprende un aminoácido no natural. En una realización, la O RS aminoacila el O tRNA con el aminoácido no natural por ejemplo, al menos el 45 %, al menos el 50 %, al menos el 60 %, al menos el 75 %, al menos el 80 %, al menos el 90 %, al menos el 95 % o incluso el 99 % o más tan eficientemente como lo hace una O RS que tiene una secuencia aminoácida, por ejemplo, como la descrita en la SEC ID Nº 86 o 45. En otra realización, la O RS comprende una secuencia aminoácida como la descrita en cualquiera de las SECs ID Nº 36 63 y / o la 86.
- 45 [0017] En ciertas realizaciones, la proteína codificada comprende una proteína terapéutica, una proteína de diagnóstico, una enzima industrial o una parte de las mismas. En una realización, la proteína producida mediante el método se modifica, además, mediante un aminoácido no natural. Por ejemplo, el aminoácido no natural se modifica mediante, por ejemplo, una reacción nucleofílica electrofílica, cicloadición [3 + 2], etc. En otra realización, la proteína producida mediante el método se modifica por, al menos, una modificación post translacional (por ejemplo, N glicosilación, O glicosilación, acetilación, acitación, modificación lipídica, palmitolación, adición de palmitato, fosforilación, modificación del enlace glicolipídico y similares) in vivo.
 - [0018] En algunas realizaciones, las composiciones y métodos de la invención incluyen células de vertebrados. Una célula de vertebrado de la invención incluye cualquiera entre, por ejemplo, una célula de mamífero, una célula de levadura, una célula de hongo, una célula de planta, una célula de insecto, etc. en la que la célula no es una célula madre no embrionaria humana o una célula humana *in vivo*. Los componentes de traducción de la invención pueden derivarse de numerosos organismos como, por ejemplo, organismos no vertebrados, como un organismo procariótico (por ejemplo, *E. coli, Bacillus steatothermophilus* o similares), o una arqueobacteria, o, por ejemplo, un organismo vertebrado.
 - **[0019]** Un codón selectivo de la invención expande la estructura del codón genético de la maquinaria biosintética de la proteína del vertebrado. Pueden utilizarse numerosos codones selectivos en la invención, incluyendo codones de terminación (por ejemplo, un codón ámbar, un codón ocre o un codón de terminación ópalo), codones sin sentido, codones raros, codones de cuatro (o más) bases y / o similares.
 - [0020] Algunos ejemplos de aminoácidos no naturales que pueden emplearse en las composiciones y métodos

descritos en la presente incluyen (de manera no limitante): una p – acetil – L – fenilalanina, una p – yodo – L – fenilalanina, una O – metil – L – tirosina, una p – propargiloxi – fenilalanina, una p – propargil – fenilalanina, una L – 3 – (2 – naftil) alanina, una 3 – metil – fenialalanina, una 0 – 4 – alil – L – tirosina, una 4 – propil – L – tirosina, una tri O – acetil – GlcNAcβ – serina, una L – Dopa, una fenialalanina fluorinada, una isopropil – L – fenilalanina, una p – azido – L – fenilalanina, una p – acil – L – fenilalanina, una p – benzoil – L – fenilalanina, una L – fosfoserina, una fosfonoserina, una fosfonotirosina, una p – bromofenilalanina, una p – amino – L – fenilalanina, una isopropil – L – fenilalanina, un análogo no natural de un aminoácido tirosina; un análogo no natural de un aminoácido glutamina; un análogo no natural de un aminoácido fenilalanina; un análogo no natural de un aminoácido serina; un análogo no natural de un aminoácido treonina; un alquilo, arilo, acilo, azido, ciano, halo, hidracina, hidracida, hidroxilo, alquenilo, alquinilo, éter, tiol, sulfonil, seleno, éster, tioácido, borato, boronato, fosfo, fosfono, fosfina, heterocíclico, enona, imina, aldehído, hidroxilamina, ceto, o aminoácido amino sustituido, o cualquier combinación de los mismos; un aminoácido con un reticulador fotoactivable; un aminoácido marcado en spin; un aminoácido fluorescente; un aminoácido de unión a metales; un aminoácido con metal; un aminoácido radioactivo; un aminoácido fotocargado y / o fotoisomerizable; un aminoácido con biotina o un análogo de biotina; un aminoácido con ceto; un aminoácido con polietilenglicol o poliéter; un aminoácido sustituido por un átomo pesado; un aminoácido químicamente escindible o fotoescindible; un aminoácido con una cadena lateral alargada; un aminoácido con un grupo tóxico; un aminoácido sustituido con azúcar: un aminoácido con azúcar con enlaces carbono: un aminoácido redox activo: un aminoácido con α – hidroxi; un aminotioácido; un aminoácido disustituido con α , α ; un aminoácido β ; un aminoácido cíclico diferente de prolina o histidina, un aminoácido aromático diferente de fenilalanina, tirosina o triptófano, y / o similares.

[0021] Se describen en la presente polipéptidos (O – RS) y polinucleótidos, por ejemplo polinucleótidos de O – tRNAs que codifican O – Rss o partes de las mismas (por ejemplo, el sitio activo de la sintetasa), oligonucleótidos utilizados para construir mutantes de aminoacil – ARNt sintetasa, polinucleótidos que codifican una proteína o un polipéptido de interés que comprende uno o más codones selectivos, etc. Por ejemplo, un polinucleótido incluye un polipéptido que comprende una secuencia aminoácida como la descrita en cualquiera de las SECs ID Nº 36 – 63 y / o la 86, un polipéptido que comprende una secuencia aminoácida codificada por una secuencia polinucleótida como la descrita en cualquiera de las SECs ID Nº 3 – 35, y un polipéptido que es inmunoreactivo específicamente con un anticuerpo específico de un polipéptido que comprende una secuencia aminoácida como se muestra en cualquiera de las SECs ID Nº 36 – 63 y / o la 86, o un polipéptido que comprende una secuencia aminoácida codificada por una secuencia polinucleótida como se muestra en cualquiera de las SECs ID Nº 3 – 35.

[0022] También se incluye entre los polipéptidos descritos en la presente un polipéptido que comprende una secuencia aminoácida que es indéntica en, al menos, el 90 % a la tirosil aminoacil – ARNt sintetasa (TyrRS) natural (por ejemplo la SEC ID Nº 2) y comprende dos o más aminoácidos de los grupos A - (descritos anteriormente). De manera similar, los polipéptidos también incluyen, opcionalmente, un polipéptido que comprende, al menos, 20 aminoácidos contiguos a cualquiera de las SECs ID Nº 36 – 63 y / o la 86, y dos o más sustituciones aminoácidas como de indica anteriormente en los grupos A – E. También se incluye como polipéptido de la invención una secuencia aminoácida que comprende una variación conservadora de cualquiera de los polipéptidos anteriores.

40 [0023] Los kits también son una característica de la invención. Por ejemplo, se presenta un kit para producir una proteína que comprende, al menos, un aminoácido no natural en una célula, que incluye un depósito con una secuencia polinucleótida como la descrita en la SEC ID Nº 87 u 88. En una realización, el kit incluye, además, al menos un aminoácido no natural. En otra realización, el kit comprende, además, materiales informativos de enseñanza para producir la proteína.

BREVE DESCRIPCIÓN DE LAS FIGURAS

[0024]

5

10

15

20

25

30

35

65

50 La Figura 1 muestra la expresión incrementada de hGH utilizando el ARNt híbrido.

DESCRIPCIÓN DETALLADA

[0025] Antes de describir detalladamente la presente invención, cabe señalar que esta invención no se limita a dispositivos particulares o sistemas biológicos, que pueden, por supuesto, variar. También cabe destacar que la terminología empleada en la presente tiene el objetivo de describir únicamente realizaciones particulares, y no pretende ser limitante. Las formas singulares "un", "una", "el" y "la" utilizadas en este especificación así como en las reivindicaciones, incluyen referencias plurales a menos que el contexto indique claramente lo contrario. Por lo tanto, por ejemplo, la referencia a "una célula" incluye una combinación de dos o más células; la referencia a "bacteria" incluye mezclas de bacterias, y similares.

[0026] A menos que se defina lo contrario en la presente o en la parte restante de la especificación, todos los términos técnicos y científicos empleados en la presente tienen el mismo significado que el entendido por cualquier experto en la disciplina a la que pertenece la invención.

[0027] Homólogo: Las proteínas y / o secuencias de proteínas son "homólogas" cuando derivan, natural o

artificialmente, de una proteína o secuencia de proteína común ancestral. De manera similar, los ácidos nucleicos y / o secuencias de ácido nucleico son homólogas cuando derivan, natural o artificialmente, de un ácido nucleico o secuencia de ácido nucleico común ancestral. Por ejemplo, un ácido nucleico natural puede modificarse mediante cualquier método de mutagénesis disponible para incluir uno o más codones selectivos. Cuando se expresa, este ácido nucleico mutagenizado codifica un polipéptido que comprende uno o más aminoácidos no naturales. El proceso de mutación puede, por supuesto, alterar adicionalmente uno o más codones estándares, cambiando así uno o más aminoácidos estándar en la proteína mutante resultante. La homología se deduce, generalmente, de la similitud de secuencia entre dos o más ácidos nucleicos o proteínas (o secuencias de los mismos). El porcentaje exacto de similitud entre secuencia útil para establecer la homología varía con el ácido nucleico y la proteína, pero se utiliza una similitud de secuencia rutinaria del 25 % para establecer la homología. También pueden utilizarse para establecer la homología niveles de similitud de secuencia más altos, por ejemplo del 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 99 % o más. Se describen en la presente métodos generalmente disponibles para determinar los porcentajes de similitud de secuencia (por ejemplo, BLAST y BLASTN utilizando parámetros por defecto).

5

10

35

40

55

60

65

15 [0028] Ortogonal: El término "ortogonal" utilizado en la presente se refiere a una molécula (por ejemplo, un ARNt ortogonal (O - tRNA) y / o a una aminoacil - ARNt sintetasa ortogonal (O - RS)) que funciona con componentes endógenos de una célula con eficacia reducida en comparación a la molécula correspondiente endógena a la célula o al sistema de traducción, o que rechaza funcionar con componentes endógenos de la célula. En el contexto de ARNts y aminoacil – ARNt sintetasas, ortogonal se refiere a incapacidad o a eficacia reducida, por ejemplo, eficaz en 20 menos del 20 %, menos del 10 %, menos del 5 % o menos del 1 % de un ARNt ortogonal para funcionar con una ARNt sintetasa endógena en comparación a un ARNt endógeno para funcionar con la ARNt sintetasa endógena, o de una amioacil - ARNt sintetasa ortogonal para funcionar con un ARNt endógeno en comparación a una ARNt sintetasa endógena para funcionar con el ARNt endógeno. La molécula ortogonal carece de molécula funcional complementaria endógena en la célula. Por ejemplo, un ARNt ortogonal en una célula es aminoacilado por cualquier 25 RS endógena de la célula con eficacia reducida o incluso cero, en comparación a la aminoacilación de un ARNt endógeno por la RS endógena. En otro ejemplo, una RS ortogonal aminoacila cualquier ARNt endógeno en una célula de interés con eficacia reducida o incluso cero, en comparación a la aminoacilación de ARNt endógeno por la RS endógena. Una segunda molécula ortogonal puede introducirse en la célula que funciona con la primera molécula ortogonal. Por ejemplo, un par ARNt / RS ortogonal incluye componentes introducidos 30 complementariamente que funcionan juntos en la célula con una eficacia (por ejemplo, con una eficacia del 50 %, el 60 %, el 70 %, el 75 %, el 80 %, el 90 %, el 95 %, el 99 % o más) a la de un par ARNt / RS endógeno correspondiente.

[0029] Complementario: El término "complementario" se refiere a componentes de un par ortogonal, O – tRNA y O – RS que pueden funcionar juntos, por ejemplo, cuando la O – RS aminoacila el O – tRNA.

[0030] Aminoacila, preferentemente: El término "aminoacila, preferentemente" se refiere a una eficacia de, por ejemplo, el 70 %, el 75 %, el 85 %, el 90 %, el 95 %, el 99 % o más, en la que una O – RS aminoacila un O – tRNA con un aminoácido no natural en comparación a la O – RS aminoacilando un ARNt natural o un material de inicio empleado para generar el O – tRNA. El aminoácido no natural se incorpora a la cadena polipeptídica en el cultivo con alta fidelidad, por ejemplo, superior al 75 % de eficacia para un codón selectivo dado, superior al 80 % de eficacia para un codón selectivo dado, superior al 95 % de eficacia para un codón selectivo dado, superior al 99 % o más de eficacia para un codón selectivo dado.

45 [0031] Codón selectivo: El término "codón selectivo" se refiere a codones reconocidos por el O – tRNA en el proceso de traducción y no reconocidos por un ARNt endógeno. El lazo anticodón de O – tRNA reconoce el codón selectivo en el ARNm e incorpora sus aminoácidos, por ejemplo, un aminoácido no natural, en este sitio del polipéptido. Los codones selectivos puede incluir, por ejemplo, codones antisentido, como, codones de terminación, por ejemplo codones ámbar, ocre y ópalo; codones de cuatro o más bases; codones raros; codones derivados de pares de bases naturales o no y / o similares.

[0032] ARNt supresor: Un ARNt supresor es un ARNt que altera la lectura de un ARN mensajero (ARNm) en un sistema de traducción dado, por ejemplo, proporcionando un mecanismo para incorporar un aminoácido en una cadena polipéptida en respuesta a un codón selectivo. Por ejemplo, un ARNt supresor puede leer a través de, por ejemplo, un codón de terminación, un codón de cuatro bases, un codón raro y / o similares.

[0033] ARNt reciclable: El término "ARNt reciclable" se refiere a un ARNt que está aminoacilado y puede reaminoacilarse repetidamente con un aminoácido (por ejemplo un aminoácido no natural) para la incorporación del aminoácido (por ejemplo el aminoácido no natural) en una o más cadenas polipeptídicas durante la traducción.

[0034] <u>Sistema de traducción</u>: El término "sistema de traducción" se refiere al conjunto de componentes que incorporan un aminoácido natural en una cadena polipéptida cultivada (proteína). Los componentes de un sistema de traducción pueden incluir, por ejemplo, ribosomas, ARNts, sintetasas, ARNm, aminoácidos y similares. Los componentes de la invención (por ejemplo, O – RS, O – tRNAs, aminoácidos no naturales, etc) pueden añadirse a un sistema de traducción *in vitro* o *in vivo* como, por ejemplo, una célula de vertebrado, por ejemplo una célula de levadura, una célula de mamífero, una célula de planta, una célula de alga, una célula de hongo, una célula de

insecto y / o similares.

5

35

50

55

60

- [0035] Aminoácido no natural: El término "aminoácido no natural" utilizado en la presente se refiere a cualquier aminoácido, aminoácido modificado y / o análogo de aminoácido que no es uno de los 20 aminoácidos esenciales naturales, seleno cisteína o pirrolisina.
- [0036] <u>Derivado de</u>: El término "derivado de" utilizado en la presente se refiere a un componente que es aislado de o hecho utilizando información de una molécula u organismo específicos.
- 10 **[0037]** RS inactiva: El término "RS inactiva" utilizado en la presente se refiere a una sintetasa que ha sido mutada de manera que no puede aminoacilar más su ARNt semejante natural con un aminoácido.
- [0038] Marcador de selección o cribado positivo: El término "marcador de selección o cribado positivo" utilizado en la presente se refiere a un marcador que, cuando está presente, por ejemplo, expresado, activado o similares, da como resultado la identificación de una célula con el marcador de selección positiva de aquellos sin el marcador de selección positiva.
- [0039] Marcador de selección o cribado negativo: El término "marcador de selección o cribado negativo" utilizado en la presente se refiere a un marcador que, cuando está presente, por ejemplo, expresado, activado o similares, permite la identificación de una célula que no posee la propiedad deseada (por ejemplo en comparación a una célula que sí posee la propiedad deseada).
- [0040] Indicador: El término "indicador" utilizado en la presente se refiere a un componente que puede utilizarse para seleccionar componentes diana de un sistema de interés. Por ejemplo, un indicador puede incluir un marcador fluorescente (por ejemplo, proteína verde fluorescente), un marcador luminiscente (por ejemplo, una proteína luciferasa "firefly"), un marcador basado en afinidad, o genes marcadores de selección como his3, ura3, leu2, lys2, lacZ, β gal / lacZ (β-galactosidasa), Adh (alcohol deshidrogenasa) o similares.
- [0041] <u>Vertebrado</u>: El término "vertebrado" utilizado en la presente se refiere a organismos que pertenecen al dominio Eucarya como animales, por ejemplo mamíferos, reptiles, aves, etc.
 - [0042] No eucariota: El término "no eucariota" utilizado en la presente se refiere a organismos no vertebrados. Por ejemplo, un organismo no vertebrado puede pertenecer al dominio filogenético de las Eubacterias (por ejemplo Escherichia coli, Thermus thermophilus, Bacillus stearothermophilus, etc.), o al dominio filogenético de las Arqueobacterias (por ejemplo, Methanococcus jannaschii, Methanobacterium thermoautotrophicum, Halobacterium como las especies NRC I Haloferax volcanii y Halobacterium, Archaeoglobus fulgidus, Pyrococcus furiosus, Pyrococcus horikoshii, Aeuropyrum pernix.
- [0043] Anticuerpo: El término "anticuerpo" utilizado en la presente incluye, de manera no limitante, un polipéptido sustancialmente codificado por el / los gen o genes de la inmunoglobulina o fragmentos del mismo, que se une específicamente y reconoce un analito (antigen). Algunos ejemplos incluyen anticuerpos policionales, monoclonales, quiméricos y monocatenarios, y similares. Los fragmentos de inmunoglobulinas, incluyendo los fragmentos Fab y los fragmentos producidos por una biblioteca de expresión, incluyendo la manifestación del fago, también se incluyen en el término "anticuerpo" utilizado en la presente. Véase, por ejemplo, Paul, Fundamental Immunology, 4ª Ed., 1999, Raven Press, New York, para la estructura y terminología sobre anticuerpos.
 - [0044] <u>Variante conservadora</u>: El término "variante conservadora" se refiere a un componente de la traducción, por ejemplo una variante conservadora de O tRNA o una variante conservadora de O RS, que actúa funcionalmente como el componente en el que se basa la variante conservadora, por ejemplo un O tRNA o una O RS, pero tiene variaciones en la secuencia. Por ejemplo, una O RS aminoacilará un O tRNA complementario o una variante conservadora de O tRNA con un aminoácido no natural, aunque el O tRNA y la variante conservadora de O tRNA no tengan la misma secuencia. La variante conservadora puede tener, por ejemplo, una variación, dos variaciones, tres variaciones, cuatro variaciones, o cinco o más variaciones en la secuencia, siempre y cuando la variante conservadora sea complementaria al O tRNA o a la O RS correspondientes.
 - [0045] Agente de selección o de cribado: El término "agente de selección o cribado" utilizado en la presente se refiere a un agente que, cuando está presente, permite la selección / cribado de ciertos componentes de una población. Por ejemplo, un agente de selección o cribado incluye, de manera no limitante, por ejemplo un nutriente, un antibiótico, una longitud de onda de luz, un anticuerpo, un polinucleótido expresado (por ejemplo una proteína moduladora transcripcional), o similares. El agente de selección puede variar, por ejemplo, por la concentración, la intensidad, etc.
 - [0046] <u>Sustancia detectable</u>: El término "sustancia detectable" utilizado en la presente se refiere a un agente que, cuando está activo, alterado, expresado o similares, permite la selección / cribado de ciertos componentes de una población. Por ejemplo, la sustancia detectable puede ser un agente químico, por ejemplo, ácido 5 fluorótico (5 FOA) que, en ciertas condiciones, por ejemplo la expresión de un indicador URA3, se vuelve detectable, por ejemplo

un producto tóxico que mata las células que expresan el indicador URA3.

- **[0047]** La capacidad de modificar genéticamente las estructuras de proteínas directamente en células de vertebrados, más allá de las limitaciones químicas impuestas por el código genético, proporcionaría una potente herramienta molecular para sondear y manipular procesos celulares. La invención proporciona componentes de traducción que expanden el número de aminoácidos genéticamente codificados en células de vertebrados. Estos incluyen ARNts (por ejemplo ARNts ortogonales (O tRNAs)), aminoacil ARNt sintetasas (por ejemplo sintetasa ortogonal (O RS)), pares de O tRNA / O RSs y aminoácidos naturales.
- 10 [0048] Típicamente, los O tRNAs de la invención se expresan y procesan eficazmente, y funcionan en la traducción en una célula de vertebrado, pero no se aminoacilan significativamente por las aminoacil ARNt sintetasas del huésped. En respuesta a un codón selectivo, un O tRNA de la invención emite un aminoácido no natural que no codifica ninguno de los veinte aminoácidos esenciales, a una cadena polipéptida cultivada durante la traducción de ARNm.
 - **[0049]** Una O RS aminoacila, preferentemente, un O tRNA de la invención con un aminoácido no natural en una célula de vertebrado, pero no aminoacila ninguno de los ARNts de huésped citoplasmático. Además, la especificidad de una aminoacil ARNt sintetasa presenta la aceptación de un aminoácido no natural mientras excluye cualquier aminoácido endógeno. También se describen en la presente polipéptidos que incluyen secuencias aminoácidas de O RSs, o partes de las mismas. Además, también se describen en la presente polinucleótidos que codifican componentes de traducción, O tRNAs, O RSs y partes de los mismos.
- [0050] La invención también proporciona los métodos para producir una proteína en una célula de vertebrado, donde la proteína comprende un aminoácido no natural. La proteína se produce empleando los componentes de traducción de la invención.
 - [0051] También son una característica de la invención los kits para producir una proteína o un polipéptido con un aminoácido no natural.
- 30 Aminoacil ARNt sintetasas ortogonales (O RS)

5

15

20

35

- [0052] Para incorporar específicamente un aminoácido no natural en una proteína o polipéptido de interés en una célula de vertebrado, la especificidad del sustrato de la sintetasa se altera de manera que solo el aminoácido no natural, pero no ninguno de los 20 aminoácidos esenciales, están cargados al ARNt. Si la sintetasa ortogonal es promiscua, dará como resultado proteínas mutantes con una mezcla de aminoácidos naturales y no naturales en la posición diana. La invención proporciona las composiciones de, y los métodos para, producir aminoacil ARNt sintetasas ortogonales que tienen especificidad de sustrato modificada para un aminoácido no natural específico.
- [0053] Se describe una célula o linea celular de un vertebrado que incluye una aminoacil ARNt sintetasa ortogonal (O RS). La O RS aminoacila, preferentemente, un ARNt ortogonal con una secuencia como la establecidad en la SEC ID Nº 87 (O tRNA) con un aminoácido no natural en la célula de vertebrado. En algunas realizaciones, la O RS emplea más de un aminoácido no natural, por ejemplo dos o más, tres o más, etc. Por lo tanto, una O RS de la divulgación puede tener la capacidad de aminoacilar, preferentemente, un O tRNA con una secuencia como la establecida en la SEC ID Nº 87 con diferentes aminoácidos naturales. Esto permite un nivel de control adicional seleccionando qué aminoácidos no naturales o qué combinación de aminoácidos no naturales se van a introducir en la célula y / o seleccionando las diferentes cantidades de aminoácidos no naturales que se van a introducir en la célula para su incorporación.
- [0054] Una O RS tiene, opcionalmente, una o más propiedades aumentadas o mejoradas para el aminoácido no natural en comparación a un aminoácido natural. Estas propiedades incluyen, por ejemplo, un Km más alto, un Km más bajo, un kcat más alto, un kcat más bajo, un kcat / km más bajo, un kcat / km más alto, etc. para el aminoácido no natural en comparación a un aminoácido natural, por ejemplo, uno de los 20 aminoácidos esenciales conocidos.
- [0055] Opcionalmente, la O RS puede presentarse en la célula de vertebrado mediante un polipéptido que incluye una O RS y / o mediante un polinucleótido que codifica una O RS o una parte de la misma. Por ejemplo, una O RS o una parte de la misma, se codifica mediante una secuencia polinucleótida como la descrita en cualquiera de las SECs ID Nº 3 35 o una secuencia polinucleótida complementaria de las mismas. En otro ejemplo, una O RS comprende una secuencia aminoácida como la descrita en cualquiera de las SECs ID Nº 36 63 y / o la 86, o una variación conservadora de las mismas. Véanse, por ejemplo, las Tablas 5, 6 y 8 y el Ejemplo 6 descritos en la presente para las secuencias de moléculas de O RS ejemplares.
 - [0056] Una O RS también puede incluir una secuencia aminoácida que es idéntica en, por ejemplo, al menos el 90 %, al menos el 95 %, al menos el 98 %, al menos el 99 % o, incluso, al menos el 99,5 % a la de una tirosil aminoacil ARNt sintetasa natural (TyrRS) (por ejemplo, como la descrita en la SEC ID N° 2) y comprende dos o más aminoácidos del grupo A E. El grupo A incluye valina, isoleucina, leucina, glicina, serina, alanina o treonina en una posición correspondiente a Tyr37 de una TyrRS de *E. coli*. El grupo B incluye aspartato en una posición

correspondiente a Asn126 de una TyrRS de *E. coli*. El grupo C incluye treonina, serina, arginina, asparagina o glicina en una posición correspondiente a Asp182 de una TyrRS de *E. coli*. El grupo D incluye metionina, alanina, valina o tirosina en una posición correspondiente a Phe183 de una TyrRS de *E. coli*. El grupo E incluye serina, metionina, valina, cisteína, treonina o alanina en una posición correspondiente a Leu186 de una TyrRS de *E. coli*.

[0057] Además de la O – RS, una célula de vertebrado de la divulgación puede incluir componentes adicionales, por ejemplo, un aminoácido no natural. La célula de vertebrado también incluye un ARNt ortogonal (O – tRNA) (por ejemplo, derivado de un organismo no vertebrado, como *Escherichia coli, Bacillu stearothermophilus* y / o similares), en el que el O – tRNA reconoce un codón selectivo y se aminoacila, preferentemente, con el aminoácido no natural mediante la O – RS. También pueden estar presente en la célula un ácido nucleico que comprende un polinucleótido que codifica un polipéptido de interés, en el que el polinucleótido comprende un codón selectivo reconocido por el O – tRNA, o una combinación de uno o más de estos.

[0058] En un aspecto, el O – tRNA regula la incorporación del aminoácido no natural en una proteína con una eficacia de, al menos, el 45 %, al menos el 50 %, al menos el 60 %, al menos el 75 %, al menos el 80 %, al menos el 90 %, al menos el 95 % o el 99 % del ARNt que comprende o se procesa a partir de una secuencia polinucleótida como la descrita en la SEC ID Nº 65.

ARNts ortogonales

5

10

15

20

25

30

35

40

45

50

55

65

[0059] La invención presenta una célula o línea celular de vertebrado que incluye un ARNt ortogonal (O – tRNA) con una secuencia como la descrita en las SECs ID Nº 87 u 88 en la que la célula no es una célula madre embrionaria humana o una célula humana *in vivo* y la línea celular no es una línea celular embrionaria humana. El ARNt ortogonal regula la incorporación de un aminoácido no natural en una proteína que se codifica mediante un polinucleótido que comprende un codón selectivo reconocido por el O – tRNA *in vivo*. En algunas realizaciones, un O – tRNA de la invención regula la incorporación de un aminoácido no natural en una proteína con una eficacia de, al menos, el 40 %, al menos el 45 %, al menos el 50 %, al menos el 60 %, al menos el 75 %, al menos el 80 % o incluso el 90 % o más que el ARNt que comprende o se procesa en una célula a partir de una secuencia polinucleótida como la descrita en la SEC ID Nº 65. *Véase* la Tabla 5 descrita en la presente.

Fidelidad, eficacia y rendimiento

[0060] Fidelidad se refiere a la exactitud con la que una molécula deseada, por ejemplo, un aminoácido no natural o un aminoácido, se incorpora en un polipéptido cultivado en una posición deseada. Los componentes de traducción de la invención incorporan aminoácidos no naturales, con alta fidelidad, en proteínas en respuesta a un codón selectivo. Por ejemplo, empleando los componentes de la invención, la eficacia de incorporación de un aminoácido no natural deseado en una cadena polipéptida cultivada en una posición deseada (por ejemplo, en respuesta a un codón selectivo) es eficaz en más de un 75 %, más de un 85 %, más de un 95 %, o incluso más de un 99 % en comparación a la incorporación no deseada de un aminoácido natural específico que se está incorporando en la posición deseada de la cadena polipéptida cultivada.

[0061] La eficacia puede también referirse al grado con el que la O – RS aminoacila el O – tRNA con el aminoácido no natural en comparación al control relevante. Las O – RSs pueden definirse por su eficacia. En ciertas realizaciones de la divulgación, se compara una O – RS a otra O – RS. Por ejemplo, una O – RS de la divulgación aminoacila un O – tRNA con un aminoácido no natural, por ejemplo, al menos al 40 %, al menos al 50 %, al menos al 60 %, al menos al 75 %, al menos al 80 % al menos al 90 %, al menos al 95 % o incluso al 99 % o más, tan eficazmente como una O – RS con una secuencia aminoácida como, por ejemplo, la descrita en las SECs ID Nº 86 o 45 (u otra RS específica en la Tabla 5) aminoacila un O – tRNA. En otra realización, una O – RS de la divulgación aminoacila el O – tRNA con el aminoácido no natural al menos 10 veces, al menos 20 veces, al menos 30 veces, etc. más eficazmente que la O – RS aminoacila el O – tRNA con un aminoácido no natural.

[0062] Empleando los componentes de traducción de la divulgación, el rendimiento del polipéptido de interés que comprende el aminoácido no natural es, por ejemplo, al menos 5 %, al menos 10 %, al menos 25 %, al menos 30 %, al menos 40 %, 50 % o más, del obtenido con el polipéptido natural de interés procedente de una célula en la que el polinucleótido carece de codón selectivo. En otro aspecto, la célula produce el polipéptido de interés en ausencia del aminoácido no natural, con un rendimiento que es, por ejemplo, menor del 30 %, menor del 20 %, menor del 15 %, menor del 10 %, menor del 5 %, menor del 2,5 %, etc. del rendimiento del polipéptido en presencia del aminoácido no natural.

Organismos fuente y huésped

[0063] Los componentes ortogonales de traducción de la divulgación derivan, normalmente, de organismos no vertebrados para su uso en células de vertebrado o sistemas de traducción. Por ejemplo, el O – tRNA ortogonal puede derivarse de un organismo no vertebrado, por ejemplo de Eubacterias como *Escherichia coli, Thermus thermophilus, Bacillus stearothermophilus* o similares, o de Arqueobacterias, como *Methanococcus jannaschii, Methanobacterium thermoautotrophicum, Halobacterium* como las especies *NRC - I Haloferax volcanii y*

Halobacterium, Archaeoglobus fulgidus, Pyrococcus furiosus, Pyrococcus horikoshii, Aeuropyrum pernix o similares, mientras la O – RS ortogonal puede derivarse de un organismo no vertebrado, como una Eubacteria, como Escherichia coli, Thermus thermophilus, Bacillus stearothermophilus o similares, o de Arqueobacterias, como Methanococcus jannaschii, Methanobacterium thermoautotrophicum, Halobacterium como las especies NRC - I Haloferax volcanii y Halobacterium, Archaeoglobus fulgidus, Pyrococcus furiosus, Pyrococcus horikoshii, Aeuropyrum pernix o similares. De manera alternativa, también pueden emplearse las fuentes de vertebrados, por ejemplo plantas, algas, protistas, hongos, levaduras, animales (por ejemplo mamíferos, insectos, artrópodos, etc) o similares, por ejemplo cuando los componentes son ortogonales a una célula o a un sistema de traducción de interés, o cuando se modifican (por ejemplo, mutan) para ser ortogonales a la célula o al sistema de traducción.

10

5

[0064] Los componentes individuales de un par O – tRNA / O – RS pueden derivarse del mismo organismo o de organismos diferentes. En una realización, el par O – tRNA / O – RS procede del mismo organismo. Por ejemplo, el par O – tRNA / O – RS puede derivar de un par tirosil – ARNt sintetasa / ARNt_{CUA} de *E. coli.* Alternativamente, el O – tRNA y la O – RS del par O – tRNA / O – RS proceden, opcionalmente, de organismos diferentes.

15

[0065] El O – tRNA, la O – RS o el par O – tRNA / O – RS ortogonales pueden seleccionarse o cribarse y / o usarse en una célula de vertebrado para producir un polipéptido con un aminoácido no natural. Una célula de vertebrado, en la que la célula no es una célula madre embrionaria humana ni una célula humana *in vivo*, puede proceder de diversas fuentes, por ejemplo de cualquier animal vertebrado (por ejemplo un mamífero, un anfibio, aves, reptiles, peces, etc.) o similares. También son una característica de la divulgación las composiciones de células de vertebrado con componentes de traducción de la divulgación.

25

20

[0066] La divulgación también proporciona el cribado eficaz en una especie para el uso opcional en esa especie y / o una segunda especie (opcionalmente sin selección / cribado adicional). Por ejemplo, los componentes del par O – tRNA / O – RS se seleccionan o se criban en una especie, por ejemplo, una especie manipulada fácilmente (como una célula de levadura, etc.) y se introduce en una segunda especie de vertebrado, por ejemplo, una planta (por ejemplo plantas complejas como monocotiledóneas o dicotiledóneas), un alga, un protista, un hongo, una levadura, un animal (por ejemplo un mamífero, un insecto, un artrópodo, etc) o similares, para su uso en la incorporación *in vivo* de un aminoácido no natural en la segunda especie.

30

35

40

[0067] Por ejemplo, Saccharomyces cerevisiae (S. cerevisiae) puede elegirse como primera especie de vertebrado, ya que es unicelular, tiene un tiempo de generación rápido y genética relativamente bien caracterizada. Véase, por ejemplo, D. Burke, et al., (2000) Methods in yeast genetics, Cold Spring Harbor Laboratiry Press, Cold Spring Harbor, NY. Además, ya que la maquinaria de traducción de eucariotas se conserva altamente (véanse, por ejemplo, (1996) Translational Control. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; Y. Kwok, & J.T. Wong, (1980), Evolutionary relationship between Halobacterium cutirubrum and eukaryotes determined by use of aminoacyl-tRNA synthetases as phylogenetic probes, Canadian Journal of Biochemistry 58: 213 - 218; y, (2001) The Ribosome. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), los genes aaRSs para la incorporación de un aminoácido no natural descubiertos en S. cerevisiae pueden introducirse en organismos vertebrados superiores y usarse, junto con ARNts semejantes (véase, por ejemplo, K. Sakamoto, et al., (2002) Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells, Nucleic Acids Res. 30: 4692 - 4699; y, C. Kohrer, et al., (2001), Import of amber and ochre suppressor tRNA's into mammalian cells: a general approach to site-specific insertion of amino acid analogues into proteins, Proc. Natl. Acad. Sci. U. S. A. 98: 14310 – 14315) para incorporar aminoácidos no naturales.

45

50

55

60

65

[0068] En un ejemplo, el método para producir O – tRNA / O – RS en una primera especie como la descrita en la presente incluye la introducción de un ácido nucleico que codifica el O – tRNA y un ácido nucleico que codifica la O – RS en una célula de vertebrado de una segunda especia (por ejemplo un mamífero, un insecto, un hongo, un algo, una planta y similares). En otro ejemplo, un método para producir una aminoacil – ARNt sintetasa ortogonal (O – RS) que aminoacila, preferentemente, un ARNt ortogonal con un aminoácido no natural en una célula de vertebrado incluye: (a) someter a selección positiva, en presencia de un aminoácido no natural, a una población de células de vertebrado de una primera especia (por ejemplo, levadura y similares). Cada célula de vertebrado comprende: i) un miembro de una biblioteca de aminoacil - ARNt sintetasas (RSs), ii) un ARNt ortogonal (O - tRNA), iii) un polinucleótido que codifica un marcador de selección positiva, y iv) un polinucleótido que codifica un marcador de selección negativa. Las células que sobrevivan a la selección positiva comprenden una RS activa que aminoacila al ARNt ortogonal (O - tRNA) en presencia de un aminoácido no natural. Las células que sobrevivan a la selección positiva se someten a selección negativa en ausencia del aminoácido no natural para eliminar las RSs activas que aminoacilan al O – tRNA con un aminoácido natural. Esto presenta una O – RS que aminoacila, preferentemente, el O - tRNA con el aminoácido no natural. Se introducen un ácido nucleico que codifica el O - tRNA y un ácido nucleico que codifica la O - RS (o los componentes O - tRNA y / u O - RS) en una célula de vertebrado de una segunda especie, por ejemplo un mamífero, un insecto, un hongo, un alga, una planta y / o similares. Normalmente, el O - tRNA se obtiene sometiendo a selección negativa a una población de células de vertebrado de una primera especie, en la que las células de vertebrado comprenden un miembro de una biblioteca de ARNts. La selección negativa elimina células que comprenden un miembro de la biblioteca de ARNts que se aminoacilan por una aminoacil - ARNt sintetasa (O - RS) endógena a las células de vertebrado, que proporcionan un grupo de ARNts ortogonales a la célula de vertebrado de la primera y segunda especie.

Codones selectivos

[0069] Los codones selectivos de la divulgación expanden la estructura genética del codón de la maquinaria biosintética de la proteína. Por ejemplo, un codón selectivo incluye, por ejemplo, un único codón de tres bases, un codón antisentido, como un codón de terminación, por ejemplo un codón ámbar (UAG), un codón ópalo (UGA), un codón no natural, al menos un codón de cuatro bases, un codón raro o similares. Pueden introducirse numerosos codones selectivos en un gen deseado, por ejemplo, uno o más, dos o más, más de tres, etc. Un gen puede incluir múltiples copias de un codón selectivo dado, o puede incluir múltiples codones selectivos diferentes, o cualquier combinación de los mismos.

10

15

20

25

60

65

5

[0070] En una realización, los métodos implican el uso de un codón selectivo que es un codón de terminación para la incorporación de aminoácidos no naturales *in vivo* en una célula de vertebrado. Por ejemplo, se produce un O – tRNA que reconoce el codón de terminación, por ejemplo UAG, y se aminoacila mediante una O – RS con un aminoácido no natural deseado. Este O – tRNA no es reconocido por las aminoacil – ARNt sintetasas naturales del huésped. La mutagénesis digirida al sitio convencional puede utilizarse para introducir el codón de terminación, por ejemplo, TAG, en el sitio de interés en un polipéptido de interés. *Véase, por ejemplo*, Sayers, J.R., et al. (1988), 5',3' Exonuclease in phosphorothioate - based oligonucleotide - directed mutagenesis. Nucleic Acids Res, 791 – 802. Cuando la O – RS, el O – tRNA y el ácido nucleico que codifica el polipéptido de interés se combinan *in vivo*, el aminoácido no natural se incorpora en respuesta al codón UAG para dar un polipéptido con el aminoácido no natural en la posición especificada.

[0071] La incorporación de aminoácidos no naturales *in vivo* puede llevarse a cabo sin perturbación significante de la célula huésped de vertebrado. Por ejemplo, ya que la eficacia de supresión del codón UAG depende de la competición entre el O – tRNA, por ejemplo, el ARNt supresor ámbar, y un factor de liberación de vertebrado (por ejemplo, eRF) (que se une a un codón de terminación e inicia la liberación del péptido cultivado desde el ribosoma), la eficacia de supresión puede modularse mediante, por ejemplo, el incremento del nivel de expresión de O – tRNA, por ejemplo, el ARNt supresor.

[0072] Los codones selectivos también comprenden codones extendidos, por ejemplo, codones de cuatro o más 30 bases, como codones de cuatro, cinco, seis o más bases. Algunos ejemplos de codones de cuatros bases incluyen, por ejemplo, AGGA, CUAG, UAGA, CCCU y similares. Algunos ejemplos de codones de cinco bases incluyen, por ejemplo, AGGAC, CCCCU, CCCUC, CUAGA, CUACU, UAGGC y similares. Una característica de la invención incluye el uso de codones extendidos basados en la supresión del marco de lectura. Los codones de cuatro o más bases pueden insertar, por ejemplo, uno o más aminoácidos no naturales en la misma proteína. Por ejemplo, en 35 presencia de O - tRNAs mutados, un ARNt supresor del marco de lectura especial, con lazos anticodones, por ejemplo con lazos anticodones de, al menos, de 8 a 10 nt, el codón de cuatro o más bases se lee como un aminoácido único. En otras realizaciones, los lazos anticodón pueden descodificar, por ejemplo, un codón de, al menos, cuatro bases, un codón de, al menos, cinco bases o un codón de, al menos, seis bases o más. Ya que hay 256 posibles codones de cuatro bases, pueden codificarse múltiples aminoácidos no naturales en la misma célula 40 utilizando un codón de cuatro o más bases. Véanse, Anderson et al., (2002) Exploring the Limits of Codon and Anticodon Size, Chemistry and Biology, 9: 237 - 244; Magliery, (2001) Expanding the Genetic Code: Selection of Efficient Suppressors of Four - base Codons and Identification of "Shifty" Four - base Cordons with a Library Approach in Escherichia coli, J. Mol. Biol. 307: 755 - 769.

[0073] Por ejemplo, los codones de cuatro bases se han utilizado para incorporar aminoácidos no naturales en proteínas utilizando métodos biosintéticos *in vitro*. *Véanse, por ejemplo*, Ma et al., (1993) Biochemistry, 32: 7939; y Hohsaka et al., (1999) J. Am. Chem. Soc., 121:34. Se utilizaron CGGG y AGGU para incorporar simultáneamente 2 – naftilalanina y un NBD derivado de lisina en estreptavidina *in vitro* con dos ARNts supresores del marco de lectura químicamente acilados. *Véase, por ejemplo*, Hohsaka et al., (1999) J. Am. Chem. Soc., 121: 12194. En un estudio *in vivo*, Moore et al., examinaron la capacidad de derivados de ARNtLeu con anticodones NCUA para eliminar codones UCUA (N puede ser U, A, G o C), y encontraron que el cuádruplo UAGA puede descodificarse mediante un ARNtLeu con un anticodón UAGN con una eficacia del 13 al 26 % con una pequeña descodificación en el marco 0 o – 1. *Véase* Moore et al., (2000) J. Mol. Biol., 298: 195. En una realización, pueden utilizarse codones extendidos basados en codones raros o codones antisentido, que pueden reducir la translectura sin sentido y la supresión del marco de lectura en otros sitios no deseados.

[0074] Para un sistema dado, un codón selectivo también puede incluir uno de los codones de tres bases naturales, en los que el sistema endógeno no emplea (o lo hace rara vez) el codón de base natural. Por ejemplo, esto incluye un sistema que carece de un ARNt que reconozca el codón de tres bases natural, y / o un sistema en el que el codón de tres bases es un codón raro.

[0075] Los codones selectivos incluyen, opcionalmente, pares de bases no naturales. Estos pares de bases no naturales expanden el alfabeto genético existente. Un par de bases extra incrementa el número de codones triples de 64 a 125. Las propiedades de los pares de la tercera base incluyen el apareamiento de bases estable y selectivo, la incorporación enzimática eficaz en ADN con alta fidelidad mediante una polimerasa, y la extensión continuada eficaz tras la síntesis del par de bases no natural emergente. Las descripciones de pares de bases no naturales que

pueden adaptarse para métodos y composiciones incluyen, por ejemplo Hirao, et al., (2002) *An unnatural base pair for incorporating amino acid analogues into protein*, Nature Biotechnology, 20: 177 – 182. Se enumeran a continuación otras publicaciones relevantes.

[0076] Para su uso in vivo, el nucleósido no natural es una membrana permeable y está fosforilada para formar el correspondiente trifosfato. Además, la información genética incrementada es estable y no se destruye por enzimas celulares. Los intentos anteriores de Benner y otros se aprovecharon de los patrones de adhesión a hidrógeno que son diferentes de los de los pares canónicos Watson - Crick, el ejemplo más notable es el par iso - C: iso - G. Véanse, por ejemplo, Switzer et al., (1989) J. Am. Chem. Soc., 111: 8322; y Piccirilli et al., (1990) Nature, 343: 33; Kool, (2000) Curr. Opin. Chem. Biol., 4: 602. Estas bases, en general, se malaparean en cierto grado con las bases naturales y no pueden replicarse enzimáticamente. Kool y sus compañeros demostraron que el material hidrofóbico que interacciona entre bases puede reemplazar la adhesión de hidrógeno para conducir a la formación de pares de bases. Véanse Kool, (2000) Curr. Opin. Chem. Biol., 4: 602; y Guckian and Kool, (1998) Angew. Chem. Int. Ed. Engl., 36, 2825. En un intento por desarrollar un par de bases no naturales que satisficiera todas las necesidades anteriores, Schultz, Romesberg y sus compañeros sintetizaron y estudiaron sistemáticamente una serie de bases hidrofóbicas no naturales. Se descubrió que un par PICS : PICS es más estable que los pares de bases naturales, y que pueden incorporarse eficazmente en ADN mediante un fragmento Klenow de la ADN polimerasa I de Escherichia coli (KF). Véanse, por ejemplo, McMinn et al., (1999) J. Am. Chem. Soc., 121: 11586; y Ogawa et al., (2000) J. Am. Chem. Soc., 122: 3274. Un par 3MN: 3MN puede sintetizarse mediante KF con eficacia y selectividad suficientes para la función biológica. Véase, por ejemplo, Ogawa et al., (2000) J. Am. Chem. Soc., 122: 8803. Sin embargo, ambas bases actúan como un terminador de cadena para otras replicaciones. Se ha desarrollado recientemente una ADN polimerasa mutante que puede emplearse para replicar el par PICS. Además, un par 7AI puede replicarse. Véase, por ejemplo, Tae et al., (2001) J. Am. Chem. Soc., 123: 7439. También se ha desarrollado un nuevo par metalobase, Dipic: Py, que forma un par estable en cuanto a su unión Cu (II). Véase, Meggers et al., (2000) J. Am. Chem. Soc., 122: 10714. Debido a que los codones extendidos y los codones no naturales son intrínsecamente ortogonales a los codones naturales, los métodos de la invención pueden aprovecharse de esta propiedad para generar ARNts ortogonales para ellos.

[0077] También puede emplearse un sistema de derivación de traducción para incorporar un aminoácido no natural en un polipéptido deseado. En un sistema de derivación de traducción, se inserta una gran secuencia en un gen, pero no se traduce en la proteína. La secuencia contiene una estructura que funciona como señal para inducir al ribosoma a recorrer la secuencia y continuar la traducción aguas abajo de la inserción.

Aminoácidos no naturales

5

10

15

20

25

30

35

40

45

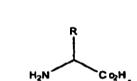
50

55

60

65

[0078] Tal y como se utiliza en la presente, un aminoácido no natural se refiere a cualquier aminoácido, aminoácido modificado o análogo de aminoácido diferente a selenocisteína y / o pirrolisina y a los siguientes aminoácidos alfa genéticamente codificados: alanina, arginina, aspargina, ácido aspártico, cisteína, glutamina, ácido glutámico, glicina, histidina, isoleucina, leucina, lisina, metionina, fenilalanina, prolina, serina, treonina, triptófano, tirosina, valina. La estructura genética de un aminoácido alfa se ilustra mediante la Fórmula I:



[0079] Un aminoácido no natural es, normalmente, cualquier estructura que tenga la Fórmula I en la que el grupo R es cualquier sustituyente diferente del utilizado en los veinte aminoácidos esenciales. Véase, por ejemplo, Biochemistry by L. Stryer, 3ª ed. 1988, Freeman and Company, New York, para las estructuras de los veinte aminoácidos esenciales. Cabe destacar que los aminoácidos no naturales de la invención pueden ser compuestos naturales diferentes de los veinte aminoácidos alfa mencionados anteriormente.

[0080] Debido a que los aminoácidos no naturales de la divulgación difieren, normalmente, de los aminoácidos naturales de la cadena lateral, los aminoácidos no naturales forman enlaces amida con otros aminoácidos, por ejemplo, naturales o no, de la misma manera en la que se forman en las proteínas naturales. Sin embargo, los aminoácidos no naturales tienen grupos de cadena lateral que los distinguen de los aminoácidos naturales. Por ejemplo, R en la Fórmula I comprende, opcionalmente, un alquil -, aril -, acil -, ceto -, azida -, hidroxil -, hidrazina -, ciano -, halo -, hidracida -, alquenil -, alquinil -, éter -, tiol -, seleno -, sulfonil -, borato, boronato, fosfo, fosfono, fosfina, hetercíclico, enona, imina, aldehído, éster, tioácido, hidroxialamina, amina y similares, o cualquier combinación de los mismos. Otros aminoácidos no naturales de interés incluyen, de manera no limitante, aminoácidos que comprenden un reticulador fotoactivable; aminoácidos marcados en spin; aminoácidos fluorescentes; aminoácidos de unión a metales; aminoácidos con metal; aminoácidos radioactivos; fotocargados y / o fotoisomerizables; aminoácidos con biotina o análogos de biotina; aminoácidos químicamente escindibles o fotoescindibles; aminoácidos con una cadena lateral alargada en comparación a los aminoácidos naturales (por

ejemplo, poliéteres o hidrocarbonos de cadena larga, por ejemplo de más de 5, de más de 10 carbonos, etc.); aminoácidos con azúcar con enlaces carbono; aminoácidos redox activos; aminoácidos con aminoácidos y aminoácidos con uno o más grupos tóxicos. En algunas realizaciones, los aminoácidos no naturales tienen un reticulador fotoactivable que se utiliza, por ejemplo, para unir una proteína a un soporte sólido. En una realización, los aminoácidos no naturales tienen un grupo sacárido unido a la cadena lateral aminoácida (por ejemplo, aminoácidas glicosilados) y / u otra modificación de carbohidrato.

[0081] Además de los aminoácidos no naturales que contienen cadenas laterales nuevas, los aminoácidos no naturales comprenden, opcionalmente, estructuras de segmento principal modificadas, por ejemplo, como la ilustrada por las estructuras de Fórmula II y III:

5

10

30

35

40

45

50

55

60

65

25 H₂N C o₂i

[0082] en las que Z comprende, normalmente, OH, NH₂, SH, NH - R' o S - R'; X e Y, que pueden ser iguales o diferentes, incluyen, normalmente, S u O, y R y R', que pueden ser iguales o diferentes, se seleccionan normalmente de la misma lista de constituyentes para el grupo R descrito anteriormente en los aminoácidos no naturales con Fórmula I, así como hidrógeno. Por ejemplo, los aminoácidos no naturales de la invención comprenden, opcionalmente, sustituciones en el grupo amino o carboxilo, como se ilustra en las Fórmulas II y III. Los aminoácidos no naturales de este tipo incluyen, de manera no limitante, aminoácidos α - hidroxi, α - tioácidos α aminotiocarboxilados, por ejemplo, con las cadenas laterales correspondientes a los veinte aminoácidos naturales esenciales o las cadenas laterales no naturales. Además, las sustituciones en el carbono α incluyen, opcionalmente, aminoácidos L, D o α - α - disustituidos, como D - glutamato, D - alanina, D - metil - O - tirosina, ácido aminobutírico y similares. Otras alternativas estructurales incluyen aminoácidos cíclicos, como análogos de prolina, así como análogos de prolina con un anillo de 3, 4, 6, 7, 8 y 9 miembros, aminoácidos β y y como β – alanina y ácido y - amino butírico. Por ejemplo, muchos aminoácidos no naturales se basan en aminoácidos naturales, como tirosina, glutamina, fenilalanina y similares. Los análogos de tirosina incluyen tirosinas para - sustituidas, tirosinas orto - sustituidas y tirosinas meta sustituidas, en los que la tirosina sustituida comprende, por ejemplo, un grupo ceto (por ejemplo un grupo acetil), un grupo benzoilo, un grupo amino, una hidrazina, una hidroxiamina, un grupo tiol, un grupo carboxi, un grupo isopropilo, un grupo metilo, un hidrocarbono C₆ - C₂₀ de cadena lineal o ramificada, unn hidrocarbono saturado o insaturado, un grupo O - metilo, un grupo poliéter, un grupo nitro, un grupo alquinilo o similares. Además, también se contemplan anillos arilo sustituidos. Los análogos de glutamina de la invención incluyen, de manera no limitante, derivados α – hidroxi, derivados sustituidos por γ, derivados cíclicos y derivados de glutamina sustituidos por amida. Algunos ejemplos de análogos de fenilalanina incluyen, de manera no limitante, fenilalaninas para - sustituidas, fenilalaninas orto - sustituidas y fenilalaninas meta - sustituidas, en las que los sustituyentes incluyen, por ejemplo, un grupo hidroxi, un grupo metoxi, un grupo metilo, un grupo alilo, un grupo aldehído, un grupo azido, un grupo yodo, un grupo bromo, un grupo ceto (por ejemplo, un grupo acetilo), un grupo benzoilo, un grupo alquinilo o similares. Alqunos ejemplos específicos de aminoácidos no naturales incluyen, de manera no limitante, una p – acetil – L - fenilalanina, una p – propargiloxifenilalanina, O – metil – L – tirosina, una L – 3 – (2 – naftil) alanina, una 3 – metil – fenilalanina, una O – 4 – alil – L – tirosina, una 4 – propil – L – tirosina, una tri - O - acetil - GlcNAcβ - serina, una L - Dopa, una fenilalanina fluorinada, una isopropil - L - fenilalanina, una p azido - L - fenilalanina, p - acil - L - fenilalanina, p - benzoil - L - fenilalanina, una L - fosfoserina, una fosfonoserina, una fosfonotirosina, p – yodo – fenilalanina, p – bromofenilfenilalanina, p – amino – L – fenilalanina, p – azido – L – fenilalanina y una isopropil – L – fenilalanina y similares. Se describen otras estructuras de una variedad de aminoácidos no naturales en, por ejemplo, las Figuras 16, 17, 18, 19, 26 y 29 de la patente WO 2002 / 085923 titulada "In vivo incorporation of unnatural amino acids". Véanse también, las estructuras 2 – 5 de la Figura 1 de Kiick et al., (2002) Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligitation, PNAS 99: 19-24, para análogos de metionina adicionales.

[0083] Se describen en la presente composiciones que incluyen un aminoácido no natural (como p – (propargiloxi) – fenilalanina). También se presentan varias composiciones que comprenden p – (propargiloxi) – fenilalanina y, por ejemplo, proteínas y / o células. En un aspecto, una composición que incluye el aminoácido no natural p – (propargiloxi) – fenilalanina también incluye un ARNt ortogonal. El aminoácido no natural puede unirse (por ejemplo, covalentemente) al ARNt ortogonal, por ejemplo, unirse covalentemente al ARNt mediante un enlace amino – acilo, o

unirse covalentemente a un 3'OH o a un 2'OH de un azúcar ribosa terminal del ARNt ortogonal, etc.

[0084] Los grupos químicos que pueden incorporarse mediante un aminoácido no natural en proteínas ofrecen numerosas ventajas y manipulaciones de la proteína. Por ejemplo, la reactividad única de un grupo ceto funcional permite la modificación selectiva de proteína con cualquier número de reactivos con hidrazina — o hidroxilamina *in vivo* e *in vitro*. Un aminoácido de un átomo pesado, por ejemplo, puede ser práctico para los datos estructurales de rayos X por fases. La introducción específica en un sitio de átomos pesados mediante aminoácidos no naturales también proporciona selectividad y flexibilidad en cuanto a la elección de posiciones para los átomos pesados. Los aminoácidos no naturales fotoreactivos (por ejemplo aminoácidos con cadenas laterales benzofenona y arilazidos (por ejemplo fenilazido)), por ejemplo, permiten la fotoreticulación de proteínas *in vivo* e *in vitro*. Algunos ejemplos de aminoácidos no naturales fotoreactivos incluyen, de manera no limitante, p — azido — fenilalanina y p — benzoil — fenilalanina. La proteína con los aminoácidos no naturales fotoreactivos pueden reticularse mediante excitación del grupo fotoreactivo que proporciona control temporal (y / o espacial). En un ejemplo, el grupo metilo de un aminoácido no natural puede sustituirse con, por ejemplo, un grupo metilo isotópicamente marcado, como una sonda de estructura local y dinámica, por ejemplo, con el uso de resonancia magnética nuclear y espectroscopia vibracional. Los grupos funcionales alquinilo o azido, por ejemplo, permiten la modificación selectiva de proteínas con moléculas mediante una reacción de cicloadición [3 + 2].

Síntesis química de aminoácidos no naturales

5

10

15

20

25

30

35

40

45

50

55

60

65

[0085] Muchos de los aminoácidos presentados anteriormente se encuentran disponibles comercialmente, por ejemplo, en Sigma (EEUU) o Aldrich (Milwaukee, WI, EEUU). Los que no se encuentran disponibles comercialmente se sintetizan, opcionalmente, como se describe en la presente o como se ha descrito en varias publicaciones o empleando métodos estándares conocidos por los expertos en la disciplina. Para las técnicas de síntesis orgánica, véanse, por ejemplo, Organic Chemistry by Fessendon and Fessendon, (1982, Second Edition, Willard Grant Press, Boston Mass.); Advanced Organic Chemistry by March (Third Edition, 1985, Wiley and Sons, New York); y Advanced Organic Chemistry by Carey and Sundberg (Third Edition, Parts A and B, 1990, Plenum Press, New York). Otras publicaciones que describen la síntesis de aminoácidos no naturales incluyen, por ejemplo, la patente WO 2002 / 085923 titulada "In vivo incorporation of Unnatural Amino Acids;" Matsoukas et al., (1995) J. Med. Chem., 38, 4660 -4669; King, F.E. & Kidd, D.A.A. (1949) A New Synthesis of Glutamine and of γ - Dipeptides of Glutamic Acid from Phthylated Intermediated. J. Chem. Soc., 3315 - 3319; Friedman, O.M. & Chatterrji, R. (1959) Synthesis of Derivatives of Glutamine as Model Substrates for Anti - Tumor Agents. J. Am. Chem. Soc. 81, 3750 - 3752; Craig, J.C. et al. (1988) Absolute Configuration of the Enantiomers of 7 - Chloro - 4[[4 - (diethylamino) - 1 - methylbutyl] amino] quinoline (Chloroquine). J. Org. Chem. 53, 1167 - 1170; Azoulay, M., Vilmont, M. & Frappier, F. (1991) Glutamine analogues as Potential Antimalarials, Eur. J. Med. Chem. 26, 201 - 5; Koskinen, A.M.P. & Rapoport, H. (1989) Synthesis of 4 - Substituted Prolines as Conformationally Constrained Amino Acid Analogues. J. Org. Chem. 54, 1859-1866; Christie, B.D. & Rapoport, H. (1985) Synthesis of Optically Pure Pipecolates from L - Asparagine, Application to the Total Synthesis of (+) - Apovincamine through Amino Acid Decarbonylation and Iminium Ion Cyclization. J. Org. Chem. 1989: 1859 - 1866; Barton et al., (1987) Synthesis of Novel a - Amino - Acids and Derivatives Using Radical Chemistry: Synthesis of L - and D - a - Amino - Adipic Acids, L - a - aminopimelic Acid and Appropriate Unsaturated Derivatives. Tetrahedron Lett. 43: 4297 - 4308; y, Subasinghe et al., (1992) Quisqualic acid analogues: synthesis of beta - heterocyclic 2 - aminopropanoic acid derivatives and their activity at a novel quisqualate - sensitized site. J. Med. Chem. 35: 4602 - 7.

Absorción celular de aminoácidos no naturales

[0086] Un aminoácido no natural absorbido por una célula de vertebrado es un tema considerado normalmente cuando se diseñan y seleccionan aminoácidos no naturales, por ejemplo, para su incorporación en una proteína. Por ejemplo, la densidad de alta carga de aminoácidos α sugiere que estos compuestos no son probablemente permeables a células. Los aminoácidos no naturales se introducen en la célula de vertebrado mediante una colección de sistemas de transporte basados en proteínas. Puede realizarse un control rápido que evalúe qué aminoácido, si lo hay, está siendo tomado por las células. *Véanse, por ejemplo*, los ensayos de toxicidad de, por ejemplo, la solicitud de patente titulada "Protein Arrays", número de expediente P1001US00 publicado el 22 de diciembre del 2002; y Liu, D.R. & Schultz, P.G. (1999) Progress toward the evolution of an organism with an expanded genetic code. PNAS United States 96: 4780 – 4785. Aunque la absorción se analiza fácilmente con varios ensayos, una alternativa para diseñar aminoácidos no naturales susceptibles a las vías de absorción celular es proporcionar vías biosintéticas para crear aminoácidos *in vivo*.

Biosíntesis de aminoácidos no naturales

[0087] Existen numerosas vías biosintéticas en células para la producción de aminoácidos y otros compuestos. Ya que un método biosintético para un aminoácido no natural en particular puede no existir en la naturaleza, por ejemplo, en una célula de vertebrado, la invención proporciona dichos métodos. Por ejemplo, las vías biosintéticas para aminoácidos no naturales se generan, opcionalmente, en células huésped añadiendo enzimas nuevas o modificando las vías de la célula huésped existentes. Puede haber otras enzimas nuevas naturales o artificialmente involucradas. Por ejemplo, la biosíntesis de *p* – aminofenilalanina (como se presenta en un ejemplo de la patente

WO 2002 / 085923 titulada ("In vivo incorporation of unnatural amino acids") relaciona la adición de una combinación de enzimas conocidas de otros organismos. Los genes para estas enzimas pueden introducirse en una célula de vertebrado transformando la célula con un plásmido que comprende los genes. Los genes, cuando se expresan en la célula, proporcionan una vía enzimática para sintetizar el compuesto deseado. Algunos ejemplos de los tipos de enzimas que se añaden opcionalmente se presentan en los ejemplos descritos a continuación. Otras secuencias de enzimas se encuentran, por ejemplo, en Genbank. También pueden añadirse, opcionalmente, enzimas involucradas artificialmente en una célula de la misma manera. De este modo, la maquinaria y los recursos celulares de una célula se manipulan para producir aminoácidos no naturales.

10 [0088] Se encuentran disponibles numerosos métodos para producir enzimas nuevas para su uso en vías biosintéticas o para la evolución de vías va existentes. Por ejemplo, la recombinación recursiva, por ejemplo, como la desarrollada por Maxygen, INC (disponible en la página web www.maxygen.com) se emplea, opcionalmente, para desarrollar nuevas enzimas y vías. Véanse, por ejemplo, Stemmer (1994), Rapid evolution of a protein in vitro by DNA shuffling, Nature 370(4): 389 - 391; y, Stemmer, (1994), DNA shuffling by random fragmentation and 15 reassembly: In vitro recombination for molecular evolution, Proc. Natl. Acad. Sci. USA., 91:10747 - 10751. Igualmente, DesignPathTM, desarrollada por Genencor (disponible en la página web genencor.com) se emplea, opcionalmente, para modificar vías metabólicas, por ejemplo, para modificar una vía para crear O - metil - L tirosina en una célula. Esta tecnología reconstruye vías existentes en organismos huésped utilizando una combinación de nuevos genes, por ejemplo, identificados a través de genómica funcional, y diseño y evolución 20 molecular. Diversa Corporation (disponible en la página web diversa.com) también proporciona tecnología para el cribado rápido de bibliotecas genéticas y vías génicas, por ejemplo, para crear nuevas vías.

[0089] Normalmente, el aminoácido no natural producido con una vía biosintética modificada de la invención se produce en una concentración suficiente para la biosíntesis eficaz de proteínas, por ejemplo, una cantidad celular natural, pero no a un nivel que afecte la concentración de otros aminoácidos o que agote recursos celulares. Las concentraciones normales producidas in vivo de esta manera son de 10 mM a 0,05 mM. Una vez que la célula ha sido transformada con un plásmido que comprende los genes empleados para producir las enzimas deseadas para una vía específica y se genera un aminoácido deseado, se emplean, opcionalmente, selecciones in vivo para optimizar la producción del aminoácido no natural tanto para la síntesis de proteína ribosómica como para el crecimiento celular.

Polipéptidos con aminoácidos no naturales

[0090] Se describen en el presente documento proteínas o polipéptidos de interés con al menos un aminoácido. La divulgación incluye asimismo polipéptidos o proteínas con, al menos, un aminoácido no natural producido utilizando las composiciones y procedimientos de la invención. Un excipiente (por ejemplo, un excipiente farmacéuticamente aceptable) puede además estar presente en la proteína.

[0091] Mediante la producción de proteínas o polipéptidos de interés con, al menos, un aminoácido no natural en células de vertebrados, las proteínas o los polipéptidos incluirán modificaciones post - traduccionales de vertebrados. Una proteína puede incluir, al menos, un aminoácido no natural y, al menos, una modificación post traduccional realizada in vivo por una célula de vertebrado, en el que la modificación post - traduccional no es realizada por una célula procariota. Por ejemplo, la modificación post - traduccional incluye, por ejemplo, acetilación, acilación, modificación de lípidos, palmitoilación, adición de palmitato, fosforilación, modificación del enlace de glucolípidos, glicosilación y similares. En un aspecto, la modificación post - traduccional incluye el acoplamiento de un oligosacáridos (por ejemplo, GlcNAc - Man)₂ - Man - GlcNAc - GlcNAc)) en una asparagina por un enlace de GIcNAc – asparagina. Véase también, la Tabla 7, que enumera algunos ejemplos de N - oligosacáridos de proteínas de vertebrados (pueden estar presentes residuos adicionales, aunque no se muestran). En otro aspecto, la modificación post - traduccional incluye el acoplamiento de un oligosacárido (por ejemplo, Gal - GalNAc, Gal -GlcNAc, etc) a una serina o treonina por un enlace de GalNAc - serina o GalNAc - treonina, o un enlace de GlcNAc serina - o GlcNAc - treonina.

TABLA 7: EJEMPLOS DE OLIGOSACÁRIDOS A TRAVÉS DEL ENLACE GICNAC

55 Tipo Estructura base Mana1-6 Manosa elevada Manα1-6. Mana1-3 Manß1-4GlcNAcß1-4GlcNAcß1-Asn <u> Manα1-3</u> Híbrido Manα1-6 Man 61-4GICNAC 61-4GICNAC 61-Asn GICNACB1-2 Manα1-3

60 65

5

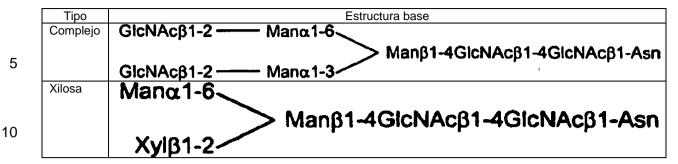
25

30

35

40

45



[0092] En otro aspecto, la modificación post – traduccional incluye el procesamiento proteolítico de precursores (por ejemplo, precursor de la calcitonina, precursor del péptido relacionado con el gen de la calcitonina, hormona preproparatiroidea, preproinsulina, proinsulina, preproopiomelanocortina, proopiomelanocortina y similares), el ensamble en una proteína de múltiples subunidades o el ensamble de macromoléculas, la traducción a otro sitio en la células (por ejemplo, en orgánulos, como el retículo endoplásmico, el aparato de Golgi, el núcleo, liposomas, peroxisomas, mitocondrias, cloroplastos, vacuolas, etc, o a través de la vía secretora). En algunas realizaciones, la proteína comprende una secuencia de secreción o localización, un marcador de epítopo, un marcador FLAG, un marcador de polihistidina, una fusión de GST o similares.

15

20

25

30

35

40

45

50

55

60

65

[0093] Una ventaja del aminoácido no natural es que presenta restos químicos adicionales que pueden ser utilizados para añadir moléculas adicionales. Estas modificaciones pueden realizarse in vivo en una célula de vertebrado, o in vitro. De este modo, en algunas realizaciones, la modificación post - traduccional se realiza a través del aminoácido no natural. Por ejemplo, la modificación post - traduccional puede realizarse a través de una reacción nucleófilica - electrofílica. La mayoría de las reacciones utilizadas actualmente para la modificación selectiva de proteínas incluyen la formación de enlaces covalentes entre pares de reacción nucleofílica y electrofílica, por ejemplo, la reacción de α - halocetonas con cadenas laterales de histidina o cisteína. La selectividad en estos casos se determina por el número y accesibilidad de los residuos nucleofílicos en la proteína. En las proteínas de la invención, pueden utilizarse otras reacciones más selectivas, como la reacción de un cetoaminoácido no natural con compuestos hidrazida o aminooxi, in vitro e in vivo. Véanse, por ejemplo, Cornish, et al., (1996) Am. Chem. Soc., 118: 8150 - 8151; Mahal, et al., (1997) Science, 276: 1125 -1128; Wang, et al., (2001) Science 292: 498 - 500; Chin, et al., (2002) Am. Chem. Soc. 124: 9026 - 9027; Chin, et al., (2002) Proc. Natl. Acad. Sci., 99: 11020 - 11024; Wang, et al., (2003) Proc. Natl. Acad. Sci., 100: 56 - 61; Zhang, et al., (2003) Biochemistry, 42: 6735 - 6746; y, Chin, et al., (2003) Science, en prensa. Esto permite el marcaje selectivo de prácticamente cualquier proteína con numerosos reactivos que incluyen fluóroforos, agentes reticulantes, derivados de sacáridos y moléculas citotóxicas. Véase, también, la Solicitud de la Patente USSN 10/686.944 titulada "Glycoprotein synthesis" presentada el 15 de octubre de 2003. Las modificaciones post - traduccionales, por ejemplo, a través de un aminoácido azida, pueden realizarse también a través de la ligadura de Staudinger (por ejemplo, con reactivos de triarilfosfina). Véase, por ejemplo, Kiick et al., (2002) Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation, PNAS 99: 19 - 24.

[0094] Una célula de vertebrado de la divulgación proporciona la capacidad para la sintetización de proteínas que comprenden aminoácidos no naturales en grandes cantidades útiles. En un aspecto, la composición incluye opcionalmente, por ejemplo, al menos 10 microgramos, al menos 50 microgramos, al menos 75 microgramos, al menos 100 microgramos, al menos 200 microgramos, al menos 250 microgramos, al menos 500 microgramos, al menos 1 miligramo, al menos 10 miligramos o más de la proteína que comprende un aminoácido no natural, o una cantidad que puede alcanzarse con los procedimientos de producción in vivo de proteínas (se proporcionan en la presente, los detalles de la producción y purificación de proteínas recombinantes). En otro aspecto, la proteína está opcionalmente presente en la composición de una concentración, por ejemplo, al menos 10 microgramos de una proteína por litro, al menos 50 microgramos de una proteína por litro, al menos 75 microgramos de una proteína por litro, al menos 100 microgramos de una proteína por litro, al menos 200 microgramos de una proteína por litro, al menos 250 microgramos de una proteína por litro, al menos 500 microgramos de una proteína por litro, al menos 1 miligramos de una proteína por litro, al menos 10 miligramos de una proteína por litro o más, en, por ejemplo, un lisado celular, un tampón, un tampón farmacéutico, u otra suspensión líquida (por ejemplo, en un volumen de, por ejemplo, que oscila entre 1 nl a 100 L). Es una característica de la invención, la producción de grandes cantidades (por ejemplo, superiores a los normalmente posibles con otros procedimientos, por ejemplo, traducción in vitro) de una proteína en una célula de vertebrado incluye, al menos, un aminoácido no natural.

[0095] La incorporación de un aminoácido no natural puede realizarse para, por ejemplo, adaptar los cambios en la estructura de la proteína y / o de la función, por ejemplo, para cambiar el tamaño, la acidez, la nucleofilicidad, los enlaces de hidrógeno, la hidrofobicidad, la accesibilidad de los sitios diana de la proteasa, la diana de un resto (por ejemplo, para una matriz de proteína), etc. Las proteínas que incluyen aminoácidos no naturales pueden tener propiedades físicas o catalíticas mejoradas o incluso completamente nuevas. Por ejemplo, las siguientes propiedades están opcionalmente modificadas por la inclusión de un aminoácido no natural en una proteína: toxicidad, biodistribución, propiedades estructurales, propiedades espectroscópicas, propiedades químicas y / o

fotoquímicas, capacidad catalítica, vida media (por ejemplo, vida media del suero), capacidad de reacción con otras moléculas, por ejemplo, covalente o no covalente, y similares. Las composiciones que incluyen proteínas, incluyen al menos un aminoácido no natural útil para, por ejemplo, nuevos terapéuticos, diagnósticos, enzimas catalíticas, enzimas industriales, proteínas de unión (por ejemplo, anticuerpos) y por ejemplo, el estudio de la estructura y función de la proteína. Véase, por ejemplo, Dougherty, (2000) Unnatural Amino Acids as Probes of Protein Structure and Function, Current Opinion in Chemical Biology, 4: 645 - 652.

5

10

15

30

35

40

45

50

55

60

65

[0096] En un aspecto descrito en el presente documento, una proteína incluye al menos uno, por ejemplo, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve o al menos diez o más aminoácidos no naturales. Los aminoácidos no naturales pueden ser iguales o diferentes, por ejemplo, puede haber 1, 2, 3, 4, 5, 6, 7, 8, 9, o 10 o más sitios diferentes en la proteína que comprende 1, 2, 3, 4, 5, 6, 7, 8, 9, o 10 o más aminoácidos no naturales diferentes. En otro aspecto, se describe en la presente, una proteína que tiene al menos, aunque no todos, un aminoácido particular presente en la proteína que se sustituye con un aminoácido no natural. Para una proteína determinada con más de un aminoácido no natural, los aminoácidos no naturales pueden ser idénticos o diferentes (por ejemplo, la proteína puede incluir dos o más tipos diferentes de aminoácidos no naturales, o puede incluir dos aminoácidos no naturales iguales). Para una proteína determinada con más de dos aminoácidos no naturales, los aminoácidos pueden ser iguales, diferentes o una combinación múltiples aminoácidos no naturales del mismo tipo con al menos un aminoácido no natural diferente.

20 [0097] Esencialmente, cualquier proteína (o parte de la misma) que incluye un aminoácido no natural (y cualquier correspondiente ácido nucleico codificador, por ejemplo, que incluye uno o más codones selectivos) puede producirse utilizando las composiciones y procedimientos de la presente invención. No se intenta identificar los cientos de miles de proteínas conocidas, cualquiera de ellas pueden ser modificadas para incluir uno o más aminoácidos no naturales, por ejemplo, adaptando cualquier procedimiento de mutación disponible para incluir uno o más codones selectivos apropiados en un sistema de traducción pertinente. Los depósitos de secuencias comunes para las proteínas conocidas incluyen GenBank EMBL, DDBJ y NCBI. Otros depósitos pueden identificarse fácilmente realizando búsquedas en Internet.

[0098] Las proteínas son, por ejemplo, idénticas al menos un 60 %, al menos un 70 %, al menos un 75 %, al menos un 80 %, al menos un 90 %, al menos un 95 %, o al menos un 99 % o más para cualquier proteína disponible (por ejemplo, una proteína terapéutica, una proteína de diagnóstico, una enzima industrial, o parte de las mismas, y similares), y comprenden uno o más aminoácidos no naturales. Los ejemplos de proteínas terapéuticas, de diagnóstico y otras pueden modificarse para comprender uno o más aminoácidos no naturales que incluyen, de manera no limitante, por ejemplo, Alfa - 1 antitripsina, angiostatina, factor antihemolítico, anticuerpos (a continuación, se encuentran más detalles sobre anticuerpos), apolipoproteína, apoproteína, factor natriurético atrial, polipéptido natriurético atrial, péptidos atriales, quimiocinas C -X - C (por ejemplo, T39765, NAP- 2, ENA - 78, Gro a, Gro -b, Gro -c, IP- 10, GCP- 2, NAP- 4, SDF - 1, PF4, MIG), calcitonina, quimiocinas CC (por ejemplo, proteína quimioatrayente de monocitos - 1, proteína quimiotáctica de monocitos - 2, proteína quimioatrayente de monocitos -3, proteína - 1 alfa inflamatoria de monocitos, proteína - 1 beta inflamatoria de monocitos, RANTES, I309, R83915, R91733, HCC1, T58847, D31065, T64262), ligando CD40, ligando C -kit, colágeno, factor estimulante de colonias (CSF), factor del complemento 5a, inhibidor del complemento, receptor 1 del complemento, citocinas (por ejemplo, péptido epitelial activador de los neutrófilos 78, GROαMGSA, GROβ, GROγ, MIP - 1α, MIP - 1δ, MCP - I), factor de crecimiento epidérmico (EGF), eritropoyetina ("EPO", que representa una diana preferida para la modificación mediante la incorporación de uno o más aminoácidos no naturales), toxinas exfoliantes A y B, factor IX, factor VII, factor VIII, factor X, factor de crecimiento de fibroblastos (FGF), fibrinógeno, fibronectina, G-CSF, GM-CSF, glucocerebrosidasa, gonadotropina, factores de crecimiento, proteínas Hedgehog (por ejemplo, Sonic, Indian, Desert), hemoglobina, factor de crecimiento de hepatocitos (HGF), hirudina, albúmina de suero humano, insulina, factor de crecimiento de tipo insulina (IGF), interferones (por ejemplo, IFN $-\alpha$, IFN $-\beta$, IFN $-\gamma$), interleucinas (por ejemplo, IL - 1, IL - 2, IL - 3, IL - 4, IL - 5, IL - 6, IL - 7, IL - 8, IL - 9, IL - 10, IL - 11, IL - 12, etc), factor de crecimiento de queratinocitos (KGF), lactoferrina, factor inhibidor de leucemia, luciferasa, neurturina, factor inhibidor de neutrófilos (NIF), oncostatina M, proteína osteogénica, hormona paratiroidea, PD - ECSF, PDGF, hormonas peptídicas (por ejemplo, hormona del crecimiento humano), pleyotropina, proteína A, proteína G, exotoxinas pirogénica A, B, y C, relaxina, renina, SCF, receptor I soluble del complemento, soluble I - CAM 1, receptores solubles de interleucina (IL - 1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15), receptor soluble de TNF, somatomedina, somatostatina, somatotropina, estreptoquinasa, superantígenos, es decir; enterotoxinas estafilocócicas (SEA, SEB, SEC1, SEC2, SEC3, SED, VER), superóxido dismutasa (SOD), toxina del síndrome de shock tóxico (TSST - 1), timosina alfa 1, activador plasminógeno tisular, factor de necrosis tumoral beta (TNF beta), receptor del factor de necrosis tumoral (TNFR), factor de necrosis tumoral - alfa (TNF alfa), factor de crecimiento endotelial vascular (VEGEF), uroquinasa, y muchos otros.

[0099] Una clase de proteínas que pueden hacerse utilizando las composiciones y procedimientos para la incorporación *in vivo* de aminoácidos no naturales descritos en el presente documento, incluye moduladores transcripcionales o partes de los mismos. Un ejemplo de moduladores transcripcionales incluye genes y proteínas moduladoras de la transcripción que modula el crecimiento celular, la diferenciación, la regulación, o similares.

Los moduladores transcripcionales se encuentran en procariotas, virus y eucariotas, incluyendo hongos, plantas, levaduras, insectos y animales, incluyendo mamíferos, que proporcionan una amplia variedad de objetivos

terapéuticos. Se apreciará que los activadores de la expresión y la transcripción regulan la transcripción mediante muchos mecanismos, por ejemplo, mediante la unión a receptores estimulando una cascada de transducción de señales, regulando la expresión de factores de transcripción, uniéndose a promotores y potenciadores, uniéndose a proteínas que se vinculan a promotores y potenciadores, desenrollando ADN, splicing pre- ARNm, poliadenilando ARN y degradando ARN. Por ejemplo, las composiciones de la proteína GAL4 o parte de la misma en una célula de vertebrado es también una característica de la invención. La proteína GAL4 o parte de la misma comprende, al menos, un aminoácido no natural. Véase también la sección del presente documento titulado "aminoacil – ARNt sintetasa ortogonal".

5

30

35

- [0100] Una clase de proteinas descritas en el presente documento (por ejemplo, proteínas con uno o más aminoácidos no naturales) incluye activadores de expresión como citocinas, moléculas inflamatorias, factores de crecimiento, sus receptores, productos de oncogenes, por ejemplo, interleucinas (por ejemplo, IL 1, IL 2, IL 8, etc), interferones, FGF, IGF 1, IGF II, FGF, PDGF, TNF, TGF α, TGF β, EGF, KGF, SCF / c Kit, CD40L / CD40, VLA 4 / VCAM 1, ICAM 1 / LFA 1, hialurina / CD44; moléculas de transducción de señales y sus correspondientes productos de oncogenes, por ejemplo; Mos, Ras, Raf y Met; activadores transcripcionales y supresores, por ejemplo; p53, Tat, Fos, Myc, Jun, Myb, Rel, y receptores de hormonas esteroideas tales como el estrógeno, progesterona, testosterona, aldosterona, ligando del receptor LDL y corticosterona.
- [0101] Se proporcionan asimismo en la invención enzimas (por ejemplo, enzimas industriales), o partes de las mismas con, al menos, un aminoácido no natural. Ejemplos de enzimas incluyen, de manera no limitante, por ejemplo, amidasas, aminoácidos racemasas, acilasas, deshalogenasas, dioxigenasas, diarilpropano peroxidasa, epimerasas, epóxido hidrolasas, esterasas, isomerasas, quinasas, glucosa isomerasas, glucosidasas, glucosil transferasas, haloperoxidasas, monooxigenasas (por ejemplo, p450s), lipasas, lignina peroxidasa, nitrilo hidratasas, proteasas, fosfatasas, subtilisinas, transaminasas y nucleasas.
 - **[0102]** Algunas de estas proteínas se encuentran disponibles comercialmente (*véase*, por ejemplo, el catálogo de Sigma BioSciences 2002 y la lista de precios), sus correspondientes secuencias proteícas y génicas y algunas variantes de las mismas ya conocidas (*véase*, por ejemplo, Genbank). Pueden modificarse cualquiera de ellas mediante la inserción de uno o más aminoácidos no naturales según la presente invención, por ejemplo, para alterar la proteína con respecto a una o más propiedades terapéuticas, de diagnóstico o enzimáticas de interés. Ejemplos de propiedades terapéuticamente pertinentes incluyen vida media del suero, vida media útil, estabilidad, inmunogenicidad, actividad terapéutica, detectabilidad (por ejemplo, por la inclusión de grupos indicadores (por ejemplo, marcadores o sitios de unión de marcadores) en aminoácidos no naturales), reducción de LD₅₀ u otros efectos secundarios, capacidad para entrar en el cuerpo a través del tracto gástrico (por ejemplo, disponibilidad oral), o similares. Ejemplos de propiedades de diagnóstico incluyen vida media útil, estabilidad, actividad de diagnóstico, detectabilidad o similares. Ejemplos de propiedades enzimáticas pertinentes incluyen vida media útil, estabilidad, actividad de producción o similares.
- [0103] Puede modificarse otra variedad de proteínas para incluir uno o más aminoácidos no naturales. Por ejemplo, sustituir uno o más aminoácidos naturales en una o más proteínas de vacuna con un aminoácido no natural, por ejemplo; en proteínas fúngicas infecciosas, por ejemplo; Aspergillus, especies Candida; bacterias, particularmente como E. coli, que sirve de modelo para las bacterias patógenas, así como bacterias importantes como Staphylococci (por ejemplo, aureus), o Streptococci (por ejemplo, pneumoniae); protozoos como esperozoos (por ejemplo, Plasmodia), rizópodos (por ejemplo, Entamoeba) y flagelados (Trypanosoma, Leishmania, Trichomonas, Giardia, etc); virus como virus ARN (+), (los ejemplos incluyen poxvirus, por ejemplo, vaccinia; picornavirus, por ejemplo, polio; togavirus, por ejemplo; rubella; flavivirus, por ejemplo; VHC; y coronavirus), virus ARN (-) (por ejemplo, rabdovirus, por ejemplo; VSV; paramixovirosis, por ejemplo; RSV; ortomixovirus; por ejemplo, influenza; bunyavirus; arenavirus), virus de ADN bicatenario (reovirus, por ejemplo), virus ARN en ADN, es decir; retrovirus, por ejemplo, VIH y HTLV, y se describen en el presente documento algunos virus ADN en ARN como la Hepatitis B.
 - [0104] Las proteínas relacionadas con la agricultura como las proteínas de resistencia contra los insectos (por ejemplo, las proteínas Cry), almidón y las enzimas de producción de lípidos, toxinas de plantas e insectos, proteínas de resistencia a la toxina, proteínas de detoxificación de micotoxinas, enzimas de crecimiento de la planta (por ejemplo, ribulosa 1, 5 biofosfato carboxilasa / oxigenasa, "RUBISCO"), lipoxigenasa (LOX), y fosfoenolpiruvato (PEP) carboxilasa ,son también dianas para la modificación de aminoácidos no naturales.
- [0105] La invención proporciona asimismo procedimientos para producir en una célula de vertebrado, al menos, una proteína que comprende, al menos, un aminoácido no natural (y proteínas producidas por dichos procedimientos). Por ejemplo, un procedimiento incluye: el crecimiento en un medio apropiado, una célula de vertebrado que comprende un ácido nucleico que comprende, al menos, un codón selectivo y codifica a la proteína. La célula de vertebrado comprende también: un ARNt ortogonal que tiene una secuencia de nucleótidos tal como se describe en la SEC ID Nº 87 (O tRNA) que funciona en las células y reconoce al codón selectivo; y una aminoacil ARNt sintetasa ortogonal (O RS) que preferentemente aminoacila el O tRNA con un aminoácido no natural, y el medio comprende un aminoácido no natural.

[0106] En una realización, el procedimiento incluye la incorporación de la proteína en el aminoácido no natural, en el que el aminoácido no natural comprende un primer grupo reactivo; y la puesta en contacto de la proteína con la molécula (por ejemplo; un colorante, un polímero, por ejemplo; un derivado de polietilenglicol, un fotoreticulador, un compuesto citotóxico, un marcador de afinidad, un derivado de la biotina, una resina, una segunda proteína o polipéptido, un quelante de metales, un cofactor, un ácido graso, un carbohidrato, un polinucleótido (por ejemplo, ADN, ARN, etc), y similares) que comprende un segundo grupo reactivo. El primer grupo reactivo reacciona con el segundo grupo reactivo para unir la molécula al aminoácido no natural a través de la cicloadición [3 + 2]. En una realización, el primer grupo reactivo es un resto alquinilo o azida y el segundo grupo reactivo es un resto azida o alquinilo. Por ejemplo, el primer grupo reactivo es un resto alquinilo (por ejemplo, en un aminoácido no natural de p – propargiloxifenilalanina) y un segundo grupo reactivo es un resto azida. En otro ejemplo, el primer grupo reactivo es un resto azida (por ejemplo, en un aminoácido no natural de p – azido – L – fenillalanina) y el segundo grupo reactivo es un resto alquinilo.

5

10

60

- [0107] En una realización, la O RS aminoacila el O tRNA con un aminoácido no natural de, al menos, el 50 % tan eficiente como una O RS que tiene una secuencia de aminoácidos, por ejemplo, tal y como se describe en la SEC ID Nº 86 o 45. En otra realización, la O RS comprende un aminoácido que se describe en cualquiera de las SECs ID Nº 36 63 y / o la 86.
- [0108] La proteína codificada puede comprender, por ejemplo, una proteína terapéutica, una proteína de diagnóstico, una enzima industrial o una parte de las mismas. Opcionalmente, la proteína producida por el procedimiento, es además, modificada a través de un aminoácido no natural. Por ejemplo, la proteína producida por el procedimiento es opcionalmente modificada por, al menos, una modificación post traduccional *in vivo*.
- [0109] Se describen además procedimientos de producción de un cribado o de selección de una proteína 25 moduladora transcripcional (y el cribado o la selección de proteínas moduladoras transcripcionales producidas por dichos procedimientos). Por ejemplo, un procedimiento incluye: la selección de una primera secuencia de polinucleótidos, en la que la secuencia de polinucleótidos codifica un dominio de unión del ácido nucleico; la mutación de una primera secuencia de polinucleótidos para incluir, al menos, un codón selectivo. Esto proporciona un cribado o selección de una secuencia de polinucleótidos. El procedimiento incluye asimismo: la selección de una 30 segunda secuencia de polinucleótidos, en la que la segunda secuencia de polinucleótidos codifica el domino de activación transcripcional; proporciona un constructo que comprende el cribado o la selección de una secuencia de polinucleótidos operativamente unida a la segunda secuencia de polinucleótidos; y la introducción del constructo, un aminoácido no natural, una ARNt sintetasa ortogonal (O - RS) y un ARNt ortogonal (O - tRNA) en una célula. Con estos componentes, la O - RS aminoacila el O - tRNA con un aminoácido no natural y el O - tRNA reconoce el 35 codón selectivo y se incorpora al aminoácido no natural en el dominio de unión del ácido nucleico en respuesta del codón selectivo en el cribado o la selección de la secuencia de polinucleótidos, proporcionando, de este modo, el cribado o la selección de la proteína moduladora transcripcional.
- [0110] En algunas realizaciones, la proteína o el polipéptido de interés (o parte del mismo) en los procedimientos de la invención se codifica mediante un ácido nucleico. El ácido nucleico comprende, al menos un codón selectivo, al menos dos codones selectivos, al menos tres codones selectivos, al menos cuatro codones selectivos, al menos cinco codones selectivos, al menos selectivos, al menos ocho codones selectivos, al menos nueve codones selectivos, diez o más codones selectivos.
- 45 [0111] Los genes que codifican las proteínas o polipéptidos de interés pueden mutagenizarse utilizando procedimientos ya conocidos para los expertos en la disciplina y se describen en el presente documento titulado "Mutagenesis and Other Molecular Biology Techniques" para incluir, por ejemplo, uno o más codones selectivos para incorporar un aminoácido no natural. Por ejemplo, se mutageniza un ácido nucleico para una proteína de interés para incluir uno o más codones selectivos proporcionando la inserción de uno o más aminoácidos no naturales. La divulgación incluye cualquier variante, por ejemplo; un mutante, versiones de cualquier proteína, por ejemplo, incluyendo al menos, un aminoácio no natural. De manera similar, la invención incluye además los correspondientes ácidos nucleicos, es decir; cualquier ácido nucleico con uno o más codones selectivos que codifican uno o más aminoácidos no naturales.
- 55 **[0112]** La purificación de proteínas recombinantes comprende aminoácidos no naturales.
 - [0113] Las proteínas que comprenden aminoácidos no naturales, los anticuerpos de proteínas que comprenden aminoácidos no naturales, etc, pueden purificarse, ya sea parcial o sustancialmente en homogeneidad, según los métodos estándar conocidos y utilizados por los expertos en la disciplina. En consecuencia, los polipéptidos pueden recuperarse y purificarse por diversos procedimientos ya conocidos en la disciplina, incluyendo por ejemplo, la precipitación de sulfato de amonio o etanol, la extracción de ácido o de base, la cromatografía en columna, la cromatografía en columna de afinidad, la cromatografía de intercambio aniónico o catiónico, la cromatografía en fosfocelulosa, la cromatografía de interacción hidrofóbica, la cromatografía de hidroxiapatita, la cromatografía de lectinas, la electroforesis en gel y similares. Las etapas de repliegue de las proteínas puede utilizarse como se desee fabricando correctamente las proteínas plegadas maduras. La cromatografía líquida de alto rendimiento (HPLC), la cromatografía de afinidad u otros procedimientos adecuados pueden utilizarse en las etapas finales de purificación

cuando se desee una alta pureza. Los anticuerpos hechos contra los aminoácidos no naturales (o proteínas que comprenden aminoácidos no naturales) se utilizan como reactivos de purificación, por ejemplo, purificación basada en la afinidad de las proteínas que comprenden uno o más aminoácido(s) no natural(es). Una vez purificados, parcial u homogéneamente, según se desee, los polipéptidos se utilizan opcionalmente por ejemplo, como componentes de ensayo, reactivos terapéuticos o como inmunógenos para la producción de anticuerpos.

[0114] Además de otras referencias ya citadas anteriormente, se conoce en la disciplina una variedad de procedimientos de purificación / plegamiento proteico, incluyendo, aquellos descritos en R. Scopes, *Protein Purification*, Springer-Verlag, N.Y. (1982); Deutscher, *Methods in Enzymology* Vol. 182; *Guide to Protein Purification*, Academic Press, Inc. N.Y. (1990); Sandana (1997) *Bioseparation of Proteins*, Academic Press, Inc.; Bollag et al. (1996) *Protein Methods*, 2ª Ed. Wiley-Liss, NY; Walker (1996) *The Protein Protocols Handbook Humana Press*, NJ, Harris and Angal (1990) *Protein Purification Applications: A Practical Approach IRL Press at Oxford*, Oxford, England; Harris and Angal *Protein Purification Methods: A Practical Approach IRL Press at Oxford*, Oxford, England; Scopes (1993) *Protein Purification: Principles and Practice* 3ª Ed. Springer Verlag, NY; Janson and Ryden (1998) *Protein Purification: Principles, High Resolution Methods and Applications*, 2ª Ed. Wiley-VCH, NY; y Walker (1998) *Protein Protocols on CD-ROM Humana Press*, NJ; y las referencias citadas en los mismos.

[0115] Una de las ventajas de producir una proteína o polipéptido de interés con un aminoácido no natural en una célula de un vertebrado, es que las proteínas o polipéptidos se plegarán en sus conformaciones nativas. Sin embargo, los expertos en la materia reconocerán que, tras la síntesis, expresión y / o purificación, las proteínas pueden poseer una conformación diferente a las conformaciones deseadas de los polipéptidos pertinentes. En un aspecto de la invención, la proteína expresada está opcionalmente desnaturalizada y, a continuación, se renaturaliza. Esto se consigue, por ejemplo, añadiendo una chaperorina a la proteína o polipéptido de interés, y / o solubilizando las proteínas en un agente caotrópico tal como una guanidina HCl, etc.

[0116] En general, es deseable de vez en cuando desnaturalizar y reducir los polipéptidos expresados y, a continuación, provocar los polipéptidos para que se replieguen en la conformación preferida. Por ejemplo, puede añadirse guanidina, urea, DTT, DTE, y / o chaperonina al producto de traducción de interés. Son ya conocidos en la disciplina los procedimientos de reducción, desnaturalización y renaturalización de proteínas (véanse, las referencias anteriores, y Debinski, et al. (1993) J. Biol. Chem., 268: 14065 -14070; Kreitman and Pastan (1993) Bioconjug. Chem.,4: 581 - 585; y Buchner, et al., (1992) Anal. Biochem.. 205: 263 - 270). Debinski, et al., describe, por ejemplo, la desnaturalización y reducción de proteínas de cuerpos de inclusión en guanidina — DTE. Las proteínas pueden volver a plegarse en una solución de tampón redox que contiene, por ejemplo, glutatión oxidado y L — arginina. Los reactivos de repliegue pueden fluir o, de otra manera, moverse al contacto con uno o más polipéptidos u otro producto de expresión o viceversa.

Anticuerpos

5

10

15

20

25

30

35

45

50

55

60

65

[0117] Se describen en la presente invención anticuerpos de moléculas de la divulgación, por ejemplo, sintetasas, ARNt y proteínas que comprenden aminoácidos no naturales. Los anticuerpos de moléculas de la divulgación se utilizan como reactivos de purificación, por ejemplo, para purificar las moléculas de la divulgación. Además, los anticuerpos pueden utilizarse como reactivos indicadores para indicar la presencia de una sintetasa, un ARNt o una proteína que comprende un aminoácido no natural, por ejemplo, para seguir la presencia o ubicación (por ejemplo, *in vivo* o *in situ*) de una molécula.

[0118] Un anticuerpo puede ser una proteína que comprende uno o más polipéptidos sustancial o parcialmente codificados por genes de inmunoglobulina o fragmentos de genes de inmunoglobulina. Los genes de inmunoglobulina reconocidos incluyen los kappa, lambda, alfa, gamma, delta, épsilon y genes de región constante mu, así como la miríada de genes de región variable de inmunoglobulina. Las cadenas ligeras se clasifican como kappa o lambda. Las cadenas pesadas se clasifican como gamma, mu, alfa, delta o épsilon, lo que, a su vez, define las clases de inmunoglobulina, IgG, IgM, IgA, igD e IgE, respectivamente. Una unidad estructural de inmunoglobulina (por ejemplo, anticuerpo) comprende un tetrámero. Cada tetrámero está compuesto de dos pares idénticos de cadenas polipeptídicas; cada par tiene una cadena "ligera" (25 kD) y una cadena "pesada" (50 – 70 kD). El extremo N - terminal de cada cadena define una región variable que oscila entre 100 y 110 o más aminoácidos responsables principalmente del reconocimiento del antígeno. Los términos cadena ligera variable (VL) y cadena pesada variable (V_H) se refieren a estas cadenas ligera y pesada, respectivamente.

[0119] Los anticuerpos existen como inmunoglobulinas intactas o como varios fragmentos ya caracterizados producidos por la digestión con diversas peptidasas. De este modo, por ejemplo, la pepsina digiere un anticuerpo por debajo de los enlaces disulfuro en la región de bisagra para producir F(ab´)₂, un dímero de Fab que, en sí mismo, es una cadena ligera unida a V_H – C_Hl por un enlace disulfuro. El F(ab´)₂ puede reducirse en condiciones suaves para romper el enlace disulfuro en la región de bisagra convirtiendo de este modo el dímero F(ab´)₂ en un monómero Fab´. El monómero Fab´ es esencialmente un Fab con parte de la región de bisagra (véase, *Fundamental Immunology*, 4ª Ed., W.E. Paul, ed., Raven Press, N.Y. (1999), para una descripción más detallada de otros fragmentos de anticuerpos). Si bien se definen varios fragmentos de anticuerpos en términos de la digestión de un anticuerpo intacto, un experto se dará cuenta de que dichos fragmentos de Fab´, etc, pueden sintetizarse *de novo*

bien químicamente, bien utilizando una metodología del ADN recombinante. Así pues, el término anticuerpo, tal y como se utiliza en la presente, opcionalmente incluye también fragmentos de anticuerpos ya sea por la modificación de anticuerpos completos o por sintetización de novo utilizando metodologías del ADN recombinante. Los anticuerpos incluyen anticuerpos de cadena única, por ejemplo anticuerpos Fv de cadena única (sFv o scFv) en los que una cadena de variable pesada y una cadena de variable ligera se juntan (directamente o a través de un conector peptídico) para formar un polipéptido continuo. Los anticuerpos de la invención pueden ser, por ejemplo, policlonales, monoclonales, quiméricos, humanizados, de cadena única, de fragmentos Fab, de fragmentos producidos por una biblioteca de expresión de Fab, o similares.

5

- [0120] En general, los anticuerpos son valiosos, tanto como reactivos generales como reactivos terapéuticos en diversos procedimientos biológicos o farmacéuticos moleculares. Los procedimientos para producir anticuerpos monoclonales y policlonales están disponibles, y pueden aplicarse para fabricar los anticuerpos de la invención. Una serie de téxtos básicos describen los procedimientos de producción de anticuerpos estándar, incluyendo, por ejemplo, Borrebaeck (ed) (1995) Antibody Engineering, 2ª Ed. Freeman and Company, NY (Borrebaeck); McCafferty et al. (1996) Antibody Engineering, A Practical Approach IRL at Oxford Press, Oxford, England (McCafferty), y Paul (1995) Antibody Engineering Protocols Humana Press, Towata, NJ (Paul); Paul (ed.), (1999) Fundamental Immunology, 5ª Ed. Raven Press, N.Y.; Coligan (1991) Current Protocols in Immunology Wiley/Greene, NY; Harlow and Lane (1989) Antibodies: A Laboratory Manual Cold Spring Harbor Press, NY; Stites et al. (eds.) Basic and Clinical Immunology (4ª ed.) Lange Medical Publications, Los Altos, CA, y referencias en; Goding (1986) Monoclonal Antibodies: Principles and Practice (2ª ed.) Academic Press, New York, NY; y Kohler and Milstein (1975) Nature 256: 495 497.
- [0121] Se han desarrollado numerosas técnicas recombinantes para la preparación de anticuerpos que no se basan en, por ejemplo, la inyección de un antígeno en un animal y puede utilizarse en el contexto de la presente invención. Por ejemplo, es posible generar y seleccionar bibliotecas de anticuerpos recombinantes en fagos o vectores similares. Véase, por ejemplo, Winter et al. (1994) Making Antibodies by Phage Display Technology Annu. Rev. Immunol. 12: 433 55 y las referencias citadas en el mismo para una revision. Véanse también, Griffiths and Duncan (1998) Strategies for selection of antibodies by phage display Curr Opin Biotechnol 9: 102 8; Hoogenboom et al. (1998) Antibody phage display technology and its applications Immunotechnology 4: 1 20; Gram et al. (1992) in vitro selection and affinity maturation of antibodies from a naïve combinatorial immunoglobulin library PNAS 89: 3576 3580; Huse et al. (1989) Science 246: 1275 1281; y Ward, et al. (1989) Nature 341: 544 546.
- [0122] Las bibliotecas de anticuerpos pueden incluir depósitos de genes V (por ejemplo, las recogidas a partir de las poblaciones de linfocitos o ensambles *in vitro*) que se clonan para la visualización de los dominios variables de las cadenas ligeras y pesadas asociadas a la superficie del bacteriófago filamentoso. Los fagos se seleccionan mediante la unión a un antígeno. Los anticuerpos solubles se expresan a partir de bacterias infectadas del fago y el anticuerpo puede mejorarse, por ejemplo, a través de la mutagénesis. *Véanse*, por ejemplo, Balint and Larrick (1993) *Antibody Engineering by Parsimonious Mutagenesis Gene* 137:109 118; Stemmer et al. (1993) *Selection of an Active Single Chain Fv Antibody From a Protein Linker Library Prepared by Enzymatic Inverse PCR Biotechniques* 14(2): 256 65; Crameri et al. (1996) *Construction and evolution of antibody-phage libraries by DNA shuffling Nature Medicine* 2:100 103; y Crameri and Stemmer (1995) *Combinatorial multiple cassette mutagenesis creates all the permutations of mutant and wildtype cassettes BioTechniques* 18: 194 195.
- [0123] Los kits para la clonación y la expresión de sistemas de fagos de anticuerpos recombinantes son también conocidos y disponibles, por ejemplo, el "sistema de anticuerpos de fagos recombinantes, Mouse ScFv Module", de Amersham Pharmacia Biotechnology (Uppasala, Suecia). Las bibliotecas de anticuerpos bacteriófagos se han producido también para realizar anticuerpos humanos con alta afinidad mezclando cadenas (véase, por ejemplo, Marks et al. (1992) By Passing Immunization: Building High Affinity Human Antibodies by Chain Snuffling Biotechniques 10: 779 782. Se reconocerá también que los anticuerpos pueden ser preparados por diversos servicios comerciales (por ejemplo, Bethyl Laboratories (Montgomery, TX), Anawa (Suiza), Eurogentec (Bélgica y EE.UU en Filadelfía, PA, etc.) y muchos otros.
- [0124] En algunos casos, es útil "humanizar" anticuerpos, por ejemplo, cuando se administran terapéuticamente los anticuerpos. El uso de anticuerpos humanizados tiende a reducir la incidencia de las respuestas inmunes no deseadas contra los anticuerpos terapéuticos (por ejemplo, cuando el paciente sea un ser humano). Las referencias de anticuerpos anteriores describen estrategias de humanización. Aparte de los anticuerpos humanizados, los anticuerpos humanos son también una característica de la invención. Los anticuerpos humanos consisten en secuencias de inmunoglobulina humana. Los anticuerpos humanos pueden ser producidos utilizando una amplia diversidad de procedimientos (*véase, por ejemplo*, Larrick et al., U.S. Pat. No. 5,001,065, para una revisión). Un enfoque general para producir anticuerpos humanos mediante la tecnología de trioma se describe en Ostberg et al. (1983), *Hybridoma* 2: 361 367, Ostberg, U.S. Pat. Nº. 4.634.664, y Engelman et al., U.S. Pat. Nº. 4.634.666.
 - [0125] Se conocen varios procedimientos de utilización de anticuerpos para purificar y detectar proteínas y pueden aplicarse para la detección y purificación de proteínas que comprenden aminoácidos no naturales descritos anteriormente. En general, los anticuerpos son reactivos útiles para ELISA, transferencia de Western, inmunoquímica, procedimientos de cromatografía de afinidad, SPR y cualquier otro procedimiento. Las referencias

citadas anteriormente proporcionan detalles sobre cómo realizar ensayos ELISA, membranas de Western, resonancia de plasmones superficiales (SPR) y similares.

[0126] En un aspecto, los mismos anticuerpos incluyen aminoácidos no naturales, proporcionan anticuerpos con propiedades de interés (por ejemplo, vida media mejorada, estabilidad, toxicidad o similares). Véase asimismo, la sección del presente documento titulada "polipéptidos con aminoácidos no naturales". Los anticuerpos representan actualmente cerca del 50 % de todos los compuestos en ensayos clínicos (Wittrup, (1999) *Phage on display Tibtech* 17: 423 - 424 y los anticuerpos se utilizan como reactivos de diagnóstico. En consecuencia, la capacidad para modificar anticuerpos con aminoácidos no naturales proporciona una importante herramienta de modificación de estos valiosos reactivos.

[0127] Por ejemplo, hay muchas aplicaciones de mAbs en el campo del diagnóstico. Los ensayos van desde simples pruebas puntuales a procedimientos más complicados como el Mab radiomarcado a NR – LU – 10 de DuPont Merck CO utilizado para la imagen tumoral (Rusch et al. (1993) NR-LU-10 monoclonal antibody scanning. A helpful new adjunct to computed tomography in evaluating non – samll – cell lung cancer. J Thorac Cardiovasc Surg 106: 200 - 4). Como se ha señalado anteriormente, los mAbs son reactivos centrales para ELISA, transferencia Western, inmunoquímica, procedimientos de cromatografía de afinidad y similares. Cualquier anticuerpo de diagnóstico puede ser modificado para incluir uno o más aminoácidos no naturales; modificar, por ejemplo, la especificidad o avidez del Ab para una diana o alterar una o más propiedades detectables, por ejemplo, incluyendo un marcador detectable (por ejemplo, espectrográfico, fluorescente, luminiscente, etc) en un aminoácido no natural.

[0128] Una clase de valiosos reactivos de anticuerpos son los Abs terapéuticos. Por ejemplo, los anticuerpos pueden ser anticuerpos monoclonales específicos de tumores que detienen el crecimiento tumoral para dirigir las células tumorales a su destrucción por citotoxicidad dependiente de anticuerpos mediada por células (ADCC) o lisis mediada por el complemento (CML) (estos tipos generales de Abs se refieren a veces como "balas mágicas "). Un ejemplo es Rituxan, un MAb anti- CD20 para el tratamiento del linfoma no Hodgkin (Scott (1998) Rituximab: a new therapeutic monoclonal antibody for non-Hodgkin's lymphoma Cancer Pract 6: 195 - 7). Un segundo ejemplo se refiere a anticuerpos que interfieren con un componente crítico del crecimiento tumoral. Herceptin es un anticuerpo monoclonal anti - HER- 2 para el tratamiento contra el cáncer de mama metastásico, y proporciona un ejemplo de un anticuerpo con este mecanismo de acción (Baselga et al. (1998) Recombinant humanized arrti-HER2 antibody (Herceptin) enhances the antitumor activity of paclitaxel and doxorubicin against HER2/neu overexpressing human breast cancer xenografts [la errata publicada aparece en Cancer Res (1999) 59(8): 2020], Cancer Res 58: 2825 -31). Un tercer ejemplo se refiere a anticuerpos para la entrega de compuestos citotóxicos (toxinas, radionucleidos, etc) directamente a un tumor u otro sitio de interés. Por ejemplo, una aplicación de Mab es CYT - 356, un anticuerpo ligado a 90Y que dirige la radiación directamente a las células del tumor de próstata (Deb et al. (1996) Treatment of hormonerefractory prostate cancer with 90Y-CYT-356 monoclonal antibody Clin Cancer Res 2: 1289 - 97. Una cuarta aplicación es la terapia enzimática de profármacos dirigida con anticuerpos, donde una enzima colocalizada en un tumor activa un profármaco sistémicamente administrado en la proximidad del tumor. Por ejemplo, un anticuerpo anti - Ep - CAM1 vinculado a la carboxipeptidasa A está siendo desarrollado para el tratamiento contra el cáncer colorrectal (Wolfe et al. (1999) Antibody-directed enzyme prodrug therapy with the T268G mutant of human carboxypeptidase A1: in vitro and in vivo studies with prodrugs of methotrexate and the thymidylate synthase inhibitors GW1031 and GW1843 Bioconjug Chem 10: 38 - 48). Otros Mabs (por ejemplo, antagonistas) están diseñados específicamente para inhibir funciones celulares normales para el beneficio terapéutico. Un ejemplo es Orthoclone OKT3, un MAb anti -CD3 ofrecido por Johnson and Johnson para reducir el rechazo agudo del trasplante de órganos (Strate et al. (1990) Orthoclone OKT3 as first-line therapy in acute renal allograft rejection Transplant Proc 22: 219 - 20. Otra clase de productos de anticuerpos son los agonistas. Estos Mabs están diseñados específicamente para mejorar las funciones celulares normales para el beneficio terapéutico. Por ejemplo, agonistas basados en Mab de los receptores de acetilcolina para la neuroterapia que se encuentra en desarrollo (Xie et al. (1997) Direct demonstration of MuSK involvement in acetylcholine receptor clustering through identification of agonist ScFv Nat. Biotechnol. 15: 768 - 71. Cualquiera de estos anticuerpos pueden ser modificados para incluir uno o más aminoácidos no naturales para mejorar una o más propiedades terapéuticas (especificidad, avidez, vida media del suero, etc).

[0129] Otra clase de productos de anticuerpos ofrecen funciones novedosas. Los principales anticuerpos de este grupo son anticuerpos catalíticos, tales como secuencias Ig que han sido diseñadas para imitar las capacidades catalíticas de las enzimas (Wentworth and Janda (1998) Catalytic antibodies Curr Opin Chem Biol 2: 138 - 44. Por ejemplo, una aplicación interesante implica la utilización del anticuerpo catalítico mAb 15A10 para hidrolizar la cocaína in vivo para la terapia de adicción (Mets et al. (1998) A catalytic antibody against cocaine prevents cocaine's reinforcing and toxic effects in rats Proc Natl Acad Sci U S A 95: 10176 - 81). Los anticuerpos catalíticos también pueden ser modificados para incluir uno o más aminoácidos no naturales para mejorar una o más propiedades de interés.

Acido nucleico y secuencia de polipéptidos y variantes

5

10

15

20

25

30

35

40

45

50

[0130] Como se describió anteriormente y como se describe a continuación, la descripción proporciona secuencias de polinucleótidos de ácido nucleico, secuencias de aminoácidos de polipéptidos, por ejemplo, O - tRNAs y O - RSs,

y, por ejemplo, composiciones y procedimientos que comprenden dichas secuencias. Se revelan en el presente documento ejemplos de dichas secuencias, por ejemplo, O - tRNAs y O - RSs (véase, la Tabla 5, por ejemplo, las SECs ID N° 3 - 65, 86, 87 y 88, y las SECs ID N° 1 y 2).

5 Mutagénesis y otras técnicas de biología molecular

10

15

20

25

30

35

40

45

50

55

60

65

[0131] Los textos generales que describen las técnicas de biología molecular incluyen a Berger and Kimmel, *Guide to Molecular Cloning Techniques,_Methods in Enzymology* volumen 152 Academic Press, Inc., San Diego, CA (Berger); Sambrook et al., *Molecular Cloning - A Laboratory Manual* (2ª Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989 ("Sambrook") y *Current Protocols in Molecular Biology*, F.M. Ausubel et al., eds., *Current Protocols, a joint venture between Greene Publishing Associates*, Inc. and John Wiley & Sons, Inc., (añadido en 1999) ("Ausubel")). Estos textos describen la mutagénesis, el uso de vectores, los promotores y muchos otros temas relevantes relacionados con, por ejemplo, la generación de genes que incluyen codones selectivos para producir proteínas que incluyen aminoácidos no naturales, ARNts ortogonales, sintetasas ortogonales y pares de los mismos.

[0132] Se utiliza en la divulgación numerosos tipos de mutagénesis, por ejemplo, para producir bibliotecas de ARNts, para producir bibliotecas de sintetasas, para insertar codones selectivos que codifican los aminoácidos no naturales en una proteína o polipéptido de interés. Incluyen, de manera no limitante, mutagénesis aleatoria dirigida, recombinación homóloga, barajado de ADN o cualquier otro procedimiento de mutagénesis recursiva, construcción quimérica, mutagénesis que utiliza plantillas que contienen uracilo, mutagénesis dirigida a oligonucleótidos, mutagénesis de ADN modificado con fosforotioato, mutagénesis que utiliza ADN bicatenario ahuecado o similares, o cualquier combinación de los mismos. Los procedimientos adecuados adicionales incluyen la reparación de malapareamientos de punto, mutagénesis que utiliza cepas huésped deficientes en la reparación, mutagénesis de restricción – selección, y mutagénesis restricción – purificación, mutagénesis por deleción, mutagénesis mediante síntesis génica total, reparación por recombinación y similares. La mutagénesis, por ejemplo, implica que las construcciones quiméricas sean también incluidas en la presente divulgación. En una realización, la mutagénesis puede ser guiada por la información ya conocida de las moléculas de origen natural o alterado o de origen natural mutado, por ejemplo, una secuencia, una comparación de secuencias, propiedades físicas, estructura cristalina o similares.

[0133] Los textos y ejemplos anteriores descubiertos en el presente documento describen estos métodos. La información adicional se halla en las siguientes publicaciones y referencias citadas en: Ling et al., Approaches to DNA mutagenesis: an overview, Anal Biochem. 254(2): 157 - 178 (1997); Dale et al., Oligonucleotide-directed random mutagenesis using the phosphorothioate method, Methods Mol. Biol. 57: 369 - 374 (1996); Smith, In vitro mutagenesis, Ann. Rev. Genet. 19: 423 – 462 (1985); Botstein & Shortle, Strategies and applications of in vitro mutagenesis, Science 229: 1193 - 1201(1985); Carter, Site-directed mutagenesis, Biochem. J. 237: 1 - 7 (1986); Kunkel, The efficiency of oligonucleotide directed mutagenesis, in Nucleic Acids & Molecular Biology (Eckstein, F. and Lilley, D.M.J. eds., Springer Verlag, Berlin)) (1987); Kunkel, Rapid and efficient site-specific mutagenesis without phenotypic selection, Proc. Natl. Acad. Sci. USA 82: 488 - 492 (1985); Kunkel et al., Rapid and efficient site-specific mutagenesis without phenotypic selection, Methods in Enzymol. 154, 367 - 382 (1987); Bass et al., Mutant Trp repressors with new DNA-binding specificities, Science 242: 240 - 245 (1988); Methods in Enzymol. 100: 468 - 500 (1983); Methods in Enzymol. 154: 329 - 350 (1987); Zoller & Smith, Oligonucleotide-directed mutagenesis using M13derived vectors: an efficient and general procedure for the production of point mutations in any DNA fragment, Nucleic Acids Res. 10: 6487 - 6500 (1982); Zoller & Smith, Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors, Methods in Enzymol. 100: 468 - 500 (1983); Zoller & Smith, Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template, Methods in Enzymol. 154: 329 - 350 (1987); Taylor et al., The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA, Nucl. Acids Res. 13: 8749 - 8764 (1985); Taylor et al., The rapid generation of oligonucleotidedirected mutations at high frequency using phosphorothioate-modified DNA, Nucl. Acids Res. 13: 8765 - 8787 (1985); Nakamaye & Eckstein, Inhibition of restriction endonuclease Nci I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis, Nucl. Acids Res. 14: 9679 - 9698 (1986); Sayers et al., Y-T Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis, Nucl. Acids Res. 16: 791 -802 (1988); Sayers et al., Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in the presence of ethidium bromide, (1988) Nucl. Acids Res. 16: 803 - 814; Kramer et al., The gapped duplex DNA approach to oligonucleotide-directed mutation construction, Nucl. Acids Res. 12: 9441 - 9456 (1984); Kramer & Fritz Oligonucleotidedirected construction of mutations via gapped duplex DNA, Methods in Enzymol. 154: 350 - 367 (1987); Kramer et al., Improved enzymatic in vitro reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations, Nucl. Acids Res. 16: 7207 (1988); Fritz et al., Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions in vitro, Nucl. Acids Res. 16: 6987 - 6999 (1988); Kramer et al., Point Mismatch Repair, Cell 38: 879 - 887 (1984); Carter et al., Improved oligonucleotide site-directed mutagenesis using M13 vectors, Nucl. Acids Res. 13: 4431 -4443 (1985); Carter, Improved oligonucleotide-directed mutagenesis using M13 vectors, Methods in Enzymol. 154: 382 - 403 (1987); Eghtedarzadeh & Henikoff, Use of oligonucleotides to generate large deletions, Nucl. Acids Res. 14: 5115 (1986); Wells et al., Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin, Phil. Trans. R. Soc. Lond. A 317: 415 - 423 (1986); Nambiar et al., Total synthesis and cloning of a gene coding for

the ribonuclease S protein, Science 223: 1299 - 1301 (1984); Sakamar and Khorana, Total synthesis and expression of a gene for the α-subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin), Nucl. Acids Res. 14: 6361 - 6372 (1988); Wells et al., Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites, Gene 34: 315 - 323 (1985); Grundström et al., Oligonucleotide-directed mutagenesis by microscale 'shot-gun'gene synthesis, Nucl. Acids Res. 13: 3305 - 3316 (1985); Mandecki, Oligonucleotide-directed doublestrand break repair in plasmids of Escherichia coli: a method for site-specific mutagenesis, Proc. Natl. Acad. Sci. USA, 83: 7177 - 7181 (1986); Arnold, Protein engineering for unusual environments, Current Opinion in Biotechnology 4: 450 - 455 (1993); Sieber, et al., Nature Biotechnology, 19: 456 - 460 (2001). W. P. C. Stemmer, Nature 370, 389 - 91 (1994); y, I. A. Lorimer, I. Pastan, Nucleic Acids Res. 23, 3067 - 8 (1995). Se pueden encontrar detalles adicionales en los procedimientos anteriores en Methods in Enzymology Volumen 154, que describe asimismo los controles útiles para solucionar problemas con respecto a diversos procedimientos de mutagénesis.

5

10

15

20

- **[0134]** La divulgación se refiere también a las células huésped de vertebrados, en el que las células no son células humanas embrionarias o células humanas *in vivo* y organismos para la incorporación *in vivo* de aminoácidos no naturales a través de los pares de ARNt / RS ortogonales. Las células huésped son técnicas genéticas (por ejemplo, transformadas, transducidas o transfectadas) con polinucleótidos de la divulgación o constructos que incluyen un polinucleótido de la divulgación, por ejemplo, un vector de la divulgación, que puede ser, por ejemplo, un vector de clonación o un vector de expresión. El vector puede estar, por ejemplo, en forma de plásmido, bacteria, virus, polinucleótido desnudo o un polinucleótido conjugado. Los vectores se introducen en células y / o microorganismos por procedimientos estándar que incluyen electroporación (From et al., Proc. Natl. Acad. Sci. USA 82, 5824 (1985), infección por vectores virales, penetración balística de alta velocidad por pequeñas partículas con un ácido nucleico bien en la matriz de pequeñas perlas o partículas, bien en la superficie (Klein et al., *Nature* 327, 70 73 (1987)).
- [0135] Las células huésped genéticas pueden ser cultivadas en medios nutrientes convencionales modificados según sea apropiado para actividades como, por ejemplo, la fase de cribado, la activación de promotores o selección de transformantes. Estas células pueden cultivarse opcionalmente en organismos transgénicos. Otras referencias útiles, por ejemplo, para el aislamiento celular y cultivo (por ejemplo, para su posterior aislamiento del ácido nucleico) incluyen Freshney (1994) Culture of Animal Cells, a Manual of Basic Technique, 3ª ed. Wiley- Liss, New York y las referencias citadas en; Payne et al. (1992) Plant Cell and Tissue Culture in liquid Systems John Wiley & Sons, Inc.
 New York, NY; Gamborg and Phillips (eds) (1995) Plant Cell, Tissue and Organ Culture; Fundamental Methods Springer Lab Manual, Springer-Verlag (Berlin Heidelberg New York) y Atlas and Parks (eds) The Handbook of Microbiological Media (1993) CRC Press, Boca Raton, FL.
- [0136] La invención se refiere además a líneas celulares de vertebrados, en la que la línea celular no es una célula madre embrionaria humana con capacidad para incorporarse a un aminoácido o aminoácidos no naturales a través de pares de ARNt / RS ortogonales, en el que el ARNt comprende una secuencia tal como se describe en la SEC ID Nº 87. Estas líneas celulares pueden establecerse utilizando técnicas de cultivo celular conocidas en la disciplina, en el que las células huésped han sido transformadas, transducidas o transfectadas con polinucleótidos de la invención o constructos que incluyen un polinucleótido de la invención. Los procedimientos de introducción de ácidos nucleicos exógenos en las células huésped son ya conocidos en la disciplina, y variarán según la célula huésped utilizada. Las técnicas incluyen, de manera no limitante, transfección mediada por dextrano, precipitación fosfato de calcio, tratamiento con cloruro de calcio, transfección mediada por polibreno, fusión de protoplastos, electroporación, infección viral o de fago, encapsulación de polinucleótido(s) en liposomas y microinyección directa.
- 45 [0137] Las células pueden ser transformadas o transfectadas de manera que permitan la incorporación bien transitoria, bien estable del ADN. A largo plazo, para la producción de alto rendimiento de proteínas recombinantes, se prefiere la expresión estable. Por ejemplo, las líneas celulares que expresan de forma estable la molécula del anticuerpo pueden ser genéticas. En lugar de utilizar vectores de expresión que contienen orígenes de replicación virales, las células huésped pueden transformarse con ADN controlado por elementos adecuados de control de la 50 expresión (por ejemplo, promotor, potenciador, secuencias, terminadores de la transcripción, sitios de poliadenilación, etc), y un marcador seleccionable. Tras la introducción de ADN extraño, las células modificadas pueden cultivarse durante 1 – 2 días en un medio enriquecido, y a continuación se cambian a un medio selectivo. El marcador seleccionable en el plásmido recombinante confiere resistencia a la selección y permite que las células se integren de forma estable en el plásmido en sus cromosomas y crezcan para formar focos que, a su vez, pueden 55 clonarse y expandirse en las líneas celulares. Este procedimiento puede ser utilizado ventajosamente en las líneas celulares genéticas que expresan una molécula de anticuerpo. Dichas líneas celulares genéticas pueden ser particularmente útiles en el cribado y evaluación de compuestos que interaccionan directa o indirectamente con la molécula del anticuerpo. Alternativamente, otras técnicas, como las técnicas de transfección mediadas por vectores virales, ya conocidas en la disciplina, pueden permitir la transfección transitoria de las células. 60
 - **[0138]** Están disponibles diversos procedimientos ya conocidos para la introducción de ácidos nucleicos diana en las células, cualquiera de ellos puede utilizarse en la presente invención. Estos incluyen: fusión de células receptoras con protoplastos bacterianos que contienen ADN, electroporación, bombardeo de proyectiles, (para una expresión más estable), e infección con vectores virales (que pueden utilizarse para la transfección estable o transitoria y que se discute además, a continuación), etc. Las células bacterianas pueden utilizarse para amplificar el número de plásmidos que contienen constructos de ADN de la presente invención. Las bacterias se cultivan hasta la fase log y

los plásmidos en la bacteria pueden aislarse gracias a diversos procedimientos ya conocidos en la disciplina (véase, por ejemplo, Sambrook). Además, un exceso de kits está disponible comercialmente para la purificación de plásmidos de bacterias, (*véanse*, por ejemplo, EasyPrepTM, FlexiPrepTM, ambos de Pharmacia Biotech; StrataCleanTM, de Stratagene; y QIAprepTM de Qiagen). Los plásmidos aislados y purificados son manipulados para producir otros plásmidos, utilizados para transfectar células o incorporados en vectores relacionados con los organismos infectados. Los vectores contienen los terminadores de transcripción y traducción, las secuencias de transcripción e iniciación de traducción, y los promotores útiles para la regulación de la expresión de un ácido nucleico diana particular. Los vectores comprenden opcionalmente casetes de expresión genérica que contienen, al menos, una secuencia terminadora independiente, secuencias que permiten la replicación de la casete en eucariotas, procariotas, o ambas, (por ejemplo, vectores lanzadera) y marcadores de selección tanto para los sistemas procariotas y de vertebrados. Los vectores son adecuados para la replicación y la integración en procariotas, eucariotas, o preferiblemente en ambas. Véanse, Giliman & Smith, *Gene* 8:81 (1979); Roberts, et al., Nature, 328: 731 (1987); Schneider, B., el a/., Protein Expr. Purif. 6435: 10 (1995); Ausubel, Sambrook, Berger (all supra). Se proporciona un catálogo de bacterias y bacteriófagos útiles para la clonación, por ejemplo., mediante ATCC, por ejemplo., The ATCC Catalogue of Bacteria and Bacteriophage (1992) Ghema et al. (eds) publicado por la ATCC. Se encuentran métodos básicos adicionales para la secuenciación, clonación y otros aspectos sobre biología molecular y consideraciones teóricas subyacentes en Watson et al. (1992) Recombinant DNA 2ª Ed. Scientific American Books, NY. Además, cualquier ácido nucleico esencial (y prácticamente cualquier ácido nucleico marcado, ya sea estándar o no estándar), pueden pedirse por encargo los estándares de cualquiera de las numerosas fuentes comerciales, como The Midland Certified Reagent Company (Midland, TX mcrc.com), The Great American Gene Company (Ramona, CA available on the World Wide Web at genco.com), ExpressGen Inc. (Chicago, IL disponible en expressgen.com), Operon Technologies Inc. (Alameda, CA), y muchas otras.

Kits

5

10

15

20

25

30

[0139] Los kits son también una característica de la divulgación. Por ejemplo, se proporciona un kit de producción de proteínas que comprende, al menos, un aminoácido no natural en una célula, en el que el kit incluye un recipiente que contiene una secuencia de polinucleótidos que codifica un O – tRNA en la secuencia descrita en la SEC ID Nº 87, y / o un O – tRNA en la secuencia descrita en la SEC ID Nº 87, y / o una secuencia de polinucleótidos que codifica un O – RS, y / o una O – RS. En una realización, el kit incluye además, al menos, un aminoácido no natural. En otra realización, el kit comprende además materiales informativos para la producción de proteínas.

EJEMPLOS

- 35 Ejemplo 1: Métodos de producción y composiciones de aminoacil ARNt sintetasas que incorporan aminoácidos no naturales en células de vertebrados.
- [0140] La expansión del código genético de los vertebrados para incluir aminoácidos no naturales con nuevas propiedades físicas, químicas o biológicas podría proporcionar poderosas herramientas para el análisis y control de 40 la función proteica en estas células. Con tal propósito, se describe un enfoque general para el aislamiento de aminoacil - ARNt sintetasas que incorporan aminoácidos no naturales con alta fidelidad en las proteínas en respuesta a un codón ámbar en Saccharomyces cerevisiae (S. cerevisiae). El procedimiento se basa en la activación de los genes indicadores sensibles GAL4, HIS3, URA3 o LacZ, suprimiendo ambos codones ámbar entre el dominio de unión de ADN y el dominio de activación transcripcional de GAL4. Se describe la optimización de un indicador 45 GAL4 para una selección positiva de variantes activas de tirosil - ARNt sintetasa de Escherichia coli (EcTyRS). Una selección negativa de variantes inactivas de EcTyrRs se ha desarrollado también con un indicador URA3 utilizando una pequeña molécula (acido 5-fluoroorótico (5-FOA)) añadida al medio de crecimiento como un "alelo tóxico". Es importante destacar que tanto las selecciones positivas como negativas pueden realizarse en una única célula y con un amplio rango de severidad. Esto puede facilitar el aislamiento de una diversidad de actividades de aminoacil -50 ARNt sintetasa (aaRS) de las grandes bibliotecas de las sintetasas mutantes. La potencia del procedimiento para el aislamiento de fenotipos de aaRS deseados se demuestra en las selecciones de modelos.

Eiemplo 2

- 55 Construcción de ARNt híbrido a partir de Tyr tRNA de E. coli y B. stearothermophilus
- [0141] Se conoce gracias a los trabajos sobre Saccharomyces cerevisiae, que los pares de Tyr tRNA / RS de E. coli son ortogonales en los ARNt / RS endógenos y apoya la supresión de aminoácidos no naturales. Sin embargo, los esfuerzos para transcribir ARNt^{TYR} de E. coli funcional *in vivo* en células de mamíferos ha supuesto todo un reto.
 Debido a ello, el interés se ha dirigido hacia B. sterothermophilus como fuente de una secuencia de ARNt que puede soportar la supresión de aminoácidos no naturales en las células mamíferas. Aunque el ARNt de B. sterothermophilus es un sustrato para la ARNt^{TYR} sintetasa de E. coli, se requiere además la modificación del ARNt para mejorar la eficacia de la aminoacilación del ARNt. La aminoacilación mejorada del ARNt, mejorará la eficacia de la supresión. El tallo aceptor del ARNt es una clave determinante para el reconocimiento de la ARNt sintetasa. En este ejemplo, se construyó un ARNt híbrido combinando diferentes componentes estructurales del ARNt^{Tyr} de E. coli y B. sterothermophilus. Este ARNt híbrido tiene el tallo aceptor del ARNt^{Tyr} de E. coli, el brazo D, el brazo TψC, un

lazo variable y un tallo anticodón de ARNt^{Tyr} de *B. estearothermophilus*. El nuevo ARNt híbrido, que tiene un tallo aceptor que deriva de *E. coli*, es un sustrato mejor para la ARNt^{Tyr} sintetasa de *E. coli*. A continuación, se muestra en el experimento que se obtuvo una eficacia de supresión de ámbar mejorada cuando se utilizó el nuevo híbrido de ARNt supresor de ámbar. En comparación, se testó el ARNt híbrido junto al ARNt^{Tyr} de *B. stearothermophilus* del que se deriva.

Estudio experimental:

5

10

20

30

40

45

60

65

[0142] Construcción de un plásmido que codifica un ARNt híbrido:

Una expresión de ARNt híbrido supresor de ámbar de única copia que incluye sitios de restricción 5´ (EcoR I y Bgl II), secuencia flanqueadora 5´de ARNt^{Tyr} humano

15 (GGATTACGCATGCTCAGTGCAATCTTCGGTTGCCTGGACTAGCGCTCCGGTTTT TCTGTGCTGAACCTCAGGGGACGCCGACACACGTACACGTC (SEQ ID NO: 89)),

el ARNt híbrido supresor de ámbar mutante que carece de 3´ CCA (la secuencia de nucleótidos del ARNt híbrido es la siguiente:

GGUGGGGUAGCGAAGUGGCUAAACGCGGCGGACUCUAAAUCCGCUCCCUUUG GGUUCGGCGGTUCGAAUCCGUCCCCCUCCACCA (SEQ ID NO:87),

y la secuencia del ADN que codifica el ARNt es la siguiente:

GGTGGGGTAGCGAAGTGGCTAAACGCGGCGGACTCTAAATCCGCTCCCTTTGG GTTCGGCGGTTCGAATCCGTCCCCCA (SEQ ID NO: 88)),

la secuencia flanqueadora 3´ del ARNt^{Tyr} humano (GACAAGTGCGGTTTTTTCTCCAGCTCCCGATGACTTATGGC (SEC ID Nº 90)) y los sitios de restricción 3´ (BamH I y Hind III), se construyeron mediante PCR por solapamiento utilizando cebadores:

35 FTam 73: cebador directo con el sitio de EcoR I y Bgl II

GTACGAATTCCCGAGATCTGGATTACGCATGCTCAGTGCAATCTTCGGTTGCCT GGACTAGCGCTCCGGTTTTTCTGTGC (SEQ ID NO: 91)

FTam 115: cebador inverso:

AGTCCGCCGCGTTTAGCCACTTCGCTACCCCACCGACGTGTACGTGTCGGCG TCCCCTGAGGTTCAGCACAGAAAAACCGGAGCGC (SEQ ID NO: 92)

FTam 116: cebador directo para parte 2:

50 GTGGCTAAACGCGGCGGACTCTAAATCCGCTCCCTTTGGGTTCGGCGGTTCGAA TCCGTCCCCCACCAGACAAGTG (SEQ ID NO: 93)

FTam 117: cebador inverso para parte 2:

55 GATGCAAGCTTGATGGATCCGCCATAAGTCATCGGGAGCTGGAGAAAAAAACCC GCACTTGTCTGGTGGGGGACGG (SEQ ID NO: 94).

Esto se ligó a pUC19 en los sitios EcoR I y Hind III.

[0143] El experimento de supresión de ámbar con ARNt híbrido (Figura 1):

Los plásmidos que codifican el mutante ámbar hGH E88, la ARNt sintetasa de *E. coli* y bien un ARNt híbrido supresor de ámbar de única copia de *B. stearothermophilus*, bien un ARNt híbrido supresor de ámbar de única copia, se transfectaron a células CHO K1. Se analizó la expresión de hGH durante 42 horas

ES 2 457 527 T3

tras la transfección. Cuando se utilizó el ARNt híbrido (hb1), la eficacia de supresión de ámbar aumentó un 30 % respecto a la obtenida cuando se utilizó un ARNt supresor de ámbar de *B. stearothermophilus*.

Ejemplo 3

- [0144] Adición de moléculas a proteínas con un aminoácido no natural.
- En un aspecto, la invención proporciona procedimientos y composiciones relacionadas de proteínas que comprenden aminoácido no natural acoplado a moléculas de sustitución adicionales.
- 10 **[0145]** Se entiende que los ejemplos y realizaciones descritos en la presente tienen propósitos meramente ilustrativos y a tenor de esto, los expertos en la disciplina sugerirán diversas modificaciones o cambios que han de incluirse en el espíritu, en el ámbito de esta solicitud, así como dentro del alcance de las reivindicaciones anexas.
- [0146] Aunque se ha descrito la invención anterior detalladamente con fines de claridad y comprensión, a partir de la lectura será evidente para los expertos en la materia, que se pueden realizar diversos cambios en la forma y los detalles sin apartarse del verdadero alcance de la invención. Por ejemplo, pueden ser utilizados en diversas combinaciones todas las técnicas y aparatos descritos en la presente.

			1
5		GGTA AGGG GGGG GGGG GGGG GGGG GGGG GGGG	HLGH SEWY SNOW SSDQ YKKTS YYVL VEM SDRLF
10		GGGGGCTT CCGGCGCA ACGGCTTC CCAGCAGA SCCGTTCC CAACTATG TTCTGGTA ACGGTGGA ACGGTGTA ACGGGAA ACGGCGAA ACGGCGAA ACGGCAAA CCGCCAAA ACGCGAAA ACGCGAAA ACGCGAAA ACGCGAAA ACGCGAAA ACGCCAAA ACGCCAAA ACGCCAAA ACGCCAAA ACGCCAAA ACGCCAAA	SEPPTADSL SLINTEETVC SUNLQIGG GGAVWLDI SGRAPRAC LAQDGVPM DPEYFFKEI
15		GCAAGAGG AAACGCTT AAACGCTT GCGACGG CTGAACGCTT GCGCGAA TGCGCGAA TGCTTTT AACATCG TTACCGCT TTACCGCT TTACCGCT TTACCGCT GCCCTGGA TTACCGCT TTACCGCT TTACCGCT TTACCGCT TTACCGCT TTACCGCT TTACCGCT TTACCGCT TTACCGCT TTACCGCT TTACCGCT	QGPIALYCC PSFKAAERI NMNVLTFL NMNVLTFL STKFGKTE STKFGKTE VALEEEDKI LSEADFEQI
20		AAACAATTO GCGTTAGG TTCGATCC TATGCCTG TAGCCGC TAGCGTAGG TCCGTAGG TCCGTAGG TCCGTAGG TCCGAATT AAAACCAGG GCGCCCAG GATCAAC GCGCCCAG GATCAAC GCGCCCAG GATCAAC GCGCCCAG GCGCCCAG GCGCCCAG TTCGTTTG GCGCCCAG TGATCATCC TGATCATCC TGATCATCC TGATCATCC TGATCATCC TGATCATCC CTATCGCC	EALAERLA GATGLIGDI INNYDWFGI YNLLQGYD YVPLITKADI FTFMSIEEIN CLFSGSLSA IRKTIASNA
25		ACTTGATT/ ACGAGGAA ATTGCGGG TCCATTGT TGCGCTGG GGAAAACT GAAAAACT CTTCGCCT TCTGACCA TCCGAAG TCCGAAG TCCGAAG TCCGAAG TCCGAAG TCCGAAG TCCGAAG TCCGAAG TCCGAAG TCCGAAG TCCGAAG TCCGAAG TCCGAAG TCCGAAG	LVAQVTDE HKPVALVG GGNSFTEFS GGNSFTEFS DVYRFLKF QAAKRITE QAAKRITE LICWK
30	NCIA	ATGCCAGGGGGGCTGGTAACTTGATTAAACAATTGCCAGGGGGGGCTGGTAGGCCCAAGGGCCCAGGGGCCAAGGGCCCAGGGCCAGGGCCCAGGGCCCAGGGCCCAGGGCCCAGGGCCCAGGGCCCAGGCCAGGCCCCATCCCGCTGCAGCCTTCCATGCCGCTTCGATCCTACCGCTGCCAGCGCTTCCATGCCGCTTCCAGCGCCTTCCAGCGGCTTCCAGCGCTTCCAGCGCCTTCCAGCGGCTTCCAGCCGCTTCCAGCGGCTTCCAGCGGCTTCCAGCGGCTTCCAGCGGCTTCCAGCGGCTTCCAGCGGCTTCCAGGCGGTTCCAGAGCTTCCAGAGCTTCCAGAGCTTCCAGAGCTTCCAGAGCTTCCAGAGCTTCCAGAGCTTCCTGCCTTCCGGGGTTGCCCTTCTTCGCGGGGTTGCCCGTTCCTGCTTTCGGCCGTTTCTTC	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALYCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYDFACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
35	TABLA 5	ATGGCA GCCCAG GCCCAG GCCACA ATTTCG GGTTCG CTTCTO CTGATC GCAGTC GCAGTC GCAGTC GCAGTC GCAGTC GCAGTC AAAAC AAAAC GTGATC GCAGTC GTGATC GCAGTC TTCTGG ATTACC ACTTCG ACTTCG ACTTCG ACTTCG ACTTCG ACTTCG ACTTCG ACTTCG ACTTCG ACTTCG	MASSNL LVPLLCI DKIRKQ' INKEAVI WGNITS PYKFYQ AEQVTR EKGADL GRFTLLI
40		de E. coli de	
45			sintetasa)
50		o TyrRS (sin	(aa) TyrRS (tipo salvaje
55	Marcador	Polinucleótido TyrRS (sintetasa tipo salvaje	Aminoácido (aa) TyrRS (sintetasa) de E. coli de tipo salvaje
60	SEC ID Nº	SEC ID Nº 1	SEC ID Nº 2
65	S	Ū	S

5		SGTA AGGC SGAC SGATT GGTT GGTT GGTT GGTT GGTT GGTT GGT
10		GGGGGCT CTGGCGCA AGCAGGC AGCAGGC CATTGG AGCACCTG CTACCAC AGCACCTG AGCAGGC AGCACCTG AGCAGGC AGCACCTG AGCAGGC ATTAACGG ATCGTC AGCTCTG AGCAGGT AGCTCC ATTAACGG
15		SCAAGAGA ACGCTTCC ACGCTTCC SACGGGTC SACGGTTCC SGTTGCCC SGTTTTCCTA SGTTTCCTA SGTTCCTA SGTTCCTA SGTTCCTA SGTTCCTA SGTTCCTA SGTTCCTA SGTTCCTA SGTTCCTA SGTTCCTA SGTTCCTA SGTTCCTA SGTTCCTA SGTTCCTA SGTTCCTA SGTTCCTA SGTTCCTA SGTTCCCA SGTTCCA SGTTCCA SGTTCCA SGTTCCA SGTTCCA SGTTCCA SGTTCCA SGTTCCA SGTTCCA SGTTCCA SGTTCCA SGTTCA SGTTCCA SGTTCCA SGTTCCA SGTT
20		AACAATTC GCGTTAGC TCGATCCT TCGATCCT GCGGCGC GCGGCGC GTAACAT AAGAACAT GGTAACAT AAGACAT GGTAACAT GGTAACAT AGGTTCT CAGGCG TACCGATG TCGATGCCT TCGATGCCT TCGATGCCT TCGATGCCT TCGATGCCT
25		CTTGATTA CGAGGAA CGAGGAA CATGTTA CAAAATC CAAAAAC CACACACA CACACACA CACACACA
30	(continuación) ENCIA	ATGGCAAGCAGTAACTTGATTAAACAATTGCAAGAGCGGGGGCTGGTA GCCCAGGTGACGGACGAGGCGTTAGCAGGCGCACGGGGCCGGGGCCCGGTTGCATTGGCATCGTACCGCTGACAGCGCGCGC
35	(contin	ATGGCAAGG GCCCAGGTG GCGATCGCA GGGGCATCT CACAAGCC CAGAGCTTC CAGAGCTTC CAGAGCTTC CAGAGCTTC CAGAGCTTC CAGAGCTTC CAGAGCTTC CAGAGCTTC CAGCATTATA AATTGGTTC AATTGGTTC AACCCGTCAAA CGTCTGGTTC AACCGGTAAA CGTCTGGTTC AACAGCTC AACAGCTC AAACAGTC AAAAAAAAAA
40		
45		
50		o sintetasa
55	Marcador	pOMe-I Polinucleótido sintetasa
60	SEC ID Nº	SEC ID Nº 3
65	SEC	SEC

5		GGGC GGGC GGGC GGGC GGGC GGGC GGGC GGCA GGGC G
10		GGGGGCTG AGCTGCAAACT CGTTCCTC CGTTCCTC CGTTCCTC CGTTCCTC CGTTCCTC CGTTCCTC CGTTCCTC CGTTCCTC CAACCTGC CGTTCCTC CAACCTGC CGTTCCTC CAACCTGC CGTTCCTC CAACCTGC CGTTCCCC CAACCTGC CGTTCCCC CAACCTGC CGTTCCCC CAACCTGC CGTTCCCC CGTTCCCCC CTACCAGT CACCTTCCCC CGCTTCCCCC CTTCCCCC CTTCCCCC CGTTCCCCC CGTTCCCCC CTTCCCCC CGTTCCCCC CTTCCCCC CTTCCCC CTTCCCCC CTTCCCC CTTCCCCC CTTCCCC CTTCCCC CTTCCCC CTTCCCC CTTCCCC CTTCCCC CTTCCCCC CTTCCCCC CTTCCCCC CTTCCCCC C
15		SCAAGAGG AGAGGGAC ACGCTTCC ACGCTTCC GACGGTTCCC GGTTAAGC GGTTAAGC JTTTCCTA GGTTAAGC JTTTCCTA AAAACTGA AAAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAAACTGA AAAACTGA AAAACTGA AAAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAACTGA AAAACTGA
20		ATGGCAAGCAGTAACTTGATTAAACAATTGCAAGAGCGGGGGGCTGGTA gCCCAGGTGACGGAGGAAGCGTTAGCAGGCGACTGGCCAAGGC CCGATCGCACTCACTTGTGCTTAGCCTGAAACCGCTGACAGCCATTGGGGACACTTGTTATTGCCTGAAACGCTTCCAGCAGGCGGGCG
25		ACTTGATTA ACGAGGAA ACGAGGAA CGCTGGTA CGCTGGTA ACAAAAT AAACTCTG GGATTTCG GGATTTCG GGATTTCG GGATTCG AAGAAA GATGCCA GGATTCG GGCACCA GGCACCA GGCACCA GGCACCA GGCACCA GGCACCA GGCACCA GGCACCA GGCACCA GGCACCA GGCACCA GGCACCA GGCACCA CCCGCCCA GGCACCCA CCCCCA GGCACCCA CCCCCA GGCACCCA CCCCCCA GGCACCCA CCCCCCA CCCCCCCA CCCCCCCC
30 G	CIA	AGCAGTA GCGCACCGG GCCGGTTG GCCGGTTG GCCGGTTG GCCGGTG GCCGGTG GCGCCGG AGCAGA AGCAGA AGCAGA AGCACGG CACCG CACCGG CACCGG CACCGG CACCGG CACCG CACCGG CACCGG CACCGG CACCGG CACCGG CACCGG CACCGG CACCGG CACCG CACCGG CACCGG CACCGG CACCGG CACCGG CACCGG CACCGG CACCGG CACCGG CACCGG CACCGG CACCGG CACCGG CACCGG CACCG CACCG CACCGG CACCGG CACCG CACCGG CACCGG CACCGG CACCGG CACCGG CACCG CACCGG CACCG CACCGG CACCG CACCG CACCGG CACCGG CACCGG CACCGG CACCGG CACCGG CACCGG CACCG
35	SECUENCIA	ATGGCAAGC GCGATCGCA GCGATCGCA GCGAGCTCT CACAGAGTTG CAGGAGTTATA CAGGATTATA AATTGGTGG AATTGGTGG ACCCGTCGTTAAAG CGTGAAAGA AATTGGTTGG ACCCGTCGTTAAAG ACCCGTCGTTAAAG ACCCGTCGTTAAAG AATTGGTTGG AATTGATGGCTG AACAGCTGG AACAGCTGG AACAGCTGG AACAGCTGG AACAGCTGA AACAGCTGA AACAGCTGA AACAGCTGA TCAGGCACG AACAGCTGA TCAGGCACG
40		
45		ā
50	L	poMe-2 Polinucleótido sintetasa
55	Marcador	pOMe-2 Polinucleó
60	SEC ID Nº	SEC ID Nº 4
65	SE	Ü,

5		SGTA AGGC SGAC SGAC SGAC SGAC SGAC SGAC STGA STGA SCAG SCAG SCAG SCAG SCAG SCAG SCAG SC
10		GGGGGCTC CAGCTTGC CAGCTTGCC TGATTGGC CGTTCCTC CGTTCCTC CGTTCCTC CGTTCCTC CGTTCCTC CGTTCCTC CGTTCCCC CGTTCCCC CGTTCCCC CGTTCCCC CGTTCCCC CGTTCCCC CGTTCCCC CGTTCCCC CGTTCCCC CGTTCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCCCCC
15		SCAAGAG ACCGCTGA ACCGCTGA ACCGCTGA GAACACCG GATTGCCA SCCAATAT GCGTTCTC TACAATTCCTA ACCGCGA ACCGCGA ACCGCGA ACCGCGA ACCGCGA ACCTGA ACCGCGA ACCGCGA ACCGCGA ACCGCGA ACCGCGA ACCGCGA ACCGCGA ACCGCA ACCGCGCA ACCGCGA ACCGCGCA ACCGCCA ACCGCA ACCGCCA ACCGCCA ACCGCCA ACCGCCA ACCGCCA ACCGCCA ACCGCCA ACCGCCA ACCGCA ACCGCCA ACCGCCA ACCGCCA ACCGCCA ACCGCA A
20		AACAATT GCGTTAGG TCGATCCT TCGATCCT TCGATCCT TGCTGAGCA TATCGCG TTCCTGC TTCCTGC TACGTT TCCTGC TCGATTCT TCCATG TCGATTCT TCCATG TCCATG TCGATTCT TCCATG TCGATTCT TCCATG
25		CCTGATTA CCGAGGAA CCGAGGAA CCTGGTAC CCTGGTAC CCCAAATC CCCAACACA AACTCTGC TGCAGACACA AACCCGTTT TGCAGAACACA AAGCAACA AAGCAACA AAGCAACA AAGCAACA AAGCACAA AAGCACAA AAGCACAA AAGCACAA AAGCACAA AAGCACAA AAGCACAA AAGCACCAAA AAGCACCAAA AAGCACCAAA AAGCACCAAA AAGCACCAAA AAGCACCAAAA AAGCACCAAAA AAGCACCAAAA AAGCACCAAAAA AAGCACCAAAAAAAA
30	cion)	ATGGCAAGCAGTAACTTGATTAAACAATTGCAAGAGCGGGGCTGGTA GCCCAGGTGACGGGGCGAGGCGTTAGCAAGAGCGACGGGGCCAGGGCCCAGGTTGCATTGGCATCGCTCACCGCTGACGCTTGCATTTGGCGTTCCATTGTTATGCCTGAACGCTTCCAGCAGCTGGCGGGCCCATCTTGTTTCGTTGCTGGCGGCGCGCGGGCGCGGGCGCGCGGGCGCGGGCGCGGGG
35	(continuacion)	ATGGCAAGG GCCCAGGTG CCGATCGCA GGGGCATCT CCGAGGTGC CCGAGGTGC CCGAGGTGC CCGAGGTGT CCGTTAAC CGTGAAGAT AGGGTTATA AATTGGTGG AATTGGTGG ACCGTCGT TCAGGTTGG GATCACAC TCAGGTTGG GATCACAC TCAGGTTGG AACAGCTGG AACAGCTGG AACAGCTGG AACAGCTGG AACAGCTGG AACAGCTGG AACAGCTGG AACAGCTGG AACAGCTGG AACAGCTGG AACAGCTGG AACAGCTGG AACAGCTGG AACAGCTGG AACAGCTGG AACAGCTGG AACAGCTCG AACAGCTCG
40		
45		
50		lo sintetasa
55	Marcador	pOMe-3 Polinucleótido sintetasa
60	SEC ID Nº	SEC ID Nº 5
65	SEC	SEC

5		GGTA AAGGC GGGGC GGGGC GGGGC GGGAC CCTGTT TGGTT TGGTT TGGTT ACTTC
10		ATGGCAAGCAGTAACTTGATTAAACAATTGCAAGAgCGGGGGGCTGGTAGCCCAGGGCGCAAGGCCCCAGGGCGCCAGGGCGCAAGGCCCCAGGGCGCCCAGGCGGC
15		TGCAAGA GCAGAGCG TACCGCTTC NACGCTTC TGAACGCTTC CGGCAATA GGCCAATA GGCGATA ACGTTCC ACGTTCC TACCAATA TACCCTTC TACCCTTC TACCCTTC TACCCTTC TGGAAGAA TGGAAGAA TGGAAGAA TGGAAGAA
20		TAAACAAT AAGCGTTAA CTTTCGATG GCGTAAGG GCGTAAGG GCGTAAGG GCGTAAGG GCTATCGC GCTATCGC GCTATCGC GCAGGTAA AAATTTGG AAGTTTAG AAGTTTAG AAGTTTAG ACGTTTAG ACGTTTAG CCACCCC CCACCCC CCACCCCC CCACCCCC ACGTTTAG ACGTTTAG ACGTTTAG ACGTTTAG ACGTTTAG ACGTTTAG CCACCCCCCCCCC
25		AACTTGAT GACGAGGA GACGAGGA GCCATTGT GCCATTGT GACAAAA AAAACTCT AAAACTCT AGGCAGG GGATTCA ATGCCGG GGATTCA AGGCACG GGACGGG AGGCACT AGGACGG AGCCGCG AGCCGCG AGCCGCG AGCCGCC AGCCGCC AGCCGCCC AGCCGCCC AGCCGCCC AGCCGCCC AGCCGCCC
30 G	NCIA	ATGCAAGCAGT SCCCAGGTGACG SCCCAGGTGACG SGGGATCTTGT SCCAGGCGGTT SCGAGGTGGGTG SCGAGTGGGTG SGGCATATGAA SGGGTTAAGGAG AATTGGTGGTTC AATTGGTGGTTC SCCCTCTGGTTC SCCCTCTGGTTC AACAGCTGGTTCAC GAATGGTTCAC GAATGGTTCAC SGCCTGGTTCAC GAATGGTTCAC GAATGGTTCAC GAATGGTTCAC GAATGCTGTTCAC GAACAGCTGGTTCAC GAACAGCTGGTTCAC GGAATTAA
35	SECUENCIA	4100C CCGA1 CCGA1 CCGAC CCGAC CCGAC CCGAC CCGAC AACAC CAGAC CAGAC CAGAC CAGAC CAGAC CAGAC
40		
45		œ S
50	1	poMe-4 Polinucleótido sintetasa
55	Marcador	Polinucleá
60	SEC ID Nº	SEC ID Nº 6
65	SE	- S

5		TOGTA AAAGGC SCATT SCATT TOGAT	CGCTG CGCTG CGCTG CGGGT CGGGT ACACC TTTGCC CAATA GATAT GTACG
10		CGGGGGC CCGGGGGC CCGGAGGAA CCGAAGAA CCGAAGAA CCGAAA CGGGGGG GGGGGGG GGAGGGG GAGGGT GAAGGGT CGAAGAAA AAAACGT CGAAGAAA AAAACGT CGAAGAAA AAAACGT CGAAGAAAA	STTAGCAC SATCCTAC CCTGAAA CGCGGCGA AAGCTGA TCGCGGC TCCTGCGC AGAAGCG AGAAGCG
15		GCAAGAG AAACGCTT AAACGCTT CGACGGG TGAACGCTA CGCCAAT CGCCAAT TACCCTT TACCCTC TACCTC TAC	AGGAAGC TGGCTTCC TGGTTATG TGGTAGGC TGGTAGGC TGAGCGT TCAACAA ATTTCGTT CTGTCTG
20		AACAATT GGGTTAGG AGGCGGCG GGGTAAGG CCGTAAGG CCGTAAGG AACATGG AACATGG AACGTTA	ACGGACG CTCACTTG TTGTTCCAT GATGCGCT VAAGCTGC GAGACAA GAGAAAA TGAAATGC CAGATGA
25		CCTTGATTY CCGAGGAA CCGCTGGTY CCCTGGTY CCCTGGTY CAACTCTC CTGCTGAC GGGATTTC GGGATTTC GAGGAAA AGAGGAAA AGAGGAAA GAGGAAA GAGGAAA CCGCCGC GTGAAAA CCGCCCCCCC CCCCCCCCCC	CCAGGTG GGGCATCI CGGCATCI CGGCTTC/ GGGGTGG GGGCTTC/ CCGTTAAC GGGTTAT/
30	ación) CIA	ATGGCAAGCAGTAACTTGATTAAACAATTGCAAGAGCGGCGGGGCTGGTA gCCCAGGTGACGGAGGAAGCGTTAGCAAGACCGTGACAGGCCCGGTGCACGGCGCAGGCCCGTGACACGCTGCAGGCCCGTGACAGGCCCGTGACAGGCTGCGTGCTTGCT	CGGGGGCTGGTAGCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGAAACCGCTGACCGCTGACCGCTGACCGCTGACCGCTGACCGCTGACCGCTGACCGCTGACCGCTGACCGCTGACCGCTGACCGCTGACCGCTGACCGCTGCCTGACCGCTTCCATTGTTATGCCTGAACCCTTCCATTGTTATGCCTGAACCGCTTCCACCGCTTCCACGGTTGCCCTGGTAGCCGCGCGCG
35	(continuación)	ATGGCAAGC CCGATCGCA TGGGGCATC CCCGAGGCTT TCGGCATA TCGGCATA TCGGCATA TCGGCATA TCGGCATA TCGGCATA AAATTGGTG GACCGTC GACCGTC AAATTGGTG GACCGTC AAATGCTG AAACGGTA TCGTCTGGTA GAATGCTG AAACAGTC AAACAGTC AAACAGTC AAACAGTC AAACAGTC	CGGGGG CTGGCG ACAGCT CAGCAG CTGATTG GAAGAA CCGTTCG ATTATGA TGGCAA GCAGCG TACAAG
40			
45			
50		sintetasa	activo) sintetasa
55	Marcador	poMe-5 Polinucleótido sintetasa	pOMe-6 (sitio activo) Polinucleótido sintetasa
60	SEC ID Nº	SEC ID Ne 7	SEC ID Nº 8
65	SE	S S	S S

(continuación) SECUENCIA CGGGGGCTGGTACCCCAGGTGACGAGGAGGGGGGGGGG	
AGGTGACGGACGAGGAAGCGTTAGCCGAAGCGTTCCTAGCGGGCGCGAAGCGGTTCCTTGTTCCTTGTTCTTTGCCTGAACCGGGCGCGCGAAGCGGTTCCTAGGCGGGCG	
AGGTGACGACCAGGAAGC TCGCACTCACTTGTGGCTTCC TCGCACTCACTTGTGGCTTCC TCGCACTCACTTGTGGCTTCC TCGCACTCACTTGTTATG AGGTGACGAAAACTCTGCTA AATATGAATGTGCTGACGA TTAACCAGATGACCATTGTTATG TTAACCAGATGCTGCTGCTAGC TCGCACTCACTTGTTCTTATG TTAACCAGATGCTGCTGCTAGC TCGCACTCACTTGTTCTTATG TTAACCAGATGCTGCTGCTAGC TCGCACTCACTTGTTCTTATG TTAACCAGATGCTGCTGCTAGC TCTGTGACGGACGATGCTTCC TCGCACTCACTTGTTCTTATG TTAACCAGTTGCTTCCTTC TTAACCAGTGGCTGCTAGCCTT TTAACCAGTGGCTGCTAGCCTT TTAACCAGTTGCTTCCTTC TTAACCAGTGGCTGCTGCTTC TTAACCAGTGGCTGCTGCTTC TTAACCAGTTGCTTCTTCTTT TTAACCAGTGGCTGCTGCTTC TTAACCAGTGGCTGCTGCTTC TTAACCAGTTGCTTCCTTC TTAACCAGTGGCTGCTGCTTC TTAACCAGTGGCTGCTGCTTC TTAACCAGTTGCTTCTTTTTTTTTT	
AGGTGACGGACG AGCCGGTTGCCACTG AGGTGACGGACA AGCCGGTTGCGCACTCACTG AGTGGACGGACA AATATGAATGTGCCA AGCCGGTTGCGCACTCCACTG AGTGGACGGACA AATATGAATGTGCCACTGCACTG	
A GOLDON A G	
CIA CIA CIA CIA CICAGGGCC GOCGACC CAAGGCC GOCGACC CAAGGCC CAAGGCC GOCGACC CAAGGCC CAAGCCC CAACCC CAAGCCC CAAGC	
Secuención) Secuención Secuención Cogogogoraco ATTATGACTO ATTATGACTO Cogogogoraco Cogogogoraco ATTATGACTO Cogogogoraco Cogogogoraco ATTATGACTO Cogogogoraco C	GTGTG
40	
45	
o sintetasa o sintetasa o sintetasa o sintetasa o sintetasa	
Marcador pOMe-7 (sitio activo) Polinucleótido sintetasa pOMe-9 (sitio activo) Polinucleótido sintetasa	
60 8 0 0 N 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	
65	

5		GCCGA GCTG GCTG GCTG GCCC CACC TGCC AATA TTCC TACG	GCGA GCTTC GCCTTC GGGT SACC TGCC ATAT TTCC TGC	GCGA GCTG GCTTC GGGT ACC TGCC TGCC TACA TTCC
10		TTAGCAGA ATCCTACC CTGAAACC GCGCGAAA AGCTGCGCC CTGCGCCG CTGCGCG GAAGCGGT ACTGAAACAG	TAGCAGA (TCCTACC CTGAAACC GCCCCGAC AGCTGCGCC CCCCGCCC CTGCCCCG GAAGCGGT CAAACAG	TTAGCAGA VTCCTACCO CTGACACO GCGCGACA AGCTGAC CTGCGCCA CTGCGCCA CTGCGCGA CTGCGCGA CAAACAG
15		GGAAGCG GGCTTCGG GGTTCGG GGTAGGC GACGGTA TCTGCTAI TGTCTGAAA	SGAAGCGI GGCTTCGA GGTTGGC GGTAGGC GAGCGTA TCTGCTAI TGACCTTC CAACAAA TTTCGTTC	GGAAGCG GGCTTCG GGTTCG GGTAGGC GAGGGTA ATCGTA TGACCTTC CAACAAA TTCGTTC
20		CGGGGCGA TGTCCTTGT TGTCCTTGT TGTCCCT TGGGCG TGGGCGCA TGGGCGCA TGGGGGGGGGG	CGGACGAC TCCTTTGT STTGCGCT AAGCTGCC STGGACAA SAGAAAA GAATGTG CAGATGAT CCAGATGAT	CGGACGA TCGGACGA TTGCGCT TTGCGCT TTGCGCT TGGACAA SAGAAAC GAATGTGC CAGATGAT CAGGGGA
25		CCAGGTG/ GATCGCAC GATCGCAC CCAGCCG GAGGTTCA GGCATTA CCATTAAC GCATTAAC GTGAAGAT	CAGGTGA GATCGCAC GGACTCT CAGCCG GGACTTCA GGACTGTG GGCATAT GGCATAT GGGTTATT	CCAGGTGA GATCGCAC GGATCTCA GGAGTTCA GGAGTTGG GGCATTTA GGCATTA GGCATTA GGGTTAAC
30 (continuación)	SIA	CGGGGGCTGGTAGCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGAACCGCAACGCCCAACGCCCAACGCCCACTTGTGGCTTCGATCCTTAGCACCCTACCGCTCCACTTGTGCTCTTCGATCCTACCGCTTCCTACCGCTTCCAACGCTCCCTTCGTTCCTTGTTCCTTTGTTCCTTGTTCCTTGTTCCTTGCCTTGCTTCCTTGCCCTGGAACGCTTCCTTTGTTCGGGGGGGG	CGGGGGCTGGTACCcCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGAACCCTGACCTGCCAAGGCCCAAGGCCTCCTTTGTGGCTTCGATCCTACCGCTGCTGAAGCCTTGGACGCTTCGATCCTTGTTATGCCTTCGATCCTGCTGCTGCTGAAGCCTTCGATTGGCGGCGCGCGC	CGGGGGCTGGTAGCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGAAAACGCTGACGCCTGACGCCTGTGCGTTGGTTG
35 gi	SECUENCIA	CGGGGG CTGGCG ACAGCT CAGCAG CTGATTK GAAGAA CCGTTCC ATTATGA TGGCAA GCAGCG TACAACC	CGGGGG CTGGCG ACAGCT CAGCAG CTGATTG GAAGAA CCGTTCC ATTATG, ATTATG, ATTATG, ATTATG, ATTATG, ATTATG, ATTATG, ATTATG, ATTATG,	CGGGGG CTGGCG ACAGCT CAGCAG CTGATTG GAAGAA CCGTTCG ATTATGA TGGCAA GCAGCG TACAAO
40				
45				
50		o activo) sintetasa	sintetasa	o activo) sintetasa
55	Marcador	pOMe-10 (sitio activo) Polinucleótido sintetasa	pOMe-11 (sitio activo) Polinucleótido sintetasa	pOMe-12 (sitio activo) Polinucleótido sintetasa
60	SEC ID Nº	SEC ID Nº 12	SEC ID Nº 13	SEC ID Nº 14
65	SEC	SEC	SEC	SEC

5		SAGCGA CGCTTC CGCTTC ACGCGT ACACC STTGCC SCAATA GATAT GTTAA TTTTCC	SAGCGA SCGCTTC CGCTTC ACGGGT ACGGGT ACGCC STTGCC SCAATT SGATAT GGTAA TTTTCC AGTACG	GAGCGA CGCTTC CGCTTC ACGGGT ACGGGT ACGGGT ACGGGT ACGGTTAA TTTTCC AGTACG CACTTC
10		STTAGCAK SATCCTAK CCTGAAA CGGCGCGAAAGCGCGGCGGCGGGGGGGGGGG	GTTAGCA GATCCTAG GATCCTAG CCTGAAA AAGCTGA AAGCGGG CCCTGCGG CCCTGCGG AGAAGCG AACAAAC	GTTAGCA SATCCTAC CCTGAAA CGGCGCCO AAGCTGA TTCGCGG TTCGCGG TTCTGCG TTCTGCGG TTCTGCGG TTCTGCGG TTCTGCGG TTCTGCGG TTCTGCGG TTCTGCGG TTCTGCGGG TTCTGCGGG TTCTGCGGGG
15		AGGAAGG TGGCTTC TGGTAGG TGGTAGG TGGTAGG TGAATCCG TCAACCA TTAGGTT CTGTTTCGTT	AGGAAGO TTGTTATG TTGTTATG TGGTAGG CGGTGCT AAATCCT CTGTGCT TCAACAA	AGGAAGG TGGCTTC TGGTTATG TGGTAGGG CGGTCGGT CTCTGCT CTGACCT TCAACAA ATTTCGTT CAATGGT CAATGGT AGAATCAA
20		ACGGACG CTCGTGTG TGTTCCAT GATTCCCC NATCCCC CAGACG GAGAAAA GAGAAAAA GAGAAAAA GAGAAAAAAAA	ACGGACG CTCTGGTG TGTTCCAT GTTGCGC AGGGACA GTGGACA GTGGACA GTGGACA TGAGGGG TTGCGTGC	ACGGACG CTCATTTG TGTTCCAT GTGTGCG TGGACA GAGGACA GAGGACA GAGGACA GAGGACA GAGGACA CAGGGG GTATGGC GTTCTGCCTC
25		CCAGGTG GGACATCI GGGCATCI ACAAGCCG GGGAGTGG GGCATAI GGCATAI CCGTTAAC	CCCAGGTG GGGCATCI GGGCATCI ACAGGCC GGAGTGG GGCATAI CCGTTAAC	CCCAGGTG GGGCATCI ACAGCCC ACAGCCC ACAGTGC GGACTGTG GGCACTGT ACGTTAAK AGGGTAATA AGGGTAATA ACCGTCCG
30 (ación)	CIA	CGGGGGCTGGTACCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGAAACCGCTGACCGCAACGCCAAGGCCACTCGTGTGGCTTCGATCCTACCGCTGACCGCTGACGCCTGACGCTGCACGCTGCAACGCTTCGATCCTTGTTATGCCTTCGATCCTGAACGCTTCCATTGTTATGCCTGAACGCTTCCAGGCGGCGCGCGC	CGGGGGCTGGTAGCCCAGGTGACGAGGAGGGTTAGCAGAGCGAACGCAGAGCGAAGGCCCAACGCCATCTGGTGTGGGCTTCGATCCTACCGCTGACCTGGCGTTGGTTG	CGGGGGCTGGTAGCCCAGGTGACGGAGGAAGCGTTAGCAGAGCGAACGCAACGCAAGGCCAATCCTATTTGTGGCTTCGATCCTACCGGTGAACGCTGCAACGCTTGGTATTTGTTATTGTCGTTCTACCGCTGAAACGCTTCCAATTTGTTATTGTTATTGCCTTGAAAACGCTTCCAGGCGGCGGCGCGCGGGTTGCAGGCGGCGGCGCGGGGGGGG
35 (continuación)	SECUENCIA	CGGGGG CTGGCG ACAGCT CAGCAG CTGATT GAAGAA CCGTTCO ATTATG TGGCAA GCAGCG TACAAC	CGGGGG CTGGCG ACAGCTC CAGCAG CTGATTO GAGGAA CCGTTCO GTTATG TGGCAA GCAGCG TACAAC	CGGGGG CTGGCG ACAGCT CAGCAG CTGATI GAAGAA CCGTTC ATTATG TGCAA GCAGCG TACAAC GTGTGG
40				
45				
50		io activo) o sintetasa	io activo) o sintetasa	io activo) o sintetasa
55	Marcador	pOMe-13 (sitio activo) Polinucleótido sintetasa	рОме-14 (sitio activo) Polinucleótido sintetasa	pOMe-15 (sitio activo) Polinucleótido sintetasa
60	SEC ID Nº	SEC ID Nº 15	SEC ID Nº 16	SEC ID Nº 17
65	SEC	SEC	SEC	SEC

5			GCCCC SGCCCA SGCCCA SACCC SATTCA STTCA TTCTC STCA SCAAA	STGGC TTGTT SCGGA AAAAT ACTCT GCTGA NTCAA NTCAA ATTTC CCGTG	GACTG TGACA TCCAG GTCTG GTCTG TAATT ATTGG ATTGG TAATT AGCA CCGTG
10			CAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCCGAAGGCCCGGATCGCATTTGGGATCCTACCGCTGACAGCTTGCATTTGGGGTCTGCATCCGCTGACACGCTTGCATTTGGGGTCTGATCCTTCCAGCAGCTTGCATTTGGCGATCTGATTGCCTGATCGCTTCCAGCGGGCGCGCCCAGGATTGGCGGCCCAGGTTCCAAGCTGGTCCCGGTTCCAAGCTGGTTCAGGTTGCCCCGATTGCGACCTGTTTCCGGGTTGCCCCGTTCCTGGTTTCCGGAACATTTTGCAAAATTCCTGCGGCCCAATAATTATGACTGGTTCGGGAATTTCCTCCCGTTCAACCTTCTCCCGGTTAACCAGTTCGCTTCTCCCGGTTAAGCAGCTTCTCCCGGTTTCCTCCCGGTTTCCTCCCGGTTCCTCC	GCGTTAGCAGAGCGACTGGCGCAAGGCCCGATCGCACTCGGGTGTGGC TTCGATCCTACCGCTGACAGCTTTTGGGGCATCTTGTTCCATTGTT ATGCCTGAAACGCTTCCAGCAGGCGGGCCACAGCCGGTTGCCATTGTT AGGCGGCGCGACGCTTGATTGGCGACCCGAGCTTGCTGTT AGGCGGCGCGACGCGTTCCTGGTTCGGGAGTGGCTGCCGAAAC CCGTAAGCAGGTTGCCCCGATCCTGTTCAGGAGTGGGTGAAAAT CCGTAAGCAGGTTGCCCGTTCCTCGATTTCGACTGTGGGAAAACTCT CCTTCCTGCGCCAATAATTATGACTGGTTCGGCAATATGAAATGTGCTGA CCAAAGAAGCGGTTAAGCAACACTTCTCCGTTAACCAGGTGATCAA CAAAGAAGCGGTTAAGCACTCTCCAACCTGAATGTGTCTGA GTTCACTGAGTTTTCCTACAACCTGCTGCAGGTTATGGTTATGCCTGTA TGAAAAAACAGTTCTGGTATCGACCTGCACCTGTA GGGTAACAACACTTCTGGTATCGACCTGACCAGTTG CAGGTG	GGGCTGGTAGCCCAGGTGACCGACGNAGAAGCGTTAGCAGAGCGACTG GCGCAAGGCCCGATCGTTTGTTGTTATGCTTCCTACCGCTGACA GCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTCCAG GCGTGCCGCGCCCCGATCTTGTTATGCCTGAAACGCTTCCAG CAGGCGGGCCACAAGCCTGCTGTAGGCGGCGCCACCGGTTCCAG ATTGCCGACCTTCAAAGCTGCGCTGAAGCGGCGCAACAGT TTCCTCGATTTCGACTGTGACAAATCCGTAAGCTGCCCCG TTCCTCGATTTCGACTGTGACAAAATCCGTAAGCTTGCCCCC TTCCTCGATTTCGACTGTGAGAAATCCGTAAGTTTGC CAAACACTTCTCCGTTAACCAGATGTGCTGCTACCGGCGAATATTGG CAAACACTTCTCCGTTAACCAGATGTGCTGCTTCCTGCGGTTAATTGG CAAACACTTCTCCGTTAACCAGATGTGCTGCTTCCTGCGGTTAATTGG TGCTCCTCCAACCGTGGTTATTCTATGGCCTGTTCCCTGGGTTTTCCTTGC TATCGACCTGCAACACTGTGCTTCCACAACAGTACGGTG TATCGACCTGCTGCATCTTCTTCTTCTTCTTCTTGG TATCGACCTGCAACACTTCTTCTTCTTCTTGGTTTTCTTTGCTTCTTTCT
15			GAGCGACT CCGCTGAC, ACGCTTCCA JACGGGTC JACGGGTC GGTTGCCG GGTTAAA CCATAAA GGTTAAAC GGTTAAGC GGTTAAGC GGTTAAGC	CCGATCGC TGGGGCAT CCCGAGCT TCGGGGAT TCGGCAT TC	GGGCTGGTAGCCCAGGTGACGACGNAGAAGCGTTAA GCGCAAGGCCCGATCGCACTCTTTGTGGCTTCGATC GCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTG GAGGCGGGCCACAGCCGGTTGCCCTGGTAGGCGGC ATTGGCGACCCGAGCTTTGTTGTTATGCCTG ATTGCCACCGAGGCTTCAAGCTGCTGGTAAGC GAAACTGTTCGACTGTGGAGAAACTCTGCTAAGC ATGACTGGTTCGGCAATGGTGAAAACTCCTGCTAAGAA GCGTCTCAACCGTGAAGATGCGTTCCACTAAGAAACTCCTGCTTCCTCCTGAACGTTCCTCCTTCCT
20			CGTTAGCA CGATCCTAA GCCGCGCC GTAAGCT CGTAAGCA CGTAAGCA CAAGAAGC TTCACTGCG AAGAAGC	SGCAAGGC CTTGCATT AGGCGGG ATTGGCGA TTCCTCGAT TCCTCGAT AACCTGT AACCTGT ACCTGCTG ACCTGCTG ACCTGCTG ACCTGCTG ACCTGCTG	CCTTTGTG GTCCATT TTGCGCTG TGCACACA GGACAACA AGATGATC AGATGATC AGATGATC AGATGATC AGATGATC AGATGATC AGATGATC AGATGATC AGATGATC AGATGATC AGATGATC
25			CAGGTGACGGACGAGGAAGCGTTAC ATCGCACTCGGTTGTTATGCTTCGATCC GGCATCTTGTTCCATTGTTATGCCTG CAAGCCGGTTGCCTGGTAGGCGGC GAGCTTGCAAAGCTGGTAGGCGGC GAGCTTGCAAAGCTGGTAGGCGGC GAGCTGGAAAACTCTGCTATCG GCAATATGAAAACTCTGCTATCG GCAATATGAATGTGTATCGTTCCT CGTTAACCAGATGATCATCGTTCCT CGTTAACCAGATGATCATCGTTCCT TGAAGATCAGGGGATTTCGTTCACT TGGTGGTTCTGACCAGAGAA TTGGTGGTTCTGACCAGGGGTAA	CGACTGGG GCTGACAG GCTGACAG GCTGCCGGA NGCCCGGA TGCCCGG TTCCCGG TTCCTACA TTCCTACA TTCCTACA	CAGGTGAC GACATCTT CAGCATCTT CAGCTCGG CAGTGGGG GACTGTGG GACTGTGG GACTGTGG GATTATG GGTTATTC TGGTTATTC TTGGTTGTTC
30	continuación)	CIA	ACCGGAC CTCCGTTCC CGATGC CGAGAC CCAGAT TGAATC CCAGAT TTCTCCAGAT	GCAGAC CCTACC GACTACC GCGCGA(GCGCGAC GCGCGC TCGCGC TCGCGC TCGCGC TCGCCG	GTAGCC GGCCCA GGCCCA GGTCCC GATTCC GGTTCC CCACCC CCACCC CCACCC CCACCC CCACCC CCACCC CCACCC CCACCC CCACCC CCACCC CCACCC
35	(cont	SECUENCIA	CAGGTG ATCGCATG CAAGCC GAGCTT GAAGTC GCATA TGAAGA GGTTAA TGGTG CCGTCO	DCGTTAK TTCGATC ATGCCT AGGCGG GCGTAA CCGTAA CCGTAA GCTATCC CAAAGA GTTCAC TGAACA GGGTAA	GGGCTG GCGCAA GCTTGC CAGGCC ATTGGC GAACT CAAACT CAAACT GCGTCT AACCTC TGGGCT TGGTGC
40					
45			activo) sa	o activo) sa	(o)
50		or	pBenzofenon -1 (sitio activo) Polinucleótido sintetasa	pBenzofenon -e2 (sitio activo) Polinucleótido sintetasa	pAzidoFe-1 (sitio activo) Polinucleótido sintetasa
55		Marcador	pBenzofe Polinucle	pBenzofe Polinucle	pAzidoFe Polinucle
60		SEC ID Nº	SEC ID Nº 18	SEC ID Ne 19	SEC ID Nº 20
65	Į	S	<u>v</u>	<u>N</u>	<u>v</u>

5		CTTCG TATGC AGGC GCGT FCCGT FCCTA ACCTT ACAAA ACCTT GCGA TGGG	CACTC CTTGT CAAA GGGTG GGGTG TGGAA VTGAA CCAG TCGGC GGTTC
10		TTAGCAGAGCGACTGGCGCAAGGCCCGATCGCACTCGTTTGTGGCTTCG TTAGCAGCGTGCATTTTGGGGCATCTTGTTCCATTGTTATGC CTGACCGCTTCCAGCGGCGGGCACCTGGTTGCTTGTTATGC GGCGCGACCGGGTTGCTGGTTGGCGAGCGTTGCTGGTAGGC GGCGCGACCGGGGCTTCCGGAGCTTCCGGTTGCCCGAGCGT AAGCTGAACTGTTCCGGACTTTCGGAGAGAGCTGCTGCTAATCCGT ACCCGGCCAATAATTATGCAACTTTTCGACTGTGGAGAAACTCTGCTAACCTT CCCGGGCCAATAATTATGCAACACTTTCCGTTAACCAGAGGGGATTTCGTAAAAAAAA	GACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGCCCGATCGCACTC CTGTGTGGCTTCGATCTACCGCTGACAGCTTGCATTTGGGGCATCTTGT TCCATTGTTATGCCTGAACACCTTCCAGCAGCTGCATTTGGGGCATCTTGT GCGCTGGTAGGCGGCGACGGGTCTGATTGGCGACCCGGTTT GCCTGGTAGGCGGCGACGGGTCTGATTGGCGACCCGGTTCAAA GACAAAATCCGTAAGCAGGTTGCCCCGTTCCTCGATTCAGGGGGGGG
15		TCGCACTC SCATCTTGT AGCTCGATA GCTTCAAA GCTTCAAA CATTTGGG STTAACCA STTAATCA STTATTCTC STTATTCTC	SCGCAAGG SCTTGCATT SACTGCACT TCCACCT TCCTCGAT TCCTCCACT CCTGCTCCAC ATCGACCT
20		AGGCCCGA GGGCCCGA GGGCCCGA CGACCCGA CGACCCGA CGACCCGA TGGTTCGG ACTTCTCC ACTTCTCTC ACTTCTCC ACTTCTCTC ACTTCTC ACTTCTC ACTTCTC ACTTCTC ACTTCTC ACTTCTC ACTTCTC ACTTCTC ACTTCTC ACTTCTC ACTTCT ACTTCTC ACTTCT ACTTC	GCGACTGG CCCGACAGG CGCTTCCAGG CGCGCGAAG TGCCCGAAGATA AATAATA AATAATA TAAGGAGG TCCTACAA
25		CTGGCGCA ACAGCTTG CTGATTGG CTGATTGG CGGTTCCT ATTATGAC TGGCAAC TGGCAAC TGGCAAC TGGCAAC TGGCAAC TGGCAAC TGGCAAC	STTAGCAG/ STTAGCAG/ STGCGGAG/ SGCGCGAG/ SAGCGGGG/ STGCGGGG/ STGCGG/ STGCG/ STGCG/ STGCG/ STGCGG/ STGCGG/ STGCG/ STGCGG/ STGC/ STGC/ STGCG/ STGCG/ STGC/
30 35	NCIA	CAGAGCGA AACCGCTG CGACGGGT CGACGGGT AGACACC AGGTTGC AGGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCTTAT CGCTTAT CGCTTAT CGCTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCTTAT CGCGTTAT CGCTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCGTTAT CGCTTAT CGCGCTTAT CGCGTTAT CGC	SACGAGGAAGCGTTA CTGTGTGGCTTCGATG CCCTTGTAGCGGG GCCTGGTAGGCGG GACAAATCCGTAAG AAAACTCTGCTATCG TGTCAACAANAA GGGATTTCGTTCACT TGCCTGTGGTCACAA
35	SECUENCIA	77A99 CGGAA AAGC CCTGA ACTG ACAA OTAA	0400 0404 0404 0404 0404 0404 0404 040
40			
45		(o e	(C o
50	Ē	pAzidoFe-2 (sitio activo) Polinucleótido sintetasa	pAzidoFe-3 (sitio activo) Polinucleótido sintetasa
55	Marcador	pAzidoFe- Polinucleó	pAzidoFe-: Polinucleó
60	SEC ID Nº	SEC ID Nº 21	SEC ID Nº 22
65	S	l is	<u>is</u>

	8				
5			GGCT GTTG GTTG GGTA SGAG TCTG TCTG TCTG TTCG TTCG AATC	ACTC GGTT SGTG SGTG SGTG SGAG TGAA ATTT TTCT	TACCG AACGC COACG GAAC GAAC GAGTT GGGTT CGGTT CGGTT TCACT
10			TCGTTTGT GTTCCATT TTGCGCTC AGGCTGC AGATGTGC AGATGTGC AGGTGAT AGGCGCT TGCGGCCT TGCGGCCT TGCGGCCT	CGATCGC SGGGCATC SGGGCATC CGACTGT CGCATA CGCATA CGGTAGA GGGTTATA CGGTCGT	TTCGATCC VTGCCTGA GGCGGCGC GTAAGCT CGTATCGC CTATCGC CTACCTGA CTACCTGA CTACCTGA CTACCTGA CAACAAA GGTAACAAA
15			SATCGCAC SGCATCTT SGCATCTCA SGCATTCA SGCATTGG SGCATTGG SGTATGC SGTATGG SGTATGG SGTGTCT SGTCGTCT	SCAAGGC GCGGGCC GCGGGCC GCCGGCC ACTGGTTC ACTGGTTC TCCCAACC TGCTGCAACG	GTGTGGC CATTGTTA GCTTGTA GCTGGTA GCACACTC GATCTGA GATCTGA GGTTTCG GCTGTCT CAGATC
20			GCGTTAGCAGAGCGACTGGCGCAAGGCCCGATCGCACTCGTTTGTGGCTTCGTTCCATTGTGCATTTGGGGCATCTTGTTCCATTGTTGTTCCATTGTTCCATTGTTCCATTGTTCCATTGTTCCATTGTTGTTCCATTGTTCCATTGTTCCATTGTTCCATTGTTCCATTGTTCCATTGTTCCATTGTTCCATTGTTCCATTGTTCCATTGTTCCATTGTTCCATTGTTGCTTCCATTGCCGAGCTTCCAAGCTGCTGCTGAGCTGCTGCTGAGCTGTTCCGAGTGGGTGG	GACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGCCCGATCGCACTC ATTTGTGGCTTCGATCCTACCGCTGACAGCCTTGCATTTGGGGCATCTTGT TCCATTGTTATGCCTGACACGCTTCCAGCGGCCCACAGGCCGTTGT GCGCTGGTAGGCGGCGCGCGGGCCCACGGGGCCACGTTGGCGGTTCCAGAAAAAAAA	CGACTGGCGCAAGGCCCGATCGCACTCACGTGTGGCTTCGATCCTACCG CTGACAGCTTTCCATTTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGC TTCCAGCAGCGGGCCACTCTTGTTCCATTGTTATGCCTGAAACGC TTCCAGCAGCGGGCCACAGCCGGTTGCGCTGGTAGGCGGCGCGACG GGTCTGATTGGCGACCCACAGCCGGTTGCGCTAAGCCGCGCGACG ACCAAGAAACTGTTCGAGCTTCCAAAGCTGCCTAACGCTGAAC ATAATTATGACTGTTCGGCAATATGAATGTGCTGACCTTCCTGCGGTT AAATTATGACAACCTTCTCCGTTAACCAGGGTT AAATTGCAACCGTTCCGTT
25			ACTGGGGG ACAGCTTC CAGCAGG TCTGATTG TCTGATTCC ACTGTTCC ATTATGAC TGCAACC TACAACC TACAACC TACAACC TACAACC TACAACC TACAACC TACAACC	AGCAGAGG CCTACCGC GAAACGCT GCGGAACA GCGGCAACA AGCGGTTA AGCGGTTA CATCACTT	SCCCGATO TTGGGGCA SCCCGAGG TTCGGAGT TTCGGCA TTCGGCA TTCGGCA TCTCGGTA TCTCGGTA TCTCGGTA TCTCGGTA TCTCGGTA TCTCGGTA TCTCGGTA
30	ión)	A	CAGAGCO/ TACCGCTC CGACGCG CGACGGG TGACGGG TGACATA TGCGTTTCC AGTTTTCC CAGTACGC CAGTACGC	BACGAGGAAGCGTT/ ATTGTGGCTTCGAT TCCATTGTTATGCCTG GCGCTGGTAGGCGGG GCTGCTGGTAGGCGGG GACAAATCGTAAC AAAACTCTGCTATCG TGTGCTGCTATCG TGTGCTGCTACCT GGGATTTCGTTCACT GCCAGTGGGGAAAAAAAAAA	GCGCAAGG GCTTGCAT CAGGCGGG ATTGGCGA ATGAACTG CGTCTCAA CGTCTCAA ATCGCTG GGTGCTGC
35	(continuación)	SECUENCIA	GCGTTAG TCGATCC TCCATCC GGCGCC CGTAAG CGTACCC CTCTCGC CTTCCTG AAGAAG AAGAAG AAGAAG	GACGAGG ATTTGTGG CCCATTG GCCATGG GCCACAAA AAAACTC AAAACTCA ATGATCA GGGATTT GGGATTT GGCATTT	CGACTG CTGACA TTCCAGG GGTCTGA ACCGGA ATAATT TATTGGA AAGCAG CCTACA CGGTGT
40					
45					
50			sintetasa	sintetasa	tio activo) sintetasa
55		Marcador	pAzidoFe-4 (sitio activo) Polinucleótido sintetasa	pAzidoFe-5 (sitio activo) Polinucleótido sintetasa	pAzidoFe-6 (sitio activo) Polinucleótido sintetasa
60			SEC ID Nº 23 PA	SEC ID Nº 24 Po	SEC ID Nº 25
65		SEC ID Nº	SEC II	SECI	SEC ID

5		AGCGA CCGCTG CCGGGT CCGGGT	TGCC AATA ATAT TTAA TTCC TACG	GCTG GCTG GCTG GCTC GGGT CACC TGCC AATA TTCC TACG
10		CGGGGGCTGGTANCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGA CTGGCGCAAGGCCCGATCGCACTCGGGTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAAACGCTTC CAGCAGGCGGGCCACAGCCGGTTGCGCTGGTAGGCGGCGCGACGGGT CTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACCCC	GAAGAAACTGTTCAGGAGTGGGTGGACAAATCCGTAAGCAGGTTGCC CCGTTCCTCGATTTCGACTGTGGAGAAACTCTGCTATCGCGGCCAATA ATTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCCTGCGGGATAT TGGCAACACTTCTCCGTTAACCAGATGATCAACAAAGAGCGGTTAA GCAGCGTCTCAACGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCC TACAACCTGCTGCAGGGTTATTCTATGGCCTGTTTGAACAAACA	CGGGGGCTGGTAGCCCAGGTGACGAGGAAGCGTTAGCAGAGCGAAACCGTGACGGCCCAAGGCCAACGCTCACGTGTGGCTTCGATCCTACCGCTGACGCCTGACGCCTGACGCCTGACGCTGACGCTGACGCTGAAACGCTTCCAAGGCTTGGTTATGCCTGAAACGCTTCCAGGCTTGGTTATGCCTGAAACGCTTCCAAGCCGGTTGCCTGGTAGGCGGCGGGGGGGTTGAAAACTCCTTCAAAGCTGCAAAACCGTTAAAACTGTTCAGAGTTGGAGAAAACTCTTCCGTAAGCTGGAAAACCCTTCCTGCTTCCAACCAA
15		GAGGAAGC TGTGGCTTC ATTGTTATC CCTGGTAGG	AAATCCGT CTCTCCCTA CTCTCCCTA NTCAACAAA ATTTCGTTTGA CTGTTTTGA CCAGTGGGG	AGGAAGCC STGGCTTCC TTGTTATG TTGTTAGG CGGGGTAGG AAATCCGT ACTCTGCTA ATTCGTTA ATTCGTTA CCAAAATC
20		TGACGGAC CACTCGGG TCTTGTTCC XCGGTTGCC TCAAAGCT	SAAGAAACTGTTCAGGAGTGGGTGGACAAAACTCCGTAAG CCGTTCCTCGATTTCGACTGTGGAGAAACTCTGCTATCG ATTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCCT TGGCAAACACTTCTCCGTTAACCAGATGATCAACAAAGAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACT TACAACCTGCTGCAGGTTATTCTATGGCCTGTTTGAACA GTGTGGTGCTGCAAATTGGTGGTTCTGACCAGTGGGGTAA	CGGGGGCTGGTAGCCCAGGTGACGGACGAGGAAGCGTTAG CTGGCCCAAGGCCCATCTCTTGTTCCATTGTTATGCCTGA ACAGCTTGCATTTGGGCACTCTGTTCCATTGTTATGCCTGA CAGCAGGCGGCCACAGCCGGTTGCTGGTAGGCGGCG CAGCAGCGGGCCACAGCCGGTTGGTAGGCG CCGTTCCTCGACCGGTTGGAAAATCCGTAAGC ATTATGACTGTTCGACTGTGGAAAATCGCTATCGC ATTATGACTGTTCCGTTAACCAGATGATCAACAAAGAAC GCAACCTTCTCCGTTAACCAGATGATCAACAAAGAAC GCAGCGTTCAACCTGAAGATTCGGTTCACTG
25		ANCCCAGG CCCGATCG TTGGGGCA CCACAGC	CAGGAGTG TCGACTGT TCGACTGT TCGGCATA TCGGCATA TCGGCATA TCGGCATA TCGGTTA TCGGTTA TCGTC	SCCCAGGT SCGATCGC SCGACCAT CACAAGCC SCGAGCTT CAGGACTGT TCGGTAA SCGTGAAG SAGGGTTAA AAATTGGT
30	(continuación)	GGCTGGT GCGCAAGO GCTTGCAT CAGGCGGG	AAACTGTT CCTCGATT GACTGGT VAACACTT CCTCTCAA NCCTGCTGCAA NCCTGCTGCA	GCCTGGTA GCCAGGG TTGCATTI AGGCGGG ATGCGACTI AAACTGTTI AAACACTTI AAACACTTI AACACTTI AACACTTI AACACTTI AACACTTI AACACTTI AACACTTI AACACTTI AACACTTI AACACTTI AACACTTI AACACTTI AACACTTI AACACTTI
35	(conti	CTGG ACAC CTGA	0AAG CCGTT ATTAT TGGCA GCAGG TACAA TGGTG	CGGGG CTGGG CAGC/ CTGA1 CCGTT CCGTT ATTA1 TGGC/ TACA/ TACA/
40		anina		
45		giloxi fenilalar vo)	es es	asa
50	dor	pPR-EcRS-1 (propargiloxi fenilal sintetasa) (sitio activo)	Polinucleótido sintetasa	pPR-EcRS-2 (sitio activo) Polinucleótido sintetasa
55	Marcador	pPR-EcF sintetas	Polinuc	pPR-Ecf Polinucl
60	SEC ID Nº	SEC ID Nº 26		SEC ID Nº 27
65	ď	l R		<u>v</u>

	Î			1	
5			GCGA CCTG CCTTC SGGT ACC MATA TAAA TTCC TTCC	GCCA GCTG GCTTC GCTTC ACC ACC ACC TGCC TAAA TTCC CTTC	GCCA SCTTC SCTTC SCGT SCGT SCC SCGT SCC SCGT SCC SCGT SCC SCGT SCC SCGT SCC SCGT SCC SCGT SCC SCGT SCC SCGT SCC SCGT SCC SCGT SCC SCGT SCC SCGT SCC SCGT SCC SCGT SCC SCGT SCC SCC SCC SCC SCC SCC SCC SCC SCC SC
10			TTAGCAGA ATCCTACCG GCTGAACG GCGCCGACG AGCTGACG AGCGGCCA CGCGGCCA CGCGGCCA CGCGGCCA CGCGGCCA CGCGGCG CAACAGGGTT ACAAACAGGTT	TTAGCAGA ATCCTACO CTGAAACG GGCGCGCAC AGCTGAGT CGCGGGCC CTGCGGCG GAAGCGGT ACTGAGTT ACTGAGTT TAACATCA	TTAGCAGA ATCCTACC CTGAAACG GCCGCGACG AGCTGAAC CAGCGGCC CTGCGCC CTGCGCC ACTGAGTT ACTGAGTT ACTGAGTT ACTGAGTT ACTGAGTT
15			IGGAAACG TGGCTTCGA TGGCTTCGA GGTAGGC GGTAGCGTA VAATCCGTA TTCGTTC TTCGTTC TTTCGTTC TTTCGTTC TTTCGTTC TTTTCGTTC TTTTCGTTC	AGGAAGCG CGGCTTCG CGGCTTCG CGGTAGGC CGGTAGCGTA AAATCCGTA CTCTCGTA CTCTCGTA ATTCGTTC CAGTGGGG	VGGAAGCG TGGCTTCG TGTTATGC GGTAGGCC CGAGCGTA AATCCGTA CTCTGCTAT CTCTGCTAT CTGTTGTA VGAATCAGG
20			ACGGACGA CTCTCTTGT TGTTCCAT GGTTGCGCT NAGGCTGC GAGGACAA GGAAAAA TCAGATGA TCAGGGAA ACGATGGC STTCTGACC	ACGGACGA CTCGCGTG TGTTCCAT GTTGCGCT AGGCTGC AGGCAAAA GAGAAAAA GAGAAAAA GAGAAAAAAAA	ACGGACG CTCGCGTG TTGTTCCAT GTTGCGCT AAAGCTGC GTGGAAAA GAGAAAAA GAGAAAAAA GAGAAAAAAAA
25			CGGGGGCTGGTACCCCAAGTGACGGACGAGGAAACGTTAGCAGAGCGA CTGGCCCAAGGCCCCATCTCTTGTGGCTTCGATCCTACCGCTG ACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCCTGAACGCTG CAGCAGGCAGCCCACAAGCCGGTTGGTAGGCGGCGCGCTG CTGATTGGCGACCCGAGCTTCAAAGCTGCTGGTAGCCGACGGT CTGATTGGCGACCCGAGCTTCAAAGCTGCCGAAGCTGCC GAAGAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGC CCGTTCCTCGATTTCGGTGGAGAAACTCCGTAAGCAGGTTGC ATTATGACTGGTTCGCAATATGAAATGCTGCTACCTTCCTGCGGGTTAA ATTATGACTGGTTCGCTTAACCAGATGATCAACAAAGAGGGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGGTTAA GCAGCGTCTCAACCGTGAAGATGGGGGTTAA TACAACCTGCTGCAAATTGGTGGTTCTGACCATTGGTGAACATCCTTC	CGGGGGCTGGTAGCCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGAACCGAGAGCGAAGCCCAAGGCCAATCGTTCGT	CGGGGGCTGGTAGCCCAGGTGACGGAGGAAGCGTTAGCAGAGCGAAGCGAAGCGTAGCGGGGCAAGCCTGGGGGGGTGTGGCTTCGATCCTACCGCGGACGGGGCAAGCTTGGTTATGCCTTGATCCTACCGCTGCAGCGTGGGGGGCATCTTGTTATGCCTGAAAGCTTGCTGGTAGGCGGCGGCGGCTTCCTGGTTGGCGGCGGCGGGGGGGG
30	(continuación)	ICIA	GCTGGTAC GCAAGGCC TGCATTTC GGCGACCC GGCGATTTC ACTGTTC ACTGTTC ACCTTCAAC GTCTCAAC GTCTCAAC	GCTGGTAG GCAAGGCC TTGCATTT GGCGGCC FGCGGCC FGCGGTT AACTGTTC AACTGTTC AACACTTC AACACTTC AACACTTC AACACTTC AACACTTC AACACTTC AACACTTC AACACTTC	GCTGGTAG GCAAGGCC TTGCATTT GGCGGGCC TGGCGACC AACTGTTC AACTGGTT AACTGGTT AACTGGTT GTCTCAAC GTCTCCAC
35	(contin	SECUENCIA	CGGGG CTGGCA CTGGCA CAGCA CCGTTC CCGTTC CCGTTC CCGTTC CCGTTC CCGTTC CCGTTC CCGTTC CCGTTC CCGTTC CCGTTC CCGTTC CCGTTC CCGTTC	CGGGG CTGGCG CTGGCA CTGAT GAAGA CCGTTG ATTATG TGCA TGCA	C00000 C10000 ACAGC C10000 C100000 C1000000 C100000000
40					
45					
50		99	pPR-EcRS-3 (sitio activo) Polinucleótido sintetasa	pPR-EcRS-4 (sitio activo) Polinucleótido sintetasa	pPR-EcRS-5 (sitio activo) Polinucleótido sintetasa
55		Marcador	pPR-EcRS-3 Polinucleóti	pPR-EcRS-4 Polinucleóti	pPR-EcRS-5 Polinucleóti
60		SEC ID Nº	SEC ID Nº 28	SEC ID Nº 29	SEC ID Nº 30
65	ļ	S	55	5	<u>N</u>

	1		-	2	
5			GCTG GCTG GCTG GCTG GGGT TGCC AATA ATAT TTCC TACG	GATC SGGC SGGC SGAG SGGA SGGT SGGT ATTG CCCG	GCTG GCTG GCTTC GCTTC AATA ATAT TTCC CTTC
10			TTAGCAGA ATCCTACC CTGACACC GGCGCGAC AGCCGGGC CCTGCGCC CCTGCGCC CCTGCGCC GAAGCGG ACAAACAG	CAAGGCCC TGCATTTGC TGCGGGCCC TGCGGTTC ACTGGTTC ACTGGTTC ACTGGTTC TCTCACC TCTCACC TGCTGCAA	TTAGCAGA ATCCTACO CTGACAGO GGCGCGAC AGCTGAAC AGCTGAAC AGCTGAAC AGCTGAAC ACTGAGGG ACTGAGGG ACTGAGGG ACTGAGGG ACTGAGGG ACTGAGGG ACTGAGGG ACTGAGGG ACTGAGGG
15			GGGAAGCG TGGCTTCG TGTTATGC TGTTGCTA VAATCCGTA CTCTGCTA TCACAAA TTTCGTTC TTTCGTC TTTCGTC TTTCGTC	SACTGGGG TGACAGCT TCCAGCACT GGTCTGATT GGTCTGATT ATTGGCAA AGCAGCG CCTACAAT CCTGGTATT	NGGAAGCG TGGCTTCG TGTTATGC TGTTATGC TGTACGCTA TCACCATA TCACCAAA TTTCGTTC TTTCGTTC TTTCGTTC TTTCGTTC TTTCGTTC TTTCGTTC
20			ACGGACGA CTCACGTG CTCACGTG GTTGCGCT NAGCTGC GTGGACAA GAGATGA CCAGATGA TCAGGGGA ACGTTTGCC	GCAGAGCC CTACCGCT SATACGCT GCGACGC GCGCGCAA GCGGCCAA GCGGCTAA ACGGTTA AACGGTTA AACGGTTA	ACGGACGA CTCGTTTG TGTTCCAT GTTGCGCT AAGCTGC GAGAAAA GAGATGTG CGAATGTGA TCAGGGGA TCAGGGGA TCAGGGGA TCAGGGGA
25			CCCAAGTG CGATCGCA GGGCATCC ACAGCCG ACGACTTC AGGACTTC CCGTTAAC CGTTAAC CGTTAAC CGTTAAC	AAGCGTTA DCTTCGATC TTATGCCT TTATGCCTAA SACCGTAA ATCCGTAA ACCAAGA ACCAAGA ACCAAGA TCGTTCAC TCGTTCAC TCGTTCACA	CCCAGGTG CGATCGCA GGGCATCT ACAGCCG GGGCATCT AGGCTTC CGCATTA CCGTTAAC CGGTTATT AATTGGTG ACCGTCGT
30	ición)	CIA	CGGGGGCTGGTACCCCAAGTGACGGACGAGGAAGCGTTAGCAGAGCGAAACCGTGACGGCCAAGGCCCAAGGCCACTCACCGTGTGGCTTCGATCCTACCGCTGACGCTGCGCTGACGCTGCGCTGACGCTGCTGCCTACCGCTGCAAGGCCGCTGCAAGGCCGCTTCCATTGTTATGCCTTGAACGCTTCCAAGGCGGGCG	GTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCCGAAGGCCCGATC GCACTCACGTGTGGCTTCGATCCTACCGCTGACAGCCTTGCATTTGGGG GCACTCACGTGTGGCTTCGATCCTACCGCTGACAGCTTTGCATTTTGGGC ATCTTGTTCCATTGTTATGCCTGAAACGCTTCCAGCAGGCGGGCCCGAG GCCGGTTGCGCTGGTAGCCGCGCTCCAGCAGAACTGTTCAGG CTTCCAAAGCTGCTGGTAAGCTGAACACGCTTGCTGCTGAGA GTGGGTGGACAAATCCGTAAGCAGGTTGCCCGTTCCTCGGCA ATATGAAACTCTGCTACGCGGCCCAATAATTATGACTGGTTCGGCA ATATGAAACTCTCCTGCGGCGGTTAAGCAACCTTCTCCGT TAACCAGATGTGCTTCACTGAGTTTTCCTACAACCTTCTCCGT AATTCGGCTGCTTTAACAAAGAAGGGGTTAAGCAGTTCTGCTGCAAAATTG GTGGTTCTGACCTGTTTAACAAACAGTACGTTCTGGTACCGG TCGTCTGCAAATCTGGTAACAAATTG	CGGGGGCTGGTAGCCCAGGTGACGGAGGAAGCGTTAGCAGAGGCGAACGCTGGCGCAAGGCCCAAGGCCATCGTTTGTTGGTTCGATCCTACCGCTGGAAGCCTTGGATCCTACCGCTGGCTG
35	(continuación)	SECUENCIA	CGGGGG CTGGCC ACAGCA CAGCA CTGATT CCGTTC ATTATC ATTATC TGGCA GCAGCG TACAAC	916ACC 9CACTC 9CCGGG 9CTTCAA 9TGGGG 1ATTCC 1ATTCC 9TGGT 1CTTCC 1ATTCC 1CTTCC	CGGGGC CTGGCC ACAGCT CAGCATT CAGCATT CGGTTC ATTATG TGCAAC GCAGCC TACAAC
40					
45					
50			pPR-EcRS-6 (sitio activo) Polinucleótido sintetasa	pPR-EcRS-7 (sitio activo) Polinucleótido sintetasa	(sitio activo) lo sintetasa
55		Marcador	pPR-EcRS-6 Polinucleótic	pPR-ECRS-7 Polinucleótik	pPR-EcRS-8 (sitio activo) Polinucleótido sintetasa
60		SEC ID Nº	SEC ID Nº 31	SEC ID Nº 32	SEC ID Nº 33
65		SE	S	SE	S

5		GAGCGA CCGCTG CGCTTC ACGGGT ACGGGT ACGGGT ACGGGT ACTTGC AGTACG AGTACG CACTTC	GAGCGA CCGCTG ACGCTTC ACGCGT ACGCGT ACGCGT ACGCGT ACGCTTC CCAATA CGATAT GGTTAA HTTTCC AGTACG	SSLHLGH FYQEWV IGGSDQ LDPKKTS RAQYVL PMVEM KEEDRLF
10		GTTAGCA GATCCTA GCTGAAA GCTGCGG TAAGCTGA TAGCGGG ATCGCGG ATCGCGG ACCTGAG GTAACATG	GTTAGCA GGTCCTGAA/ GCCTGAA/ GCGCGCG TAAGCTG/ TACGCGG ATCGCGG ATCGCGG TCCTGCGG AGAAGC TCCTGCGG	CGFDPTAI RKLNTEE FLRDIGKH QYGVVLQ TEGGAVW KNSGKAP QCDPEYFF
15		SGGGAAGG GTGGCTTG TTGTTATG TTGTTAGG TGGAGGG NAATTCGT ATCAACA/ ATCAACA/ ATCAACA/ ATCAACA/ ATCAACA/ ATCAACA/ ATCAACA/ ATCAACA/ CTGTTGG	SAGGAAGG GTGGCTTG CTGGCTAGG CCGAGCG AACTCTGCI GCTGACCC ATCACCC ATCACCC CCTGTAGG CCTGTAGG	AQGPIALN GNMNYL1 SGNMNYL1 SGTKFGK ENALEEEI SALSEADFI AAITINGEK
20		GACGGACC ACTCAGTT GGTTGCGC GGTTGCGC GGTGGACA GGAGAAA NTGAATGT NCCAGATG ATCAGATG ATCAGATG	GACGGACC ACTCACGI CITTGTTCCC CIGGTGCGC CIGGTGGAAA CIGAGAAAA CIGAGAAAAAAAAAAAAAAAAAAA	GEALAERI GGATGLIG GNANYDWE SYNLLQGY LTVPLITKA FFTFMSIEI ECLFSGSL QARKTIASN
25		ACCCAAGT GGGGCAT GGGGCAT CACAGCCT CACAGCT CACACT CACACT CACAT CACAT ACCATA ACCATA ACCATA	SCCCAGGT CGGGCCAT CGGGCCAT CACAGCCT CGGCAT CGGCAT TCCGTTAA TCCGTTAA CGGCAT CGGCAT AGGGTTAA	SLVAQVTD SHKPVALV SCGENSALV DQGISFTEF HQNQVFG ADVYRFLK LQAAKRIT SELQPSRQC
30 (continuación)	NCIA	CGGGGGCTGGTANCCCAAGTGACGGACGGGGAAGCGTTAGCAGAGCGAACGCTGACGCGCGAACGCCTCACTTGTTGTTGCTTTCGATCCTACCGCTGAACGCTTCGAACGCTTCGAACGCTTCGCAAGGCCGAACGCTTCCAAGGCGGCGCGCTTCCAAGGCGGGCG	CGGGGGCTGGTAGCCCAGGTGACGGAGGAAGCGTTAGCAGAGCGAACGCAACGCGAAGCCAAGGCCAAGGCCAACGCTCTCACGTGTGGCTTCGATCCTACCGCTGAACGCTGCAAGGCCAAGGCCGATCGTGGCTTCGTTATTATGCCTGAAACGCTTCCATGCTTAGTTATTATGCCTGAAACGCTTCCAGGCGGGGGGGG	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALVCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGVSYACLNKQVGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEDKNSGKAPRAQYVL AEQVTRLVHGEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF
35	SECUENCIA	CGGGG CTGGC ACAGC CAGCA CTGAT GAAGA CCGTT ATTAT TGGCA TACAA TACAA	CGGGG CTGGC ACAGC CTGAT CTGAT GAAGA CCGTTA TGGCA TGGCA TACAA TGGTA	MASSN LVPLLC DKIRKC INKEAN WGNITY PYKFYC AEQVT EKGAD GRFTLI
40				
45		G m	(o, e	aa)
50		pPR-EcRS-9 (sitio activo) Polinucleótido sintetasa	pPR-EcRS-10 (sitio activo) Polinucleótido sintetasa	p-yodofeRS-1 Aminoácido sintetasa (aa)
55	Marcador	pPR-EcRS-	pPR-EcRS- Polinucleót	p-yodofeRS-1 Aminoácido s
60	SEC ID Nº	SEC ID Nº 34	SEC ID Nº 35	SEC ID Nº 36
65	SE	S	S	SE

			·		2-2
5		SLHLGH VQEWV FSVNQM GGSDQ DPKKTS AQYVL PMVEM	SLHLGH VQEWV FSVNQM IGGSDQ DPKKTS AQYVL PMVEM EEDRLF	SLHLGH VQEWV VQEWV GGSDQ DPKKTS AQYVL PMVEM	SLHLGH SVQEWV SSVNQM GGSDQ DPKKTS AQYVL PMVEM
10		GFDPTADS RKLNTEET LKDIGKHI 2YGVVLQI EGGAVWL KNSGKAPR QLAQDGVI	CGFDPTAD RKLNTEET PLRDICKHI QYGVVLQI EGGAVWL KNSGKAPR QLAQDGV	CGFDPTAD RKLNTEET LRDIGKHI 2YGVVLQI EGGAVWL KNSGKAPR QLAQDGV	GFDPTAD RKLNTEET LRDIGKHI 2YGVVLQI EGGAVWL ENSGKAPR CNSGKAPR QLAQDGVI
15		AQGPIALIC SPSFKAAEI SNMNVLTF SMACLNKC OGTKFGKT INALEEEDI ALSEADFE	VQGPIALVO DPSFKAAEI SINMINULTI SIMACANKO DGTKFGKT INALEEEDI ALSEADFE	NQGPIALVO PPSFKAAEI SNMNVLTF SMACLNKO SQTKFGKT INALEEEDI ALSEADFE	VQQPIALTO DPSFKAAEI SINMINULTE SINMACLINK OGTKEGKT INALEEED ALSEADFE
20		SEALAERL/ SGATGLIGI ANNYDWFG SYNLLGGY TYPLITKAI FFTFMSIEE SCLFSGSLS ARKTIASN	SEALAERLA SIGATGLIGI STANYDWFG SYNLLQGY SYNLLQGY SYNLLQGY SYNLLGGY SYNLLGGY STANSIEE SCLFSQSLS ARKTIASN	BALAERL/ SGATGLIGI SGATGLIGI SYNLLGGY TYPLITKAI FTTFMSIEE SCLFSGSLS ARKTIASN	BALAERLA BGATGLIGI BGATGLIGI BYNLLQGY TYPLITKAI FTFMSIEEI FTFMSIEEI CCLFSGSLS
25		LVAQVTDE HKPVALVC CGENSALA CGENSTEFS QQNQVFGL DVYRFLKI QAAKRITE ELQPSRGQ	LVAQVTDE HKPVALVC CGENSAIA. QGISFTEFS QQNQVFGL QDVYRFLKI QAAKRITE ELQPSRGQ	LVAQVTDE HKPVALVC CGENSAIA QGISFTEFS QQNYRFLKI QAAKRITE ELQPSRGQ	LVAQVTDE HKPVALVC CGENSALV QGISFTEFS QQUSFTEFS QDVYRFLKI QAAKRITE ELQPSRGQ
30 ación)	CIA	MASSNLIKQLQERGLVAQVTDEĘALAERLAQGPIALICGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSMACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDFEYFFKEEDRLF GRFTLLRRGKKNYCLICWK	MASSNLÍKQLQERGLVAQVTDEEALAERLAQGPIALVCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM NKEAVKQRLNREDQGISFTEFSYNLLQGYSMACANKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALVCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM NKEAVKQRLNREDQGISFTEFSYNLLQGYSMACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK	MASSNLÍKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYTMACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEDKNSGKAPRAQYVL AEQVTRLVHGEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF
35 (continuación)	SECUENCIA	MASSNL LVPLLCI DKIRKQ INKEAVI WGNITS PYKFYQ AEQVTR EKGADL GRFTLLI	MASSNL LVPLLCI DKIRKQ INKEAVI WGNITS PYKFYQ AEQVTR EKGADL GRFTLLL	MASSNL LVPLLCI DKIRKQ INKEAVI WGNITS PYKFYQ AEQVTR EKGADL GRFTLLI	MASSNL LVPLLCI DKIRKQ INKEAVI WGNITS PYKFYQ AEQVTR EKGADL GRFTLLI
40					
45		G.	6	0	(e
50		p-yodofeRS-2 Aminoácido sintetasa (aa)	p-yodofeRS-3 Aminoácido sintetasa (aa)	OMeTyrRS-1 Aminoácido sintetasa (aa)	OMeTyrRS-2 Aminoácido sintetasa (aa)
55	Marcador	p-yodofeRS-2 Aminoácido s	p-yodofeRS-3 Aminoácido s	OMeTyrRS-1 Aminoácido	OMeTyrRS-2 Aminoácido
60	SEC ID Nº	SEC ID Nº 37	SEC ID Nº 38	SEC ID Nº 39	SEC ID Nº 40
65	SE	S	SE	S	- S

						7
5			HLGH VNQM SSDQ PKKTS QYVL AVEM EDRLF	HLGH QEWV VNQM SSDQ PKKTS QYVL TVEM EDRLF	HLGH OEWV VNQM GSDQ PKKTS QYVL TVEM EDRLF	HLGH QEWV VNQM GSDQ PKKTS QYVL TVEM EDRLF
10			SFDPTADSL KLINTEETV RDIGKHFS GGVVLQIGG GGAVWLD NSGKAPRA LAODGVPN	JFDPTADSI KLNTEETV KDIGKHFS YGVVLQIG YGAVWLD NSGKAPRA LAQDGVPN IDPEYFFKE	JFDPTADSI KLNTEETV KDIGKHFS YGVVLQIG GGAVWLD NSGKAPRA LAQDGVPN IDPEYFFKE	JFDPTADSI KLNTEETV KLNTEETV ROJCKHFS YGVVLQIG GGAVWLD NSGKAPRA LAQDGVPN IDPEYFFKE
15			QGPIALTCC PSFKAAERI NMNVLTFI YACLNKQ1 GTKFGKTE GTKFGKTE AALEEEDK1 LSEADFEO	QGPIALLCC PSFKAAER NMNVLTFI NMCSNKQ GTKFGKTE VALEEDKI LSEADFEQ	QGPIALLCC PSFKAAERI NMNVLTFI NMNVLTFI GTKFGNKQ GTKFGKTE VALEEEDKI LSEADFEQ	QGPIALTCC PSFKAAERI NMNVLTFI NMACLNKQ GTKFGKTE VALEEEDKI LSEADFEQ
20			SALAERLA(SATGLIGDI NNYDWFGI NNLQGYT NLLQGYT TFMSIEEN TLFSGSLSA RKTIASNA	SALAERLAG GATGLIGDI NNYDWFGI NNLQGYSI VPLITKADI TFMSIEEN TFSGSLSA RKTIASNA	SALAERLAG SATGLIGDI NNYDWFGI NNLQGYSI VPLITKADI TTFMSIEEN SILFSGSLSA RKTIASNA	SALAERLAG SATGLIGDI NNYDWFGI NNLQGYR VPLITKADI TFMSIEEIN TFMSIEEIN
25			VAQVTDEE ENSALAG GENSALAA GENSALAA GENSFTEFS QUQVFGLT DVYRFLKFF QAAKRITEC LQPSRGQA	VAQVTDEE IKPVALVG GENSALVG GENSALAG GISFTEFSY QUQVFGLT DVYRFLKF QAAKRITEC LQPSRGQA	VAQVTDEE IRPVALVGG GENSAIAAI QGISFTEFSY QNQVFGLT QVARKUTEC LQPSRGQA	VAQVTDEE IKPVALVGG GENSAIAAI QGISFTEFSY QNQVFGLT QVYRFLKFF QAAKRITEC LQPSRGQA
30	ación)	CIA	MASSNLÍKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM NKEAVKQRLNREDQGISFTEFSYNLLQGYTYACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEOVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEOLAODGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALLCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSMACSNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK	MASSNLÍKQLQERGLVAQVTDEEALAERLAQGPIALLCGFDFTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM NNKEAVKQRLNREDQGISFTEFSYNLLQGYSMACANKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF	MASSNLÍKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYRMACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF
35	(continuación)	SECUENCIA	MASSNLI LVPLLCL DKIRKQV INKEAVK WGNITSC PYKFYQF AEOVTRI EKGADLI GRFTLLR	MASSNLJ LVPLLCL DKIRKOV INKEAVK WGNITSC PYKFYQI AEQVTRI EKGADL GRFTLLR	MASSNLI LVPLLCL DKIRKQN INKEAVK WGNITSC PYKFYQI AEQVTRI EKGADL GRFTLLR	MASSNLI LVPLLCL DKIRKQN INKEAVK WGNITSC PYKFYQI AEQVTRI EKGADL
40						
45						
50			OMeTyrRS-3 Aminoácido sintetasa (aa)	OMeTyrRS-4 Aminoácido sintetasa (aa)	OMeTyrRS-5 Aminoácido sintetasa (aa)	OMeTyrRS-6 Aminoácido sintetasa (aa)
55		Marcador	OMeTyrRS-3 Aminoácido	OMeTyrRS-4 Aminoácido	OMeTyrRS-5 Aminoácido	OMeTyrRS-6 Aminoácido
60		ēN QI	SEC ID Nº 41	SEC ID Nº 42	SEC ID Nº 43	SEC ID Nº 44
65	(3	SEC	SEC	SEC	SEC	SEC

			M.				
5			HLGH VNQM GSDQ PKKTS QYVL IVEM EDRLF	HLGH VNOM SSDQ PKKTS QYVL IVEM EDRLF	LHLGH YQEWY SVNQM GGSDQ OPKKTS AQYVL MVEM	LHLGH /QEWV SVNQM	SSDQ PKKTS QYVL IVEM EDRLF
10			FDPTADSLI TINTEETVC RDIGKHFS YGVVLQIG GGAVWLDI SGKAPRA LAQDGVPN	ifDPTADSL (LNTEETV (LNTEETV GVVLQIGG GGAVWLDI ISGKAPRA LAQDGVPN DPEYFFKE	GFDPTADS UKLNTEETY LRDIGKHF QYGVVLQI GGGAVWLI NSGKAPR QLAQDGVP SDPEYFFKI	GFDPTADS KLNTEETV LRDIGKHF	GVVLQIGG SGAVWLDI SGKAPRA(AQDGVPN
15			GPIALICGI SFKAAERK MANVLTFL MACANKQ' STKFGKTE ALEEEDKN ALEEEDKN TINGEKQS	GPIALGCC SFKAAERN AMNVLTFL FACANKQY FACANKQY FACANKQY FACANKQY TINGEEDKN TINGEKQS	QGPIALGC PPSFKAAER SNMNVLTF SYACMNK GTKFGKT NALEEEDK NLSEADFEC	QGPIALLC PSFKAAER INMNVLTF	TACANKOY TKFGKTEC ALEEEDKN SEADFEQI TINGEKQSI
20			ALAERLAC IATOLIGDP INYDWFGD NLLQGYGI VPLITKADC TFMSIEEN LFSGSLSAI KTIASNAI	ALAERLAG NATGLIGDP NAYDWFGD NLLQGYGI VPLITKADG TFMSIEEIN LFSGSLSAI	EALAERLA GATGLIGD INNYDWFG YNLLQGYG YNLLQGYG YNLTKAD FTFMSIEEI CLFSGSLS/	EALAERLA GATGLIGD NNYDWFG	NLLQGYSN VPLITKADC TFMSIEEIN LFSGSLSAI RKTIASNAI
25			VAQVTDEE KPVALVGG GENSAIAAN GISFTEFSY 2NQVFGLT VYRFLKFF 2AAKRITEC LQPSRGQAI	VAQVTDEE KPVALVGC GENSAIAAN GISFTEFSY 2NQVFGLT VVYRFLKFF AAKRITEC LQPSRGQAI	LVAQVTDE HKPVALVG CGENSAIAA QGISFTEFS IQNQVFGL, DVYRFLKF QAAKRITE ELQPSRGQ,	LVAQVTDE HKPVALVC CGENSAIAA	GISFTEFSY NVYRFLKFF NVARFLKFF NAAKRITEC LQPSRGQA:
30	ión)	Ι	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALICGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSALANYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYGMACANKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALGCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYGFACANKQYGVVLQIGGSDQ WGNITSGIDLTRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK	MASSNLÍKQLQERGLVAQVTDEEALAERLAQGPIALGCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYGVACMNKQYGVVLQIGGSDQ WGNITSGIDLTRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALLCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM	INKEAVKQRLNREDQGISFTEFSYNLLQGYSMACANKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
35	(continuación)	SECUENCIA	MASSNLII LVPLLCLI DKIRKQV INKEAVK WGNITSG PYKFYQF AEQVTRL EKGADLIN GRFTLLRI	MASSNLII LVPLLCLI DKIRKQV INKEAVK WGNITSG PYKFYQF AEQVTRL EKGADLA GRFTLLRI	MASSNL LVPLLCI DKIRKQ INKEAVI WGNITS PYKFYQ AEQVTR EKGADL GRFTLLI	MASSNL LVPLLCI DKIRKQ	INKEAVK WGNITSG PYKFYQF AEQVTRL EKGADLM GRFTLLR
40							
45			2000	673-2	-		
50			p-acetilFeRS-1 Aminoácido sintetasa (aa)	p-benzoilFeRS -1 Aminoácido sintetasa (aa)	p-benzoilFeRS -2 Aminoácido sintetasa (aa)	p-azidoFeRS-1 Aminoácido sintetasa (aa)	
55		Marcador	p-acetilFeRS-1 Aminoácido sir	p-benzoilFeRS - 1 Aminoácido sinte	p-benzoilFeRS -2 Aminoácido sinte	p-azidoFeRS-1 Aminoácido sii	
60		SEC ID Nº	SEC ID Nº 45	SEC ID Nº 46	SEC ID Nº 47	SEC ID Nº 48	
65		SE	SEC	SEC	SEC	SEC	

				= = = = = = = = = = = = = = = = = = = =	
5		LHLGH QEWV SYNQM GSDQ PRKTS QYYVL MVEM	CHLGH QEWV QEWV GSDQ GSDQ PKKTS QYVL MVEM	LHLGH QEWV VYNQM GSDQ PRKTS QYYL MVEM	HLGH QEWV VNQM GSDQ PKKTS QYVL AVEM
10		JFDPTADS CLNTEETV RDIGKHFS TGVVLQIG GGAVWLC SGKAPRA LAQDGVPI DPEYFFKE	FDPTADSI CLNTEETV RDIGKHFS GGAVWLD GGAVWLD SGKAPRA LAQDGVPI DPEYFFKE	JEDPTADS CLNTEETV RDIGKHFS GGAVWLQIG GGAVWLQIG SGKAPRA LAQDGVPI DPEYFFKE	FDPTADSL CLNTEETV RDIGKHFS 'GVVLQIG GGAVWLD SGKAPRA LAQDGVPI DPEYFFKE
15		GPIALVCC SFKAERI IMNVLTFL ACANKQY TKFGKTE ALEEDKY TINGEKQS	GPIALLCG SFKAAERI IMNVLTFL IACANKQY ITKFGKTE ALEEBKN SEADFEQI	KOPIALVCC SFKAERI MNVLTFL ACVNKQY STKFGKTE ALEEDKN SEADFEQI	GPIALICGI SFKAAERK SFKAAERK IMNVLTFL SACVNKQY STKFGKTE STKFGKTE SEADFEQI TINGEKQS
20		ALAERLAG SATGLIGDP NILYDWFGN NILQGYSA VPLITKADC TFMSIEEIN TFSSSLSAI	ALAERLAG SATGLIGDP SATGLIGDP NILLQGYSA VPLITKADG TFMSIEEN TFMSIEEN RKTIASNAI	ALAERLAC SATGLIGDP NILLGGYSA NILLGGYSA VPLITKADC TFMSIEEIN ILFSGSLSAI	ALAERLAQ DATGLIGDP DAYDWFGN NLLQGYNF VPLITKADG TFMSIEEIN LFSGSLSAL
25		VAQVTDEE KPVALVGG SENSAIAA) GISFTEFSY NQVFGLT VYRFLKFF AAKRITEC LQPSRGQAI	VAQVTDEE KPVALVGG GENSAIAA) GISFTEFSY NQVFGLT VYRFLKFF AAKRITEC QPSRGQAI	VAQVTDEE KPVALVGC GENSAIAA) GISFTEFSY NQVFGLT VYRFLKFF AAKRITEC QPSRGQAI	VAQVTDEE KPVALVGG JENSAIAAN GISFTEFSY ZNQVFGLTF VYRFLKFF VYRFLKFF CAAKRITEC QAAKRITEC
30 (r.)	4	MASSNLÍKQLQERGLVAQVTDEEALAERLAQGPIALVCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSAACANKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALLCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM NKEAVKQRLNREDQGISFTEFSYNLLQGYSAACANKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALVCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSAACVNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALICGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANDYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYNFACVNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEGGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF
35 (continuación)	SECUENCIA	MASSNLIK LVPLLCLK DKIRKQV, INKEAVK(WGNITSGI PYKFYQFI AEQVTRL EKGADLM GRFTLLRF	MASSNLIK LVPLLCLK DKIRKQV, NKEAVK WGNITSGI PYKFYQFY AEQVTRL EKGADLM	MASSNLIK LVPLLCLK DKIRKQV, INKEAVK WGNITSGI PYKFYQFY AEQVTRL EKGADLM GRFTLLRF	MASSNLIK LVPLLCLK DKIRKQV, INKEAVK WGNITSGI PYKFYQFY AEQVTRL EKGADLM GRFTLLRF
40					
45					
50		ntetasa (aa)	ntetasa (aa)	ntetasa (aa)	ntetasa
55	Marcador	p-azidoFeRS-2 Aminoácido sintetasa (aa)	p-azidoFeRS-3 Aminoácido sintetasa (aa)	p-azidoFeRS-4 Aminoácido sintetasa (aa)	p-azidoFeRS-5 Aminoácido sintetasa ,
60			45	to the second se	
65	SEC ID Nº	SEC ID Nº 49	SEC ID Nº 50	SEC ID Nº 51	SEC ID Nº 52

5		LHLGH YQEWV SVNQM GSDQ DPKKTS AQYVL MVEM EEDRLF	ILHLGH YQEWV SVNQM GGSDQ OPKKTS AQYVL MVEM	LHLGH VQEWV SVNQM SGSDQ DPKKTS AQYVL MVEM EEDRLF	LHLGH VQEWV SVNQM GGSDQ DPKKTS AQYVL MVEM EEDRLF
10		GFDPTADS KLNTEETV LRDIGKHFI YGVVLQIG GGAVWLI NSGKAPR NLAQDGVP SDPEYFKI	GFDPTADS KLNTEETA LRDIGKHF YGVVLQIC SGGAVWLI INSGKAPR QLAQDGVP SDPEYFFK	GFDPTADS KLNTEET TRDIGKHF PYGVVLQIC EGGAVWL CNSGKAPR QLAQDGVP SDPEYFFK	GFDPTADS WELNTEET TRDIGKHF 2YGVVLQII EGGAVWLI CNSGKAPR QLAQDGVP SDPEYFFKI
15		QGPIALTC PPSFKAAER BINMNVLTF SAACLINKQ SGTKFGKTI NALEEEDK ALSEADFEC	QGPIALGC SPSFKAAER SIMMIVLTF SMACLINKQ SQTKFGKTI NALEEEDK ALSEADFEC	VQGPIALTC SPSFKAAEI SNMNVLTF SAACLNKQ SGTKFGKT NALEEED! ALSEADFE	QGPIALSC PPSFKAAEI SINMNVLTF SITKFGKT NALEEEDI ALSEADFE
20		EALAERLA SGATGLIGE SGATGLIGE STANLLGGYS TYPLITKAL FTFMSIEEI CCLFSGSLS/	EALAERLA SGATGLIGE SGATGLIGE SYNLLQGY SYNLLQGY SYNLLQGY SYNLLQGY SYNLL SCLESGSLS ARKTIASN	EALAERLA SGATGLIGI SGATGLIGI SYNLLQGY YYNLLQGY TYPLITKAI TYPLITKAI TYPLITKAI TYPLITKAI TYPLITKAI TYPLITKAI	EALAERLA SGATGLIGI SGATGLIGI STANYDWFG YNLLQGY TVPLITKAI TYPLITKAI TTFMSIEEI CLFSGSLS/ ARKTIASN/
25		LVAQVTDE HKPVALVC CGENSAIA/ QGISFTEFS IQNQVFGL DVYRFLKF QAAKRITE ELQPSRGQ,	LVAQVTDE HKPVALVC CGENSAIA/ QGISFTEFS HQNQVFGL DVYRFLKF QAAKRITE ELQPSRGQ,	LVAQVTDE HKPVALVC CGENSAIAV QGISFTEFS IQNQVFGL DVYRFLKF QAAKRITE ELQPSRGQ,	LVAQVTDE HKPVALVG CGENSALA QGISFTEFS HQNQVFGL DVYRFLKF QAAKKITE CAAKKITE
30 (noión)	CIA	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSAACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALGCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSMACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSAACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALSCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYTMACVNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK
(continuación)	SECUENCIA	MASSNL LVPLLCI DKIRKQY INKEAV WGNITSK PYKFYQI AEQVTR EKGADL GRFTLLE	MASSNL LVPLLCI DKIRKQ' INKEAVI WGNITSG PYKFYQI AEQVTR EKGADL GRFTLLI	MASSNL LVPLLCI DKIRKQY INKEAVY WGNITS(PYKFYQI AEQVTR EKGADL GRFTLLK	MASSNL LVPLLCL DKIRKQ INKEAVI WGNITSG PYKFYQ AEQVTR EKGADL GRFTLLI
40			œ		
45		0	a) na sintetasa	(1)	2
50		s intetasa (aa	intetasa (ag ip henilalani	intetasa (aa	intetasa (aa
55	Marcador	p-azidoFeRS-6 Aminoácido sintetasa (aa)	pPR-EcRS-1 Aminoácido sintetasa (aa) p-propargiloxip henilalanina sintet	pPR-EcRS-2 Aminoácido sintetasa (aa)	pPR-EcRS-3 Aminoácido sintetasa (aa)
60	SEC ID Nº	SEC ID Nº 53	SEC ID Nº 54	SEC ID Nº 55	SEC ID Nº 56
65	SE	SE	SE	SE	SEC

5		LHLGH QEWV VNQM 3SDQ PKKTS QYVL AVEM EDRLF	LHLGH QEWV SYNQM GSDQ PPKKTS QYVL AVEM EDRLF	JHLGH QEWV WNQM GSDQ PKKTS QYVL AVEM EDRLF	JHLGH QEWV YNQM 3SDQ PKKTS QYVL AVEM EDRLF
10		JFDPTADSI KLNTEETV KDIGKHFS (GVVLQIGG GGAVWLD NSGKAPRA LAQDGVPA	GFDPTADSI KLNTEETV RDIGKHFS YGVVLQIG GGAVWLD NSGKAPRA LAQDGVPN	JFDPTADSI KLNTEETV RDIGKHFS YGVVLQIG GGAVWLD NSGKAPRA LAQDGVPN IDPEYFFKE	JFDPTADSI KLNTEETV KDIGKHFS GGVVLQIGG GGAVWLD NSGKAPRA LAQDGVPN IDPEYFFKE
15		QGPIALACO PSFKAAERI NMNVLTFL YACLNKQY GTKFGKTE NALEEEDKI LSEADFEQ	QGPIALACO PSFKAAER NMNVLTFI NACCNKQ GTKFGKTE JALEEEDKI LSEADFEQ	QGPIALTCC PSFKAAERI NMNVLTFI NMNVLTFI STKFGKTE GTKFGKTE IALEEEDKI LSEADFEQ	QGPIALTCC SFKAAERI NMNVLTFI NMNVLTFI STKFGKTE JALEEDKI LSEADFEQ
20		SALAERLAG SATGLIGDI NNYDWFGI NNLQGYS' VPLITKADI TFMSIEEN TFMSIEEN KKTIASNA	SALAERLAG SATGLIGDI NNYDWFGI NNLQGYT VPLITKADI TFMSIEEIN ILFSGSLSA RKTIASNA	SALAERLAG SATGLIGDI NNYDWFGI NNLQGYT VPLITKADI TFMSIEEN LFSGSLSA RKTIASNA	SALAERLAC SATGLIGDI NNYDWFGI NNLQGYS' VPLITKADI TFMSIEEN TFSGSLSA RKTIASNA
25		VAQVTDEE IKPVALVG GENSAIAAI GISFTEFSY QUSFTEFSY OVYRFLKF QAAKRITEC LQPSRGQA LCWK	VAQVTDEI IKPVALVG GENSAIAA QGISFTEFSY QNQVFGLT DVYRFLKFI QAAKRITEC LQPSRGQA	VAQVTDEE IRPVALVG GENSAIAA JGISFTEFSY QNQVFGLT DVYRFLKFF QAAKRITEC LQPSRGQA ICWK	VAQVTDEE IKPVALVG GENSAIAAI GENSAIAAI GISFTEFS QUQVFGLT QVYRFLKF QAAKRITEC LQPSRGQA
30 continuación	SIA	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALACGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSYACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWNTADADVYRFLKFFTFMSIEEINALEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALACGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM NKEAVKQRLNREDQGISFTEFSYNLLQGYTMACCNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF	MASSNLÍKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYTFACMNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK	MASSNLÍKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKINTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSVACLNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF
35	SECUENCIA	MASSNLI LVPLLCL DKIRKQV INKEAVK WGNITSG PYKFYQF AEQVTRL EKGADLA GRFTLLR	MASSNLI LVPLLCL DKIRKQV INKEAVK WGNITSG PYKFYQF AEQVTRI EKGADLN GRFTLLR	MASSNLI LVPLLCL DKIRKQV INKEAVK WGNITSG PYKFYQF AEQVTRL EKGADLA GRFTLLR	MASSNLI LVPLLCL DKIRKQV INKEAVK WGNITSQ PYKFYQF AEQVTRL EKGADLA GRFTLLR
40					
45				_	
50		pPR-EcRS-4 Aminoácido sintetasa (aa)	pPR-EcRS-5 Aminoácido sintetasa (aa)	pPR-EcRS-6 Aminoácido sintetasa (aa)	pPR-EcRS-7 Aminoácido sintetasa (aa)
55	Marcador	pPR-EcRS-4 Aminoácido :	pPR-EcRS-5 Aminoácido	pPR-EcRS-6 Aminoácido	pPR-EcRS-7 Aminoácido
60	SEC ID Nº	SEC ID Nº 57	SEC ID Nº 58	SEC ID Nº 59	SEC ID Nº 60
65	SE	S	- S	S	SE

		6								
5			LHLGH YQEWV SVNQM GSSDQ DPKKTS AQYVL MVEM	LHLGH YQEWV SVNQM GSDQW KKTSP QYVLA IVEME EDRLF	LHLGH YQEWV SVNQM GSDQ OPKKTS AQYVL MVEM	TGGTG	SUGGU			
10			GFDPTADS KLNTEETY LRDIGKHF YGVVLQIC GGGAVWLI NSGKAPR NSGKAPR SDPEYFFK	SFDPTADS KLNTEETY CROVLQIG GAVWLDP SGKAPRA AQDGVPN	GEDPTADS KLNTEETA REDIGKHF GGVVLQIG GGGAVWLI NSGKAPR NSGKAPR SDPEYFFK	TACCCCG CTGCCGT	NUACCCC AUCUGCC			
15			QGPIALVC PSFKAAER NMNVLTFI MACTNKQ GTKFGKTE NALEEEDK LSEADFEG	QGPIALSO PSFKAAER NMNVLTFI FACLNKQ) TKFGKTEG ALEEEDKN SEADFEQL TINGEKQSI	QGPIALTO PSFKAAER NMNVLTFI NGTKFGKTF NALEEEDK LISEADFEC	AAAGCAT STCTAAAT NCCA	AAAAGCAI ACUCUAAA ACCACCA	3CG-3	AAG-3'	TG-3
20			EALAERLA GATGLIGD INNYDWFG YNLLQGYS IVPLITKAD FTFMSIEEI CLFSGSLSA NRKTIASNA	EALAERLA GATGLIGD NNYDWFG YNLLQGYS PLITKADG TFMSIEEIN LFSGSLSAL KTIASNAI	EALAERLA GATGLIGD INNYDWFG YNLLQGYT IVPLITKAD FTFMSIEED CLFSGSLSA	GCCAGTAA GAGCAGA(CCCCCACCA	GCCAGUA GGAGCAG UUCCCCC	ACAAGCATG	GACTTAAA	втстссвс
25			JVAQVTDE HKPVALVG GENSATAA QGISFTEFS QNQVFGL DVYRFLKF QAAKRITE GLQPSRGQA	LVAQVTDE HKPVALVG CGENSAIAA QGISFTEFS NQVFGLTV VYRFLKFF AAKRITECI QPSRGQAI	LVAQVTDE HKPVALVG CGENSAIAA CGISFTEFS QQUSFTEFS QQNYRFLKF QAAKRITE SLQPSRGQA	GGAGCAG SCCAAAGG	3GCCAAAG 3GCCAAAG JCGAAUCC	TTCTATCGA	ATTAGTGC	ATAGAAAAG
30	(continuación)	CIA	MASSNLÍKQLQERGLVAQVTDEEALAERLAQGPIALVCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSMACTNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK	MASSNLÍKQLQERGLVAQVTDEEALAERLAQGPIALSCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREDQGISFTEFSYNLLQGYSFACLNKQYGVVLQIGGSDQW GNITSGIDLTRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTSP YKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVLA EQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK	MASSNLÍKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM NKEAVKQRLNREDQGISFTEFSYNLLQGYTFACTNKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLF	AGCTTCCCGATAAGGGAGCAGGCCAGTAAAAAGCATTACCCCGTGGTG GGGTTCCCGAGGGCAAAGGGAGCAGACTCTAAATCTGCCGTCATCG ACCTCGAAGGTTCGAATCCTTCCCCCACCACCA	AGCUUCCCGAUAAGGGAGCAGGCCAGUAAAAAGCAUUACCCCGUGGU GGGGUUCCCGAGCGGCCAAAGGGAGCAGACUCUAAAUCUGCCGUCAU CGACCUCGAAGGUUCGAAUCCUUCCCCCACCACCA	5-ATGAAGTAGCTGTCTTCTATCGAACAAGCATGCG-3	5'-CGAACAAGCATGCGATTAGTGCCGACTTAAAAAG-3'	5'-CGCTACTCCCCAAATAGAAAAGGTCTCCGCTG-3
35	(contin	SECUENCIA	MASSNLJ LVPLLCL LVPLLCL DKIRKQV INKEAVK WGNITSC PYKFYQF AEQVTRU EKGADLJ GRFTLLR	MASSNLJ LVPLLCL DKIRKQV INKEAVK GNITSGII YKFYQFV EQVTRLV KGADLM GRFTLLR	MASSNLI LVPLLCL LVPLLCL DKIRKQV INKEAVK WGNITSC PYKFYQF AEQVTRI EKGADLI GRFTLLR	AGCTTC(GGGTTC) ACCTCG/	AGCUUC GGGGUU CGACCU	5'-ATGAAG	5'-CGAACA	5-CGCTAC
40										
45										
50			pPR-EcRS-8 Aminoácido sintetasa (aa)	pPR-EcRS-9 Aminoácido sintetasa (aa)	pPR-EcRS-10 Aminoác do sintetasa (aa)	nucleótido		ıbar L3TAG	ıbar 113TAG	ıbar T44TAG
55		Marcador	pPR-EcRS-8 Aminoácido s	pPR-EcRS-9 Aminoácido s	pPR-EcRS-10 Aminoác do s	ARNt/Tyr polinucleótido	ARNt/Tyr	Mutantes ámbar L3TAG	Mutantes ámbar 113TAG	Mutantes ámbar T44TAG
60		SEC ID Ne	SEC ID Nº 61	SEC ID Nº 62	SEC ID Nº 63	SEC ID Nº 64	SEC ID Nº 65	SEC ID Nº 66	SEC ID Nº 67	SEC ID Nº 68
65		SE	SEC	SEC	SEC	SEC	SEC	SEC	SEC	SEC

5																				LHLGH QEWV SVNQM GGSDQ OPKKTS NQYVL MVEM	concc
10																				FDPTADSI KLNTEETY RDIGKHFI YGVVLQI GGAVWLI NSGKAPR LAQDGVP	JAAAUCO
15			3,	TC-3,	ර-3	6-3	CGG-3'	'AGTAG-3'	660-3	TG-3	TG-3'	:TG-3:	.TG-3	TG-3'	TG-3,	TG-3	CTG-3'	7TG-3	TG-3.	QGPIALICO SFKAAER NMNVLTFI MACANKQ GTKFGKTE IALEEDKI LSEADFEQ	CGGACUCI
20			тттсстсе-	GTGGAGAC	TGCTCTAA	AATAATGC	CATCATCAT	SAAGAGAGT	ATTGACTC	стстссес	втстссвс	зетстссес	втстссво	втстссвс	GTGGAGAC	GTGGAGAC	AGTGGAGAG	GTGGAGAC	GTGGAGAC	ALAERLAC SATGLIGDI SATGLIGDI NLLQGYGI VPLITKADO TFMSIEEN LFSGSLSA RKTIASNAI	AACGCGG
25			CTACTGATT	тсестс⊿	AGACTGATA	ACAGCATAG	GAGTGCGA	AGTCATCG	TAGGTATCO	ATTTAAAAG	ATATAAAAG	ATGGAAAAG	AGATAAAAG	AAAAAAG	TEGCTTCA	гтевсттся	теестс/	TTGGCTTCA	тевсттс⊿	VAQVTDEE KPVALVGC GENSALAAI GISFTEFSY NQVFGLT VYRFLKFF VARKRITEC LQPSRGQAI	GUGGCUA
30	ación)	٨	5-CTGGAACAGCTATAGCTACTGATTTTTCCTCG-3	5'-GCCGTCACAGATTACTTGGCTTCAGTGGAGACTC-3'	5-GATTGGCTTCATAGGAGACTGATATGCTCTAAC-3	5-GCCTCTATAGTTGAGACAGCATAGAATAATGCG-3	5-GAGACAGCATAGATAGAGTGCGACATCATCGG-3	5-GAATAAGTGCGACATAGTCATCGGAAGAGAGTAGTAG-3	5-GGTCAAAGACAGTTGTAGGTATCGATTGACTCGGC-3	5-CGCTACTCTCCCCAAATTTAAAAGGTCTCCGCTG-3	5-CGCTACTCTCCCCAAATATAAAAGGTCTCCGCTG-3	5-CGCTACTCTCCCCAAATGGAAAAGGTCTCCGCTG-3	5'-CGCTACTCTCCCCAAAGATAAAAGGTCTCCGCTG-3"	5-CGCTACTCTCCCCAAAAAAAAAAGGTCTCCGCTG-3	5-GCCGTCACAGATTTTTGGCTTCAGTGGAGACTG-3	5-GCCGTCACAGATTATTTGGCTTCAGTGGAGACTG-3	5'-GCCGTCACAGATTGGTTGGCTTCAGTGGAGACTG-3'	5-GCCGTCACAGATGATTTGGCTTCAGTGGAGACTG-3	5-GCCGTCACAGATAAATTGGCTTCAGTGGAGACTG-3	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALICGFDPTADSLHLGH LVPLLCLKRFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWV DKIRKQVAPFLDFDCGENSAIAANNYDWFGNMNVLTFLRDIGKHFSVNQM INKEAVKQRLNREGQGISFTEFSYNLLQGYGMACANKQYGVVLQIGGSDQ WGNITSGIDLTRRLHQNQVFGLTVPLITKADGTKFGKTEGGAVWLDPKKTS PYKFYQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGKAPRAQYVL AEQVTRLVHGEEGLQAAKRITECLFSGSLSALSEADFEQLAQDGVPMVEM EKGADLMQALVDSELQPSRQQARKTIASNAITINGEKQSDPEYFFKEEDRLF GRFTLLRRGKKNYCLICWK	GGUGGGGUAGCGAAGUGGCUAAACGCGGCGGACUCUAAAUCCGCUCC CUUUGGGUUCGGCGGTUCGAAUCCGUCCCCCUCCACCA
35	(continuación)	SECUENCIA	5'-CTGGAAC	5'-GCCGTC/	5'-GATTGG	5'-GCCTCTA	5'-GAGACA	5'-GAATAAG	5'-GGTCAA	5'-CGCTACT	5'-CGCTACT	5'-CGCTACT	5'-CGCTACT	5'-CGCTACT	5'-GCCGTC/	s'-GCCGTC/	5'-GCCGTC/	5'-GCCGTC/	5'-GCCGTC	MASSNLIK LVPLLCLK DKIRKQV/ DKIRKQV/ NKNTSGI WGNITSGI PYKFYQF/ AEQVTRL EKGADLM	CUUUGGG
40																					1
45										io T44F	en sitio T44Y	io T44W	en sitio T44D	io T44K	io R110F	io R110Y	en sitio R110W	io R110D	en sitio R110K	œ	
50			ar F68TAG	ar R110TAG	ámbar V114TAG	ar T121TAG	ar I127TAG	ar S131TAG	ámbar T145TAG	antes en sit	antes en sit	antes en sit		antes en sit	antes en sit	antes en sit		antes en sit		ıtetasa (aa)	
55		Marcador	Mutantes ámbar F68TAG	Mutantes ámbar R110TAG	Mutantes ámb	Mutantes ámbar T121TAG	Mutantes ámbar I127TAG	Mutantes ámbar S131TAG	Mutantes ámb	Mutantes tolerantes en sitio T44	Mutantes tolerantes	Mutantes tolerantes en sitio T44W	Mutantes tolerantes	Mutantes tolerantes en sitio T44K	Mutantes tolerantes en sitio R11	Mutantes tolerantes en sitio R11	Mutantes tolerantes	Mutantes tolerantes en sitio R11	Mutantes tolerantes	p-acetilFeRS-1 Aminoácido sintetasa (aa)a	ARNt híbrido
60			7		71		73	1					79	-	81		83	Const		Tages	
65		SEC ID Nº	SEC ID Nº 69	SEC ID Nº 70	SEC ID Nº	SEC ID Nº 72	SEC ID Ne	SEC ID Nº 74	SEC ID Nº 75	SEC ID Nº 76	SEC ID Nº 77	SEC ID Nº 78	SEC ID Ng	SEC ID Nº 80	SEC ID Nº	SEC ID Nº 82	SEC ID Nº	SEC ID Nº 84	SEC ID Nº 85	SEC ID Nº 86	SEC ID Nº 87

5		сестесе	CGCTCCG		гсттсбет	сототот	GGT
10		CTAAATC	GGACTAG	360	GTGCAAT	ACGTGTA AGCGC	SETTCGGC
15		GCGGACT	GTTGCCT	TGACTTATG	CATGCTCA	AAACCGG	TCCCTTTGG
20		GGTGGGGTAGCGAAGTGGCTAAACGCGGGGGGGGCTCTAAATCCGCTCCC TTTGGGTTCGGCGGTTCGAATCCGTCCCCCA	GGATTACGCATGCTCAGTGCAATCTTCGGTTGCCTGGACTAGCGCTCCG GTTTTTCTGTGCTGAACCTCAGGGGACGCCGACACACGTACACGTC	GACAAGTGCGGTTTTTTTCTCCAGCTCCCGATGACTTATGGC	GTACGAATTCCCGAGATCTGGATTACGCATGCTCAGTGCAATCTTCGGT TGCCTGGACTAGCGCTCCGGTTTTTCTGTGC	AGTCCGCCGCGTTTAGCCACTTCGCTACCCCACCGACGTGTACGTGTGT CGGCGTCCCCTGAGGTTCAGCACAGAAAACCGGAGCGC	GTGGCTAAACGCGGCGGACTCTAAATCCGCTCCCTTTGGGTTCGGCGGT
25		AAGTGGC:	CTCAGTGC GAACCTC	тттстсса	SAGATCTO	TAGCCAC AGGTTCAG	CGGACTCT
30	(continuación)	GGTAGCG	ACGCATG ICTGTGCT	тесестт	AATTCCC GGACTAG	оссосот тссссто	AAACGCGG
35	SECUENCIA	GGTGG TTTGG	GGATT	GACAAG	GTACG TGCCT	AGTCC	GTGGCT
40			ar	YR humano			
45		2000	ssor de ámb		0	rso	to
50		RNt híbrido	ARNt supre	anqueadora	sador direct	ebador inve	bador direc
55	Marcador	ADNc para ARNt híbrido	Expresión de ARNt supresor de ámbar	Secuencia flanqueadora de ARNtT	FTam 73: cebador directo	FTam 115 : cebador inverso	FTam116: cebador directo
60	ēN QI ⊃	Ne 88 ∪	68 ∘N QI	06 aN QI	ID Ne 91	ID Nº 92	ID Ne 93

LISTADO SECUENCIAL

I	0	1	4	7	1

5	<110> Feng, Tian Norman, Thea chu, Stephanie
	<120> ARNT HÍBRIDO SUPRESOR PARA CÉLULAS DE VERTEBRADOS
10	<130> AMBX-0125.00US
10	<150> 60/843,092 <151> 2006-09-08
15	<160> 95
10	<170> Patente en versión 3.4
20	<210> 1 <211> 1275 <212> ADN <213> Escherichia coli
	<400> 1
25	atggcaagca gtaacttgat taaacaattg caagagcggg ggctggtagc cc
	gacgaggaag cgttagcaga gcgactggcg caaggcccga tcgcgctcta tt

caggtgacg 60 120 tgcggcttc 180 gatcctaccg ctgacagctt gcatttgggg catcttgttc cattgttatg cctgaaacgc 30 ttccagcagg cgggccacaa gccggttgcg ctggtaggcg gcgcgacggg tctgattggc 240 300 gacccgagct tcaaagctgc cgagcgtaag ctgaacaccg aagaaactgt tcaggagtgg 360 gtggacaaaa tccgtaagca ggttgccccg ttcctcgatt tcgactgtgg agaaaactct 35 gctatcgcgg cgaacaacta tgactggttc ggcaatatga atgtgctgac cttcctgcgc 420 qatattqqca aacacttctc cqttaaccaq atgatcaaca aagaagcggt taagcagcgt 480 ctcaaccgtg aagatcaggg gatttcgttc actgagtttt cctacaacct gttgcagggt 540 40 600 tatgacttcg cctgtctgaa caaacagtac ggtgtggtgc tgcaaattgg tggttctgac cagtggggta acatcacttc tggtatcgac ctgacccgtc gtctgcatca gaatcaggtg 660 45 720 tttggcctga ccgttccgct gatcactaaa gcagatggca ccaaatttgg taaaactgaa ggcggcgcag tctggttgga tccgaagaaa accagcccgt acaaattcta ccagttctgg 780 atcaacactg cggatgccga cgtttäccgc ttcctgaagt tcttcacctt tatgagcatt 840 50 gaagagatca acgccctgga agaagaagat aaaaacagcg gtaaagcacc gcgcgcccag 900 960 tatgtactgg cggagcaggt gactcgtctg gttcacggtg aagaaggttt acaggcggca 1020 aaacgtatta ccgaatgcct gttcagcggt tctttgagtg cgctgagtga agcggacttc 55 1080 gaacagctgg cgcaggacgg cgtaccgatg gttgagatgg aaaagggcgc agacctgatg 1140 caggcactgg tcgattctga actgcaacct tcccgtggtc aggcacgtaa aactatcgcc 1200 tccaatgcca tcaccattaa cggtgaaaaa cagtccgatc ctgaatactt ctttaaagaa 60 1260 gaagatcgtc tgtttggtcg ttttacctta ctgcgtcgcg gtaaaaagaa ttactgtctg 1275 atttgctgga aataa

-	<2 <2	10> 2 11> 42 12> PI 13> Es	RT	chia co	oli											
5	<4	00> 2														
10	Met 1	Ala	Ser	Ser	Asn S	Leu	Ile	Lys	Gln	Leu 10	Gln	Glu	Arg	Gly	Leu 15	val
	Ala	Gln	val	Thr 20	Asp	G1u	G1u	Ala	Leu 25	Аlа	Glu	Arg	Leu	A7a 30	Gln	GΊy
15	Pro	Ile	Ala 35	Leu	Туг	Cys	Gly	Phe 40	Asp	Pro	Thr	Ala	Asp 45	Ser	Leu	His
20	Leu	G1y 50	His	Leu	va1	Pro	Leu 55	Leu	cys	Leu	Lys	Arg 60	Phe	G1n	Gln	Ala
	G]y 65	His	Lys	Pro	val	Ala 70	Leu	val	Gly	Gly	А1а 75	Thr	Gly	Leu	ıle	61y 80
25	Asp	Pro	Ser	Phe	Lys 85	Ala	Αla	G1u	Arg	Lys 90	Leu	Asn	Thr	Glu	Glu 95	Thr
30	val	Gln	Glu	Trp 100	va1	Asp	Lys	Ile	Arg 105	Lys	Gìn	val	Ala	Pro 110	Phe	Leu
35	Asp	Phe	ASP 115	Cys	Gly	Glu	Asn	Ser 120	Ala	Ile	Ala	Ala	Asn 125	Asn	Туг	Asp
	Тгр	Phe 130	Gly	Asn	Met	Asn	va1 135	Leu	Thr	Phe	Leu	Arg 140	Asp	Ile	Gly	Lys
40	His 145	Phe	Ser	val	Asn	G]n 150	Met	Ile	Asn	Lys	Glu 155	Ala	val	Lys	G1n	Arg 160
45	Leu	Asn	Arg	Glu	ASP 165	Gln	ĢΊy	Ile	ser	Phe 170	Thr	Glu	Phe	Ser	Tyr 175	Asr
	Leu	Leu	Gln	Gly 180	Tyr	Asp	Phe	Ala	Cys 185	Leu	Asn	Lys	Gln	Tyr 190	Gly	Val
50	Val	Leu	G]n 195	Ile	Gly	Gly	Ser	ASP 200	Gln	Тгр	Gly	Asn	11e 205	Thr	Ser	Gly
55	Ile	Asp 210	Leu	Thr	Arg	Arg	Leu 215	His	GÌn	Asn	Gln	Va1 220	Phe	GТу	Leu	Thr
	va1 225	Pro	Leu	Ile	Thr	Lys 230	ΑΊа	Asp	Gly	Thr	Lys 235	Phe	Gly	Lys	Thr	G1u 240

	Gly	Gly	Ala	Val	Trp 245	Leu	ASP ,	Pro	Lys	Lys 250	Thr	Ser	Pro	Tyr	Lys 255	Phe	
5	туг	Gìn	Phe	Trp 260	Ile	Asn	Thr	Аlа	ASP 265	Ala	Asp	val	Туг	Arg 270	Phe	Leu	
10	Lys	Phe	Phe 275	Thr	Phe	Met	Ser	11e 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	Glu	
	Glu	Asp 290	Lys	Asn	Ser	GТу	Lys 295	Ala	Pro	Arg	Ala	G1n 300	Tyr	Val	Leu	Ala	
15	G] u 305		va1	Thr	Arg	Leu 310	val	нis	Gly	Glu	Glu 315	Gly	Leu	Gln	Ala	A1a 320	
20	Lys	Arg	Ile	Thr	G1u 325	Cys	Leu	Phe	Ser	G]y 330	Ser	Leu	Ser	Ala	Leu 335	Ser	
	Glu	ΑΊа	Asp	Phe 340		Gln	Leu	Аlа	G1n 345	Asp	Gly	val	Pro	Met 350	val	Glu	
25	Met	Glu	Lys 355	Gly	Ala	Asp	Leu	Met 360	Gln	ΑΊa	Leu	val	Asp 365	Ser	Glu	Leu	
30	Gln	Pro 370	ser	Arg	Gly	Gln	A1a 375	Arg	Lys	Thr	Ile	A1a 380	ser	Asn	Ala	Ile	
35	Thr 385	Ile	Ąsn	Gly	Glu	Lys 390	Gln	Ser	Asp	Pro	G]u 395	туг	Phe	Phe	Lys	G]u 400	
00	Glu	Asp	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys	
40	Asn	Tyr	Cys	Leu 420	Ile	Cys	Trp	Lys									
45	<21 <21	10> 3 1> 12 12> AI 13> Ar	NC														
50	<22 <22	20> 23> sir	ntetasa	a artific	ial												
	<40	00> 3															
55	atggcaag	jca g	taac	ttga	t ta	aaca	attg	caa	gagc	999	ggct	ggta	gc c	cagg	tgac	9	60
00	gacgagga	ag c	gtta	gcag	a go	gact	ggcg	caa	ggcc	cga	tcgc	actc	gt g	tgtg	gctt	c	120
	gatcctac	cg c	tgac	agct	t gc	attt	9999	cat	cttg	ttc	catt	gtta	tg c	ctga	aacg	c	180
60	ttccagca	igg c	9990	caca	a go	cggt	tgcg	ctg	gtag	gcg	gcgc	gacg	gg t	ctga	ttgg	c	240
	gacccgag	ct t	caaa	gctg	c cg	agcg	taag	ctg	aaca	ccg	aaga	aact	gt t	cagg	agtg	g	300

	gtggacaaaa	tccgtaagca	ggttgccccg	ttcctcgatt	tcgactgtgg	agaaaactct	360
	gctatcgcgg	ccaataatta	tgactggttc	ggcaatatga	atgtgctgac	cttcctgcgc	420
5	gatattggca	aacacttctc	cgttaaccag	atgatcaaca	aagaagcggt	taagcagcgt	480
	ctcaaccgtg	aagatcaggg	gatttcgttc	actgagtttt	cctacaacct	gctgcagggt	540
	tatagtatgg	cctgtttgaa	caaacagtac	ggtgtggtgc	tgcaaattgg	tggttctgac	600
10	cagtggggta	acatcacttc	tggtatcgac	ctgacccgtc	gtctgcatca	gaatcaggtg	660
	tttggcctga	ccgttccgct	gatcactaaa	gcagatggca	ccaaatttgg	taaaactgaa	720
	ggcggcgcag	tctggttgga	tccgaagaaa	accagcccgt	acaaattcta	ccagttctgg	780
15	atcaacactg	cggatgccga	cgtttaccgc	ttcctgaagt	tcttcacctt	tatgagcatt	840
	gaagagatca	acgccctgga	agaagaagat	aaaaacagcg	gtaaagcacc	gcgcgcccag	900
00	tatgtactgg	cggagcaggt	gactcgtctg	gttcacggtg	aagaaggttt	acaggcggca	960
20	aaacgtatta	ccgaatgcct	gttcagcggt	tctttgagtg	cgctgagtga	agcggacttc	1020
	gaacagctgg	cgcaggacgg	cgtaccgatg	gttgagatgg	aaaagggcgc	agacctgatg	1080
25	caggcactgg	tcgattctga	actgcaacct	tcccgtggtc	aggcacgtaa	aactatcgcc	1140
20	tccaatgcca	tcaccattaa	cggtgaaaaa	cagtccgatc	ctgaatactt	ctttaaagaa	1200
	gaagatcgtc	tgtttggtcg	ttttacctta	ctgcgtcgcg	gtaaaaagaa	ttactgtctg	1260
30	atttgctgga	aataa					1275
35	<220>	1275 ADN Artificial					
	<223> \$	sintetasa artificial					
40	<400>	4					
	atggcaagca	gtaacttgat	taaacaattg	caagagcggg	ggctggtagc	ccaggtgacg	60
45	gacgaggaag	cgttagcaga	gcgactggcg	caaggcccga	tcgcactcac	ttgtggcttc	120
40	gatcctaccg	ctgacagctt	gcatttgggg	catcttgttc	cattgttatg	cctgaaacgc	180
	ttccagcagg	cgggccacaa	gccggttgcg	ctggtaggcg	gcgcgacggg	tctgattggc	240
50	gacccgagct	tcaaagctgc	cgagcgtaag	ctgaacaccg	aagaaactgt	tcaggagtgg	300
	gtggacaaaa	tccgtaagca	ggttgccccg	ttcctcgatt	tcgactgtgg	agaaaactct	360
	gctatcgcgg	ccaataatta	tgactggttc	agcaatatga	atgtgctgac	cttcctgcgc	420
55	gatattggca	aacacttctc	cgttaaccag	atgatcaaca	aagaagcggt	taagcagcgt	480
	ctcaaccgtg	aagatcaggg	gatttcgttc	actgagtttt	cctacaacct	gctgcagggt	540
	tatacgtatg	cctgtctgaa	caaacagtac	ggtgtggtgc	tgcaaattgg	tggttctgac	600
60	cagtggggta	acatcacttc	tggtatcgac	ctgacccgtc	gtctgcatca	gaatcaggtg	660

	tttggcctga	ccgttccgct	gatcactaaa	gcagatggca	ccaaatttgg	taaaactgaa	720
	ggcggcgcag	tctggttgga	tccgaagaaa	accagcccgt	acaaattcta	ccagttctgg	780
5	atcaacactg	cggatgccga	cgtttaccgc	ttcctgaagt	tcttcacctt	tatgagcatt	840
	gaagagatca	acgccctgga	agaagaagat	aaaacagcg	gtaaagcacc	gcgcgcccag	900
	tatgtactgg	cggagcaggt	gactcgtctg	gttcacggtg	aagaaggttt	acaggcggca	960
10	aaacgtatta	ccgaatgcct	gttcagcggt	tctttgagtg	cgctgagtga	agcggacttc	1020
	gaacagctgg	cgcaggacgg	cgtaccgatg	gttgagatgg	aaaagggcgc	agacctgatg	1080
	caggcactgg	tcgattctga	actgcaacct	tcccgtggtc	aggcacgtaa	aactatcgcc	1140
15	tccaatgcca	tcaccattaa	cggtgaaaaa	cagtccgatc	ctgaatactt	ctttaaagaa	1200
	gaagatcgtc	tgtttggtcg	ttttacctta	ctgcgtcgcg	gtaaaaagaa	ttactgtctg	1260
00	atttgctgga	aataa					1275
20							
	<210>						
	<211> : <212> :						
25	<213>	Artificial					
	<220> <223> s	sintetasa artificia	I				
30	<400>	5					
	atggcaagca	gtaacttgat	taaacaattg	caagagcggg	ggctggtagc	ccaggtgacg	60
	gacgaggaag	cgttagcaga	gcgactggcg	caaggcccga	tcgcactcgt	gtgtggcttc	120
35	gatcctaccg	ctgacagctt	gcatttgggg	catcttgttc	cattgttatg	cctgaaacgc	180
	ttccagcagg	cgggccacaa	gccggttgcg	ctggtaggcg	gcgcgacggg	tctgattggc	240
40	gacccgagct	tcaaagctgc	cgagcgtaag	ctgaacaccg	aagaaactgt	tcaggagtgg	300
40	gtggacaaaa	tccgtaagca	ggttgccccg	ttcctcgatt	tcgactgtgg	agaaaactct	360
	gctatcgcgg	ccaataatta	tgactggttc	ggcaatatga	atgtgctgac	cttcctgcgc	420
45	gatattggca	aacacttctc	cgttaaccag	atgatcaaca	aagaagcggt	taagcagcgt	480
10	ctcaaccgtg	aagatcaggg	gatttcgttc	actgagtttt	cctacaacct	gctgcagggt	540
	tatagtatgg	cctgtttgaa	caaacagtac	ggtgtggtgc	tgcaaattgg	tggttctgac	600
50	cagtggggta	acatcacttc	tggtatcgac	ctgacccgtc	gtctgcatca	gaatcaggtg	660
	tttggcctga	ccgttccgct	gatcactaaa	gcagatggca	ccaaatttgg	taaaactgaa	720
	ggcggcgcag	tctggttgga	tccgaagaaa	accagcccgt	acaaattcta	ccagttctgg	780
55	atcaacactg	cggatgccga	cgtttaccgc	ttcctgaagt	tcttcacctt	tatgagcatt	840
	gaagagatca	acgccctgga	agaagaagat	aaaaacagcg	gtaaagcacc	gcgcgcccag	900
	tatgtactgg	cggagcaggt	gactcgtctg	gttcacggtg	aagaaggttt	acaggcggca	960
60	aaacgtatta	ccgaatgcct	gttcagcggt	tctttgagtg	cgctgagtga	agcggacttc	1020
	gaacagctgg	cgcaggacgg	cgtaccgatg	gttgagatgg	aaaagggcgc	agacctgatg	1080

	caggcactgg	tcgattctga	actgcaacct	tcccgtggtc	aggcacgtaa	aactatcgcc	1140
	tccaatgcca	tcaccattaa	cggtgaaaaa	cagtccgatc	ctgaatactt	ctttaaagaa	1200
5	gaagatcgtc	tgtttggtcg	ttttacctta	ctgcgtcgcg	gtaaaaagaa	ttactgtctg	1260
	atttgctgga	aataa		*			1275
10	<210> <211> <212> <213>	1275					
15	<220> <223>	Sintetasa artificia	I				
	<400>	6					
20	atggcaagca	gtaacttgat	taaacaattg	caagagcggg	ggctggtagc	ccaggtgacg	60
20	gacgaggaag	cgttagcaga	gcgactggcg	caaggcccga	tcgcactcgt	gtgtggcttc	120
	gatcctaccg	ctgacagctt	gcatttgggg	catcttgttc	cattgttatg	cctgaaacgc	180
25	ttccagcagg	cgggccacaa	gccggttgcg	ctggtaggcg	gcgcgacggg	tctgattggc	240
	gacccgagct	tcaaagctgc	cgagcgtaag	ctgaacaccg	aagaaactgt	tcaggagtgg	300
	gtggacaaaa	tccgtaagca	ggttgccccg	ttcctcgatt	tcgactgtgg	agaaaactct	360
30	gctatcgcgg	ccaataatta	tgactggttc	ggcaatatga	atgtgctgac	cttcctgcgc	420
	gatattggca	aacacttctc	cgttaaccag	atgatcaaca	aagaagcggt	taagcagcgt	480
	ctcaaccgtg	aagatcaggg	gatttcgttc	actgagtttt	cctacaacct	gctgcagggt	540
35	tatagtatgg	cctgtttgaa	caaacagtac	ggtgtggtgc	tgcaaattgg	tggttctgac	600
	cagtggggta	acatcacttc	tggtatcgac	ctgacccgtc	gtctgcatca	gaatcaggtg	660
	tttggcctga	ccgttccgct	gatcactaaa	gcagatggca	ccaaatttgg	taaaactgaa	720
40	ggcggcgcag	tctggttgga	tccgaagaaa	accagcccgt	acaaattcta	ccagttctgg	780
	atcaacactg	cggatgccga	cgtttaccgc	ttcctgaagt	tcttcacctt	tatgagcatt	840
4-	gaagagatca	acgccctgga	agaagaagat	aaaaacagcg	gtaaagcacc	gcgcgcccag	900
45	tatgtactgg	cggagcaggt	gactcgtctg	gttcacggtg	aagaaggttt	acaggcggca	960
	aaacgtatta	ccgaatgcct	gttcagcggt	tctttgagtg	cgctgagtga	agcggacttc	1020
5 0	gaacagctgg	cgcaggacgg	cgtaccgatg	gttgagatgg	aaaagggcgc	agacctgatg	1080
50	caggcactgg	tcgattctga	actgcaacct	tcccgtggtc	aggcacgtaa	aactatcgcc	1140
	tccaatgcca	tcaccattaa	cggtgaaaaa	cagtccgatc	ctgaatactt	ctttaaagaa	1200
55	gaagatcgtc	tgtttggtcg	ttttacctta	ctgcgtcgcg	gtaaaaagaa	ttactgtctg	1260
00	atttgctgga	aataa					1275
60	<210> <211> <212> <213>	1275					
65	<220> <223> <400>	Sintetasa artificia	al				

E	atygcaagca	gtaactigat	Laaacaacty	caayaycygy	ggctggtage	ccaygryacy	00
5	gacgaggaag	cgttagcaga	gcgactggcg	caaggcccga	tcgcactcac	gtgtggcttc	120
	gatcctaccg	ctgacagctt	gcatttgggg	catcttgttc	cattgttatg	cctgaaacgc	180
10	ttccagcagg	cgggccacaa	gccggttgcg	ctggtaggcg	gcgcgacggg	tctgattggc	240
.0	gacccgagct	tcaaagctgc	cgagcgtaag	ctgaacaccg	aagaaactgt	tcaggagtgg	300
	gtggacaaaa	tccgtaagca	ggttgccccg	ttcctcgatt	tcgactgtgg	agaaaactct	360
15	gctatcgcgg	ccaataatta	tgactggttc	ggcaatatga	atgtgctgac	cttcctgcgc	420
	gatattggca	aacacttctc	cgttaaccag	atgatcaaca	aagaagcggt	taagcagcgt	480
	ctcaaccgtg	aagatcaggg	gatttcgttc	actgagtttt	cctacagcct	gctgcagggt	540
20	tatacgatgg	cctgtctgaa	caaacagtac	ggtgtggtgc	tgcaaattgg	tggttctgac	600
	cagtggggta	acatcacttc	tggtatcgac	ctgacccgtc	gtctgcatca	gaatcaggtg	660
	tttggcctga	ccgttccgct	gatcactaaa	gcagatggca	ccaaatttgg	taaaactgaa	720
25	ggcggcgcag	tctggttgga	tccgaagaaa	accagcccgt	acaaattcta	ccagttctgg	780
	atcaacactg	cggatgccga	cgtttaccgc	ttcctgaagt	tcttcacctt	tatgagcatt	840
00	gaagagatca	acgccctgga	agaagaagat	aaaaacagcg	gtaaagcacc	gcgcgcccag	900
30	tatgtactgg	cggagcaggt	gactcgtctg	gttcacggtg	aagaaggttt	acaggcggca	960
	aaacgtatta	ccgaatgcct	gttcagcggt	tctttgagtg	cgctgagtga	agcggacttc	1020
35	gaacagctgg	cgcaggacgg	cgtaccgatg	gttgagatgg	aaaagggcgc	agacctgatg	1080
00	caggcactgg	tcgattctga	actgcaacct	tcccgtggtc	aggcacgtaa	aactatcgcc	1140
	tccaatgcca	tcaccattaa	cggtgaaaaa	cagtccgatc	ctgaatactt	ctttaaagaa	1200
40	gaagatcgtc	tgtttggtcg	ttttacctta	ctgcgtcgcg	gtaaaaagaa	ttactgtctg	1260
	atttgctgga	aataa	10				1275
45	<210> <211> <212> <213>	540					
50	<220> <223>	Sintetasa artificia	al				
	<400>	8					
55	cgggggctga	tagcccaggt	gacggacgag	gaagcgttag	cagagcgact	ggcgcaaggc	60
	AND		cttcgatcct	Search Country Set Automotive			120
60			acgcttccag				180
00			tggcgacccg		commendate de la comp		240

	accgaagaaa	ctgttcagga	gtgggtggac	aaaatccgta	agcaggttgc	cccgttcctc	300
5	gatttcgact	gtggagaaaa	ctctgctatc	gcggccaata	attatgactg	gttcagcaat	360
	atgaatgtgc	tgaccttcct	gcgcgatatt	ggcaaacact	tctccgttaa	ccagatgatc	420
	aacaaagaag	cggttaagca	gcgtctcaac	cgtgaagatc	aggggatttc	gttcactgag	480
10	ttttcctaca	acctgctgca	gggttatacg	tatgcctgtc	tgaacaaaca	gtacggtgtg	540
15	<210> <211> <212> <213>	540					
20	<220> <223> <400>	Sintetasa artificia	al				
			gacggacgag	gaagcottag	cagagcgact	gocgcaaggc	60
25					gcttgcattt		120
20					acaagccggt		180
		CONTRACTOR AND A CONTRACTOR			ctgccgagcg		240
30			wild in Samme		agcaggttgc		300
	gatttcgact	gtggagaaaa	ctctgctatc	gcggccaata	attatgactg	gttcagcaat	360
	atgaatgtgc	tgaccttcct	gcgcgatatt	ggcaaacact	tctccgttaa	ccagatgatc	420
35	aacaaagaag	cggttaagca	gcgtctcaac	cgtgaagatc	aggggatttc	gttcactgag	480
	ttttcctaca	acctgctgca	gggttatacg	tatgcctgtc	tgaacaaaca	gtacggtgtg	540
40	<210> <211> <212> <213>	540					
45	<220> <223>	Sintetasa artificia	al				
	<400>	10					
50	cgggggctgg	tagcccaggt	gacggacgag	gaagcgttag	cagagcgact	ggcgcaaggc	60
					gcttgcattt		120
55					acaagccggt		180
55	e of the positional tensor from the contract	and the same of the	a to the term of term of term of the term of the term of term of term of term of term of term of term		ctgccgagcg		240
					agcaggttgc		300
60					attatgactg		360
		V/38			tctccgttaa		420
					aggggatttc		480
65	ttttcctaca	acctgctgca	gggttatacg	tatgeetgee	tgaacaaaca	gtacggtgtg	540

	<210>	11					
5	<211> <212> <213>						
	<220> <223>	Sintetasa artificia	al				
10	<400>	11					
	cgggggctgg	tagcccaggt	gacggacgag	gaagcgttag	cagagcgact	ggcgcaaggc	60
	ccgatcgcac	tcacttgtgg	cttcgatcct	accgctgaca	gcttgcattt	ggggcatctt	120
15	gttccattgt	tatgcctgaa	acgcttccag	caggcgggcc	acaagccggt	tgcgctggta	180
	ggcggcgcga	cgggtctgat	tggcgacccg	agcttcaaag	ctgccgagcg	taagctgaac	240
20	accgaagaaa	ctgttcagga	gtgggtggac	aaaatccgta	agcaggttgc	cccgttcctc	300
20	gatttcgact	gtggagaaaa	ctctgctatc	gcggccaata	attatgactg	gttcggcaat	360
	atgaatgtgc	tgaccttcct	gcgcgatatt	ggcaaacact	tctccgttaa	ccagatgatc	420
25	aacaaagaag	cggttaagca	gcgtctcaac	cgtgaagatc	aggggatttc	gttcactgag	480
	ttttcctaca	acctgctgca	gggttattcg	tatgcctgtg	cgaacaaaca	gtacggtgtg	540
30	<210> <211> <212> <213>	540					
35		Sintetasa artificia	al				
	<400>						
40		tagcccaggt					60
		tcacttgtgg					120
		tatgcctgaa					180
45		cgggtctgat					240
		ctgttcagga					300
	Control of the Control of the Control	gtggagaaaa					360
50		tgaccttcct					420
		cggttaagca					480
	ttttcctaca	acctgctgca	gggttatacg	tatgcctgtc	tgaacaaaca	gtacggtgtg	540
55 60	<210> <211> <212> <213>	540					
- -	<220> <223>	Sintetasa artificia	al				
G E	<400>	13					

	cgggggctgg	taccccaggt	gacggacgag	gaagcgttag	cagagcgact	ggcgcaaggc	50
_	ccgatcgcac	tcctttgtgg	cttcgatcct	accgctgaca	gcttgcattt	ggggcatctt	120
5	gttccattgt	tatgcctgaa	acgcttccag	caggcgggcc	acaagccggt	tgcgctggta	180
	ggcggcgcga	cgggtctgat	tggcgacccg	agcttcaaag	ctgccgagcg	taagctgaac	240
10	accgaagaaa	ctgttcagga	gtgggtggac	aaaatccgta	agcaggttgc	cccgttcctc	300
10	gatttcgact	gtggagaaaa	ctctgctatc	gcggccaata	attatgactg	gttcggcaat	360
	atgaatgtgc	tgaccttcct	gcgcgatatt	ggcaaacact	tctccgttaa	ccagatgatc	420
15	aacaaagaag	cggttaagca	gcgtctcaac	cgtgaagatc	aggggatttc	gttcactgag	480
	ttttcctaca	acctgctgca	gggttattct	attgcctgtt	cgaacaaaca	gtacggtgtg	540
20	<210> <211> <212> <213>	540					
25	<220> <223>	Sintetasa artificia	al				
	<400>	14					
	cgggggctgg	tagcccaggt	gacggacgag	gaagcgttag	cagagcgact	ggcgcaaggc	60
30	ccgatcgcac	tcgtgtgtgg	cttcgatcct	accgctgaca	gcttgcattt	ggggcatctt	120
	gttccattgt	tatgcctgaa	acgcttccag	caggcgggcc	acaagccggt	tgcgctggta	180
35	ggcggcgcga	cgggtctgat	tggcgacccg	agcttcaaag	ctgccgagcg	taagctgaac	240
00	accgaagaaa	ctgttcagga	gtgggtggac	aaaatccgta	agcaggttgc	cccgttcctc	300
	gatttcgact	gtggagaaaa	ctctgctatc	gcggccaata	attatgactg	gttcggcaat	360
40	atgaatgtgc	tgaccttcct	gcgcgatatt	ggcaaacact	tctccgttaa	ccagatgatc	420
	aacaaagaag	cggttaagca	gcgtctcaac	cgtgaagatc	aggggatttc	gttcactgag	480
	ttttcctaca	acctgctgca	gggttatagt	attgcctgtt	tgaacaaaca	gtacggtgtg	540
45	<210> <211> <212>	540 ADN					
50	<220>	Artificial Sintetasa artificia	al				
E E	<400>	15					
55	cgggggctgg	taccccaggt	gacggacgag	gaagcgttag	cagagcgact	ggcgcaaggc	60
	ccgatcgcac	tcgtgtgtgg	cttcgatcct	accgctgaca	gcttgcattt	ggggcatctt	120
60	gttccattgt	tatgcctgaa	acgcttccag	caggcgggcc	acaagccggt	tgcgctggta	180
	ggcggcgcga	cgggtctgat	tggcgacccg	agcttcaaag	ctgccgagcg	taagctgaac	240
	accgaagaaa	ctgttcagga	gtgggtggac	aaaatccgta	agcaggttgc	cccgttcctc	300
65	gatttcgact	gtggagaaaa	ctctgctatc	gcggccaata	attatgactg	gttcggcaat	360

	atgaatgtgc	tgaccttcct	gcgcgatatt	ggcaaacact	tctccgttaa	ccagatgatc	420
5	aacaaagaag	cggttaagca	gcgtctcaac	cgtgaagatc	aggggatttc	gttcactgag	480
	ttttcctaca	acctgctgca	gggttatagt	attgcctgtt	tgaacaaaca	gtacggtgtg	540
10	<210> <211> <212> <213>	540					
15	<220> <223>	Sintetasa artificia	al				
	<400>	16					
20	cgggggctgg	tagcccaggt	gacggacgag	gaagcgttag	cagagcgact	ggcgcaaggc	60
20	ccgatcgcac	tctggtgtgg	cttcgatcct	accgctgaca	gcttgcattt	ggggcatctt	120
	gttccattgt	tatgcctgaa	acgcttccag	caggcgggcc	acaagccggt	tgcgctggta	180
25	ggcggcgcga	cgggtctgat	tggcgacccg	agcttcaagg	ctgccgagcg	taagctgaac	240
20	accgaagaaa	ctgttcagga	gtgggtggac	aaaatccgta	agcaggttgc	cccgttcctc	300
	gatttcgact	gtggagaaaa	ctctgctatc	gcggccaatt	gttatgactg	gttcggcaat	360
30	atgaatgtgc	tgaccttcct	gcgcgatatt	ggcaaacact	tctccgttaa	ccagatgatc	420
	aacaaagaag	cggttaagca	gcgtctcaac	cgtgaagatc	aggggatttc	gttcactgag	480
	ttttcctaca	acctgctgca	gggttatatg	cgtgcctgtg	agaacaaaca	gtacggtgtg	540
3540	<210> <211> <212> <213>	624					
	<223>	Sintetasa artificia	al				
45	<400>		0300030030	oaancottan	cagagcgact	000000000	60
					gcttgcattt		120
	on moneyers: Eron, those woo				acaagccggt		180
50			Water and the control of the control		ctgccgagcg		240
					agcaggttgc		300
55	27.040.0400.0000.0000.000.000.000.00				attatgactg		360
55		umana comproduca accidence			tctccgttaa	u meno econologico con accessor	420
					aggggatttc		480
60	000000000000000000000000000000000000000				ctaacaaaca	er madicaenstawe satini eco	540
					cttctggtat		600
		atcagaatca		39			624
65	-99.0090		23.2				

5	<210> <211> <212> <213>	609					
5	<220> <223>	Sintetasa artificia	al				
10	<400>	18					
	caggtgacgg	acgaggaagc	gttagcagag	cgactggcgc	aaggcccgat	cgcactcggt	60
4-5	tgtggcttcg	atcctaccgc	tgacagcttg	catttggggc	atcttgttcc	attgttatgc	120
15	ctgaaacgct	tccagcaggc	gggccacaag	ccggttgcgc	tggtaggcgg	cgcgacgggt	180
	ctgattggcg	acccgagctt	caaagctgcc	gagcgtaagc	tgaacaccga	agaaactgtt	240
20	caggagtggg	tggacaaaat	ccgtaagcag	gttgccccgt	tcctcgattt	cgactgtgga	300
20	gaaaactctg	ctatcgcggc	caataattat	gactggttcg	gcaatatgaa	tgtgctgacc	360
	ttcctgcgcg	atattggcaa	acacttctcc	gttaaccaga	tgatcaacaa	agaagcggtt	420
25	aagcagcgtc	tcaaccgtga	agatcagggg	atttcgttca	ctgagttttc	ctacaacctg	480
	ctgcagggtt	atggttttgc	ctgtttgaac	aaacagtacg	gtgtggtgct	gcaaattggt	540
	ggttctgacc	agtggggtaa	catcacttct	ggtatcgacc	tgacccgtcg	tctgcatcag	600
30	aatcaggtg						609
35	<210> <211> <212> <213>	591					
40		Sintetasa artificia	al				
	<400>	19					
45	gcgttagcag	agcgactggc	gcaaggcccg	atcgcactcg	ggtgtggctt	cgatcctacc	60
	gctgacagct	tgcatttggg	gcatcttgtt	ccattgttat	gcctgaaacg	cttccagcag	120
	gcgggccaca	agccggttgc	gctggtaggc	ggcgcgacgg	gtctgattgg	cgacccgagc	180
50	ttcaaagctg	ccgagcgtaa	gctgaacacc	gaagaaactg	ttcaggagtg	ggtggacaaa	240
	atccgtaagc	aggttgcccc	gttcctcgat	ttcgactgtg	gagaaaactc	tgctatcgcg	300
	gccaataatt	atgactggtt	cggcaatatg	aatgtgctga	ccttcctgcg	cgatattggc	360
55	aaacacttct	ccgttaacca	gatgatcaac	aaagaagcgg	ttaagcagcg	tctcaaccgt	420
	gaagatcagg	ggatttcgtt	cactgagttt	tcctacaacc	tgctgcaggg	ttatggttat	480
60	gcctgtatga	acaaacagta	cggtgtggtg	ctgcaaattg	gtggttctga	ccagtggggt.	540
UU		ctontatena	cctnacccot	catctocatc	agaatcaggt	•	591

	<210> <211> <212> , <213>	621					
5	<220> <223>	Sintetasa artificia	al				
10	<222>	misc_feature (26)(26) n es a, c, g, o t					
15	<222>	misc_feature (612)(612) n es a, c, g, o t					
20	<222>	misc_feature (618)(618) n es a, c, g, o t					
25	<400>	20					
20	gggctggtag	cccaggtgac	ggacgnagaa	gcgttagcag	agcgactggc	gcaaggcccg	60
	atcgcactcc	tttgtggctt	cgatcctacc	gctgacagct	tgcatttggg	gcatcttgtt	120
30	ccattgttat	gcctgaaacg	cttccagcag	gcgggccaca	agccggttgc	gctggtaggc	180
	ggcgcgacgg	gtctgattgg	cgacccgagc	ttcaaagctg	ccgagcgtaa	gctgaacacc	240
	gaagaaactg	ttcaggagtg	ggtggacaaa	atccgtaagc	aggttgcccc	gttcctcgat	300
35	ttcgactgtg	gagaaaactc	tgctatcgcg	gccaataatt	atgactggtt	cggcaatatg	360
	aatgtgctga	ccttcctgcg	cgatattggc	aaacacttct	ccgttaacca	gatgatcaac	420
	aaagaagcgg	ttaagcagcg	tctcaaccgt	gaagatcagg	ggatttcgtt	cactgagttt	480
40	tcctacaacc	tgctgcaggg	ttattctatg	gcctgtgcga	acaaacagta	cggtgtggtg	540
	ctgcaaattg	gtggttctga	ccagtggggt	aacatcactt	ctggtatcga	cctgacccgt	600
45	cgtctgcatc	anaatcangt	9				621
	<210> <211>						
50	<212>						
50	_	Artinolal					
	<220> <223>	Sintetasa artificia	I				
55	<400>	21					

	ttagcagagc	gactggcgca	aggcccgatc	gcactcgttt	gtggcttcga	tcctaccgct	60
5	gacagcttgc	atttggggca	tcttgttcca	ttgttatgcc	tgaaacgctt	ccagcaggcg	120
	ggccacaagc	cggttgcgct	ggtaggcggc	gcgacgggtc	tgattggcga	cccgagcttc	180
	aaagctgccg	agcgtaagct	gaacaccgaa	gaaactgttc	aggagtgggt	ggacaaaatc	240
10	cgtaagcagg	ttgccccgtt	cctcgatttc	gactgtggag	aaaactctgc	tatcgcggcc	300
	aataattatg	actggttcgg	caatatgaat	gtgctgacct	tcctgcgcga	tattggcaaa	360
	cacttctccg	ttaaccagat	gatcaacaaa	gaagcggtta	agcagcgtct	caaccgtgaa	420
15	gatcagggga	tttcgttcac	tgagttttcc	tacaacctgc	tgcagggtta	ttctgcggcc	480
	tgtgcgaaca	aacagtacgg	tgtggtgctg	caaattggtg	gttctgacca	gtggggtaac	540
00	atcacttctg	gtatcgacct	gacccgtcgt	ctgcatcaga	atcaggtg		588
20							
	<210> <211>						
25	<212>	ADN					
25		Artificial					
	<220> <223>	Sintetasa artificia	al				
30	<220>						
		misc_feature					
		(403)(403) n es a, c, g, o t					
35	<220> <221>	misc_feature					
	<222>	(513)(513)					
		n es a, c, g, o t					
40	<220> <221>	misc_feature					
	<222>	(515)(515)					
		n es a, c, g, o t					
45	<220> <221>	misc_feature					
		(518)(518) n es a, c, g, o t					
		00 a, 0, 9, 0 :					
50		misc_feature					
		(531)(531) n es a, c, g, o t					
	<400>	_					
55	\4 00 <i>></i>	<u>LL</u>					
60							

```
gacgaggaag cgttagcaga gcgactggcg caaggcccga tcgcactcct gtgtggcttc
                                                                                60
      gatcctaccg ctgacagctt gcatttgggg catcttgttc cattgttatg cctgaaacgc
                                                                               120
 5
                                                                               180
      ttccagcagg cgggccacaa gccggttgcg ctggtaggcg gcgcgacggg tctgattggc
      gacccgagct tcaaagctgc cgagcgtaag ctgaacaccg aagaaactgt tcaggagtgg
                                                                               240
10
                                                                               300
      gragacaaaa teegraagea ggttgeeeeg tteetegatt tegaetgtgg agaaaactet
      gctatcgcgg ccaataatta tgactggttc ggcaatatga atgtgctgac cttcctgcgc
                                                                               360
                                                                               420
      gatattggca aacacttctc cgttaaccag atgatcaaca aanaagcggt taagcagcgt
15
                                                                               480
      ctcaaccgtg aagatcaggg gatttcgttc actgagtttt cctacaacct gctgcagggt
                                                                               540
      tattcqqctq cctqtqcqaa caaacaqtac qqnqnqqnqc tqcaaattgq nggttctgac
                                                                               600
      caggggggta acatcacttc tggtatcgac ctgacccgtc gtctgcatca aaatcaggtg
20
             <210> 23
             <211> 591
             <212> ADN
25
             <213> Artificial
             <220>
             <223> Sintetasa artificial
30
             <220>
             <221> misc_feature
             <222> (588)..(588)
             <223> n es a, c, g, o t
35
             <400> 23
       gcgttagcag agcgactggc gcaaggcccg atcgcactcg tttgtggctt cgatcctacc
                                                                                 60
                                                                                120
       qctgacagct tgcatttggg gcatcttgtt ccattgttgt gcctgaaacg cttccagcag
40
       gegggecaca ageoggttge getggtagge ggeggaegg gtetgattgg egaecegage
                                                                                180
                                                                                240
       ttcaaaqctg ccgagcgtaa gctgaacacc gaagaaactg ttcaggagtg ggtggacaaa
45
       atccgtaagc aggttgcccc gttcctcgat ttcgactgtg gagaaaactc tgctatcgcg
                                                                                300
                                                                                360
       qccaataatt atgactggtt cggcaatatg aatgtgctga ccttcctgcg cgatattggc
                                                                                420
       aaacacttct ccgttaacca gatgatcaac aaagaagcgg ttaagcagcg tctcaaccgt
50
       gaagatcagg ggatttcgtt cactgagttt tcctacaacc tgctgcaggg ttatagtgcg
                                                                                480
                                                                                540
       gcctgtgtta acaaacagta cggtgtggtg ctgcaaattg gtggttctga ccagtggggt
                                                                                591
       aacatcactt ctggtatcga cctgacccgt cgtctgcatc agaatcangt g
55
             <210> 24
             <211> 600
             <212> ADN
60
             <213> Artificial
             <223> Sintetasa artificial
65
             <400> 24
```

```
gacgaggaag cgttagcaga gcgactggcg caaggcccga tcgcactcat ttgtggcttc
                                                                                  60
       gatcctaccg ctgacagctt gcatttgggg catcttgttc cattgttatg cctgaaacgc
                                                                                 120
                                                                                 180
       ttccagcagg cgggccacaa gccggttgcg ctggtaggcg gcgcgacggg tctgattggc
 5
                                                                                 240
       gacccgagct tcaaagctgc cgagcgtaag ctgaacaccg aagaaactgt tcaggagtgg
                                                                                 300
       gtggacaaaa tccgtaagca ggttgccccg ttcctcgatt tcgactgtgg agaaaactct
10
       gctatcgcgg ccaatgatta tgactggttc ggcaatatga atgtgctgac cttcctgcgc
                                                                                 360
                                                                                 420
       gatattggca aacacttctc cgttaaccag atgatcaaca aagaagcggt taagcagcgt
       ctcaaccgtg aagatcaggg gatttcgttc actgagtttt cctacaacct gctgcagggt
                                                                                 480
15
                                                                                 540
       tataattttg cctgtgtgaa caaacagtac ggtgtggtgc tgcaaattgg tggttctgac
       cagtggggta acatcacttc tggtatcgac ctgacccgtc gtctgcatca gaatcaggtg
                                                                                 600
             <210> 25
20
             <211> 579
             <212> ADN
             <213> Artificial
             <220>
25
             <223> Sintetasa artificial
             <400> 25
       cgactggcgc aaggcccgat cgcactcacg tgtggcttcg atcctaccgc tgacagcttg
                                                                                  60
30
       catttggggc atcttgttcc attgttatgc ctgaaacgct tccagcaggc gggccacaag
                                                                                 120
       ccggttgcgc tggtaggcgg cgcgacgggt ctgattggcg acccgagctt caaagctgcc
                                                                                 180
       gagcgtaagc tgaacaccga agaaactgtt caggagtggg tggacaaaat ccgtaagcag
                                                                                 240
35
       gttgccccgt tcctcgattt cgactgtgga gaaaactctg ctatcgcggc caataattat
                                                                                 300
       gactggttcg gcaatatgaa tgtgctgacc ttcctgcgcg atattggcaa acacttctcc
                                                                                 360
       gttaaccaga tgatcaacaa agaagcggtt aagcagcgtc tcaaccgtga agatcagggg
                                                                                 420
40
       atttcgttca ctgagttttc ctacaatctg ctgcagggtt attcggctgc ctgtcttaac
                                                                                 480
       aaacagtacg gtgtggtgct gcaaattggt ggttctgacc agtggggtaa catcacttct
                                                                                 540
45
       ggtatcgacc tgacccgtcg tctgcatcag aatcaggtg
                                                                                 579
             <210> 26
             <211> 624
50
             <212> ADN
             <213> Artificial
             <223> Sintetasa artificial
55
             <220>
             <221> misc_feature
             <222> (13)..(13)
             <223> n es a, c, g, o t
60
             <220>
             <221> misc_feature
             <222> (599)..(599)
             <223> n es a, c, g, o t
65
             <400> 26
```

```
cgggggctgg tancccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc
                                                                                 60
       ccgatcgcac tcgggtgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt
                                                                                120
       gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta
                                                                                180
 5
       ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac
                                                                                240
       accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc
                                                                                300
       gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat
                                                                                360
10
       atgaatgtgc tgaccttcct qcgcgatatt ggcaaacact tctccgttaa ccagatgatc
                                                                                420
       aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag
                                                                                480
15
       ttttcctaca acctgctgca gggttattct atggcctgtt tgaacaaaca gtacggtgtg
                                                                                540
       gtgctgcaaa ttggtggttc tgaccagtgg ggtaacatca cttctggtat cgacctganc
                                                                                600
       cgtcgtctgc atcagaatca ggtg
                                                                                624
20
             <210> 27
             <211> 625
             <212> ADN
             <213> Artificial
25
             <220>
             <223> Sintetasa artificial
             <220>
30
             <221> misc feature
             <222> (600)..(600)
             <223> n es a, c, g, o t
             <400> 27
35
       cgggggctgg tagcccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc
                                                                                 60
                                                                                120
       ccgatcgcac tcacgtgtgg cttcgatcct accgctgaca gcttgcattt ggggcatctt
                                                                                180
       gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta
40
       ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac
                                                                                240
       accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc
                                                                                300
       gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat
                                                                                360
45
       atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc
                                                                                420
       aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag
                                                                                480
50
       ttttcctaca atctgctgca gggttattcg gctgcctgtc ttaacaaaca gtacggtgtg
                                                                                540
       gtgctgcaaa ttggtggttc tgaccagtgg ggtaacatca cttctggtat cgaacctgan
                                                                                600
                                                                                625
       ccgtcgtctg catcaaaatc aagtg
55
             <210> 28
             <211> 624
             <212> ADN
             <213> Artificial
60
             <220>
             <223> Sintetasa artificial
             <400> 28
65
```

	cgggggctgg	taccccaagt	gacggacgag	gaaacgttag	cagagcgact	ggcgcaaggc	60
	ccgatcgcac	tctcttgtgg	cttcgatcct	accgctgaca	gcttgcattt	ggggcatctt	120
5	gttccattgt	tatgcctgaa	acgcttccag	caggcaggcc	acaagccggt	tgcgctggta	180
	ggcggcgcga	cgggtctgat	tggcgacccg	agcttcaaag	ctgccgagcg	taagctgaac	240
40	accgaagaaa	ctgttcagga	gtgggtggac	aaaatccgta	agcaggttgc	cccgttcctc	300
10	gatttcgact	gtggagaaaa	ctctgctatc	gcggccaata	attatgactg	gttcggcaat	360
	atgaatgtgc	tgaccttcct	gcgcgatatt	ggcaaacact	tctccgttaa	ccagatgatc	420
15	aacaaagaag	cggttaagca	gcgtctcaac	cgtgaagatc	aggggatttc	gttcactgag	480
.0	ttttcctaca	acctgctgca	gggttatacg	atggcctgtg	tgaacaaaca	gtacggtgtg	540
	gtgctgcaaa	ttggtggttc	tgaccagtgg	ggtaacatca	cttctggtat	cgacctgacc	600
20	cgtcgtctgc	atcagaatca	9959				524
25	<210> <211> <212> <213>	624					
30	<220> <223>	Sintetasa artificia	al				
00	<400>	29					
25	cgggggctgg	tagcccaggt	gacggacgag	gaagcgttag	cagagcgact	ggcgcaaggc	60
35	ccgatcgcac	tcgcgtgcgg	cttcgatcct	accgctgaca	gcttgcattt	ggggcatctt	120
	gttccattgt	tatgcctgaa	acgcttccag	caggcgggcc	acaagccggt	tgcgctggta	180
40	ggcggcgcga	cgggtctgat	tggcgacccg	agcttcaagg	ctgccgagcg	taagctgaac	240
. •	accgaagaaa	ctgttcagga	gtgggtggac	aaaatccgta	agcaggttgc	cccgttcctc	300
	gatttcgact	gtggagaaaa	ctctgctatc	gcggccaata	attatgactg	gttcggcaat	360
45	atgaatgtgc	tgaccttcct	gcgcgatatt	ggcaaacact	tctccgttaa	ccagatgatc	420
	aacaaagaag	cggttaagca	gcgtctcaac	cgtgaagatc	aggggatttc	gttcactgag	480
	ttttcctaca	acctgctgca	gggttattct	tatgcctgtc	ttaacaaaca	gtacggtgtg	540
50	gtgctgcaaa	ttggtggttc	tgaccagtgg	ggtaacatca	cttctggtat	cgacctgacc	600
	cgtcgtctgc	atcagaatca	ggtg				624
55	<210> <211> <212> <213>	624					
60	<220> <223>	Sintetasa artificia	al				
	<400>	30					
65							

	cgggggctgg	tagcccaggt	gacggacgag	gaagcgttag	cagagcgact	ggcgcaaggc	60
5	ccgatcgcac	tcgcgtgtgg	cttcgatcct	accgctgaca	gcttgcattt	ggggcatctt	120
	gttccattgt	tatgcctgaa	acgcttccag	caggcgggcc	acaagccggt	tgcgctggta	180
	ggcggcgcga	cgggtctgat	tggcgacccg	agcttcaaag	ctgccgagcg	taagctgaac	240
10	accgaagaaa	ctgttcagga	gtgggtggac	aaaatccgta	agcaggttgc	cccgttcctc	300
	gatttcgact	gtggagaaaa	ctctgctatc	gcggccaata	attatgactg	gttcggcaat	360
	atgaatgtgc	tgaccttcct	gcgcgatatt	ggcaaacact	tctccgttaa	ccagatgatc	420
15	aacaaagaag	cggttaagca	gcgtctcaac	cgtgaagatc	aggggatttc	gttcactgag	480
	ttttcctaca	acctgctgca	gggttatacg	atggcctgtt	gtaacaaaca	gtacggtgtg	540
00	gtgctgcaaa	ttggtggttc	tgaccagtgg	ggtaacatca	cttctggtat	cgacctgacc	600
20	cgtcgtctgc	atcagaatca	ggtg				624
25	<210> <211> <212> <213>	624					
30	<220> <223>	Sintetasa artificia	al				
	<400>	31					
35	cgggggctgg	taccccaagt	gacggacgag	gaagcgttag	cagagcgact	ggcgcaaggc	60
	ccgatcgcac	tcacgtgtgg	cttcgatcct	accgctgaca	gcttgcattt	ggggcatctt	120
	gttccattgt	tatgcctgaa	acgcttccag	caggcgggcc	acaagccggt	tgcgctggta	180
40	ggcggcgcga	cgggtctgat	tggcgacccg	agcttcaaag	ctgccgagcg	taagctgaac	240
	accgaagaaa	ctgttcagga	gtgggtggac	aaaatccgta	agcaggttgc	cccgttcctc	300
45	gatttcgact	gtggagaaaa	ctctgctatc	gcggccaata	attatgactg	gttcggcaat	360
45	atgaatgtgc	tgaccttcct	gcgcgatatt	ggcaaacact	tctccgttaa	ccagatgatc	420
	aacaaagaag	cggttaagca	gcgtctcaac	cgtgaagatc	aggggatttc	gttcgctgag	480
50	ttttcctaca	acctgctgca	gggttatacg	tttgcctgta	tgaacaaaca	gtacggtgtg	540
00	gtgctgcaaa	ttggtggttc	tgaccagtgg	ggtaacatca	cttctggtat	cgacctgacc	600
	cgtcgtctgc	atcagaatca	ggtg				624
55							
	<210> <211> <212> <213>	606					
60	<220> <223>	Sintetasa artificia	ıl				
65	<400>	32					

```
gtgacggacg aggaagcgtt agcagagcga ctggcgcaag gcccgatcgc actcacgtgt
                                                                                 60
       ggcttcgatc ctaccgctga cagcttgcat ttggggcatc ttgttccatt gttatgcctg
                                                                                120
                                                                                180
       aaacgcttcc agcaggcggg ccacaagccg gttgcgctgg taggcggcgc gacgggtctg
 5
                                                                                240
       attggcgacc cgagcttcaa agctgccgag cgtaagctga acaccgaaga aactgttcag
                                                                                300
       gagtgggtgg acaaaatccg taagcaggtt gccccgttcc tcgatttcga ctgtggagaa
                                                                                360
       aactctgcta tcgcggccaa taattatgac tggttcggca atatgaatgt gctgaccttc
10
       ctgcgcgata ttggcaaaca cttctccgtt aaccagatga tcaacaaaga agcggttaag
                                                                                420
       cagcgtctca accgtgaaga tcaggggatt tcgttcactg agttttccta caatctgctg
                                                                                480
15
                                                                                540
       cagggttatt cggctgcctg tcttaacaaa çagtacggtg tggtgctgca aattggtggt
                                                                                600
       tctgaccagt ggggtaacat cacttctggt atcgacctga cccgtcgtct gcatcagaat
                                                                                606
       caggtg
20
             <210> 33
             <211> 624
             <212> ADN
25
             <213> Artificial
             <220>
             <223> Sintetasa artificial
30
             <400> 33
                                                                                 60
       coopgoctog tagcccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc
                                                                                120
       ccqatcqcac tcqtttqtqq cttcqatcct accgctgaca gcttgcattt ggggcatctt
35
                                                                                180
       ottccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta
                                                                                240
       ggcggcgcga cgggtctgat tggcgacccg agcttcaaag ctgccgagcg taagctgaac
       accqaaqaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc
                                                                                300
40
       gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat
                                                                                360
                                                                                420
       atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgttaa ccagatgatc
                                                                                480
       aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag
45
       ttttcctaca acctgctgca gggttattcg atggcctgta cgaacaaaca gtacggtgtg
                                                                                540
       gtgctgcaaa ttggtggttc tgaccagtgg ggtaacatca cttctggtat cgacctgacc
                                                                                600
50
       cgtcgtctgc atcagaatca ggtg
                                                                                624
             <210> 34
             <211> 624
             <212> DNA
55
             <213> Artificial
             <220>
             <223> Sintetasa artificial
60
             <220>
             <221> misc_feature
             <222> (13)..(13)
             <223> n es a, c, g, o t
65
             <400> 34
```

	cgggggctgg	tancccaagt	gacggacggg	gaagcgttag	cagagcgact	ggcgcaaggc	60
	ccgatcgcac	tcagttgtgg	cttcgatcct	accgctgaca	gcttgcattt	ggggcatctt	120
5	gttccattgt	tatgcctgaa	acgcttccag	caggcgggcc	acaagccggt	tgcgctggta	180
	ggcggcgcga	cgggtctgat	tggcgacccg	agcttcaaag	ctgccgagcg	taagctgaac	240
10	accgaagaaa	ctgttcagga	gtgggtggac	aaaatccgta	agcaggttgc	cccgttcctc	300
10	gatctcgact	gtggagaaaa	ctctgctatc	gcggccaata	attatgactg	gttcggcaat	360
	atgaatgtgc	tgaccttcct	gcgcgatatt	ggcaaacact	tctccgttaa	ccagatgatc	420
15	aacaaagaag	cggttaagca	gcgtctcaac	cgtgaagatc	aggggatttc	gttcactgag	480
	ttttcctaca	acctgctgca	gggttatagt	tttgcctgtc	tgaacaaaca	gtacggtgtg	540
	gtgctgcaaa	ttggtggttc	tgaccagtgg	ggtaacatca	cttctggtat	cgacctgacc	600
20	cgtcgtctgc	atcagaatca	ggtg	4.540			624
25	<210> <211> <212> <213>	624					
30	<220> <223>	Sintetasa artificia	al				
00	<400>	35					
35		tagcccaggt					60
						ggggcatctt	120
						tgcgctggta	180
40						taagctgaac	240
		ctgttcagga					300
	196	gtggagaaaa					360
45	9534 OFFIS (FG	tgaccttcct					420
	Company State of the	cggttaagca					480
5 0						gtacggtgtg	540
50		ttggtggttc		ggtaacatca	cttctggtat	cgacctgacc	600
	cgtcgtctgc	atcagaatca	ggtg				624
55	<210> <211> <212> <213>	424					
60	<220> <223>	Sintetasa artificia	ıl				
	<400>	36					
65							

	Me 1	t	Ala	Ser	Ser	Asn 5	Leu	Ile	Lys	Gln	Leu 10	Gln	Glu	Arg	G1y	Leu 15	val
5	AT	a	Gln	val	Thr 20	Asp	Glu	Glu	Ala	Leu 25	Ala	Glu	Arg	Leu	A1a 30	Gln	Gly
10	Pr	0	Ile	A1a 35	Leu	va1	Cys	Gly	Phe 40	Asp	Pro	Thr	Ala	ASP 45	ser	Leu	нis
	Le	u	G]y 50	His	Leu	Va1	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	Gln	Gln	Ala
15	G1 65	У	His	Lys	Pro	val	Ala 70	Leu	val	Gly	Gly	Ala 75	Thr	Gly	Leu	Ile	G]y 80
20	As	p	Pro	Ser	Phe	Lys 85	Αla	Ala	Glu	Arg	Lys 90	Leu	Asn	Thr	Glu	G1u 95	The
	Va	1	G1n	Glu	Тгр	val	Asp	Lys	Ile	Arg	Lys	Gln	Val	Ala	Pro	Phe	Leu
25																	
30																	
35																	
33																	
40																	
45																	
50																	
55																	
60																	
65																	

				100					105			77.2		110		
5	Asp	Phe	ASP 115	Cys	Gly	Glu	Asn	Ser 120	Ala	Ile	Ala	Ala	Asn 125	Asn	Туг	AS
10	Trp	Phe 130	Gly	Asn	Met	Asn	va1 135	Leu	Thr	Phe	Leu	Arg 140	Asp	Ile	Gly	Ly:
	ніs 145	Phe	Ser	val	Asn	G]n 150	Met	Ile	Asn	Lys	Glu 155	Аlа	va1	Lys	GÌn	16
15	Leu	Asn	Arg	Glu	Asp 165	Gln	Gly	Ile	Ser	Phe 170	Thr	Glu	Phe	Ser	Туг 175	Ası
20	Leu	Leu	Gln	G]y 180	Туг	ser	туг	Ala	Cys 185	Leu	Asn	Lys	Gln	Tyr 190	G1y	va
	val	Leu	Gln 195	Ile	Gly	Gly	Ser	Asp 200	Gln	Тгр	Gly	Asn	11e 205	Thr	Ser	G1;
25	Ile	Asp 210	Leu	Thr	Arg	Arg	Leu 215	His	Gln	Asn	Gln	Va1 220	Phe	Gly	Leu	Th
30	va1 225	Pro	Leu	ile	Thr	Lys 230	Аlа	ASP	Gly	Thr	Lys 235	Phe	Gly	Lys	Thr	G1(
35	Gly	GТу	Ála	val	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	Ser	Pro	Туг	Lys 255	Pho
	туг	Gln	Phe	Trp 260	Ile	Asn	Thr	Аlа	ASP 265	Ala	Asp	val	туг	Arg 270	Phe	Lei
40	Lys	Phe	Phe 275	Thr	Phe	Met	Ser	11e 280	Glu	Glu	Ile	Asn	A1a 285	Leu	G1u	Gl
45	Glu	Asp 290	Lys	Asn	Ser	Gly	Lys 295	Ala	Pro	Arg	Ala	G1n 300	Туг	Val	Leu	Ala
	G1u 305	GÌn	val	Thr	Arg	Leu 310	val	His	Gly	Glu	G)u 315	Gly	Leu	G1n	ΑΊа	A1:
50	Lys	Arg	Ile	Thr	G1u 325	Cys	Leu	Phe	Ser	Gly 330	Ser	Leu	Ser	Ala	Leu 335	Se
55	G1u	Ala	Asp	Phe 340	Glu	Gln	Leu	Αla	G1n 345	Asp	Gly	Val	Pro	Met 350	va1	GT
	Met	Glu	Lys 355	g1y	Αla	Asp	Leu	Met 360	Gln	Ala	Leu	Val	Asp 365	Ser	Glu	Le

	G1	n Pro 370		Arg	Gly	G1r	Ala 375	Arg	l Lys	Thr	Ile	380		ASI	n Ala	Ile
5	Th 38	r Ile 5	e Asr	G1y	/ Glu	1 Lys 390	G]r	Ser	· Asp	Pro	G]u 395	Туг	Phe	Phe	Ly:	400
10	Gli	u Asp) Arg	Leu	405		Arg	Phe	Thr	- Leu 410		Arg	Arg	G)	Ly 415	Lys
15	As	п Туі	· Cys	420		cys	Tr	Lys	•	ā.						
20	</td <td>210> 3 211> 4 212> F 213> <i>F</i></td> <td>24 PRT</td> <td>ıl</td> <td></td>	210> 3 211> 4 212> F 213> <i>F</i>	24 PRT	ıl												
		220> 223> S	Sinteta	sa artii	ficial											
25	<	400> 3	37													
30	Met 1	Ala	Ser	Ser	Asn 5	Leu	Ile	Lys	Gln	Leu 10	Gln	Glu	Arg	Gly	Leu 15	Va1
	Ala	G1n	val	Thr 20	Asp	Glu	Glu	Ala	Leu 25	Ala	G1u	Arg	Leu	A1a 30	Gln	Gly
35	Pro	Ile	A1a 35	Leu	Ile	Cys	Gly	Phe 40	Asp	Pro	Thr	Ala	ASP 45	Ser	Leu	His
40	Leu	Gly 50	His	Leu	Val	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	G٦n	Gln	Ala
	G] y 65	His	Lys	Pro	٧a٦	Ala 70	Leu	val	GТу	Gly	A1a 75	Thr	Gly	Leu	Ile	G]y 80
45	Asp	Pro	ser	Phe	Lys 85	Ala	Ala	Glu	Arg	Lys 90	Leu	Asn	Thr	Glu	G]u 95	Thr
50	val	G1n	Glu	Trp 100	val	Asp	Lys	Ile	Arg 105	Lys	G1n	Val	Ala	Pro 110	Phe	Leu
55	Asp	Phe	ASP 115	Cys	Gly	Glu	Asn	Ser 120	Ala	Ile	Ala	Ala	Asn 125	Asn	Туг	Asp
00	Тгр	Phe 130	Gly	Asn	Met	Asn	va] 135	Leu	Thr	Phe	Leu	Arg 140	Asp	Ile	Gly	Lys
60	His 145	Phe	Ser	val	Asn	Gln 150	Met	Ile	Asn	Lys	Glu 155	Ala	val	Lys	Gln	Arg 160
65	Leu	Asn	Arg	Glu	Asp	G1n	Gly	Ile	Ser	Phe	Thr	Glu	Phe	Ser	туг	Asn

					165					170					175	
5	Leu	Leu	Gln	Gly 180	Туг	Ser	Met	Ala	Cys 185	Leu	Asn	Lys	Gln	Tyr 190	Gly	٧a
10	val	Leu	Gln 195	Ile	Gly	Gly	Ser	ASP 200	Gln	Тгр	Gly	Asn	11e 205	Thr	Ser	G1 _y
	Ile	Asp 210	Leu	Thr	Arg	Arg	Leu 215	His	Gln	Asn	G1n	va1 220	Phe	Gly	Leu	The
15	Va1 225	Pro	Leu	Ile	Thr	Lys 230	Ala	ASP	Gly	Thr	Lys 235	Phe	Gly	Lys	Thr	G11 240
20	G1y	Gly	Αla	val	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	Ser	Pro	туг	Lys 255	Pho
	Tyr	Gln	Phe	Trp 260	Ile	ASN	Thr	Ala	Asp 265	Ala	Asp	Va1	Туг	Arg 270	Phe	Lei
25	Lys	Phe	Phe 275	Thr	Phe	Met	Ser	11e 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	G٦ι
30	Glu	Asp 290	Lys	Asn	Ser	Gly	Lys 295	Αla	Pro	Arg	Ala	G]n 300	Туг	Va1	Leu	Ala
35	G] u 305	Gln	val	Thr	Arg	Leu 310	val	ніѕ	Gly	Glu	G1u 315	Gly	Leu	G1n	Ala	A1a 320
	Lys	Arg	Ile	Thr	G]u 325	Cys	Leu	Phe	Ser	G]y 330	Ser	Leu	Ser	ΑΊa	Leu 335	Ser
40	Glu	Ala	Asp	Phe 340	Glu	Gln	Leu	Ala	G1n 345	Asp	Gly	va1	Pro	Met 350	val	G٦ι
4 5	Met	Glu	Lys 355	Gly	Ala	Asp	Leu	Met 360	Gln	Ala	Leu	va1	Asp 365	Ser	Glu	Leu
	Gln	Pro 370	Ser	Arg	Gly	GIn	A1a 375	Arg	Lys	Thr	Ile	A1a 380	Ser	Asn	Ala	Ile
50	Thr 385	Ile	Asn	Gly	Glu	Lys 390	Gln	ser	Asp	Pro	G1u 395	Туг	Phe	Phe	Lys	G71 400
55	Glu	Asp	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys
60	Asn	Туг	Cys	Leu 420	Ile	Cys	Тгр	Lys								
50	<21 <21	10> 38 11> 42 2> Pf	24 RT													
25	<21	13> Aı	tificial													

<220>

<223> Sintetasa artificial

<400> 38

5	Met 1	ΑΊа	Ser	Ser	Asn 5	Leu	Ile	Lys	G1n	Leu 10	Gln	G1u	Arg	Gly	Leu 15	val
10	Ala	Gln	val	Thr 20	Asp	Glu	Glu	Ala	Leu 25	Ala	Glu	Arg	Leu	А7а 30	Gln	Gly
	Pro	Ile	Ala 35	Leu	Va1	Cys	Gly	Phe 40	Asp	Pro	Thr	Ala	Asp 45	Ser	Leu	His
15	Leu	Gly 50	His	Leu	۷a۱	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	Gln	Gln	Ala
20	G] y 65	His	Lys	Pro	val	А1а 70	Leu	val	Gly	Gly	A]a 75	Thr	Gly	Leu	Ile	G] y 80
	Asp	Pro	Ser	Phe	Lys 85	Аlа	Αla	Glu	Arg	Lys 90	Leu	Asn	Thr	G1u	G] u 95	Thr
25	val	G1n	Glu	Trp 100	۷a٦	Asp	Lys	Ile	Arg 105	Lys	Gln	va1	Ala	Pro 110	Phe	Leu
30	Asp	Phe	ASP 115	Cys	Gly	Glu	Asn	Ser 120	Ala	Ile	Ala	Ala	Asn 125	Asn	туг	Asp
35	Тгр	Phe 130	GТу	Asn	Met	Asn	Val 135	Leu	Thr	Phe	Leu	Arg 140	Asp	Ile	Gly	Lys
	His 145	Phe	Ser	val	Asn	G]n 150	Met	Ile	Asn	Lys	Glu 155	ΑΊа	val	Lys	G]n	Arg 160
40	Leu	Asn	Arg	Glu	ASP 165	Gln	Gly	Ile	Ser	Phe 170	Thr	Glu	Phe	ser	Tyr 175	Asn
45	Leu	Leu	Gln	Gly 180	Туг	Ser	Met	Αla	Cys 185	Ala	Asn	Lys	Gln	Туг 190	Gly	val
	val	Leu	G]n 195	Ile	Gly	GТу	ser	ASP 200	G٦n	Trp	Gly	Asn	11e 205	Thr	Ser	Gly
50	Ile	Asp 210	Leu	Thr	Arg	Arg	Leu 215	His	Gln	Asn	Gln	Va1 220	Phe	Gly	Leu	Thr
55	val	Pro	Leu	Ile	Thr	Lys	Ala	Asp	G٦y	Thr	Lys	Phe	Gly	Lys	Thr	G1u

60

	225	7	230	235	240
5	Gly Gly Al	a Val Trp i 245	Leu Asp Pro 1	Lys Lys Thr Ser 250	Pro Tyr Lys Phe 255
10	Tyr Gln Ph	e Trp Ile A 260	Asn Thr Ala	Asp Ala Asp Val	Tyr Arg Phe Leu 270
	Lys Phe Ph 27		Met Ser Ile (280	Glu Glu Ile Asn	Ala Leu Glu Glu 285
15	Glu Asp Ly 290	s Asn Ser (Gly Lys Ala I 295	Pro Arg Ala Gln 300	Tyr Val Leu Ala
20	Glu Gln Va 305	Thr Arg	Leu Val His 0 310	Gly Glu Glu Gly 315	Leu Gln Ala Ala 320
	Lys Arg Il	e Thr Glu (Cys Leu Phe S	Ser Gly Ser Leu 330	Ser Ala Leu Ser 335
25	Glu Ala As	p Phe Glu (340	Gln Leu Ala 🤅	Gln Asp Gly Val	Pro Met Val Glu 350
30	Met Glu Ly 35	s Gly Ala A	Asp Leu Met 0 360	Gln Ala Leu Val	Asp Ser Glu Leu 365
35	Gln Pro Se 370	r Arg Gly (Gln Ala Arg 1 375	Lys Thr Ile Ala 380	Ser Asn Ala Ile
	Thr Ile As 385	n Gly Glu i	Lys Gln Ser A 390	Asp Pro Glu Tyr 395	Phe Phe Lys Glu 400
40	Glu Asp Ar	g Leu Phe 0 405	Gly Arg Phe	Thr Leu Leu Arg . 410	Arg Gly Lys Lys 415
45	Asn Tyr Cy	s Leu Ile (420	Cys Trp Lys		
50	<210> 39 <211> 424 <212> PRT <213> Artificia	al			
	<220> <223> Sinteta	sa artificial			
55	<400> 39				
60	1	5		Sin Leu Gin Glu / 10	15
	Ala Gln Va	1 Thr Asp 0 20	Glu Glu Ala I	Leu Ala Glu Arg 25	Leu Ala Gln Gly 30
65					

	Pro	Ile	35 35	Leu	val	Cys	GIY	40	ASP	Pro	Thr	Ala	ASP 45	ser	Leu	HIS
5	Leu	Gly 50	His	Leu	Val	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	Gln	Gln	Ala
10	G] y 65	His	Lys	Pro	val	A1a 70	Leu	val	Gly	G1y	A]a 75	Thr	Gly	Leu	Ile	G1y 80
15	Asp	Pro	Ser	Phe	Lys 85	Ala	Ala	Glu	Arg	Lys 90	Leu	Asn	Thr	Glu	G1u 95	Thr
.0	va1	Gln	Glu	Trp 100	val	Asp	Lys	Ile	Arg 105	Lys	GÌn	val	Ala	Pro 110	Phe	Leu
20	Asp	Phe	ASP 115	Cys	Gly	Glu	Asn	Ser 120	Ala	Ile	Ala	Ala	Asn 125	Asn	Tyr	Asp
25	Тгр	Phe 130	Gly	Asn	Met	Asn	Va1 135	Leu	Thr	Phe	Leu	Arg 140	Asp	Ile	Gly	Lys
	Нis 145		Ser	Val	Asn	G]n 150	Met	Ile	Asn	Lys	Glu 155	Ala	٧a٦	Lys	Gln	Arg 160
30	Leu	Asn	Arg	Glu	Asp 165	Gln	Gly	Ile	ser	Phe 170	Thr	Glu	Phe	Ser	Tyr 175	Asn
35	Leu	Leu	Gln	Gly 180	туг	Ser	Met	Ala	Cys 185	Leu	Asn	Lys	Gln	Tyr 190	Gly	Va 1
40	Val	Leu	G]n 195	Ile	Gly	Gly	Ser	Asp 200	GÌn	Тгр ;	Gly	Asn	11e 205	Thr	Ser	Gly
40	Ile	ASP 210	Leu	Thr	Arg	Arg	Leu 215	His	GÌn	Asn	G]n	Va1 220	Phe	Gly	Leu	Thr
45	Va1 225	Pro	Leu	Ile	Thr	Lys 230	Ala	Asp	Gly	Thr	Lys 235	Phe	Gly	Lys	Thr	G1u 240
50	Gly	Gly	Ala	Val	Trp 245	Leu	ASP	Pro	Lys	Lys 250	Thr	Ser	Pro	туг	Lys 255	Phe
	Tyr	Gln	Phe	Trp 260	Ile	Asn	Thr	Ala	Asp 265	Ala	Asp	Va1	Tyr	Arg 270	Phe	Leu
55	Lys	Phe	Phe 275	Thr	Phe	Met	Ser	11e 280	Glu	Ġlu	Ile	Asn	A1a 285	Leu	Glu	Glu
60	Glu	Asp	Lys	Asn	Ser	Gly	Lys	Ą٦a	Pro	Arq	Ala	G1n	Туг	val	Leu	Αla

	29	0				295					300				
5	G1u G1 305	n Val	Thr	Arg	Leu 310	va1	His	Gly	Glu	G1u 315	Gly	Leu	Gìn	Ala	A1a 320
10	Lys Ar	g Ile	Thr	Glu 325	Cys	Leu	Phe	Ser	G1y 330	Ser	Leu	Ser	Ala	Leu 335	Ser
	Glu Al	a Asp	Phe 340	G1u	Gln	Leu	Ala	G1n 345	Asp	G1y	Va1	Pro	Met 350	val	Glu
15	Met G1	u Lys 355		Ala	Asp	Leu	Met 360	G1n	Ala	Leu	val	Asp 365	Ser	Glu	Leu
20	Gln Pr 37	o Ser O	Arg	Gly	Gln	A1a 375	Arg	Lys	Thr	Ile	A1a 380	Ser	Asn	Ala	Ile
0.5	Thr 11	e Asn	Gly	Glu	Lys 390	Gln	Ser	Asp	Pro	G1u 395	Туг	Phe	Phe	Lys	G1u 400
25	Glu As	p Arg	Leu	Phe 405	GТу	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys
30	Asn Ty	r Cys	Leu 420	Ile	Cys	Тгр	Lys			*		‡	8		
35	<210> <211> <212> <213>	424	I												
40	<220> <223>		sa artifi	cial											
	<400>	40													
45	Met Al	a Ser	ser	Asn 5	Leu	Ile	Lys	Gln	Leu 10	G1n	G1u	Arg	Gly	Leu 15	Va1
50	Ala Gl	n Val	Thr 20	Asp	Glu	Glu	Ala	Leu 25	Ala	Glu	Arg	Leu	Ala 30	Gln	Gly
	Pro Il	e Ala 35	Leu	Thr	cys	G1y	Phe 40	Asp	Pro	Thr	Ala	ASP 45	Ser	Leu	His
55	Leu G1		Leu	val	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	Gln	Gln	Ala
60	Gly Hi 65	s Lys	Pro	Val	A1a 70	Leu	val	Gly	GТу	Ala 75	Thr	Gly	Leu	Ile	Gly 80
65	ASP Pr	o Ser	Phe	Lys 85	Αla	Ala	Glu	Arg	Lys 90	Leu	Asn	Thr	G1u	G] u 95	Thr

5	val	Gln	Glu	Trp 100	val	Asp	Lys	Ile	Arg 105	Lys	Gln	val	Ala	Pro 110	Phe	Leu
	Asp	Phe	ASP 115	Cys	Gly	Glu	Asn	Ser 120	Ala	Ile	Ala	Ala	Asn 125	Asn	туг	Asp
10	Trp	Phe 130	Gly	Asn	Met	Asn	va1 135	Leu	Thr	Phe	Leu	Arg 140	ASP	Ile	GТу	Lys
15	His 145	Phe	Ser	val	Asn	G]n 150	Met	Ile	Asn	Lys	Glu 155	Ala	Va1	Lys	Gln	Arg 160
	Ley	Asn	Arg	Glu	ASP 165	G1n	GТу	Ile	Ser	Phe 170	Thr	Glu	Phe	ser	Туг 175	Asn
20	Leu	Leu	Gln	Gly 180	Туг	Thr	Met	Ala	Cys 185	Leu	Asn	Lys	Gln	Tyr 190	Gly	val
25	val	Leu	Gln 195	Ile	Gly	Gly	ser	Asp 200	Gln	Тгр	Gly	Asn	11e 205	Thr	Ser	Gly
30	Ile	ASP 210	Leu	Thr	Arg	Arg	Leu 215	His	Gln	Așn	Gln	va1 220	Phe	Gly	Leu	Thr
	Va1 225	Pro	Leu	Ile	Thr	Lys 230	Ala	Asp	G1y	Thr	Lys 235	Phe	G1y	Lys	Thr	G]u 240
35	Gly	Gly	Ala	۷a٦	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	Ser	Pro	Туг	Lys 255	Phe
40	Туг	G1n	Phe	Trp 260	Ile	Asn	Thr	Ala	Asp 265	Аlа	Asp	val	туг	Arg 270	Phe	Leu
	Lys	Phe	Phe 275	Thr	Phe	Met	Ser	11e 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	G1u
45	Glu	Asp 290	Lys	Asn	Ser	Gly	Lys 295	Ala	Pro	Arg	Ala	G]n 300	Туг	val	Leu	Ala
50	G1u 305	Gln	Val	Thr	Arg	Leu 310	val	нis	Gly	Glu	Glu 315	G1y	Leu	Gln	Ala	Ala 320
	Lys	Arg	Ile	Thr	Glu 325	Cys	Leu	Phe	Ser	G]y 330	Ser	Leu	Ser	Ala	Leu 335	Ser
55	Glu	Ala	Asp	Phe 340	Glu	Gln	Leu	Ala	G1n 345	Asp	Gly	Val	Pro	Met 350	Val	Glu
60	Met	Glu	Lys	Gly	Αla	Asp	Leu	Met	Gln	Ala	Leu	va1	Asp	Ser	Glu	Leu

				355					360					365			
5	,	Gln	Pro 370	Ser	Arg	Gly	Gìn	A1a 375	Arg	Lys	Thr	Ile	Ala 380	Ser	Asn	Ala	Ile
10		Thr 385	Ile	Asn	Gly	Glu	Lys 390	G1n	Ser	Asp	Pro	G] u 395	Туг	Phe	Phe	Lys	G]u 400
		Glu	Asp	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys
15		Asn	Туг	cys	Leu 420	ıle	Cys	Тгр	Lys		8						
20		<21 <21	0> 41 1> 42 2> PF 3> Ar	4 RT													
25			20> 23> Sii 00> 41		a artifi	cial											
30					Ser	Asn 5	Leu	Ile	Lys	Gln	Leu 10	G] n	Glu	Arg	GТу	Leu 15	val
35		Ala	Gln	va1	Thr 20	Asp	Glu	Glu	Аlа	Leu 25	Ala	Glu	Arg	Leu	Ala 30	Gln	Gly
40		Pro	Ile	Ala 35	Leu	Thr	Cys	Gly	Phe 40	Asp	Pro	Thr	Ala	Asp 45	Ser	Leu	His
		Leu	Gly 50	His	Leu	val	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	Gln	Gln	Ala
45		G]y 65	His	Lys	Pro	val	А1а 70	Leu	Va 1	Gly	Gly	Ala 75	Thr	G1y	Leu	Ile	Gly 80
50		Asp	Pro	Ser	Phe	Lys 85	Ala	Ala	G1u	Arg	Lys 90	Leu	Asn	Thr	Glu	G]u 95	Thr
		val	Gln	Glu	Trp 100	Val	Asp	Lys	Ile	Arg 105	Lys	GÌn	Val	Ala	Pro 110	Phe	Leu
55		Asp	Phe	ASP 115	Cys	Gly	Glu	Asn	Ser 120	Ala	Ile	Аlа	Ala	Asn 125	Asn	туг	Asp
60		тгр	Phe 130	Gly	Asn	Met	Asn	Val 135	Leu	Thr	Phe	Leu	Arg 140	Asp	Ile	Gly	Lys
65		Нis 145	Phe	Ser	val	Asn	G]n 150	Met	Ile	Asn	Lys	Glu 155	Ala	۷a۱	Lys	Gln	Arg 160

	Leu	ASN	Arg	Glu	ASP 165	Gln	Gly	Ile	Ser	Phe 170	Thr	Glu	Phe	Ser	Tyr 175	Asn
5	Leu	Leu	Gìn	Gly 180	туг	Thr	Tyr	Ala	Cys 185	Leu	Asn	Lys	Gln	Tyr 190	Gly	val
10	val	Leu	Gln 195	Ile	Gly	Gly	Ser	ASP 200	Gln	Тгр	Gly	Asn	11e 205	Thr	Ser	Gly
15	Ile	Asp 210	Leu	Thr	Arg	Arg	Leu 215	His	Gln	Asn	Gln	Va1 220	Phe	GТу	Leu	Thr
10	Va1 225	Pro	Leu	Ile	Thr	Lys 230	Аlа	Asp	Gly	Thr	Lys 235	Phe	Gly	Lys	Thr	G1u 240
20	Gly	Gly	Ala	Val	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	Ser	Pro	Туг	Lys 255	Phe
25	Tyr	Gln	Phe	Trp 260	Ile	Asn	Thr	Ala	Asp 265	Ala	Asp	val	Туг	Arg 270	Phe	Leu
	Lys	Phe	Phe 275	Thr	Phe	Met	Ser	11e 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	Glu
30	Glu	Asp 290	Lys	Asn	Ser	Gly	Lys 295	Ala	Pro	Arg	Аlа	G]n 300	Туг	val	Leu	Аlа
35	G]u 305	GÌn	val	Thr	Arg	Leu 310	val	His	Gly	Glu	G]u 315	Gly	Leu	GÌn	Ala	A1a 320
	Lys	Arg	Ile	Thr	G1u 325	Cys	Leu	Phe	Ser	G]y 330	Ser	Leu	Ser	Ala	Leu 335	Ser
40	Glu	Ala	Asp	Phe 340	Glu	Gln	Leu	Ala	G]n 345	Asp	ςΊý	val	Pro	Met 350	val	Glu
45	Met	Glu	Lys 355	Gly	Ala	ASP	Leu	Met 360	Gln	ΑΊa	Leu	val	Asp 365	Ser	Glu	Leu
50	Gln	Pro 370	Ser	Arg	Gly	Gln	A1a 375	Arg	Lys	Thr	Ile	A1a 380	Ser	Asn	Ala	Ile
	Thr 385	Ile	Asn	Gly	Glu	Lys 390	GÌn	Ser	ASP	Pro	G1u 395	Туг	Phe	Phe	Lys	G1u 400
55	Glu	Asp	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	GТу	Lys 415	Lys
60	Asn	Туг	Cys	Leu	Ile	Cys	тгр	Lys								

84

<210> 42 <211> 424 <212> PRT

<213> Artificial

-00	Λ.
/</td <td></td>	

<223> Sintetasa artificial

<400> 42

5

10

15

20

25

30

35

40

45

50

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly
20 25 30 Pro Ile Ala Leu Leu Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly 65 70 75 80 Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu 100 105 110 Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn 165 170 175 Leu Leu Gln Gly Tyr Ser Met Ala Cys Ser Asn Lys Gln Tyr Gly Val 180 190 Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr 210 220

55

60

5	Va1 225	Pro	Leu	Ile	Thr	Lys 230	Ala	Asp	Gly	Thr	Lys 235	Phe	Gly	Lys	Thr	G1u 240
	Gly	Gly	Ala	۷a۱	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	Ser	Pro	Туг	Lys 255	Phe
10	туг	Gln	Phe	Trp 260	Ile	Asn	Thr	Ala	Asp 265	Ala	ASP	va1	туг	Arg 270	Phe	Leu
15	Lys	Phe	Phe 275	Thr	Phe	Met	Ser	I1e 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	Glu
	Glu	Asp 290	Lys	Asn	Ser	Gly	Lys 295	Аlа	Pro	Arg	Аlа	G]n 300	Туг	val	Leu	Ala
20	Glu 305	Gln	val	Thr	Arg	Leu 310	val	ніѕ	Gly	Glu	Glu 315	Gly	Leu	Gln	Αla	A1a 320
25	Lys	Arg	Ile	Thr	G1u 325	Cys	Leu	Phe	Ser	G]y 330	Ser	Leu	Ser	Ala	Leu 335	Ser
30	Glu	Ala	Asp	Phe 340	Glu	Gln	Leu	Ala	G]n 345	Asp	Gly	۷al	Pro	Met 350	val	Glu
	Met	G1u	Lys 355	Gly	A1a	Asp	Leu	Met 360	Gln	Ala	Leu	va1	ASP 365	Ser	Glu	Leu
35	Gln	Pro 370	Ser	Arg	Gly	Gln	Ala 375	Arg	Lys	Thr	Ile	Ala 380	Ser	Asn	Ala	Ile
40	Thr 385	Ile	Asn	Gly	Glu	Lys 390	GÌn	Ser	Asp	Pro	Glu 395	Туг	Phe	Phe	Lys	G1u 400
	Glu	Asp	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys
45	Asn	туг	Cys	Leu 420	Ile	Cys	Тгр	Lys					2			
50	<2°	10> 43 11> 42 12> Pl 13> Aı	24 RT													
55		20> 23> Si	intetas	a artifi	cial											
	<40	00> 43	3													
60	Met	Ala	Ser	Ser	Asn	Leu	Ile	Lys	Gln	Leu	Gln	Glu	Arg	Gly	Leu	val

	Ala	Gln	val	Thr 20	ASP	Glu	Glu	Ala	Leu 25	Ala	Glu	Arg	Leu	A1a	Gln	Gly
5	Pro	Ile	Ala 35	Leu	Leu	Cys	Gly	Phe 40	Asp	Pro	Thr	Ala	Asp 45	Ser	Leu	ніѕ
10	Leu	G]y 50	His	Leu	val	Pro	Leu SS	Leu	Cys	reu	Lys	Arg 60	Phe	Gln	Gln	Αla
15	G1y 65	His	Lys	Pro	val	Ala 70	Leu	val	Gly	Gly	Ala 75	Thr	Gly	Leu	Ile	Gly 80
	Asp	Pro	Ser	Phe	Lys 85	Ala	Ala	Glu	Arg	Lys 90	Leu	Asn	Thr	Glu	G] u 95	Thr
20	val	Gln	G1u	Trp 100	val	ASP	Lys	īle	Arg 105	Lys	Gln	val	Ala	Pro 110	Phe	Leu
25	Asp	Phe	ASP 115	Cys	Gly	Glu	Asn	Ser 120	Ala	Ile	Ala	Ala	Asn 125	Asn	Туг	Asp
30	Тгр	Phe 130	Gly	Asn	Met	Asn	va1 135	Leu	Thr	Phe	Leu	Arg 140	Asp	Ile	Gly	Lys
	His 145	Phe	Ser	Va1	Asn	G]n 150	Met	Ile	Asn	Lys	G1u 155	Ala	va1	Lys	Gln	Arg 160
35	Leu	Asn	Arg	Glu	ASP 165	GÌn	Gly	Ile	Ser	Phe 170	Thr	Glu	Phe	Ser	Tyr 175	Asn
40	Leu	Leu	GÌn	Gly 180	Tyr	Ser	Met	Ala	Cys 185	Αla	Asn	Lys	Gln	Tyr 190	Gly	val
	val	Leu	G]n 195	Ile	Gly	Gly	Ser	Asp 200	Gln	тгр	Gly	Asn	11e 205	Thr	Ser	Gly
45	Ile	ASP 210	Leu	Thr	Arg	Arg	Leu 215	нis	GÌn	Asn	Gln	va1 220	Phe	Gly	Leu	Thr
50	Va1 225	Pro	Leu	Ile	Thr	Lys 230	Ala	ASP	Gly	Thr	Lys 235	Phe	GТу	Lys	Thr	G1u 240
	Gly	Gly	Ala	Va1	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	Ser	Pro	туг	Lys 255	Phe
55	Туг	G1n	Phe	Trp 260	Ile	Asn	Thr	ΑĬα	Asp 265	Αla	Asp	Va1	туг	Arg 270	Phe	Leu
60	Lys	Phe	Phe 275	Thr	Phe	Met	ser	Ile 280	Glu	Glu	ıle	Asn	A1a 285	Leu	Glu	G٦u

	Glu	Asp 290	Lys	Asn	Ser	Gly	Lys 295	Ala	Pro	Arg	Ala	G1n 300	туг	val	Leu	Ala
5	G1u 305		val	Thr	Arg	Leu 310	Va1	His	GТу	Glu	G]u 315	Gly	Leu	Gln	Ala	A1a 320
10	Lys	Arg	Ile	Thr	G1u 325	Cys	Leu	Phe	Ser	Gly 330	Ser	Leu	Ser	Ala	Leu 335	Ser
15	Glu	Ala	Asp	Phe 340	Glu	Gln	Leu	Ala	G1n 345	Asp	G1y	Va1	Pro	Met 350	Va1	Glu
13	Met	Glu	Lys 355	Ģly	Ala	Asp	Leu	Met 360	Gln	Ala	Leu	۷a۱	Asp 365	Ser	Glu	Leu
20	Gln	Pro 370	Ser	Arg	Gly	Gln	A1a 375	Arg	Lys	Thr	Ile	A1a 380	Ser	Asn	Ala	Ile
25	Thr 385	Ile	Asn	Gly	Glu	Lys 390	Gìn	Ser	Asp	Pro	G1u 395	Tyr	Phe	Phe	Lys	G1u 400
	Glu	Asp	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys
30	Asn	Туг	Cys	Leu 420	Ile	Cys	Тгр	Lys								
35	<21 <21	10> 44 11> 42 12> PI 13> Ai	24 RT													
40	<22 <22	20> 23> Si	ntetas	a artifi	cial											
	<40	00> 44	1													
45	Met 1	Ala	Ser	Ser	Asn 5	Leu	Ile	Lys	Gln	Leu 10	Gìn	Glu	Arg	G1y	Leu 15	Val
50	Ala	G1n	Val	Thr 20	Asp	Glu	Glu	Ala	Leu 25	Ala	Glu	Arg	Leu	Ala 30	G] n	Gly
	Pro	Ile	Ala 35	Leu	Thr	Cys	Gly	Phe 40	ASP	Pro	Thr	Ala	ASP 45	ser	Leu	His
55	Leu	G1y 50	His	Leu	val	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	Gln	Gln	Ala
60	G]y 65	His	Lys	Pro	val	A1a 70	Leu	val	Gly	Gly	Ala 75	Thr	Gly	Leu	Ile	G]y 80

5	Asp	Pro	Ser	Phe	Lys 85	Ala	Ala	Glu	Arg	Lys 90	Leu	Asn	Thr	Glu	Glu 95	Thr
	val	Gln	Glu	Trp 100	val	Asp	Lys	Ile	Arg 105	Lys	Gln	val	ΑΊа	Pro 110	Phe	Leu
10	Asp	Phe	Asp 115	Cys	Gly	Glu	Asn	Ser 120	Ala	Ile	Ala	Ala	Asn 125	Asn	Туг	Asp
15	Trp	Phe 130	Gly	Asn	Met	Asn	va1 135	Leu	Thr	Phe	Leu	Arg 140	ASP	Ile	Gly	Lys
	His 145	Phe	Ser	Val	Asn	G1n 150	Met	Ile	Asn	Lys	G]u 155	Ala	٧a٦	Lys	Gln	Arg 160
20	Leu	Asn	Arg	Glu	Asp 165	Gln	Gly	Ile	ser	Phe 170	Thr	Glu	Phe	Ser	Tyr 175	ASI
25	Leu	Leu	Gln	Gly 180	Туг	Arg	Met	Ala	Cys 185	Leu	Asn	Lys	Gln	Tyr 190	Gly	Va1
30	val	Leu	G]n 195	Ile	Gly	Gly	ser	ASP 200	Gln	тгр	Gly	Asn	11e 205	Thr	ser	Gly
	Ile	Asp 210	Leu	Thr	Arg	Arg	Leu 215	His	Gln	ASN	Gln	Va1 220	Phe	Gly	Leu	Thr
35	Va1 225	Pro	Leu	Ile	Thr	Lys 230	Ala	Asp	Gly	Thr	Lys 235	Phe	GТу	Lys	Thr	G1u 240
40	Gly	Gly	Ala	val	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	Ser	Pro	Туг	Lys 255	Phe
	Туг	Gln	Phe	Trp 260	Ile	Asn	Thr	Аlа	Asp 265	Аlа	Asp	val	Tyr	Arg 270	Phe	Leu
45	Lys	Phe	Phe 275	Thr	Phe	Met	Ser	Ile 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	Glu
50	Glu	Asp 290	Lys	Asn	Ser	Gly	Lys 295	Ala	Pro	Arg	Ala	G]n 300	Туг	val	Leu	Ala
	G1u 305	Gln	val	Thr	Arg	Leu 310	val	ніѕ	Gly	Glu	G]u 315	Gly	Leu	GÌn	Ala	A1a 320
55	Lys	Arg	Ile	Thr	G1u 325	Cys	Leu	Phe	Ser	G]y 330	Ser	Leu	Ser	Ala	Leu 335	Ser
60	Glu	Аlа	Asp	Phe 340	Glu	Gln	Leu	ΑΊa	G1n 345	Asp	Gly	Val	Pro	Met 350	va1	Glu

	Met	Glu	Lys 355	Gly	Ala	Asp	Leu	Met 360	Gln	Ala	Leu	Val	Asp 365	Ser	Glu	Leu
5	Gln	Pro 370	Ser	Arg	Gly	Gln	Ala 375	Arg	Lys	Thr	Ile	Ala 380	ser	Asn	Ala	Ile
10	Thr 385	Ile	Asn	Gly	Glu	Lys 390	GIn	Ser	Asp	Pro	G]u 395	Туг	Phe	Phe	Lys	G]u 400
15	Glu	Asp	Arg	Leu	Phe 405	G1y	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	GТу	Lys 415	Lys
15	Asn	туг	Cys	Leu 420	Ile	Cys	Trp	Lys								
20	<21	0> 45 1> 42 2> PF	24													
25	<22	3> Ar 20> 23> Si		a artifi	cial											
30	<40	0> 45														
	Met 1	ΑΊа	ser	Ser	Asn 5	Leu	Ile	Lys	G۱n	Leu 10	G1n	G1u	Arg	Gly	Leu 15	۷a٦
35	Ala	Gln	val	Thr 20	Asp	Glu	Glu	Ala	Leu 25	Ala	Glu	Arg	Leu	A1a 30	Gln	Gly
40	Pro	Ile	A1a 35	Leu	Ile	Cys	Gly	Phe 40	Asp	Pro	Thr	Ala	Asp 45	Ser	Leu	His
	Leu	G1y 50	ніѕ	Leu	Va1	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	G1n	G1n	Ala
45	G1y 65	His	Lys	Pro	val	Ala 70	Leu	val	Gly	Gly	Ala 75	Thr	Gly	Leu	Ile	GTy 80
50	Asp	Pro	ser	Phe	Lys 85	Ala	Αla	Glu	Arg	Lys 90	Leu	Asn	Thr	Glu	Glu 95	Thr
55	Va1	Gln	Glu	Trp 100	val	Asp	Lys	Ile	Arg 105	Lys	Gln	val	Ala	Pro 110	Phe	Leu
	Asp	Phe	Asp 115	Cys	Gly	Glu	Asn	Ser 120	Ala	Ile	Ala	Ala	Asn 125	Asn	Tyr	Asp
60	Trp	Phe 130	Gly	Asn	Met	Asn	Va1 135	Leu	Thr	Phe	Leu	Arg 140	Asp	Ile	GТу	Lys

	His 145	Phe	Ser	Va1	Asn	G1n 150	Met	Ile	Asn	Lys	G7u 155	Ala	Val	Lys	Gln	160
5	Leu	Asn	Arg	Glu	ASP 165	Gln	Gly	Ile	Ser	Phe 170	Thr	Glu	Phe	Ser	Tyr 175	Asn
10	Leu	Leu	Gln	Gly 180	Tyŗ	Gly	Met	Ala	Cys 185	Ala	Asn	Lys	Gln	Tyr 190	Gly	Val
15	val	Leu	Gìn 195	Ile	Gly	Gly	Ser	Asp 200	Gln	Trp	GТу	Asn	11e 205	Thr	Ser	Gly
	Ile	ASP 210	Leu	Thr	Arg	Arg	Leu 215	His	G1n	Asn	Gln	va1 220	Phe	Gly	Leu	Thr
20	va1 225	Pro	Leu	Ile	Thr	Lys 230	Ala	ASP	Gly	Thr	Lys 235	Phe	Gly	Lys	Thr	G]u 240
25	Gly	Gly	Ala	Va1	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	Ser	Pro	Туг	Lys 255	Phe
	Туг	Gln	Phe	Trp 260	Ile	Asn	Thr	Ala	A5P 265	Ala	ASP	val	туг	Arg 270	Phe	Leu
30	Lys	Phe	Phe 275	Thr	Phe	Met	Ser	11e 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	G1 u
35	Glu	Asp 290	Lys	Asn	Ser	Gly	Lys 295	Ala	Pro	Arg	Ala	G1n 300	туг	val	Leu	Ala
40	G1u 305	Gln	val	Thr	Arg	Leu 310	val	ніѕ	Gly	Glu	Glu 315	Gly	Leu	Gln	Ala	A1a 320
40	Lys	Arg	Ile	Thr	G1u 325	Cys	Leu	Phe	Ser	Gly 330	Ser	Leu	Ser	Ala	Leu 335	Ser
45	Glu	Ala	Asp	Phe 340	.G1u	G1n	Leu	Ala	G]n 345	Asp	Gly	val	Pro	Met 350	Va1	Glu
50	Met	Glu	Lys 355	61y	Ala	ASP	Leu	Met 360	GÌn	Ala	Leu	va1	ASP 365	Ser	G1u	Leu
	GÌn	Pro 370	ser	Arg	Gly	G1n	A1a 375	Arg	Lys	Thr	Ile	A1a 380	Ser	Asn	Ala	Ile
55	Thr 385	Ile	Asn	Gly	Glu	Lys 390	Gln	ser	Asp	Pro	G1u 395	Туг	Phe	Phe	Lys	G1u 400
60	Glu	Asp	Arg	Leu	405		14.00 F.	Phe		410		10000000	VSSSCHEE	G1y	Lys 415	Lys
					Asn	Tyr	Cys	Leu 420	Ile	cys	Тгр	Lys				

91

<210>46 <211>411

		<212> <213>		ial												
5	<	<220> <223> <400>		asa ar	tificial											
10	Met 1	Ala	Ser	Ser	Asn S	Leu	Ile	Lys	G]n	Leu 10	Gln	Glu	Arg	Gly	Leu 15	va1
15	Ala	Gln	val	Thr 20	Asp	G1u	G1u	Ala	Leu 25	Ala	Glu	Arg	Leu	Ala 30	Gln	Gly
	Pro	Ile	Ala 35	Leu	Gly	Cys	Gly	Phe 40	Asp	Pro	Thr	Ala	ASP 45	Ser	Leu	His
20	Leu	Gly 50	His	Leu	va1	Pro	Leu SS	Leu	Cys	Leu	Lys	Arg 60	Phe	Gln	G1n	Ala
25	G1y 65	His	Lys	Pro	val	Ala 70	Leu	va1	Gly	G1y	Ala 75	Thr	GТу	Leu	Ile	G]y 80
	Asp	Pro	Ser	Phe	Lys 85	Ala	Ala	Glu	Arg	Lys 90	Leu	Asn	Thr	Glu	G1u 95	Thr
30	val	Gln	Glu	Trp 100	va1	Asp	Lys	Ile	Arg 105	Lys	Gln	va1	Ala	Pro 110	Phe	Leu
35	ASP	Phe	ASP 115	Cys	G1y	G1u	Asn	5er 120	Ala	Ile	Ala	Ala	Asn 125	Asn	Туг	Asp
40		Phe 130														
	His 145	Phe	Ser	Val	Asn	G]n 150	Met	Ile	Asn	Lys	Glu 155	Ala	va1	Lys	Gln	Arg 160
45	Leu	Asn	Arg	G1u	Asp 165	Gln	Gly	Ile	Ser	Phe 170	Thr	Glu	Phe	Ser	Tyr 175	Asn
50	Leu	Leu	Gln	Gly 180	Туг	Gly	Phe	Ala	Cys 185	Ala	Asn	Lys	GÌn	Туг 190	Gly	Val
55	val	Leu	G]n 195	Ile	Gly	GТу	Ser	Asp 200	Gln	Trp	Gly	Asn	11e 205	Thr	Ser	Gly
60																

	Ile	ASP 210	Leu	Thr	Arg	Arg	Leu 215	His	Gln	Asn	Gln	Va1 220	Phe	Gly	Leu	Thr
5																
	225	Pro	Leu	Tie	Thr	230	AIA	ASP	GIY	inr	235	Pne	GIY	Lys	Inr	240
10	Gly	Gly	Ala	val	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	Ser	Pro	Туг	Lys 255	Phe
15	Туг	Gln	Phe	Trp 260	Ile	Asn	Thr	Ala	Asp 265	Ala	Asp	Va1	туг	Arg 270	Phe	Leu
	Lys	Phe	Phe 275	Thr	Phe	Met	Ser	11e 280	G1u	G1u	Ile	Asn	A1a 285	Leu	Glu	G1 u
20	Glu	Asp 290	Lys	Asn	Ser	Gly	Lys 295	Ala	Pro	Arg	Αla	G1n 300	Туг	۷a٦	Leu	Ala
25	G1u 305	Gln	val	Thr	Arg	Leu 310	val	His	Gly	Glu	G]u 315	Gly	Leu	Gln	Ala	A1a 320
30	Lys	Arg	Ile	Thr	G1u 325	Cys	Leu	Phe	ser	G]y 330	ser	Leu	ser	Ala	Leu 335	ser
	Glu	Ala	Asp	Phe 340	Glu	Gln	Leu	Ala	G1n 345	Asp	Gly	Val	Pro	Met 350	Val	Glu
35	Met	Glu	Lys 355	Gly	Ala	Asp	Leu	Met 360	GÌn	Ala	Leu	Va1	Asp 365	Ser	Glu	Leu
40	Gln	Pro 370	ser	Arg	Gly	Gln	A1a 375	Arg	Lys	Thr	Ile	Ala 380	ser	Asn	Ala	Ile
	Thr 385	Ile	Asn	Gly	Glu	Lys 390	G1n	Ser	Asp	Pro	G1u 395	Tyr	Phe	Phe	Lys	G1u 400
45	Glu	Asp	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys
50	Asn	туг	Cys	Leu 420	Ile	Cys	ТГР	Lys								
55	<21 <21	0> 47 1> 42 2> Pi 3> Ar	<u>2</u> 4 RT													
60		-		a artifi	icial											

5	Met 1	Ala	Ser	Ser	Asn 5	Leu	Ile	Lys	Gln	Leu 10	Gln	Glu	Arg	Gly	Leu 15	val
J	Ala	Gln	Val	Thr 20	Asp	Glu	Glu	Ala	Leu 25	Ala	Glu	Arg	Leu	A1a 30	Gln	G1y
10	Pro	Ile	A1a 35	Leu	Gly	Cys	Gly	Phe 40	Asp	Pro	Thr	Ala	ASP 45	ser	Leu	His
15	Leu	G]y 50	His	Leu	Val	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	Gln	Gln	Ala
	Gly 65	His	Lys	Pro	Va1	Ala 70	Leu	val	Gly	Gly	A1a 75	Thr	G1y	Leu	Ile	61y 80
20	ASP	Pro	ser	Phe	Lys 85	Ala	Ala	Glu	Arg	Lys 90	Leu	Asn	Thr	Glu	Glu 95	Thr
25	val	Gln	Glu	Trp 100	Val	Asp	Lys	Ile	Arg 105	Lys	Gln	val	Ala	Pro 110	Phe	Leu
30	ASP	Phe	Asp 115	Cys	Gly	Glu	Asn	Ser 120	Ala	Ile	Ala	Ala	Asn 125	Asn	Туг	Asp
30	Тгр	Phe 130	Gly	Asn	Met	Asn	Va1 135	Leu	Thr	Phe	Leu	Arg 140	Asp	Ile	Gly	Lys
35	His 145	Phe	Ser	val	Asn	G]n 150	Met	Ile	Asn	Lys	Glu 155	Ala	val	Lys	Gln	Arg 160
40	Leu	Asn	Arg	Glu	ASP 165	Gln	Gly	Ile	Ser	Phe 170	Thr	Glu	Phe	Ser	Ţŗŗ	Asn
	Leu	Leu	G1n	G1y 180	Туг	Gly	туг	Аlа	Cys 185	Met	Asn	Lys	Gln	Tyr 190	Gly	Val
45	val	Leu	G]n 195	Ile	Gly	Gly	Ser	Asp 200	Gln	Тгр	Gly	Asn	11e 205	Thr	Ser	Gly
50	Ile	ASP 210	Leu	Thr	Arg	Arg	Leu 215	His	Gln	Asn	Gln	Va1 220	Phe	Gly	Leu	Thr
55	va1 225	Pro	Leu	Ile	Thr	Lys 230	Αla	Asp	ĢΊy	Thr	Lys 235	Phe	Gly	Lys	Thr	G1L 240
55	Gly	Gly	Ala	val	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	Ser	Pro	Туг	Lys 255	Phe
60	Туг	Gln	Phe	Trp 260	Ile	Asn	Thr	Аlа	ASP 265	Ala	Asp	val	Туг	Arg 270	Phe	Leu

	Lys	Phe	Phe 275	Thr	Phe	Met	Ser	11e 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	Glu
5	Glu	Asp 290	Lys	Asn	ser	Gly	Lys 295	Ala	Pro	Arg	Ala	G]n 300	туг	val	Leu	Ala
10	G1u 305	G1n	val	Thr	Arg	Leu 310	val	His	Gly	Glu	Glu 315	GТу	Leu	Gln	Αla	A1a 320
15	Lys	Arg	Ile	Thr	G1u 325	Cys	Leu	Phe	Ser	G1y 330	Ser	Leu	Ser	Ala	Leu 335	Ser
	Glu	Ala	ASP	Phe 340	Glu	G1n	Leu	Ala	G1n 345	Asp	Gly	val	Pro	Met 350	val	G1 u
20	Met	G1u	Lys 355	GТу	Αla	Asp	Leu	Met 360	Gln	Ala	Leu	۷al	Asp 365	Ser	Glu	Leu
25	Gln	Pro 370	Ser	Arg	Gly	Gln	A7a 375	Arg	Lys	Thr	Ile	A1a 380	Ser	Asn	Ala	Ile
	Thr 385	Ile	Asn	Gly	Glu	Lys 390	Gln	Ser	Asp	Pro	G]u 395	Туг	Phe	Phe	Lys	G1u 400
30	Glu	Asp	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys
35		Туг	Cys	Leu 420		Cys	ТГР	Lys								
40	<21 <21	0> 48 1> 42 2> PF 3> Art	4 RT													
45				a artifi	cial											
50	Met 1	Ala	Ser	Ser	Asn 5	Leu	Ile	Lys	Gln	Leu 10	Gln	Glu	Arg	GТу	Leu 15	۷a۱
	Ala	G1n	Val	Thr 20	Asp	G1u	Glu	Ala	Leu 25	Ala	G1u	Arg	Leu	Ala 30	Gln	Gly
55	Pro	Ile	A1a 35	Leu	Leu	Cys	Gly	Phe 40	Asp	Pro	Thir	Ala	Asp 45	Ser	Leu	ніѕ
60	Leu	Gly 50	His	Leu	val	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	Gln	Gln	Ala

5	Gly 65	His	Lys	Pro	Val	A1a 70	Leu	Val	Gly	Gly	A1a 75	Thr	Gly	Leu	Ile	61y 80
	Asp	Pro	Ser	Phe	Lys 85	Αla	Ala	G۱u	Arg	Lys 90	Leu	Asn	Thr	Glu	G1u 95	Thr
10	val	Gln	Glu	Trp 100	val	Asp	Lys	Ile	Arg 105	Lys	GÌn	val	Ala	Pro 110	Phe	Leu
15	Asp	Phe	ASP 115	Cys	Gly	Glu	Asn	Ser 120	Аlа	Ile	Ala	Ala	Asn 125	Asn	туг	Asp
	Trp	Phe 130	ĢΊy	Asn	Met	Asn	Va 1 135	Leu	Thr	Phe	Leu	Arg 140	Asp	Ile	Gly	Lys
20	His 145	Phe	Ser	val	Asn	G]n 150	Met	Ile	Asn	Lys	G]u 155	Ala	val	Lys	Gln	Arg 160
25	Leu	Asn	Arg	Glu	ASP 165	Gln	Gly	Ile	Ser	Phe 170	Thr	Glu	Phe	Ser	Tyr 175	Asn
30	Leu	Leu	G1n	Gly 180	туг	\$er	Met	Аlа	Cys 185	Ala	Asn	Lys	Gln	Tyr 190	Gly	yal
	val	Leu	G]n 195	rle	Ģ٦y	ĢΊy	Ser	Asp 200	Gln	Trp	GТу	Asn	11e 205	Thr	Ser	Gly
35	Ile	Asp 210	Leu	Thr	Arg	Arg	Leu 215	His	Gln	Asn	Gln	va1 220	Phe	Gly	Leu	Thr
40	va1 225	Pro	Leu	Ile	Thr	Lys 230	Ala	Asp	Gly	Thr	Lys 235	Phe	Gly	Lys	Thr	G1u 240
	Gly	Gly	Ala	Va1	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	Ser	Pro	Туг	Lys 255	Phe
45	туг	Ģìn	Phe	Trp 260	Ile	Aşn	Thr	Ala	Asp 265	Ala	Asp	Val	туг	Arg 270	Phe	Leu
50	Lys	Phe	Phe 275	Thr	Phe	Met	Ser	11e 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	Glu
	Glu	ASP 290	Lys	Asn	ser	Gly	Lys 295	Аlа	Pro	Arg	Ala	G]n 300	Tyr	val	Leu	Ala
55	G] u 305	Gìn	Val	Thr	Arg	Leu 310	val	His	GТу	Glu	G]u 315	Gly	Leu	Gln	Ala	A1a 320
60	Lys	Arg	Ile	Thr	Glu 325	Cys	Leu	Phe	Ser	Gly 330	Ser	Leu	Ser	Ala	Leu 335	Ser

Glu	Ala	Asp	Phe 340	Glu	Gln	Leu	Ala	G1n 345	ASP	Gly	val	Pro	Met 350	val	Glu
Met	Glu	Lys 355	Gly	Ala	ASP	Leu	Met 360	Gln	Аlа	Leu	Va1	Asp 365	Ser	Glu	Leu
Gln	Pro 370	Ser	Arg	Gly	G1n	Ala 375	Arg	Lys	Thr	Ile	Ala 380	ser	Asn	Ala	Ile
Thr 385	Ile	Asn	Gly	Glu	Lys 390	Gln	Ser	Asp	Pro	G1u 395	Туг	Phe	Phe	Lys	G1u 400
Glu	ASP	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys
Asn	Туг	Cys	Leu 420	Ile	Cys	Тгр	Lys								
<21 <21	1> 42 2> PF	:4 RT													
		ntetas	a artifi	cial											
<40	00> 49)													
Met 1	Ala	ser	Ser	Asn 5	Leu	Ile	Lys	G] n	Leu 10	Gln	Glu	Arg	Gly	Leu 15	Val
Ala	Gln	val	Thr 20	Asp	G1u	Glu	Ala	Leu 25	Ala	Glu	Arg	Leu	Ala 30	Gln	Gly
Pro	Ile	Ala 35	Leu	va1	Cys	Gly	Phe 40	Asp	Pro	Thr	Ala	Asp 45	Ser	Leu	His
Leu	G]y	ніѕ	Leu	val	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	G1n	G1n	Ala
G] y 65	His	Lys	Pro	val	A1a 70	Leu	val	Gly	Gly	А1а 75	Thr	Gly	Leu	Ile	G]y 80
Asp	Pro	Ser	Phe	Lys 85	Ala	Ala	Glu	Arg	Lys 90	Leu	Asn	Thr	Glu	G]u 95	Thr
Val	Gln	Glu	Trp 100	val	Asp	Lys	Ile	Arg 105	Lys	GÌn	val	Ala	Pro 110	Phe	Leu
Asp	Phe	ASP 115	Cys	Gly	Glu	Asn	Ser 120	Ala	Ile	Ala	Ala	Asn 125	Asn	Туг	Asp
	Met Gln Thr 385 Glu Asn <21 <21 <21 <21 <21 <1 Ala Pro Leu Gly 65 Asp val	Met Glu Gln Pro 370 Thr Ile 385 Glu Asp Asn Tyr <210> 49 <211> 42 <212> PF <213> Ar <220> <223> Sir <400> 49 Met Ala 1 Ala Gln Pro Ile Leu Gly Gly His 65 Asp Pro Val Gln	Met Glu Lys 355 Gln Pro Ser 370 Thr Ile Asn 385 Glu Asp Arg Asn Tyr Cys <210> 49 <211> 424 <212> PRT <213> Artificial <220> <223> Sintetas <400> 49 Met Ala Ser 1 Ala Gln Val Pro Ile Ala 35 Leu Gly His 50 Gly His Lys 65 Asp Pro Ser Val Gln Glu	Met Glu Lys Gly Gln Pro Ser Arg Thr Ile Asn Gly 370 Glu Asp Arg Leu Asn Tyr Cys Leu 420 4210> 49 4211> 424 4212> PRT 4213> Artificial 420> 420> 420> 420> 7223> Sintetasa artifi 4400> 49 Met Ala Ser Ser Ala Gln Val Thr 20 Pro Ile Ala Leu 35 Leu Gly His Leu Gly His Lys Pro Asp Pro Ser Phe Val Gln Glu Trp 100	Met Glu Lys Gly Ala Gln Pro Ser Arg Gly Thr Ile Asn Gly Glu Glu Asp Arg Leu Phe 405 Asn Tyr Cys Leu Ile <210> 49 <211> 424 <212> PRT <213> Artificial <220> <223> Sintetasa artificial <400> 49 Met Ala Ser Ser Asn Ala Gln Val Thr Asp Pro Ile Ala Leu Val 35 Leu Gly His Leu Val Gly His Leu Val Asp Pro Ser Phe Lys 85 Val Gln Glu Trp Val	Met Glu Lys Gly Ala Asp Gln Pro Ser Arg Gly Gln Thr Ile Asn Gly Glu Lys 385 Glu Asp Arg Leu Phe Gly Asn Tyr Cys Leu Ile Cys 4210> 49 <210> 49 <211> 424 <212> PRT <213> Artificial <220> <223> Sintetasa artificial <400> 49 Met Ala Ser Ser Asn Leu 1 Ala Gln Val Thr Asp Glu 20 Pro Ile Ala Leu Val Cys Leu Gly His Leu Val Pro Gly His Lys Pro Val Ala 65 Asp Pro Ser Phe Lys Ala 85 Val Gln Glu Trp Val Asp	Met Glu Lys Gly Ala Asp Leu Gln Pro Ser Arg Gly Gln Ala 370 Thr Ile Asn Gly Glu Lys Gln 385 Glu Asp Arg Leu Phe Gly Arg Asn Tyr Cys Leu Ile Cys Trp <210> 49 <211> 424 <212> PRT <213> Artificial <220> <223> Sintetasa artificial <400> 49 Met Ala Ser Ser Asn Leu Ile 1 Ala Gln Val Thr Asp Glu Glu 20 Pro Ile Ala Leu Val Cys Gly Leu Gly His Leu Val Pro Leu 50 Gly His Lys Pro Val Ala Leu 70 Asp Pro Ser Phe Lys Ala Ala Val Gln Glu Trp Val Asp Lys	Met Glu Lys Gly Ala Asp Leu Met 360 Gln Pro Ser Arg Gly Gln Ala Arg 375 Thr Ile Asn Gly Glu Lys Gln Ser 385 Glu Asp Arg Leu Phe Gly Arg Phe 405 Asn Tyr Cys Leu Ile Cys Trp Lys 211> 424	Met Glu Lys Gly Ala Asp Leu Met Gln Gln Pro Ser Arg Gly Gln Ala Arg Lys 370 Thr Ile Asn Gly Glu Lys Gln Ser Asp Glu Asp Arg Leu Phe Gly Arg Phe Thr Asn Tyr Cys Leu Ile Cys Trp Lys 210> 49 211> 424 212> PRT 213> Artificial 220> 2223> Sintetasa artificial 400> 49 Met Ala Ser Ser Asn Leu Ile Lys Gln Ala Gln Val Thr Asp Glu Glu Ala Leu 20 Pro Ile Ala Leu Val Cys Gly Phe Asp 40 Leu Gly His Leu Val Pro Leu Leu Cys Gly His Lys Pro Val Ala Leu Val Gly Asp Pro Ser Phe Lys Ala Ala Glu Arg Val Gln Glu Trp Val Asp Lys Ile Arg 100	Met Glu Lys Gly Ala Asp Leu Met Gln Ala Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu 405 Asn Tyr Cys Leu Ile Cys Trp Lys 210> 49 211> 424 212> PRT 213> Artificial 220> 223> Sintetasa artificial 400> 49 Met Ala Ser Ser Asn Leu Ile Lys Gln Leu 1 Ala Gln Val Thr Asp Glu Glu Ala Leu Ala 20 Pro Ile Ala Leu Val Cys Gly Phe Asp Pro Leu Gly His Leu Val Pro Leu Cys Leu 50 Gly His Lys Pro Val Ala Leu Val Gly Gly Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys 85 Val Gln Glu Trp Val Asp Lys Ile Arg Lys 105 Val Gln Glu Trp Val Asp Lys Ile Arg Lys 105	Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu 355 Gly Ala Asp Leu Met Gln Ala Leu 375 Gly Gln Ala Arg Lys Thr Ile Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu 385 Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu 405 Gly Arg Phe Thr Leu Leu 405 Asn Tyr Cys Leu Ile Cys Trp Lys 2210> 49 2211> 424 2212> PRT 2213> Artificial 220> 2223> Sintetasa artificial 400> 49 Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln 10 Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu 20 Pro Ile Ala Leu Val Cys Gly Phe Asp Pro Thr 40 Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys 50 Gly His Lys Pro Val Ala Leu Val Gly Gly Ala 75 Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Gln 100 Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln 100	Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala 375 Arg Lys Thr Ile Ala 380 Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr 385 Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Asn Tyr Cys Leu Ile Cys Trp Lys 210> 49 211> 424 212> PRT 213> Artificial 220> 223> Sintetasa artificial 400> 49 Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg 20 Pro Ile Ala Leu Val Cys Gly Phe Asp Pro Thr Ala Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg 65 Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr 75 Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Cln Val Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val	Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp 365 Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser 370 Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe 385 Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Asn Tyr Cys Leu Ile Cys Trp Lys Asn Tyr Cys Leu Ile Cys Trp Lys	Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser 365 Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn 377 Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe 385 Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Asn Tyr Cys Leu Ile Cys Trp Lys 2210> 49 2210> 49 2210> PRT 2220> 2220> 2223> Sintetasa artificial 400> 49 Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala 30 Pro Ile Ala Leu Val Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gly His Lys Pro Val Ala Clu Val Gly Gly Ala Thr Gly Leu Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro 100 Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro 110	Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu 365 Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys 405 Asn Tyr Cys Leu Ile Cys Trp Lys 210> 49 211> 424 212> PRT 213> Artificial 220> 220> 223> Sintetasa artificial 400> 49 Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Pro Ile Ala Leu Val Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe

5	Trp	Phe 130	Gly	Asn	Met	Asn	va1 135	Leu	Thr	Phe	Leu	Arg 140	ASP	Ile	Gly	Lys
	His 145	Phe	Ser	val	Asn	Gln 150	Met	Ile	Asn	Lys	Glu 155	Ala	val	Lys	Gln	Arg 160
10	Leu	Asn	Arg	Glu	Asp 165	Gln	Gly	Ile	ser	Phe 170	Thr	Glu	Phe	Ser	Tyr 175	Asn
15	Leu	Leu	Gln	Gly 180	Tyr	Ser	Аlа	Ala	Cys 185	Ala	ASN	Lys	Gln	Tyr 190	G] y	val
20	val	Leu	Gln 195	Ile	Gly	Gly	Ser	Asp 200	Gln	Тгр	Gly	Asn	11e 205	Thr	Ser	Gly
	Ile	Asp 210	Leu	Thr	Arg	Arg	Leu 215	His	G1n	Asn	ĢÌn	Va1 220	Phe	Gly	Leu	Thr
25	Va1 225	Pro	Leu	Ile	Thr	Lys 230	Ala	Asp	Gly	Thr	Lys 235	Phe	Gly	Lys	Thr	G1u 240
30	Gly	Gly	Ala	val	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	ser	Pro	туг	Lys 255	Phe
	Туг	Gln	Phe	Trp 260	Ile	Asn	Thr	Ala	Asp 265	Ala	Asp	Val	Туг	Arg 270	Phe	Leu
35	Lys	Phe	Phe 275	Thr	Phe	Met	Ser	Ile 280	Glu	Glu	īle	Asn	A1a 285	Leu	Glu	Glu
40	Glu	Asp 290	Lys	Asn	Ser	Gly	Lys 295	Ala	Pro	Arg	Ala	G]n 300	Туг	val	Leu	Ala
45	G1u 305	Gln	Val	Thr	Arg	Leu 310	val	His	Gly	Glu	G]u 315	Gly	Leu	GÌn	Ala	A1a 320
	Lys	Arg	Ile	Thr	G1u 325	Cys	Leu	Phe	Ser	Gly 330	Ser	Leu	Ser	Аlа	Leu 335	Ser
50	G1u	Ala	Asp	Phe 340	Glu	Gln	Leu	Ala	G]n 345	Asp	Gly	val	Pro	Met 350	val	Glu
55	Met	Glu	Lys 355	Gly	Ala	Asp	Leu	Met 360	Gln	Ala	Leu	val	Asp 365	Ser	Glu	Leu
	Gln	Pro 370		Arg	GТу	Gln	A1a 375	Arg	Lys	Thr	Ile	A1a 380	Ser	Asn	Ala	Ile
60	Thr 385	Ile	Asn	Gly	Glu	Lys 390	G In	ser	Asp	Pro	G1u 395	туг	Phe	Phe	Lys	G1u 400

5	Glu	Asp	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys
3	Asn	туг	Cys	Leu 420	Ile	Cys	Trp	Lys								
10	<21 <21	0> 50 1> 42 2> PF 3> Ar	4 RT													
15	<22 <22	20> 23> Sii	ntetasa	a artifio	cial											
	<40	00> 50														
20	Met 1	Ala	Ser	Ser	Asn 5	Leu	Ile	Lys	Gln	Leu 10	Gln	Glu	Arg	Gly	Leu 15	val
25	Ala	Ģln	val	Thr 20	Asp	Glu	Glu	Αla	Leu 25	Ala	Glu	Arg	Leu	A1a 30	G1n	GТу
30	Pro	Ile	A]a 35	Leu	Leu	Cys	GΊy	Phe 40	Asp	Pro	Thr	Ala	ASP 45	Ser	Leu	His
	Leu	G1y 50	His	Leu	val	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	GÌn	Gln	Ala
35	Gly 65	His	Lys	Pro	Val	A1a 70	Leu	Va1	Gly	Gly	Ala 75	Thr	Gly	Leu	Ile	61y 80
40	Asp	Pro	Ser	Phe	Lys 85	Ala	Ala	Glu	Arg	Lys 90	Leu	Asn	Thr	Glu	Glu 95	Thr
	Va1	Gln	Glu	Trp 100	val	Asp	Lys	Ile	Arg 105	Lys	Gln	Va1	Ala	Pro 110	Phe	Leu
45	Asp	Phe	ASP 115	Cys	Gly	Glu	Asn	ser 120	Ala	Ile	Ala	Àlа	Asn 125	Asn	Туг	Asp
50	Тгр	Phe 130	Gly	Asn	Met	Aşn	va1 135	Leu	Thr	Phe	Leu	Arg 140	Asp	Ile	Gly	Lys
EE	His 145	Phe	Ser	va1	Asn	G]n 150	Met	Ile	Asn	Lys	Glu 155	Ala	val	Lys	G1n	Arg 160
55	Leu	Asn	Arg	Glu	Asp 165	Gln	Gly	Ile	ser	Phe 170	Thr	Glu	Phe	Ser	Tyr 175	Asn
60	Leu	Leu	GÌn	Gly 180	туг	ser	ΑΊа	ΑΊа	Cys 185	Аlа	Asn	Lys	G1n	Туг 190	Gly	va1

5	val	Leu	Gln 195	Ile	Gly	Gly	Ser	ASP 200	Gln	Trp	Gly	Asn	11e 205	Thr	Ser	Gly
	īle	Asp 210	Leu	Thr	Arg	Arg	Leu 215	His	GÌn	Asn	Gln	va1 220	Phe	Gly	Leu	Thr
10	Va1 225	Pro	Leu	Ile	Thr	Lys 230	Ala	ASP	Gly	Thr	Lys 235	Phe	Gly	Lys	Thr	G1u 240
15	Gly	Gly	Ala	Va1	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	Ser	Pro	Туг	Lys 255	Phe
	туг	Gln	Phe	Trp 260	Ile	Asn	Thr	Ala	ASP 265	Ala	Asp	val	туг	Arg 270	Phe	Leu
20	Lys	Phe	Phe 275	Thr	Phe	Met	Ser	Ile 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	Glu
25	Ģlu	Asp 290	Lys	Asn	Ser	Gly	Lys 295	Ala	Pro	Arg	Ala	G]n 300	туг	val	Leu	Ala
30	G] u 305	Gln	val	Thr	Arg	Leu 310	val	нis	Gly	Glu	Glu 315	Gly	Leu	Gln	Ala	A1a 320
	Lys	Arg	Ile	Thr	G1u 325	Cys	Leu	Phe	Ser	G]y 330	Ser	Leu	Ser	Ala	Leu 335	Ser
35	Glu	Ala	Asp	Phe 340	Glu	Gln	Leu	Ala	G1n 345	Asp	Gly	val	Pro	Met 350	va1	Glu
40	Met	Glu	Lys 355	Gly	Ala	Asp	Leu	Met 360	Gln	Ala	Leu	va1	Asp 365	Ser	Glu	Leu
	Gln	Pro 370	Ser	Arg	Gly	Gln	Ala 375	Arg	Lys	Thr	Ile	Ala 380	Ser	Asn	Ala	Ile
45	Thr 385	Ile	Asn	Gly	Glu	Lys 390	Gln	Ser	Asp	Pro	G]u 395	Туг	Phe	Phe	Lys	G1u 400
50	Glu	Asp	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys
	ASN	туг	Cys	Leu 420	Ile	cys	тгр	Lys								
55	<2	10> 5 ⁻ 11> 42 12> Pl	24	720												
60	<2:	13> A 20> 23> S		sa artif	icial											
65	<4	00> 5°	1													

	Туг	Gln	Phe	7rp 260	Ile	Asn	Thr	Ala	Asp 265	Ala	Asp	Val	Tyr	Arg 270	Phe	Leu
5	Lys	Phe	Phe 275	Thr	Phe	Met	Ser	11e 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	Glu
10	Glu	Asp 290	Lys	Asn	Ser	Gly	Lys 295	Ala	Pro	Arg	Ala	G]n 300	Туг	٧a٦	Leu	Ala
15	G1u 305	Gln	val	Thr	Arg	Leu 310	val	His	Gly	Glu	Glu 315	Gly	Leu	Gln	Ala	Ala 320
. •	Lys	Arg	Ile	Thr	G1u 325	Cys	Leu	Phe	Ser	G]y 330	Ser	Leu	Ser	Ala	Leu 335	Ser
20	Glu	Ala	Asp	Phe 340	Glu	Gln	Leu	Ala	G]n 345	Asp	Gly	va1	Pro	Met 350	۷al	Glu
25	Met	Glu	Lys 355	Gly	Ala	Asp	Leu	Met 360	Gln	Ala	Leu	va1	Asp 365	Ser	Glu	Leu
	Gln	Pro 370	Ser	Arg	Gly	GÌn	A1a 375	Arg	Lys	Thr	Ile	A1a 380	Ser	Asn	Ala	Ile
30	Thr 385	Ile	Asn	Gly	Glu	Lys 390	G1n	ser	Asp	Pro	Glu 395	Туг	Phe	Phe	Lys	G]u 400
35	Glu	Asp	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys
40	Asn	туг	Cys	Leu 420	Ile	Cys	Тгр	Lys								
45	<2 <2	210> 5 211> 4 212> P 213> A	24 PRT	I												
	<2	220> 223> S			icial											
50	<4	100> 5	2													
55	Met 1	Ala	Ser	Ser	Asn 5	Leu	Ile	Lys	Gln	Leu 10	ĢÌn	Glu	Arg	Gly	Leu 15	val
	Ala	Gln	val	Thr 20	Asp	Glu	Glu	Ala	Leu 25	Ala	Glu	Arg	Leu	Ala 30	Gln	G1y
60	Pro	Ile	A1a 35	Leu	Ile	Cys	Gly	Phe 40	Asp	Pro	Thr	Ala	Asp 45	Ser	Leu	His
	Leu	G]v	His	Leu	Val	Pro	Leu	Leu	CVS	Leu	Lvs	Ara	Phe	G]n	Gln	Αla

50 55 60

5 Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr 10 Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu 100 110 15 Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asp Tyr Asp 20 Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg 25 Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn 165 170 175 30 Leu Leu Gln Gly Tyr Asn Phe Ala Cys Val Asn Lys Gln Tyr Gly Val Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205 35 Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr 210 220 40 Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu 230 235 240 Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe 45 Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu 260 265 50 Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu 275 280 55 Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala 290 300 Glu Gln val Thr Arg Leu val His Gly Glu Glu Gly Leu Gln Ala Ala 305 310 320 60

_	Ly	S AI	rg 1	ie ii	3	25	cys	Leu	Pne	Ser	330	Ser	Leu	Ser	Ald	335	36
5	G1	u A	la A		he G	lu	Gln	Leu	Ala	G1n 345	Asp	Gly	Val	Pro	Met 350	۷a٦	G1
10	ме	t Ģ	lu L	ys G	1y A	la .	Asp	Leu	Met 360	Gln	ΑΊa	Leu	va1	ASP 365	ser	Glu	Le
15	G1	n P	ro 5	er A	rg G	ìу	Gln	Ala 375	Arg	Lys	Thr	Ile	Ala 380	ser	Asn	Ala	11
	Th 38	r I	le A	sn G	1y 6	ilu	Lys 390	Gln	ser	Asp	Pro	G]u 395	Туг	Phe	Phe	Lys	G1 40
20	G1	u A:	sp A	rg L		he 05	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Ly
25	As	n T	yr C	ys L 4	eu 1 20	le:	Cys	Тгр	Lys								
30	<	210> 211> 212> 213>	424	cial													
35	<	220> 223> 400>	Sinte	tasa a	rtificia	al											
40	Met /	Ala	Ser	Ser	Asn 5	Lei	u Il	e Ly	s G1	n Le	iu G1	n G1	u Ar	g G1	y Le 15	u Va	11
45	Ala	Gln	val	Thr 20	Asp	G1:	u G1	u Al	a Le 25	u Al	a G1	и Аг	g Le	1A U	a Gl	n G	ly
	Pro	Ile	A7a 35	Leu	Thr	Cy	s G1	y Ph 40	ie As	ip Pr	o Th	וא חו	a As 45	p Se	r Le	u H	is
50	Leu	61y 50	His	Leu	۷a٦	Pro	o Le 55	u Le	u Cy	/S Le	eu Ly	s Ar 60	g Ph	ie G1	n Gl	n A	la
55	G] y	His	Lys	Pro	۷a٦	A1.	a Le	u Va	il Gl	y G	ly A1	a Th	r G1	y Le	iu I1	e G	ly 0
60	Asp	Pro	5er	Phe	Lys 85	Al	a Al	a G1	u Ar	g Ly 90	/S LE	u As	n Th	ır Gl	u G]	u Ti	hr
60	Val	G1n	Glu	Trp 100	۷a٦	As	p Ly	s IT	le A1	5 L)	/s G1	n Va	1 A1	a Pr 11	o Ph	e L	eu
65	Asp	Phe	Asp	Cys	Gly	G1	u As	n Se	r A	a I	le Al	a Al	a As	n As	in Ty	r A:	sp

115	120	
115	170	125

5	Тгр	Phe 130	Gly	Asn	Met	Asn	Va1 135	Lev	Thr	Phe	Leu	Arg 140	Asp	Ile	Gly	Ly
10	His 145	Phe	Ser	Val	Asn	G]n 150	Met	Ile	Asn	Lys	Glu 155	Ala	val	Lys	Gln	Arg 160
15	Leu	Asn	Arg	G1u	Asp 165	Gln	Gly	Ile	ser	Phe 170	Thr	Glu	Phe	Ser	īyç	ASI
	Leu	Leu	Gln	G]y 180	Туг	Ser	Аlа	Ala	Cys 185	Leu	Asn	Lys	G1n	Tyr 190	Gly	va [*]
20	Val	Leu	G]n 195	Ile	Gly	Gly	Ser	Asp 200	Gln	Trp	Gly	Asn	11e 205	Thr	Ser	G1 _y
25	Ile	ASP 210	Leu	Thr	Arg	Arg	Leu 215	His	Gln	Asn	G1n	Va1 220	Phe	Gly	Leu	Thi
	Va1 225	Pro	Leu	Ile	Thr	Lys 230	Ala	Asp	GТу	Thr	Lys 235	Phe	Gly	Lys	Thr	G10 240
30	Gly	Gly	Ala	val	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	ser	Pro	Туг	Lys 255	Phe
35	Tyr	GÌn	Phe	Trp 260	Ile	ASN	Thr	Ala	Asp 265	Ala	ASP	va1	туг	Arg 270	Phe	Lei
40	Ļys	Phe	Phe 275	Thr	Phe	Met	ser	Ile 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	Glu
10	Glu	Asp 290	Lys	Asn	Ser	Gly	Lys 295	Ala	Pro	Arg	Αla	G]n 300	Туг	Val	Leu	Αla
45	Glu 305	Gln	val	Thr	Arg	Leu 310	va1	His	Gly	Glu	Glu 315	Gly	Leu	Gln	Ala	A1a 320
50	Lys	Arg	Ile	Thr	G]u 325	cys	Leu	Phe	Ser	G]y 330	Ser	Leu	ser	Ala	Leu 335	sei
	Glu	Ala	Asp	Phe 340	G1u	G1n	Leu	Αla	G]n 345	Asp	GТу	Va1	Pro	Met 350	Va1	G٦ι
55	Met	G1u	Lys 355	Gly	Аlа	Asp	Leu	Met 360	Gln	Ala	Leu	vàl	Asp 365	Ser	Glu	Let
60	Gln	Pro 370	Ser	Arg	GТу	Gln	A1a 375	Arg	Lys	Thr	Ile	A1a 380	Ser	Asn	ΑΊa	rle

5	Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro 385	Glu Tyr Phe Phe Lys Gl 395 40
	Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu 405 410	Leu Arg Arg Gly Lys Ly 415
10	Asn Tyr Cys Leu Ile Cys Trp Lys 420	×
15	<210> 54 <211> 424 <212> PRT <213> Artificial	
20	<220> <223> Sintetasa artificial	
	<400> 54	
25	Met Ala Ser Ser Asn Leu Ile Lys Gln Leu G 1 5 10	in Glu Arg Gly Leu Val 15
30	Ala Gln Val Thr Asp Glu Glu Ala Leu Ala G 20 25	lu Arg Leu Ala Gln Gly 30
	Pro Ile Ala Leu Gly Cys Gly Phe Asp Pro T 35 40	hr Ala Asp Ser Leu His 45
35	Leu Gly His Leu Val Pro Leu Leu Cys Leu L 50 55	ys Arg Phe Gln Gln Ala 60
40	Gly His Lys Pro Val Ala Leu Val Gly Gly A 65 70 7	la Thr Gly Leu Ile Gly 5 80
4-	Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys L 85 90	eu Asn Thr Glu Glu Thr 95
45	val Gln Glu Trp Val Asp Lys Ile Arg Lys G 100 105	ln Val Ala Pro Phe Leu 110
50	Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile A 115 120	la Ala Asn Asn Tyr Asp 125
55	Trp Phe Gly Asn Met Asn Val Leu Thr Phe L 130	eu Arg Asp Ile Gly Lys 140
	His Phe Ser Val Asn Gln Met Ile Asn Lys G 145	ilu Ala Val Lys Gln Arg 55 160
60	Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe T 165	hr Glu Phe Ser Tyr Asn 175
65	Leu Leu Gln Gly Tyr Ser Met Ala Cys Leu A	sn Lys Gln Tyr Gly Val

		180		185	=	190
5	val Leu Glr 19	ile Gly	Gly ser Asp 200	Gln Trp Gl	y Asn Ile 205	Thr Ser Gly
10	Ile Asp Leo 210	Thr Arg	Arg Leu His 215	Gln Asn Gl	n Val Phe 220	Gly Leu Thr
	Val Pro Lei 225	lle Thr	Lys Ala Asp 230	Gly Thr Ly 23	s Phe Gly	Lys Thr Glu 240
15	Gly Gly Ala	val Trp 245	Leu Asp Pro	Lys Lys Th	r Ser Pro	Tyr Lys Phe
20	Tyr Gln Pho	260 Trp 11e	Asn Thr Ala	ASP Ala AS 265	p val Tyr	Arg Phe Leu 270
	Lys Phe Phe 27	Thr Phe	Met Ser Ile 280	Glu Glu Il	e Asn Ala 285	Leu Glu Glu
25	Glu Asp Lys 290	s Asn Ser	Gly Lys Ala 295	Pro Arg Al	a Gln Tyr 300	Val Leu Ala
30	Glu Gln Va 305	Thr Arg	Leu Val His 310	s Gly Glu Gl 31	u Gly Leu .S	Gln Ala Ala 320
35	Lys Arg Ile	Thr Glu 325	Cys Leu Phe	ser Gly Se 330	r Leu Ser	Ala Leu Ser 335
	Glu Ala As	Phe Glu 340	Gln Leu Ala	Gln Asp Gl 345	y Val Pro	Met Val Glu 350
40	Met Glu Ly 35	Gly Ala	Asp Leu Met 360	Gln Ala Le	u Val Asp 365	Ser Glu Lei
45	Gln Pro Se 370	r Arg Gly	Gln Ala Arg 375	Lys Thr Il	e Ala Ser 380	Asn Ala Ile
	Thr Ile Ass 385	n Gly Glu	Lys Gln Ser 390	r Asp Pro Gl		Phe Lys Glu
50	Glu Asp Ar	g Leu Phe 405	Gly Arg Phe	Thr Leu Le 410	eu Arg Arg	Gly Lys Lys 415
55	Asn Tyr Cy	s Leu Ile 420	Cys Trp Ly	5		
60	<210> 55 <211> 424 <212> PRT <213> Artificia	al				
65	<220> <223> Sinteta	ısa artificial				

5	Met 1	Ala	Ser	Ser	Asn 5	Leu	Ile	Lys	Gln	Leu 10	Gln	Glu	Arg	Gly	Leu 15	val
	Ala	Gln	Val	Thr 20	Asp	Glu	Glu	Ala	Leu 25	Αla	G1u	Arg	Leu	Ala 30	Gln	Gly
10	Pro	Ile	Ala 35	Leu	Thr	Cys	Gly	Phe 40	Asp	Pro	Thr	Ala	ASP 45	ser	Leu	нis
15	Leu	Gly 50	His	Leu	val	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	Gln	Gln	Ala
20	G] y 65	His	Lys	Pro	val	Ala 70	Leu	٧a٦	Gly	Gly	Ala 75	Thr	Gly	Leu	Ile	G] y 80
	Asp	Pro	Ser	Phe	Lys 85	Ala	Ala	Glu	Arg	Lys 90	Leu	Asn	Thr	G1u	G1u 95	Thr
25	val	Gln	Glu	Trp 100	va1	Asp	Lys	Ile	Arg 105	Lys	GÌn	va1	Ala	Pro 110	Phe	Leu
30	Asp	Phe	Asp 115	Cys	Gly	Glu	Asn	Ser 120	Ala	Ile	Ala	Ala	Asn 125	Asn	Туг	Asp
	Trp	Phe 130	Gly	Asn	Met	Asn	val 135	Leu	Thr	Phe	Leu	Arg 140	ASP	Ile	Gly	Lys
35	His 145	Phe	Ser	val	Asn	G]n 150	Met	Ile	Asn	Lys	Glu 155	Ala	va1	Lys	Gln	Arg 160
40	Leu	Asn	Arg	G1u	ASP 165	Gln	Gly	Ile	ser	Phe 170	Thr	Glu	Phe	ser	Tyr 175	Asn
45	Leu	Leu	Gln	Gly 180	Туг	Ser	Ala	Ala	Cys 185	Leu	Asn	Lys	Gln	Tyr 190	Gly	Val
10	val	Leu	G]n 195	Ile	Gly	Gly	Ser	ASP 200	Gln	тгр	Gly	Asn	11e 205	Thr	Ser	G1y
50	Ile	Asp 210	Leu	Thr	Arg	Arg	Leu 215	His	Gln	Asn	Gln	Va1 220	Phe	Gly	Leu	Thr
55	Va1 225	Pro	Leu	Ile	Thr	Lys 230	Αla	ASP	Gly	Thr	Lys 235	Phe	GΊy	Lys	Thr	G1u 240
	Gly	Gly	Ala	va1	тгр	Leu	Asp	Pro	Lys	Lys	Thr	Ser	Pro	Tyr	Lys	Phe

					245					250	*				255	
5	Туг	G1n	Phe	Trp 260	Ile	Asn	Thr	Ala	ASP 265	Ala	Asp	va1	туг	Arg 270	Phe	Leu
10	Lys	Phe	Phe 275	Thr	Phe	Met	Ser	11e 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	Glu
15	Glu	Asp 290	Lys	Asn	Ser	Gly	Lys 295	Ala	Pro	Arg	Ala	G]n 300	Туг	val	Leu	Ala
	G1u 305	Gln	Val	Thr	Arg	Leu 310	val	His	G1y	Glu	G]u 315	Gly	Leu	G1n	Ala	A1a 320
20	Lys	Arg	Ile	Thr	G1u 325	Cys	Leu	Phe	Ser	Gly 330	Ser	Leu	Ser	Ala	Leu 335	Ser
25	Glu	Ala	Asp	Phe 340	Glu	Gln	Leu	Аlа	G1n 345	ASP	Gly	va1	Pro	Met 350	va1	Glu
	Met	G1u	Lys 355	Gly	Ala	Asp	Leu	меt 360	Gln	Ala	Leu	Val	Asp 365	Ser	Glu	Leu
30	Gln	Pro 370	Ser	Arg	GТу	Gln	A1a 375	Arg	Lys	Thr	Ile	Ala 380	Ser	Asn	Ala	Ile
35	Thr 385	Ile	Asn	G1y	G1u	Lys 390	Gln	Ser	Asp	Pro	G]u 395	Туг	Phe	Phe	Lys	G1u 400
40	Glu	Asp	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys
40	Asn	туг	Cys	Leu 420	ıle	Cys	Тгр	Lys		*	*					
45	<2	10> 56 11> 42	24													
50	<2 <2	12> PI 13> Ai 20> 23> Si	rtificial	o ortifi	cial											
55		23> 31 00> 56		a arun	Ciai											
33	Met	Ala	Ser	Ser	Asn S	Leu	Ile	Lys	Gln	Leu 10	GÌn	Glu	Arg	Gly	Leu 15	Val
60	Ala	Gln	val	Thr 20	Asp	Glu	Glu	Ala	Leu 25	Ala	Glu	Arg	Leu	A1a 30	G1n	Gly
65	Pro	Ile	Ala 35	Leu	Ser	Cys	Gly	Phe 40	Asp	Pro	Thr	Аlа	ASP 45	ser	Leu	His

5	Leu	G1y 50	His	Leu	Val	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	Gln	Gln	Ala
	G]y 65	His	Lys	Pro	va1	A]a 70	Leu	val	G1y	GТу	A1a 75	Thr	G1y	Leu	Ile	G] y 80
10	Asp	Pro	Ser	Phe	Lys 85	Ala	Ala	Glu	Arg	Lys 90	Leu	Asn	Thr	Glu	Glu 95	Thr
15	Val	Gln	Glu	Trp 100	va1	Asp	Lys	Ile	Arg 105	Lys	Gln	val	Ala	Pro 110	Phe	Leu
20	ASP	Phe	ASP 115	Cys	Gly	Glu	Asn	Ser 120	Ala	Ile	Ala	Ala	Asn 125	Asn	Туг	ASP
	Trp	Phe 130	Gly	Asn	Met	Asn	val 135	Leu	Thr	Phe	Leu	Arg 140	Asp	Ile	Gly	Lys
25	ніs 145	Phe	Ser	va1	Asn	G]n 150	Met	Ile	Asn	Lys	G] u 155	Ala	val	Lys	Gln	Arg 160
30	Leu	Asn	Arg	Glu	ASP 165	Gln	Gly	Ile	Ser	Phe 170	Thr	Glu	Phe	Ser	Tyr 175	Asn
	Leu	Leu	Gln	Gly 180	туг	Thr	Met	Αla	Cys 185	val	Asn	Lys	Gln	Tyr 190	GJY	val
35	val	Leu	G]n 195	ıle	Gly	Gly	ser	ASP 200	Gln	Тгр	Gly	Asn	11e 205	Thr	Ser	GTy
40	Ile	ASP 210	Leu	Thr	Arg	Arg	Leu 215	нis	GÌn	Asn	Gln	va1 220	Phe	Gly	Leu	Thr
45	Va1 225	Pro	Leu	Ile	Thr	Lys 230	Аlа	ASP	Gly	Thr	Lys 235	Phe	G1y	Lys	Thr	G1u 240
	Gly	Gly	Ala	va1	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	Ser	Pro	Туг	Lys 255	Phe
50	Туг	Gln	Phe	Trp 260	Ile	Asn	Thr	Аlа	Asp 265	Ala	Asp	val	Туг	Arg 270	Phe	Leu
55	Lys	Phe	Phe 275	Thr	Phe	Met	ser	11e 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	Glu
	Glu	Asp 290	Lys	Asn	ser	Gly	Lys 295	۸la	Pro	Arg	Ala	G1n 300	Туг	val	Leu	Αla
60	Ģlu	Gln	va1	Thr	Arg	Leu	val	ніѕ	Gly	Glu	Glu	Gly	Leu	G٦n	Ala	Αla

	305		310	31	5	320
5	Lys Arg I	le Thr Glu 325	Cys Leu Phe	Ser Gly Ser 330	r Leu Ser A	la Leu Ser 335
10	Glu Ala A	sp Phe Glu 340	Gln Leu Ala	Gln Asp Gly 345		et Val Glu 50
		ys Gly Ala	Asp Leu Met 360		val Asp 56 365	er Glu Leu
15	Gln Pro S 370	er Arg Gly	Gln Ala Arg 375	Lys Thr Il	e Ala Ser A 380	sn Ala Ile
20	Thr Ile A	asn Gly Glu	Lys Gln Ser 390	Asp Pro Gl	u Tyr Phe Pl	he Lys Glu 400
05	Glu Asp A	rg Leu Phe 405	Gly Arg Phe	Thr Leu Le	u Arg Arg G	ly Lys Lys 415
25	ASN TYP C	ys Leu Ile 420	Cys Trp Lys			
30	<210> 57 <211> 424 <212> PRT <213> Artifi					
35		etasa artificial				
40	<400> 57					
	Met Ala S	er Ser Asn 5	Leu Ile Lys	Gln Leu Gli 10	n Glu Arg G	ly Leu Val 15
45	Ala Gln V	al Thr Asp 20	Glu Glu Ala	Leu Ala Gli 25	u Arg Leu A	la Gln Gly O
50		ala Leu Ala IS	Cys Gly Phe 40	Asp Pro Th	r Ala Asp 50 45	er Leu His
	Leu Gly H 50	lis Leu Val	Pro Leu Leu 55	Cys Leu Ly	s Arg Phe G	ln Gln Ala
55	Gly His L	ys Pro Val	Ala Leu Val 70	Gly Gly Al	a Thr Gly L	eu Ile Gly 80
60	Asp Pro S	Ser Phe Lys 85	Ala Ala Glu	Arg Lys Le	u Asn Thr G	lu Glu Thr 95
65	Val Gln o	Slu Trp Val 100	Asp Lýs Ile	Arg Lys GT	n Val Ala P 1	ro Phe Leu 10

5	ASP	Phe	ASP 115	Cys	Gly	Glu	ASN	Ser 120	Ala	Ile	Ala	Ala	Asn 125	Asn	Туг	Asp
	Trp	Phe 130	Gly	Asn	Met	Asn	Val 135	Leu	Thr	Phe	Leu	Arg 140	Asp	Ile	Gly	Lys
10	His 145	Phe	Ser	val	Asn	Gln 150	Met	Ile	Asn	Lys	Glu 155	Ala	va1	Lys	Gln	Arg 160
15	Leu	Asn	Arg	G1u	Asp 165	G]n	Gly	Ile	ser	Phe 170	Thr	Glu	Phe	Ser	Tyr 175	Asn
	Leu	Leu	Gln	G]y 180	Туг	ser	Туг	Аlа	Cys 185	Leu	Asn	Lys	G1n	Tyr 190	Gly	۷a۱
20	val	Leu	G]n 195	Ile	Gly	Gly	Ser	Asp 200	Gln	Тгр	Gly	Asn	11e 205	Thr	Ser	Gly
25	Ile	ASP 210	Leu	Thr	Arg	Arg	Leu 215	His	Gln	Asn	Gìn	Va1 220	Phe	Gly	Leu	Thr
30	Va1 225	Pro	Leu	Ile	Thr	Lys 230	Αla	Asp	Gly	Thr	Lys 235	Phe	Gly	Lys	Thr	G1u 240
	Gly	Gly	Ala	val	Trp 245	Ļeu	Asp	Pro	Lys	Lys 250	Thr	Ser	Pro	туг	Lys 255	Phe
35	Туг	Gln	Phe	Trp 260	Ile	Asn	Thr	ΑΊa	Asp 265	Ala	Asp	val	Туг	Arg 270	Phe	Leu
40	Lys	Phe	Phe 275	Thr	Phe	Met	ser	11e 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	Glu
	Glu	Asp 290	Lys	Asn	Ser	GТу	Lys 295	Αla	Pro	Arg	Ala	G1n 300	Туг	Va1	Leu	Ala
45	Glu 305	Gln	Val	Thr	Arg	Leu 310	Val	ніѕ	Gly	Glu	Glu 315	Gly	Leu	Gln	Ala	A1a 320
50	Lys	Arg	Ile	Thr	G1u 325	Cys	Leu	Phe	Ser	Gly 330	Ser	Leu	Ser	Ala	Leu 335	Ser
	Glu	Ala	Asp	Phe 340	Glu	Gln	Leu	Ala	G1n 345	Asp	Gly	va1	Pro	Met 350	val	Glu
55	мет	Glu	Lys 355	Gly	Ala	ASP	Leu	Met 360	Gln	Ala	Leu	val	ASP 365	Ser	Glu	Leu
60	Gln	Pro	Ser	Arg	Gly	Gln	Ala	Arg	Lys	Thr	Ile	Αla	Ser	Asn	Αla	Ile

		370					375					380				
5	Thr 385	Ile	Asn	Gly	Glu	Lys 390	Gln	ser	Asp	Pro	Glu 395	Туг	Phe	Phe	Lys	G] u 400
10	Glu	Asp	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys
	Asn	Туг	Cys	Leu 420	Ile	Cys	Trp	Lys								
15	<21	10> 58 11> 42 2> PF	24													
20	<21 <22	13> Ar	tificial	a artifi	cial											
25	<40	00> 58	3													
	Met 1	Ala	Ser	Ser	Asn 5	Leu	Ile	Lys	Gln	Leu 10	Gln	Glu	Arg	Gly	Leu 15	val
30	Ala	Gln	Va1	Thr 20	Asp	Glu	Glu	Ala	Leu 25	Ala	Glu	Arg	Leu	A1a 30	Gln	Gly
35	Pro	Ile	A1a 35	Leu	Ala	Cys	Gly	Phe 40	Asp	Pro	Thr	Ala	ASP 45	ser	Leu	ніѕ
	Leu	G]y 50	His	Leu	val	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	G1n	G7n	Ala
40	Gly 65	His	Lys	Pro	va1	Ala 70	Leu	va1	Gly	Gly	Ala 75	Thr	Gly	Leu	Ile	G]y 80
45	Asp	Pro	Ser	Phe	Lys 85	Ala	Ala	Glu	Arg	Lys 90	Leu	Asn	Thr	Glu	G]u 95	Thr
50	val	Gln	Glu	Trp 100	val	ASP	Lys	Ile	Arg 105	Lys	Gìn	val	Ala	Pro 110	Phe	Leu
	Asp	Phe	ASP 115	Cys	GТу	Glu	Asn	Ser 120	Ala	Ile	Ala	Ala	Asn 125	Asn	Туг	Asp
55	тгр	Phe 130		Asn	Met	Asn	va1 135	Leu	Thr	Phe	Leu	Arg 140	Asp	Ile	GТу	Lys
60		Phe					83									
	Leu	Asn	Arg	Glu	ASP 165	Gln	Gly	Ile	ser	Phe 170	Thr	Glu	Phe	Ser	Tyr 175	Asn

	Leu	Leu	Gln	Gly 180	туг	Thr	Met	Ala	Cys 185	Cys	Asn	Lys	Gln	Tyr 190	Gly	Val
5	val	Leu	Gln 195	Ile	Gly	Gly	Ser	Asp 200	G1n	Trp	Gly	Asn	11e 205	Thr	Ser	Gly
10	Ile	ASP 210	Leu	Thr	Arg	Arg	Leu 215	His	Gln	Asn	Gln	va1 220	Phe	Gly	Leu	Thr
	va1 225	Pro	Leu	Ile	Thr	Lys 230	Αla	Asp	Gly	Thr	Lys 235	Phe	Gly	Lys	Thr	G1u 240
15	Gly	Gly	Ala	val	Trp 245	Leu	Asp	Pro	Ļys	Lys 250	Thr	ser	Pro	туг	Lys 255	Phe
20	туг	Gln	Phe	Trp 260	Ile	Asn	Thr	Ala	ASP 265	Ala	Asp	val	Туг	Arg 270	Phe	Leu
	Lys	Phe	Phe 275	Thr	Phe	Met	ser	11e 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	Glu
25	Ģ1u	ASP 290	Lys	Asn	ser	Gly	Lys 295	Ala	Pro	Arg	Ala	G1n 300	туг	val	Leu	Αla
30	G1u 305	Gln	val	Thr	Arg	Leu 310	va1	ніѕ	Ģly	Glu	Glu 315	Gly	Leu	Gln	Аlа	A1a 320
35	Lys	Arg	Ile	Thr	G]u 325	cys	Leu	Phe	Ser	G]y 330	Ser	Leu	Ser	Ala	Leu 335	Ser
	G1u	Ala	Asp	Phe 340	Ģlu	Gln	Leu	Ala	Gln 345	Asp	Gly	val	Pro	Met 350	Val	G1u
40	Met	Glu	Lys 355	Gly	Аlа	ASP	Leu	Met 360	Gln	Ala	Leu	val	ASD 365	ser	Glu	Leu
45	Gln	Pro 370	Ser	Arg	GΊy	Gln	A1a 375	Arg	Lys	Thr	Ile	A)a 380	Ser	ASI	Ala	Ile
	Thr 385	Ile	Asn	Gly	Glu	Lys 390	Gln	Ser	Asp	Pro	G1u 395	Туг	Phe	Phe	Lys	G]u 400
50	Glu	Asp	Arg	Leu	Phe 405	ĢΊy	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	G1y	Lys 415	Lys
55	Asn	Туг	Cys	Leu 420	Ile	Cys	Тгр	Lys					534			
	<21	0>	59													
60	<	<211> <212> :213>		ial												
65	<	<220> <223> <400>	Sintet	asa aı	tificial											

	Met 1	Ala	Ser	Ser	Asn 5	Leu	Ile	Lys	Gln	Leu 10	Gln	Glu	Arg	Gly	Leu 15	Va1
5	Ala	Gln	va1	Thr 20	ASP	Glu	Glu	Ala	Leu 25	Ala	G1u	Arg	Leu	A1a 30	Gln	Gly
10	Pro	Ile	A1a 35	Leu	Thr	Cys	Glу	Phe 40	Asp	Pro	Thr	Ala	ASP 45	Ser	Leu	His
15	Leu	61y 50	His	Leu	۷a۱	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	Gln	Gln	Ala
13	61 y	нis	Lys	Pro	val	Ala 70	Leu	va1	Gly	Gly	A1a 75	Thr	Gly	Leu	Ile	G]y 80
20	Asp	Pro	Ser	Phe	Lys 85	Ala	Ala	Glu	Arg	Lys 90	Leu	Asn	Thr	Glu	Glu 95	Thr
25	val	Gln	Glu	Trp 100	Val	Asp	Lys	Ile	Arg 105	Lys	Gln	Val	Ala	Pro 110	Phe	Leu
	Asp	Phe	ASP 115	Cys	Gly	Glu	Asn	Ser 120	Ala	Ile	Ala	Ala	ASN 125	Asn	Туг	Asp
30	Trp	Phe 130	Gly	Asn	Met	Asn	va1 135	Leu	Thr	Phe	Leu	Arg 140	Asp	Ile	Gly	Lys
35	нis 145	Phe	Ser	val	Asn	G]n 150	Met	Ile	Asn	Lys	G]u 155	Ala	val	Lys	Gln	Arg 160
	Leu	Asn	Arg	Glu	Asp 165	GÌn	Gly	Ile	Ser	Phe 170	Thr	Glu	Phe	Ser	Tyr 175	Asn
40	Leu	Leu	Gln	Gly 180	Туг	Thr	Phe	Ala	Cys 185	Met	Asn	Lys	Gln	Tyr 190	Gly	Val
45	val	Leu	Gln 195	Ile	Gly	Gly	Ser	Asp 200	Gln	Тгр	Gly	Asn	11e 205	Thr	Ser	Gly
50	Ile	ASP 210	Leu	Thr	Arg	Arg	Leu 215	His	Gln	Asn	Gln	Va1 220	Phe	Gly	Leu	Thr
50	va1 225	Pro	Leu	Ile	Thr	Lys 230	Ala	Asp	Gly	Thr	Lys 235	Phe	G1y	Lys	Thr	G1u 240
55																

60	Ala	Gln	val	Thr 20	Asp	Glu	Glu	Ala	Leu 25	۸٦a	Glu	Arg	Leu	Ala 30	Gln	Gly
55	Met 1	Ala	ser	Ser	Asn 5	Leu	Ile	Lys	Gln	Leu 10	Gln	Glu	Arg	Gly	Leu 15	va1
	<40	00> 60)													
50	<22 <22	20> !3> Sir	ntetasa	a artific	cial											
45	<21 <21	0> 60 1> 42 2> PF 3> Ar	4 RT													
40	Asn	туг	Cys	Leu 420	Ile	Cys	Trp	Lys								
	Glu	Asp	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys
35	Thr 385	Ile	Asn	Gly	Glu	Lys 390	Gln	Ser	Asp	Pro	G]u 395	Туг	Phe	Phe	Lys	G1u 400
30	Gln	Pro 370	Ser	Arg	Gly	Gln	Ala 375	Arg	Lys	Thr	Ile	Ala 380	Ser	Asn	Ala	Ile
	Met	Glu	Lys 355	Gly	Ala	Asp	Leu	Met 360	G1n	Ala	Leu	Va1	Asp 365	ser	Glu	Leu
25	Glu	Ala	Asp	Phe 340	Glu	Gln	Leų	Αla	G]n 345	Asp	Gly	Val	Pro	Met 350	Val	Glu
20	Lys	Arg	Ile	Thr	G1u 325	Cys	Leu	Phe	ser	G]y 330	Ser	Leu	ser	Ala	Leu 335	Ser
10	Glu 305	GÌn	Val	Thr	Arg	Leu 310	val	His	Gly	Glu	Glu 315	Gly	Leu	G]n	Ala	A1a 320
15	Glu	ASP 290	Lys	Asn	ser	Gly	Lys 295	Ala	Pro	Arg	Ala	G1n 300	Туг	val	Leu	Ala
10	Lys	Phe	Phe 275	Thr	Phe	Met	Ser	11e 280	Glu	G1u	Ile	Asn	A1a 285	Leu	Glu	Glu
5	туг	Gln	Phe	7rp 260	Ile	Asn	Thr	Ala	Asp 265	Ala	ASP	val	туг	Arg 270	Phe	Leu
	Gly	Gly	Ala	Val	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	Ser	Pro	ТУГ	Lys 255	Phe

5	Pro	Ile	Ala 35	Leu	Thr	Cys	Gly	Phe 40	ASP	Pro	Thr	Ala	ASP 45	Ser	Leu	His
	Leu	G1y 50	His	Leu	Va1	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	Gln	Gln	Ala
10	G1y 65	His	Lys	Pro	val	А1а 70	Leu	va1	GТу	Gly	Ala 75	Thr	Gly	Leu	Ile	G] y 80
15	Asp	Pro	ser	Phe	Lys 85	Ala	Ala	Glu	Arg	Lys 90	Leu	Asn	Thr	Glu	GTu 95	Thr
	val	Gln	Glu	Trp 100	val	Asp	Lys	īle	Arg 105	Lys	Gln	val	Ala	Pro 110	Phe	Leu
20	Asp	Phe	ASP 115	Cys	Gly	Glu	Asn	Ser 120	Аlа	Ile	Ala	Ala	Asn 125	Asn	Туг	Asp
25	Trp	Phe 130	Gly	ASN	Met	Asn	va1 135	Leu	Thr	Phe	Leu	Arg 140	Asp	Ile	Gly	Lys
30	His 145	Phe	Ser	va1	Asn	G]n 150	Met	Ile	Asn	Lys	Glu 155	Ala	val	Lys	Gln	Arg 160
	Leu	Asn	Arg	Glu	ASP 165	G٦n	Gly	Ile	Ser	Phe 170	Thr	G1u	Phe	Ser	Туг 175	ASI
35	Leu	Leu	Gln	Gly 180	Туг	Ser	Val	Ala	Cys 185	Leu	Asn	Lys	GÌn	Tyr 190	GТу	val
40	val	Leu	G7n 195	Ile	Gly	GТу	Ser	ASP 200	Gln	Тгр	Gly	Asn	Ile 205	Thr	ser	Gly
	Ile	Asp 210	Leu	Thr	Arg	Arg	Leu 215	His	Gln	Asn	Gln	va1 220	Phe	Gly	Leu	Thr
45	va1 225	Pro	Leu	Ile	Thr	Lys 230	Ala	Asp	Gly	Thr	Lys 235	Phe	Gly	Lys	Thr	G1u 240
50	Gly	Gly	Ala	val	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	Ser	Pro	туг	Lys 255	Phe
	Туг	G1n	Phe	Trp 260	Ile	Asn	Thr	Ala	Asp 265	Ala	Asp	Val	Туг	Arg 270	Phe	Leu
55	Lys	Phe	Phe 275	Thr	Phe	Met	Ser	11e 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	G1u
60	Glu	Asp 290	Lys	Asn	Ser	G٦y	Lys 295	Ala	Pro	Arg	Ala	G]n	туг	val	Leu	Ala

5	G1u 305	Gln	va1	Thr	Arg	Leu 310	va1	His	Gly	Glu	Glu 315	Gly	Leu	Gln	Ala	A1a 320
	Lys	Arg	Ile	Thr	G1u 325	Cys	Leu	Phe	Ser	G]y 330	Ser	Leu	Ser	Ala	Leu 335	Ser
10	Glu	Ala	Asp	Phe 340	Glu	Gln	Leu	Ala	G]n 345	ASP	Gly	val	Pro	Met 350	va1	Glu
15	Met	Glu	Lys 355	Gly	Ala	Asp	Leu	меt 360	Gln	Ala	Leu	٧a٦	Asp 365	Ser	Glu	Leu
	ĢÌn	Pro 370	Ser	Arg	GТу	Gln	Ala 375	Arg	Lys	Thr	Ile	Ala 380	Ser	Asn	Ala	Ile
20	Thr 385	Ile	Asn	Gly	G lu	Lys 390	G1n	Ser	Asp	Pro	G] u 395	Туг	Phe	Phe	Lys	G]u 400
25	Glu	Asp	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys
30	Asn	Туг	Cys	Leu 420	Ile	Cys	Trp	Lys								
35	<21 <21	0> 61 1> 42 2> PR 3> Art	RT													
40		:0> :3> Sir :0> 61	ntetasa	a artific	cial											
45	Met 1	Αla	Ser	Ser	Asn 5	Leu	Ile	Lys	G1n	Leu 10	Gln	Glu	Arg	Gly	Leu 15	va1
	Ala	Gln	va1	Thr 20	Asp	Glu	Glu	ΑΊa	Leu 25	Ala	Glu	Arg	Leu	A1a 30	G1n	Gly
50	Pro	Ile	Ala 35	Leu	val	Cys	Gly	Phe 40	Asp	Pro	Thr	Ala	ASP 45	Ser	Leu	ніѕ
55	Leu	G]y 50	His	Leu	val	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	G1n	Gln	Аlа
60	Gly 65	His	Lys	Pro	val	Ala 70	Leu	val	Gly	Gly	A1a 75	Thr	G1y	Leu	Ile	G]y 80
OO .	Asp	Pro	Ser	Phe	Lys	Αla	Ala	Glu	Arg	Lys	Leu	Asn	Thr	Glu	<u>G]u</u>	Thr

5	val	Gln	Glu	Trp 100	va1	А́SР	Lys	Ile	Arg 105	Lys	Gln	Val	Аlа	Pro 110	Phe	Leu
	ASP	Phe	ASP 115	Cys	Gly	Glu	Asn	Ser 120	Ala	Ile	Ala	Аlа	Asn 125	Asn	туг	Asp
10	Trp	Phe 130	Gly	Asn	Met	Ásn	va1 135	Leu	Thr	Phe	Leu	Arg 140	Asp	Ile	Gly	Lys
15	His 145	Phe	ser	val	Asn	G]n 150	Met	Ile	Asn	Lys	G]u 155	Ala	val	Lys	Gln	Arg
20	Leu	Asn	Arg	Glu	Asp 165	Gln	Gly	Ile	Ser	Phe 170	Thr	Glu	Phe	Ser	Tyr 175	Ası
	Leu	Leu	Gln	Gly 180	туг	Ser	Met	Αla	Cys 185	Thr	Asn	Lys	Gln	Tyr 190	Gly	Va1
25	val	Leu	Gln 195	Ile	Gly	Gly	Ser	Asp 200	Gln	Trp	Gly	Asn	11e 205	Thr	Ser	G1 _y
30	Ile	ASP 210	Leu	Thr	Arg	Arg	Leu 215	нis	Gln	Asn	Gln	va1 220	Phe	Gly	Leu	Thr
	Val 225	Pr.o	Leu	Ile	Thr	Lys 230	Ala	Asp	Gly	Thr	Lys 235	Phe	Gly	Lys	Thr	G11 240
35	Gly	Gly	Ala	Val	Trp 245	Leu	ASP	Pro	Lys	Lys 250	Thr	Ser	Pro	Туг	Lys 255	Phe
40	Tyr	Gln	Phe	Trp 260	Ile	Asn	Thr	Ala	Asp 265	Ala	Asp	Val	туг	Arg 270	Phe	Leu
45	Ļys	Phe	Phe 275	Thr	Phe	Met	ser	11e 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	Glu
	Ģlu	Asp 290	Lys	Asn	Ser	Gly	Lys 295	Ala	Pro	Arg	Ala	G]n 300	Tyr	val	Leu	Ala
50	Glu 305	Gln	val	Thr	Arg	Leu 310	val	His	Gly	Glu	G]u 315	Gly	Leu	Gln	Ala	A1a 320
55	Lys	Arg	Ile	Thr	G1u 325	Cys	Leu	Phe	ser	Gly 330	Ser	Leu	Ser	Ala	Leu 335	Ser
	Glu	Ala	Asp	Phe 340	Glu	Gln	Leu	Ala	G1n 345	Asp	Gly	Va1	Pro	Met 350	val	Gli
60	Met	G1u	Lys 355	Gly	Αla	Asp	Leu	Met 360	G1n	Ala	Leu	Va1	Asp 365	Ser	Glü	Lei

5	G1n	Pro 370	Ser	Arg	Gly	G1n	Ala 375	Arg	Lys	Thr	Ile	A1a 380	Ser	Asn	Ala	Ile
	Thr 385	Ile	Asn	Gly	Glu	Lys 390	Gìn	Ser	Asp	Pro	G1u 395	Туг	Phe	Phe	Lys	G1u 400
10	GÌu	Asp	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys
15	Asn	туг	Cys	Leu 420	Ile	Cys	Тгр	Lys								
20	<21 <21	10> 62 1> 42 12> PI 13> Ar	24 RT													
25	<22	20> 23> Sii 00> 62		a artifi	cial											
30	Met 1	Ala	ser	Ser	Asn 5	Leu	Ile	Lys	G]n	Leu 10	Gln	Glu	Arg	G1y	Leu 15	va1
	Ala	Gln	Val	Thr 20	Asp	Glu	Glu	Ala	Leu 25	Ala	Glu	Arg	Leu	A1a 30	Gln	Gly
35	Pro	Ile	A1a 35	Leu	Ser	Cys	Gly	Phe 40	Asp	Pro	Thr	Ala	ASP 45	Ser	Leu	ніѕ
40	Leu	GTy 50	His	Leu	va1	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	Gln	Gln	Ala
45	G1y 65	His	Lys	Pro	val	A1a 70	Leu	val	Gly	Gly	Ala 75	Thr	Gly	Leu	Ile	G]y 80
.0	ASP	Pro	Ser	Phe	Lys 85	Ala	Аlа	Glu	arg	Lys 90	Leu	Asn	Thr	Glu	G1u 95	Thr
50	va1	GÌn	Glu	Trp 100	۷a۱	Asp	Lys	Ile	Arg 105	Lys	Gln	۷a٦	Ala	Pro 110	Phe	Leu
55	Asp	Phe	ASP 115	Cys	Gly	Glu	Asn	ser 120	Ala	Ile	Ala	Ala	Asn 125	Asn	туг	Asp
	Тгр	Phe 130		Asn	Met	Asn	va1 135	Leu	Thr	Phe	Leu	Arg 140	ASP	Ile	G1y	Lys
60	His 145	Phe	Ser	val	Asn	G]n 150	Met	Ile	Asn	Lys	Glu 155	Ala	va1	Lys	G1n	Arg 160

	Leu	Asn	Arg	Glu	ASP 165	Gln	Gly	Ile	Ser	Phe 170	Thr	Glu	Phe	Ser	175	Asn
5	Leu	Leu	Gln	Gly 180	Туг	ser	Phe	Ala	Cys 185	Leu	Asn	Lys	G1n	Туг 190	Gly	val
10	Va1	Leu	G1n 195	Ile	Gly	Gly	Ser	ASP 200	Gln	Trp	Gly	Asn	11e 205	Thr	Ser	Gly
15	Ile	Asp 210	Leu	Thr	Arg	Arg	Leu 215	His	Gln	Asn	Gln	va1 220	Phe	Gly	Leu	Thr
13	va1 225	Pro	Leu	Ile	Thr	Lys 230	Ala	ASP	Gly	Thr	Lys 235	Phe	Gly	Lys	Thr ,	G1u 240
20	Gly	Gly	Ala	val	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	Ser	Pro	Туг	Lys 255	Phe
25	Tyr	Gln	Phe	Trp 260	Ile	Asn	Thr	Ala	Asp 265	Αla	Asp	val	туг	Arg 270	Phe	Leu
	Lys	Phe	Phe 275	Thr	Phe	Met	Ser	11e 280	Glu	Glu	Ile	Asn	Ala 285	Leu	Glu	Glu
30	Glu	Asp 290	Lys	Asn	Ser	Gly	Lys 295	Ala	Pro	Arg	Ala	G]n 300	Туг	val	Leu	Ala
35	G1u 305	Gln	Val	Thr	Arg	Leu 310	va1	His	Gly	Glu	G1u 315	Gly	Leu	Gln	Ala	A1a 320
40	Lys	Arg	Ile	Thr	G1u 325	Cys	Leu	Phe	ser	G1y 330	ser	Leu	ser	Ala	Leu 335	ser
40	Glu	Ala	Asp	Phe 340	Glu	Gln	Leu	Ala	G1n 345	Asp	Gly	Val	Pro	Met 350	val	Glu
45	Met	Glu	Lys 355	Gly	Ala	Asp	Leu	Met 360	Gln	Ala	Leu	Va1	ASP 365	Ser	Glu	Leu
50	Gln	Pro 370		Arg	Gly	GÌn	A1a 375	Arg	Lys	Thr	Ile	A7a 380	Ser	Asn	Ala	Ile
	Thr 385	Ile	Asn	Gly	Glu	Lys 390	Gln	Ser	Asp	Pro	G1u 395	Tyr	Phe	Phe	Lys	G1u 400
55	Glu	Asp	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys
60	<2° <2°	Tyr 10> 63 11> 42 12> PI	3 24	420		Cys	Тгр	Lys								
65	<2	13> Aı	tificial													

-0	റ	
<∠.	Zυ	>

<223> Sintetasa artificial

<400> 63

5

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val 10 Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly 25 30 Pro Ile Ala Leu Thr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His 15 Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala 20 Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly 65 70 75 80 Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr 25 Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu 100 105 110 30 Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp 35 Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys 130 His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg 145 150 155 160 40 Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn 165 170 175 45 Leu Leu Gln Gly Tyr Thr Phe Ala Cys Thr Asn Lys Gln Tyr Gly Val Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly 195 200 205 50 Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr 210 220

55

60

	va1 225	Pro	Leu	Ile	Thr	Lys 230	Ala	ASP	Gly	Thr	Lys 235	Phe	Gly	Lys	Thr	G1u 240	
5	Gly	Gly	Ala	val	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	ser	Pro	туг	Lys 255	Phe	
10	Tyr	Gln	Phe	Trp 260	Ile	Asn	Thr	Ala	Asp 265	Ala	Asp	val	Туг	Arg 270	Phe	Leu	
15	Lys	Phe	Phe 275	Thr	Phe	Met	ser	11e 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	Glu	
	Glu	Asp 290	Lys	Asn	Ser	Gly	Lys 295	Ala	Pro	Arg	ΑΊa	G1n 300	туг	Val	Leu	Ala	
20	G1u 305	Gln	val	Thr	Arg	Leu 310	Va1	His	Gly	Glu	Glu 315	Gly	Leu	GÌn	Ala	A1a 320	
25	Lys	Arg	Ile	Thr	G1u 325	Cys	Leu	Phe	ser	G]y 330	Ser	Leu	Ser	Ala	Leu 335	ser	
	Glu	Ala	Asp	Phe 340	Glu	Gln	Leu	Ala	G]n 345	Asp	Gly	۷a۱	Pro	Met 350	Val	Glu	
30	Met	Glu	Lys 355	Gly	Ala	Asp	Leu	меt 360	Gln	Ala	Leu	val	ASP 365	Ser	Glu	Leu	
35	Gln	Pro 370	ser	Arg	Gly	Gln	A1a 375	Arg	Lys	Thr	Ile	Ala 380	ser	ASN	Ala	Ile	
40	Thr 385	Ile	Asn	Gly	Glu	Lys 390	GÌn	Ser	Asp	Pro	Glu 395	туг	Phe	Phe	Lys	G1u 400	
40	Glu	Asp	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys	
45	Asn	Туг	Cys	Leu 420	Ile	Cys	тгр	Lys									
50	<21 <21	10> 64 11> 12 12> Al 13> Es	29 DN	chia co	oli												
55	<40	00> 64	1														
	agcttccc ggccaaag																120
60	ccaccacc	a															129
	<210> 6	5															
65		11> 12 12> ΔΙ	-														

<213> Escherichia coli <400> 65 5 agcuucccga uaagggagca ggccaguaaa aagcauuacc ccgugguggg guucccgagc 60 ggccaaaggg agcagacucu aaaucugccg ucaucgaccu cgaagguucg aauccuuccc 120 129 ccaccacca 10 <210> 66 <211> 34 <212> ADN <213> Artificial 15 <220> <223> cebador oligonucleótido <400> 66 20 atgaagtagc tgtcttctat cgaacaagca tgcg 34 <210> 67 <211> 34 <212> ADN 25 <213> Artificial <223> cebador oligonucleótido 30 cgaacaagca tgcgattagt gccgacttaa aaag 34 <210> 68 <211> 33 35 <212> DNA <213> Artificial <220> <223> cebador oligonucleótido 40 <400> 68 cgctactctc ccaaatagaa aaggtctccg ctg 33 <210> 69 45 <211> 32 <212> ADN <213> Artificial 50 <223> cebador oligonucleótido <400> 69 ctggaacagc tatagctact gatttttcct cg 32 55 <210> 70 <211> 34 <212> ADN <213> Artificial <220> 60 <223> cebador oligonucleótido

34

gccgtcacag attagttggc ttcagtggag actg

<210> 71 <211> 33

	<212> ADN <213> Artificial
5	<220> <223> cebador oligonucleótido
	<400> 71 gattggcttc ataggagact gatatgctct aac 33
10	<210> 72 <211> 33 <212> ADN <213> Artificial
15	<220> <223> cebador oligonucleótido
20	<400> 72 gcctctatag ttgagacagc atagaataat gcg 33
	<210> 73 <211> 35 <212> ADN <213> Artificial
25	<220> <223> cebador oligonucleótido
30	<400> 73 gagacagcat agatagagtg cgacatcatc atcgg 35
35	<210> 74 <211> 37 <212> ADN <213> Artificial
	<220> <223> cebador oligonucleótido
40	<400> 74 gaataagtgc gacatagtca tcggaagaga gtagtag 37 <210> 75 <211> 35
45	<212> ADN <213> Artificial <220> <223> cebador oligonucleótido ggtcaaagac agttgtaggt atcgattgac tcggc 35
50	<210> 76 <211> 34 <212> ADN <213> Artificial
55	<220> <223> cebador oligonucleótido
60	<400> 76 cgctactctc cccaaattta aaaggtctcc gctg 34
0.5	<210> 77 <211> 34 <212> ADN <213> Artificial
65	<220>

	<223> cebador oligonucleótido	
5	<400> 77 cgctactctc cccaaatata aaaggtctcc gctg	34
Š	<210> 78 <211> 34 <212> ADN <213> Artificial	
10	<220> <223> cebador oligonucleótido	
15	<400> 78 cgctactctc cccaaatgga aaaggtctcc gctg	34
20	<210> 79 <211> 34 <212> ADN <213> Artificial	
	<220> <223> cebador oligonucleótido	
25	<400> 79 cgctactctc cccaaagata aaaggtctcc gctg	34
30	<210> 80 <211> 34 <212> ADN <213> Artificial <220> <223> cebador oligonucleótido	
35	<400> 80 cgctactctc cccaaaaaaa aaaggtctcc gctg <211> 34 <212> ADN	34
40	<213> Artificial <220> <223> cebador oligonucleótido	
45	<400> 81 gccgtcacag attttttggc ttcagtggag actg	34
50	<210> 82 <211> 34 <212> ADN <213> Artificial	
	<220> <223> cebador oligonucleótido	
55	<400> 82 gccgtcacag attatttggc ttcagtggag actg	34
60	<210> 83 <211> 34 <212> ADN <213> Artificial	
65	<220> <223> cebador oligonucleótido <400> 83	

	geogreaday anggngge neagiggag aong 54
5	<210> 84 <211> 34 <212> ADN <213> Artificial
10	<220> <223> cebador oligonucleótido <400> 84
	gccgtcacag atgatttggc ttcagtggag actg 34
15	<210> 85 <211> 34 <212> ADN <213> Artificial <220>
20	<223> cebador oligonucleótido
	<400> 85 gccgtcacag ataaattgge ttcagtggag actg 34 <210> 86 <211> 424
25	<212> PRT <213> Artificial
30	<220> <223> Sintetasa artificial
	<400> 86
35	Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Va 1 10 15
	Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gl 20 25 30
40	
45	
50	
55	
60	
65	

	Pro	Ile	Ala 35	Leu	Ile	Cys	Gly	Phe 40	ASP	Pro	Thr	Ala	ASP 45	ser	Leu	His
5					000000000000000000000000000000000000000			8.7								
	Leu	Gly 50	His	Leu	Val	Pro	Leu 55	Leu	Cys	Leu	Lys	Arg 60	Phe	Gln	Gln	Ala
10	61y	His	Lys	Pro	val	Ala 70	Leu	val	Gly	Gly	A1a 75	Thr	G1y	Leu	Ile	61y 80
15	Asp	Pro	Ser	Phe	Lys 85	Ala	Ala	Glu	Arg	Lys 90	Leu	ASN	Thr	Glu	Glu 95	Thr
	val	G1n	G1u	Trp 100	Val	Asp	Lys	Ile	Arg 105	Lys	G1n	Val	Ala	Pro 110	Phe	Leu
20	Asp	Phe	ASP 115	Cys	Gly	Glu	Asn	Ser 120	Αla	Ile	Ala	Ala	Asn 125	Asn	Туг	Asp
25	Trp	Phe 130	Gly	Asn	Met	Asn	va1 135	Leu	Thr	Phe	Leu	Arg 140	Asp	Ile	Gly	Lys
30	ніs 145	Phe	Ser	val	Asn	G]n 150	Met	Ile	Asn	Lys	Glu 155	Ala	val	Lys	Gln	Arg 160
	Leu	Asn	Arg	Glu	Gly 165	Gln	G1y	Ile	Ser	Phe 170	Thr	Glu	Phe	Ser	Tyr	Asn
35	Leu	Leu	Gln	Gly 180	Туг	Gly	Met	Ala	Cys 185	Ala	Asn	Lys	Gln	Tyr 190	Gly	val
40	∨a1	Leu	Gln 195	Ile	Gly	Gly	ser	ASP 200	Gln	Trp	Gly	Asn	11e 205	Thr	ser	Gly
	Ile	Asp 210	Leu	Thr	Arg	Arg	Leu 215	His	G1n	Asn	Gln	va1 220	Phe	Gly	Leu	Thr
45	Va1 225	Pro	Leu	Ile	Thr	Lys 230	Ala	Asp	Gly	Thr	Lys 235	Phe	Gly	Lys	Thr	G1u 240
50	Gly	Gly	Аlа	val	Trp 245	Leu	Asp	Pro	Lys	Lys 250	Thr	Ser	Pro	Tyr	Lys 255	Phe
	Туг	Gln	Phe	Trp 260	Ile	Aşn	Thr	Ala	Asp 265	Ala	Asp	٧a٦	Tyr	Arg 270	Phe	Leu
55	Lys	Phe	Phe 275	Thr	Phe	Met	ser	11e 280	Glu	Glu	Ile	Asn	A1a 285	Leu	Glu	Glu
60	Glu	Asp	Lys	Asn	ser	GТу	Lys	Ala	Pro	Arg	Ala	Gln	Туг	Va1	Leu	Ala

		290					295		ŭ.			300					
5	G1u 305	Gln	val	Thr	Arg	Leu 310	val	His	Gly	Glu	G]u 315	Gly	Leu	Gln	Ala	A1a 320	
10	Lys	Arg	Ile	Thr	G]u 325	Cys	Leu	Phe	Ser	Gly 330	Ser	Leu	Ser	Ala	Leu 335	Ser	
	Ģlu	Ala	Asp	Phe 340	Glu	Gln	Leu	Αla	G] n 345	Asp	Gly	val	Pro	Met 350	val	Glu	
15	Met	Glu	Lys 355	Gly	Аlа	Asp	Leu	Met 360	Gln	Ala	Leu	Va1	Asp 365	Ser	Glu	Leu	
20	Gln	Pro 370	Ser	Arg	Gly	Gìn	A1a 375	Arg	Lys	Thr	Ile	A1a 380	Ser	Asn	Ala	Ile	
25	Thr 385	Ile	Asn	Gly	Glu	Lys 390	Gln	Ser	ASP	Pro	G1u 395	Tyr	Phe	Phe	Lys	G]u 400	
20	G1 u	Asp	Arg	Leu	Phe 405	Gly	Arg	Phe	Thr	Leu 410	Leu	Arg	Arg	Gly	Lys 415	Lys	
30	Asn	Туг	Cys	Leu 420	Ile	Cys	Тгр	Lys									
35	<21 <21	0> 87 1> 85 2> AE 3> Ar	DN														
40		3> AF	RNt art	ificial													
	<40	00> 87	•														
45	gguggggu						ggcg	gac	ucua	aau	ccgc	uccc	uu u	gggu	ucgg	C	60
	gguucgaa	uc c	gucc	cccu	c ca	cca											85
50	<21 <21	0> 88 1> 79 2> AE 3> Ar) DN														
55		23> A[ONc pa	ara AR	Nt												
	<40	88 <00	3														
60	ggtggggt					acgo	ggcg	gac	tcta	aat	ccgc	tccc	tt t	gggt	tcgg	c	60
	ggttcgaa	tc c	gtcc	ccca													79
65		0> 89 1> 95															

	<212> ADN <213> Artificial	
5	<220> <223> Inserción de plásmido artificial	
	<400> 89	
10	ggattacgca tgctcagtgc aatcttcggt tgcctggact agcgctccgg tttttctgtg	60
	ctgaacctca ggggacgccg acacacgtac acgtc	95
15	<210> 90 <211> 42 <212> ADN <213> Homo sapiens	
20	<400> 90 gacaagtgcg gtttttttct ccagctcccg atgacttatg gc 42	
25	<210> 91 <211> 80 <212> ADN <213> Artificial	
30	<220> <223> FTam 73 cebador directo <400> 91	
25	gtacgaattc ccgagatctg gattacgcat gctcagtgca atcttcggtt gcctggacta	60
35	gcgctccggt ttttctgtgc	80
40	<210> 92 <211> 88 <212> ADN <213> Artificial	
45	<220> <223> FTam 115 cebador inverso	
	<400> 92	
50	agtccgccgc gtttagccac ttcgctaccc caccgacgtg tacgtgtgtc ggcgtcccct	60
	gaggttcagc acagaaaaac cggagcgc	88
55	<210> 93 <211> 76 <212> ADN <213> Artificial	
60	<220> <223> FTam 116 cebador directo <400> 93	
65		

	grggcraaac geggeggaet craaateege teeerriggg treggeggtt egaateegte	60
	ccccaccaga caagtg	76
5		190500
10	<210> 94 <211> 75 <212> ADN <213> Artificial	
	<220> <223> FTam 117 cebador inverso	
15	<400> 94	
	gatgcaagct tgatggatcc gccataagtc atcgggagct ggagaaaaaa accgcacttg	60
20	tctggtgggg gacgg	75
25	<210> 95 <211> 11 <212> ADN <213> Artificial	
30	<220> <223> secuencia caja B <220>	
35	<221> misc_feature <222> (8)(8) <223> n es a, c, g, o t <400> 95 ggttcgantc c 11	

REIVINDICACIONES

- 1. Una célula de vertebrado o una línea celular que comprende una secuencia de nucleótidos descrita en la SEC ID Nº 87 o en la SEC ID Nº 88, en la que la célula no es una célula madre embrionaria humana o una célula humana *in vivo* y la línea celular no es una línea celular madre embrionaria humana.
- 2. La célula de la reivindicación 1, en la que la secuencia de nucleótidos codifica una molécula de ARNt que tiene una secuencia de reconocimiento de anticodones específica para un codón selectivo.
- 3. La célula de la reivindicación 2, en la que el codón selectivo se selecciona del grupo formado por: codón ámbar, codón ocre, codón ópalo, y codones de cuatro o más bases.

5

15

20

35

40

45

- **4.** La célula de la reivindicación 2, en la que la secuencia de nucleótidos codifica una molécula de ARNt capaz de ser aminoacilada con al menos un aminoácido no natural.
- 5. La célula de la reivindicación 1, en la que la secuencia de nucleótidos codifica una molécula de ARNt que es un ARNt ortogonal (O tRNA).
- La célula de la reivindicación 5, en la que el O tRNA es capaz de aminoacilarse con un aminoácido natural o un aminoácido no natural.
 - 7. La célula de la reivindicación 4, en la que el O tRNA es capaz de aminoacilarse con un aminoácido natural o un aminoácido no natural.
- 25 **8.** La célula de la reivindicación 5, en la que el O tRNA es capaz de aminoacilarse con un aminoácido no natural.
 - 9. La línea celular de la reivindicación 1, en la que la línea celular ha sido transfectada de manera estable.
- 30 10. La línea celular de la reivindicación 1, en la que la línea celular ha sido transfectada de manera transitoria.
 - **11.** Un procedimiento para producir una célula de vertebrado, en la que la célula no es una célula madre embrionaria humana o una célula humana *in vivo*, en la que al menos una proteína comprende al menos un aminoácido no natural, comprendiendo el procedimiento:

cultivar en un medio apropiado, dicha célula de vertebrado que comprende un ácido nucleico que comprende al menos un codón selectivo y codifica a la proteína; en el que el medio comprende un aminoácido no natural y la célula de vertebrado comprende: un O – tRNA que tiene una secuencia de nucleótidos descrita en la SEC ID Nº 87 o en la SEC ID Nº 88 que funciona en la célula y reconoce al codón selectivo; y

una aminoacil - ARNt sintetasa ortogonal (O – RS) que aminoacila preferentemente el O – tRNA con el aminoácido no natural.

- **12.** El procedimiento de la reivindicación 11, en el que la célula ha sido transfectada de forma estable para comprender el O tRNA y la O RS.
- **13.** El procedimiento de la reivindicación 11, en el que la célula ha sido transfectada de forma transitoria para comprender el O tRNA y la O RS.
- 50
 14. El procedimiento de la reivindicación 11, en el que la célula ha sido transfectada de forma estable para comprender el O tRNA y la O RS y ha sido transfectada de forma transitoria para comprender al O tRNA y a la O RS.
- 15. El procedimiento de la reivindicación 11 o la célula de la reivindicación 8, en el que se selecciona un 55 aminoácido no natural del grupo formado por: p - acetil -L - fenilalanina, p - yodo - L - fenilalanina, O - metil - L-tirosina, p - propargiloxifenilalanina, L - 3 - (2 - naftil) alanina, 3 - metil - fenilalanina, O - 4 -alil - L tirosina, 4 - propil – L - tirosina, tri – O - acetil - GlcNAcβ - serina, L - Dopa, fenilalanina fluorada, isopropil - L - fenilalanina, p - azido - L - fenilalanina, p -acil - L - fenilalanina, ap - benzoílo - L - fenilalanina, L fosfoserina, fosfonoserina, fosfonotirosina, p - bromofenilalanina, p - amino - L - fenilalanina, un análogo no 60 natural de un aminoácido de tirosina; un análogo no natural de un aminoácido de glutamina; un análogo no natural de un aminoácido de fenilalanina; un análogo no natural de un aminoácido de serina; un análogo no natural de un aminoácido de treonina; un alquilo, arilo, acilo, azida, ciano, halo , hidrazina, hidrazida, hidroxilo, alquenilo, alquinilo, éter, tiol, sulfonilo, selenio, éster, tioácido, borato, boronato, fosfo, fosfono, fosfina, heterociclo, enona, imina, aldehído, hidroxilamina, ceto, o aminoácidos de amino sustituido, o 65 cualquier combinación de los mismos; un aminoácido con un reticulador fotoactivable; un aminoácido marcado en spin; un aminoácido fluorescente; un aminoácido de unión a metales; un aminoácido que

contiene un metal; un aminoácido radiactivo; un aminoácido fotobloqueado y / o fotoisomerizado; una biotina o biotina - analóga que contiene aminoácidos, un aminoácido que contiene ceto; un aminoácido que comprende polietilenglicol o poliéter; un aminoácido sustituido de átomo pesado; un aminoácido químicamente escindible o fotoescindible; un aminoácido con una cadena lateral alargada; un aminoácido que contiene un grupo tóxico; un aminoácido de azúcar sustituido; un aminoácido que contiene azúcar enlazada al carbono; un amino ácido activo de redox; un ácido que contiene α - hidroxi; un ácido tioaminado; un α , ácido α amino disustituido; un ácido β - amino, un aminoácido cíclico que no sea prolina o histidina, y un aminoácido aromático distinto a la fenilalanina, tirosina o triptófano.

16. El procedimiento de la reivindicación 11, en el que la proteína comprende una proteína terapéutica, una proteína de diagnóstico, una enzima industrial o una parte de las mismas.

5

15

20

25

30

35

40

45

50

- 17. El procedimiento de la reivindicación 11, en el que la proteína o polipéptido de interés comprende una proteína o una parte de una proteína seleccionada del grupo formado por: una citocina, un factor de crecimiento, un receptor del factor de crecimiento, un interferón, una interleucina, una molécula inflamatoria, un producto oncogénico, una hormona peptídica, una molécula de transducción de señales, un receptor de la hormona esteroide, una eritropoyetina (EPO), una insulina, una hormona del crecimiento humano, una alfa - 1 antitripsina, una angiostatina, un factor antihemolítico, un anticuerpo, una apolipoproteína, una apoproteína, un factor natriurético atrial, un polipéptido natriurético atrial, un péptido atrial, una quimiocina C - X - C. T39765 , NAP - 2 , ENA -78, Gro - a, Gro - b, Gro - c, IP- 10, GCP- 2, PNA - 4, SDF - 1, PF4, MIG, una calcitonina, un ligando de c - kit, una quimiocina CC, una proteína quimiotáctica de monocitos - 1, una proteína quimiotáctica de monocitos - 2, una proteína quimioatrayente de monocitos - 3, proteína- 1 alfa inflamatoria de monocitos, proteína - 1 beta inflamatoria de monocitos, RANTES, I309, R83915, R91733, HCC1, T58847, D31065, T64262, CD40, un ligando CD40, un colágeno, un factor estimulante de colonias (CSF), un factor del complemento 5a, un inhibidor del complemento, receptor 1 del complemento, DHFR, un péptido epitelial activador de los neutrófilos 78, GROα / MGSA, GROβ, GROγ, un MIP - 1α, MIP – 1δ, MCP - I, un factor de crecimiento epidérmico (EGF), un péptido epitelial activador de neutrófilos, una toxina exfoliante, un factor IX, un factor VII, un factor VIII, un factor X, un factor de crecimiento de fibroblastos (FGF), un fibrinógeno, una fibronectina, G-CSF, GM-CSF, una glucocerebrosidasa, una gonadotropina, una proteína Hedgehog, una hemoglobina, un factor de crecimiento de hepatocitos (HGF), una hirudina, una albúmina de suero humano, ICAM – 1, un receptor ICAM – 1, un LFA – 1, un receptor LFA – 1, una insulina, un factor de crecimiento de tipo insulina (IGF), IGF – II, un IFN – α, un IFN – β, un IFN – γ, una IL - 1, una IL - 2. una IL - 3. una IL - 4. una IL - 5. una IL - 6. una IL - 7. una IL - 8. una IL - 9. una IL - 10. una IL - 11. una IL - 12, un factor de crecimiento de queratinocitos (KGF), una lactoferrina, un factor inhibidor de leucemia, una luciferasa, una neurturina, un factor inhibidor de neutrófilos (NIF), una oncostatina M, una proteína osteogénica, una hormona paratiroidea, PD - ECSF, PDGF, una pleyotropina, una proteína A, una proteína G, exotoxinas pirogénicas A, B, y C, una relaxina, una renina, SCF, un receptor I soluble del complemento, un soluble I - CAM 1, receptores solubles de interleucina, un receptor soluble de TNF, una somatomedina, una somatostatina, una somatotropina, una estreptoquinasa, superantígenos, enterotoxinas estafilocócicas SEA, SEB, SEC1, SEC2, SEC3, SED, SEE, un receptor de hormona esteroide, un superóxido dismutasa (SOD), una toxina del síndrome de shock tóxico, una timosina alfa 1, un activador plasminógeno tisular, un factor de crecimiento tumoral (TGF), un TGF $-\alpha$, un TGF $-\beta$, un factor de necrosis tumoral, un factor de necrosis tumoral - alfa, un factor de necrosis tumoral beta, un receptor del factor de necrosis tumoral (TNFR), una proteína VLA - 4, una proteína VCAM - 1, un factor de crecimiento endotelial vascular (VEGEF), uroquinasa, Mos, Ras, Raf, Met, p53, Tat, Fos, Myc, Jun, Myb, Rel, un receptor de estrógeno, un receptor de progesterona, un receptor de testosterona, un receptor de aldosterona, un receptor LDL, un SCF / c - Kit, CD40L / CD40, VLA -4 / VCAM - 1, ICAM - 1 / LFA - 1, una hialurina / CD44, y una corticosterona.
- 18. Un kit para producir una proteína que comprende, al menos, un aminoácido no natural, comprendiendo el kit: un recipiente que contiene una secuencia de polinucleótidos descrita en la SEC ID Nº 87 o en la SEC ID Nº 88.
 - 19. El kit de la reivindicación 18, en el que el kit comprende además, al menos, un aminoácido no natural.
 - **20.** El kit de la reivindicación 18, en el que el kit comprende además materiales informativos para la producción de proteínas.

Figura 1

