

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 709 217**

51 Int. Cl.:

C12P 7/40 (2006.01)
C12P 7/22 (2006.01)
C12P 5/00 (2006.01)
C12N 1/26 (2006.01)
C12N 9/88 (2006.01)
C12N 9/00 (2006.01)
C12N 15/62 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **21.06.2013** **PCT/US2013/047098**
87 Fecha y número de publicación internacional: **27.12.2013** **WO13192543**
96 Fecha de presentación y número de la solicitud europea: **21.06.2013** **E 13806939 (8)**
97 Fecha y número de publicación de la concesión europea: **17.10.2018** **EP 2864491**

54 Título: **Enzimas y métodos para la síntesis del estireno**

30 Prioridad:

22.06.2012 US 201261663500 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
15.04.2019

73 Titular/es:

PHYTOGENE, INC. (100.0%)
30111 Tomas Street
Rancho Santa Margarita, CA 92688, US

72 Inventor/es:

BHUIYA, MOHAMMAD WADUD;
CHEN, HUI;
CAI, XIANPENG;
HAN, JIXIANG y
YU, XIAODAN

74 Agente/Representante:

SALVÀ FERRER, Joan

ES 2 709 217 T3

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

DESCRIPCIÓN

Enzimas y métodos para la síntesis del estireno

INCORPORACIÓN DEL LISTADO DE SECUENCIAS

[0001] En la presente memoria se proporciona y se incorpora a modo de referencia una copia en papel del listado de secuencias y una copia en soporte informático del listado de secuencias que contiene el archivo con el nombre «32990-5_ST25.txt», el cual tiene un tamaño de 151.674 bytes (medido en MS-DOS). Este listado de secuencias consiste en las SECs. N.º ID: 1-44.

ANTECEDENTES

[0002] La tecnología objeto de esta invención por lo general se refiere a enzimas y métodos para la biosíntesis de estireno.

[0003] El estireno (vinilbenceno) es un compuesto orgánico con la fórmula química C_8H_8 . Este hidrocarburo cíclico es un líquido incoloro y oleoso que se evapora fácilmente y que tiene un olor dulce similar al del caucho. A concentraciones más elevadas, el estireno emite un olor menos agradable. El estireno tiene el mismo nombre que los árboles del género *Styrax* (*Styrax platanifolius*), de los que se puede extraer savia (un tipo de resina de benjuí). Varias especies de planta contienen estireno de manera natural en concentraciones bajas. Diversos alimentos, como las frutas, las verduras, los frutos secos, las bebidas y las carnes también contienen estireno.

[0004] En un plano industrial, el estireno es el precursor del poliestireno y de varios copolímeros. La presencia del grupo vinilo permite que el estireno polimerice. Se producen aproximadamente 6.800 millones de kilogramos (15.000 millones de libras) anualmente. La producción de estireno en los Estados Unidos aumentó de manera espectacular durante los años 40, cuando se popularizó como materia prima para la fabricación del caucho sintético. En la actualidad, entre los productos importantes desde un punto de vista comercial se incluyen el poliestireno, el acrilonitrilo butadieno estireno (ABS), el caucho de estireno-butadieno (SBR), el látex de estireno-butadieno, el estireno-isopreno-estireno (SIS), el estireno-etileno/butileno-estireno (S-EB-S), el estireno-divinilbenceno (S-DVB), la resina de estireno-acrilonitrilo (SAN) y los poliésteres insaturados. Estos materiales se utilizan en el caucho, el plástico, el aislamiento, la fibra de vidrio, las tuberías, los automóviles y las piezas de buques, los recipientes para alimentos y los refuerzos de moqueta.

[0005] El estireno se produce en cantidades industriales sobre todo a partir del etilbenceno, que a su vez se manufactura a gran escala mediante alquilación del benceno con etileno. Constituye uno de los productos petroquímicos más importantes. Existen varios métodos para producir estireno. La deshidrogenación del etilbenceno es el medio más común de producción. El etilbenceno se mezcla en fase gas con 10-15 veces su volumen en vapor a elevada temperatura, y se hace pasar sobre un lecho de catalizador sólido. La mayoría de catalizadores para la deshidrogenación del etilbenceno son de óxido de hierro (III), potenciado con óxido de potasio o carbonato de potasio en diversos porcentajes. El vapor desempeña varios papeles en esta reacción. Es la fuente de calor para alimentar a la reacción endotérmica, y elimina el coque que tiende a formarse en el catalizador de óxido de hierro por medio de la reacción de desplazamiento del vapor de agua. El potenciador de potasio acelera esta reacción de descoquización. El vapor también diluye el reactivo y los productos, desplazando la posición del equilibrio químico hacia los productos. Una planta habitual de producción de estireno consiste en dos o tres reactores en serie, que operan al vacío para aumentar la conversión y la selectividad. Las conversiones por paso habituales son de aproximadamente un 65% con dos reactores y de un 70-75% con tres reactores. La selectividad a estireno es de un 93-97%. Los principales subproductos son el benceno y el tolueno. Debido a que el estireno y el etilbenceno tienen puntos de ebullición similares (145 y 136 °C, respectivamente), su separación requiere torres de destilación altas y elevados cocientes de retorno/reflujo. A su temperatura de destilación, el estireno tiende a polimerizar. A fin de reducir este problema, las primeras plantas de producción de estireno añadían azufre elemental para inhibir la polimerización. Durante los años 70, se desarrollaron nuevos inhibidores de los radicales libres que consistían en retardantes que contenían fenoles nitrados. Más recientemente, se han desarrollado varios aditivos que muestran una mayor inhibición de la polimerización.

[0006] Dado que el estireno es un compuesto petroquímico indispensable utilizado en muchos productos químicos, urge encontrar otros métodos de producción, especialmente métodos que no requieran combustibles fósiles como materia prima de alimentación. Por lo tanto, a pesar de la disponibilidad de métodos para producir estireno, existe una necesidad continua de encontrar nuevos métodos eficientes y menos costosos para producir monómeros de estireno.

La patente china n.º 101.397.568 (A) describe un método de expresión, purificación y cristalización, elucidación estructural y el campo de tecnología de aplicación de la ferulato descarboxilasa, que pertenece al campo de la biología molecular y la aplicación de la microbiología. En ella se describe un medio para obtener la secuencia de la región

codificante de la enzima clave, la ferulato descarboxilasa, que descompone el ácido ferúlico para generar 4-vinilguayacol en *Enterobacter sp.* Px6-4 mediante el método de la clonación génica, expresar la enzima en una gran cantidad en el colibacilo (*E. coli* BL21) mediante tecnología de expresión heterogénea, obtener la ferulato descarboxilasa como productos puros mediante purificación, cristalizar la ferulato descarboxilasa por el método de cristalización de la gota colgante, y obtener la estructura y el centro activo de la enzima. Con la invención se puede obtener una gran cantidad de ferulato descarboxilasa muy activa, y aporta directrices teóricas para reconstruir bacterias obtenidas mediante técnicas de ingeniería genética, así como para mejorar la producción de 4-vinilguayacol mediante el análisis del centro activo de la enzima y, por lo tanto, tiene una buena aplicabilidad potencial.

McKenna, R. y Nielsen, D.R., 2011. Biosíntesis del estireno a partir de glucosa con *E. coli* modificada mediante técnicas de ingeniería genética. En *Metabolic engineering*, 13(5), págs. 544-554 se expone que mediante el diseño *de novo* y el desarrollo de una nueva ruta metabólica, el estireno se puede sintetizar de otra manera a partir de sustratos renovables como la glucosa. La conversión de la 1-fenilalanina sintetizada endógenamente en estireno se logró mediante la coexpresión de la fenilalanina amonio liasa y la *trans*-cinamato descarboxilasa. Las isoenzimas candidatas para cada paso se cribaron a partir de fuentes genéticas procedentes de bacterias, levaduras y plantas. La sobreexpresión de PAL2 de *Arabidopsis thaliana* y FDC1 de *Saccharomyces cerevisiae* (originariamente clasificada como ferulato descarboxilasa) en un huésped de *Escherichia coli* que sobreproducía 1-fenilalanina dio lugar a la acumulación en concentraciones de hasta 260 mg/l en cultivos en matraz agitado. Los valores que se pueden lograr ya se acercan al umbral de toxicidad del estireno (estimado en ~300 mg/l).

Rangarajan, E.S., Li, Y., Iannuzzi, P., Tocilj, A., Hung, L.W., Matte, A. y Cygler, M., 2004. *Crystal structure of a dodecameric FMN-dependent UbiX-like decarboxylase (Pad1) from Escherichia coli O157:H7*. En *Protein science*, 13(11), págs. 3006-3016 se describe la estructura cristalina de la flavoproteína Pad1 de *Escherichia coli* O157:H7 complejada con el cofactor FMN, determinada por el método de difracción anómala múltiple y refinada a una resolución de 2,0 Å. Esta proteína se describe como un paralogo de la UbiX (3-octaprenil-4-hidroxibenzoato carboxilasa, identidad de secuencia del 51%), que cataliza el tercer paso en la biosíntesis de la ubiquinona, y de la Pad1 de *Saccharomyces cerevisiae* (identidad del 54%), una enzima que confiere resistencia frente a los ácidos fenilacrilícos (compuestos antimicrobianos) por medio de la descarboxilación de estos compuestos. Cada monómero de Pad1 consiste en un pliegue clásico de Rossmann que contiene una molécula de FMN unida no covalentemente. El pliegue de la Pad1 es similar al de la MrsD, una enzima asociada a la síntesis lantibiótica; la EpiD, una peptidil-cisteína descarboxilasa; y la enzima AtHAL3a, que descarboxila la 4'-fosfopantotenoilcisteína a 4'-fosfopanteteína durante la biosíntesis de la coenzima A, todas con una ubicación similar del sitio de unión del FMN en la interfaz entre dos monómeros, a pesar de la escasa similitud entre sus secuencias.

SUMARIO

[0007] La tecnología objeto de esta invención por lo general se refiere a la biosíntesis del estireno. Algunas realizaciones de la tecnología objeto de esta invención están basadas, en parte, en el entendimiento de que la fenilalanina se puede transformar en estireno por medio de una ruta de dos pasos de desaminación y descarboxilación, utilizando el ácido *trans*-cinámico (tCA) como intermedio. Hay dos tipos de enzimas que participan directamente en este proceso, la fenilalanina amonio liasa (PAL), que transforma la fenilalanina en tCA, y la cinamato descarboxilasa, que transforma el tCA en estireno. Las células hospedadoras que expresan estos dos tipos de enzimas se pueden cultivar en un biorreactor para producir estireno a partir de sustratos renovables como la glucosa.

Según un primer aspecto de la invención se proporciona una célula hospedadora que comprende: (a) una proteína de fusión expresada recombinantemente que comprende un primer dominio que comprende una fenilalanina amonio liasa, y un segundo dominio que comprende una cinamato descarboxilasa; y (b) un transportador unido a la membrana expresado recombinantemente.

En determinadas realizaciones del primer aspecto, la cinamato descarboxilasa producida por dicha célula es una cinamato descarboxilasa mutada que comprende una mutación en una posición de residuo aminoacídico que corresponde a una posición seleccionada del grupo que consiste en: 155, 156, 159, 162, 163, 164, 172, 173, 174, 175, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 226, 227, 280, 285, 286, 287, 291, 326, 331, 360, 361, 395, 396, 398, 440, 441 de la SEC. N.º ID: 8, y combinaciones de las mismas.

En determinadas realizaciones del primer aspecto, la fenilalanina amonio liasa producida por dicha célula es una fenilalanina amonio liasa que comprende una secuencia aminoacídica seleccionada del grupo que consiste en: SEC. N.º ID: 2, SEC. N.º ID: 4, SEC. N.º ID: 6 y un fragmento funcional o variante de las mismas.

Según un segundo aspecto de la invención se proporciona un método para cristalizar una proteína de fusión que comprende un primer dominio que comprende una fenilalanina amonio liasa, y un segundo dominio que comprende una cinamato descarboxilasa expresada por la célula hospedadora de cualquier reivindicación precedente, donde el método comprende

(a) obtener la proteína de fusión expresada por la célula hospedadora de cualquier reivindicación anterior;

- (b) proporcionar una solución de la proteína de fusión a una concentración de aproximadamente 1 mg/ml a aproximadamente 50 mg/ml;
- (c) mezclar la solución de la proteína de fusión con una solución de depósito en una relación de volúmenes de aproximadamente 1:10 a aproximadamente 10:1;
- (d) añadir ácido 3-hidroxicinámico a la solución del paso (a) o del paso (b) a una concentración final de ácido 3-hidroxicinámico del 0,01% en p/v al 1% en p/v;
- (e) y mantener la mezcla del paso (c) a una temperatura de 1 °C a 20 °C adecuada para la formación del cristal de la proteína de fusión, de manera que la proteína de fusión cristalizada se encuentra formando un complejo con el ácido 3-hidroxicinámico.

[0008] La tecnología objeto de esta invención se refiere a una proteína de fusión que comprende: (a) un primer dominio que comprende una fenilalanina amonio liasa y (b) un segundo dominio que comprende una cinamato descarboxilasa.

[0009] En determinados ejemplos, la fenilalanina amonio liasa procede de un organismo seleccionado del grupo que consiste en: una *Arabidopsis*, una *Anabaena*, una *Nostoc* y una *Saccharomyces*. En un ejemplo de realización, la fenilalanina amonio liasa comprende una secuencia aminoacídica seleccionada del grupo que consiste en: SEC. N.º ID: 2, SEC. N.º ID: 4, SEC. N.º ID: 6, y un fragmento funcional o variante de las mismas.

[0010] En determinados ejemplos, la cinamato descarboxilasa se obtiene a partir de un organismo seleccionado del grupo que consiste en: una *Arabidopsis*, una *Anabaena*, una *Nostoc* y una *Saccharomyces*. En un ejemplo de realización, la cinamato descarboxilasa comprende una secuencia aminoacídica seleccionada del grupo que consiste en: SEC. N.º ID: 8; SEC. N.º ID: 10; SEC. N.º ID: 16; SEC. N.º ID: 18; SEC. N.º ID: 20; SEC. N.º ID: 22; SEC. N.º ID: 24; SEC. N.º ID: 26; SEC. N.º ID: 28; SEC. N.º ID: 30; SEC. N.º ID: 32; SEC. N.º ID: 34; SEC. N.º ID: 36; SEC. N.º ID: 38; y un fragmento funcional o variante de las mismas. Por ejemplo, la cinamato descarboxilasa puede comprender una secuencia aminoacídica seleccionada del grupo que consiste en: SEC. N.º ID: 8; SEC. N.º ID: 10; SEC. N.º ID: 16; SEC. N.º ID: 18; SEC. N.º ID: 20; SEC. N.º ID: 22; SEC. N.º ID: 24; SEC. N.º ID: 26; SEC. N.º ID: 28; SEC. N.º ID: 30; SEC. N.º ID: 32; SEC. N.º ID: 34; SEC. N.º ID: 36; y SEC. N.º ID: 38.

[0011] Una variante funcional de una cinamato descarboxilasa puede comprender una secuencia aminoacídica que es idéntica en aproximadamente un 85%, en aproximadamente un 90%, en aproximadamente un 95%, en aproximadamente un 96%, en aproximadamente un 97%, en aproximadamente un 98%, en aproximadamente un 99% o incluso en un 100% a cualquiera de las SEC. N.º ID: 8; SEC. N.º ID: 10; SEC. N.º ID: 16; SEC. N.º ID: 18; SEC. N.º ID: 20; SEC. N.º ID: 22; SEC. N.º ID: 24; SEC. N.º ID: 26; SEC. N.º ID: 28; SEC. N.º ID: 30; SEC. N.º ID: 32; SEC. N.º ID: 34; SEC. N.º ID: 36; y SEC. N.º ID: 38. Como otra posibilidad, o de manera adicional, una variante funcional de una cinamato descarboxilasa puede comprender: I en una posición que corresponde al residuo 173 de la SEC. N.º ID: 8, A en una posición que corresponde al residuo 174 de la SEC. N.º ID: 8, R en una posición que corresponde al residuo 175 de la SEC. N.º ID: 8, V en una posición que corresponde al residuo 188 de la SEC. N.º ID: 8, I en una posición que corresponde al residuo 189 de la SEC. N.º ID: 8, K en una posición que corresponde al residuo 90 de la SEC. N.º ID: 8, L en una posición que corresponde al residuo 194 de la SEC. N.º ID: 8, E en una posición que corresponde al residuo 280 de la SEC. N.º ID: 8, M en una posición que corresponde al residuo 286 de la SEC. N.º ID: 8, F en una posición que corresponde al residuo 291 de la SEC. N.º ID: 8, y F en una posición que corresponde al residuo 440 de la SEC. N.º ID: 8.

[0012] En determinados ejemplos, la cinamato descarboxilasa puede comprender una cinamato descarboxilasa mutada que comprende una mutación en una posición de residuo aminoacídico que corresponde a una de las siguientes posiciones: 155, 156, 159, 162, 163, 164, 172, 173, 174, 175, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 226, 227, 280, 285, 286, 287, 291, 326, 331, 360, 361, 395, 396, 398, 440, 441 de la SEC. N.º ID: 8, o una combinación de las mismas. Por ejemplo, la cinamato descarboxilasa mutada puede comprender una mutación en una posición de residuo aminoacídico que corresponde a una de las siguientes posiciones: 175, 190, 193 de la SEC. N.º ID: 8, y una combinación de las mismas.

[0013] En determinadas realizaciones, la cinamato descarboxilasa puede comprender una cinamato descarboxilasa mutada que comprende una eliminación, una sustitución o una adición de un residuo aminoacídico en una de las posiciones de la SEC. N.º ID: 8 seleccionadas de: 155, 156, 159, 162, 163, 164, 172, 173, 174, 175, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 226, 227, 280, 285, 286, 287, 291, 326, 331, 360, 361, 395, 396, 398, 440, 441 y una combinación de las mismas. Por ejemplo, un residuo aminoacídico en una de las siguientes posiciones: 155, 156, 159, 162, 163, 164, 172, 173, 174, 175, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 226, 227, 280, 285, 286, 287, 291, 326, 331, 360, 361, 395, 396, 398, 440 o 441 de la SEC. N.º ID: 8 se puede sustituir con otro residuo aminoacídico. En otro ejemplo, la cinamato descarboxilasa mutada comprende una eliminación, una sustitución o una adición de un residuo aminoacídico en una de las posiciones de la SEC. N.º ID: 8: 175, 190, 193 y una combinación de las mismas.

[0014] En un ejemplo, la proteína de fusión de la tecnología objeto de esta invención también comprende un conector que enlaza covalentemente el primer dominio con el segundo dominio. El conector de la proteína de fusión descrita

en la presente memoria puede ser un conector peptídico, por ejemplo, un conector peptídico que comprende de 2 a 15 aminoácidos. Los conectores peptídicos pueden incluir, por ejemplo, los que se muestran en el Cuadro 1.

[0015] También se proporcionan ácidos nucleicos que codifican las proteínas de fusión descritas en la presente memoria, un vector que comprende el ácido nucleico que codifica la proteína de fusión, y células hospedadoras que comprenden el vector descrito en la presente memoria.

[0016] La tecnología objeto de esta invención también se refiere a un método para producir estireno que comprende (a) poner en contacto la célula hospedadora con un sustrato fermentable (preferentemente un sustrato de carbono o un sustrato de nitrógeno), donde la célula hospedadora comprende una proteína de fusión que comprende: (i) un primer dominio que comprende una fenilalanina amonio liasa, y (ii) un segundo dominio que comprende una cinamato descarboxilasa; y (b) cultivar la célula en un medio de cultivo durante el tiempo suficiente para producir estireno. El método además puede comprender la recolección de estireno a partir del cultivo celular. Cualesquiera de las proteínas de fusión descritas en la presente memoria se pueden utilizar para producir estireno.

[0017] La tecnología objeto de esta invención también se refiere a una cinamato descarboxilasa que comprende (a) cualquiera de las SEC. N.º ID: 16; SEC. N.º ID: 18; SEC. N.º ID: 20; SEC. N.º ID: 22; SEC. N.º ID: 24; SEC. N.º ID: 26; SEC. N.º ID: 28; SEC. N.º ID: 30; SEC. N.º ID: 32; SEC. N.º ID: 34; SEC. N.º ID: 36; y SEC. N.º ID: 38; (b) una secuencia aminoacídica que es idéntica en aproximadamente un 85%, en aproximadamente un 90%, en aproximadamente un 95%, en aproximadamente un 97%, en aproximadamente un 98%, en aproximadamente un 99% o incluso en un 100% a cualquiera de las SEC. N.º ID: 16; SEC. N.º ID: 18; SEC. N.º ID: 20; SEC. N.º ID: 22; SEC. N.º ID: 24; SEC. N.º ID: 26; SEC. N.º ID: 28; SEC. N.º ID: 30; SEC. N.º ID: 32; SEC. N.º ID: 34; SEC. N.º ID: 36; y SEC. N.º ID: 38, con el requisito de que dicha secuencia aminoacídica no sea la SEC. N.º ID: 8; ni (c) un fragmento funcional de (a) ni de (b).

[0018] En determinados ejemplos, la cinamato descarboxilasa comprende: I en una posición que corresponde al residuo 173 de la SEC. N.º ID: 8, A en una posición que corresponde al residuo 174 de la SEC. N.º ID: 8, R en una posición que corresponde al residuo 175 de la SEC. N.º ID: 8, V en una posición que corresponde al residuo 188 de la SEC. N.º ID: 8, I en una posición que corresponde al residuo 189 de la SEC. N.º ID: 8, K en una posición que corresponde al residuo 190 de la SEC. N.º ID: 8, I en una posición que corresponde al residuo 194 de la SEC. N.º ID: 8, E en una posición que corresponde al residuo 280 de la SEC. N.º ID: 8, M en una posición que corresponde al residuo 286 de la SEC. N.º ID: 8, F en una posición que corresponde al residuo 291 de la SEC. N.º ID: 8, y F en una posición que corresponde al residuo 440 de la SEC. N.º ID: 8.

[0019] En determinados ejemplos, la cinamato descarboxilasa comprende cualesquiera de las SEC. N.º ID: 16; SEC. N.º ID: 18; SEC. N.º ID: 20; SEC. N.º ID: 22; SEC. N.º ID: 24; SEC. N.º ID: 26; SEC. N.º ID: 28; SEC. N.º ID: 30; SEC. N.º ID: 32; SEC. N.º ID: 34; SEC. N.º ID: 36; y SEC. N.º ID: 38.

[0020] La tecnología objeto de esta invención también se refiere a una cinamato descarboxilasa mutada que comprende una mutación en una posición de residuo aminoacídico que corresponde a una de las siguientes posiciones: 155, 156, 159, 162, 163, 164, 172, 173, 174, 175, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 226, 227, 280, 285, 286, 287, 291, 326, 331, 360, 361, 395, 396, 398, 440, 441 de la SEC. N.º ID: 8 y una combinación de las mismas. Por ejemplo, la cinamato descarboxilasa mutada puede comprender una mutación en una posición de residuo aminoacídico que corresponde a una de las siguientes posiciones: 175, 190, 193 de la SEC. N.º ID: 8, y una combinación de las mismas.

[0021] En determinadas realizaciones, la cinamato descarboxilasa mutada comprende una eliminación, una sustitución o una adición de un residuo aminoacídico en una de las posiciones de la SEC. N.º ID: 8: 155, 156, 159, 162, 163, 164, 172, 173, 174, 175, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 226, 227, 280, 285, 286, 287, 291, 326, 331, 360, 361, 395, 396, 398, 440, 441 y una combinación de las mismas. En determinadas realizaciones, la mutación es una sustitución.

[0022] En determinadas realizaciones, la cinamato descarboxilasa mutada comprende una eliminación, una sustitución o una adición de un residuo aminoacídico en una de las posiciones de la SEC. N.º ID: 8: 175, 190, 193 y una combinación de las mismas.

[0023] También se proporcionan ácidos nucleicos que codifican las cinamato descarboxilasas descritas en la presente memoria, células hospedadoras que comprenden las cinamato descarboxilasas descritas en la presente memoria y células hospedadoras que comprenden los ácidos nucleicos descritos en la presente memoria.

[0024] La tecnología objeto de esta invención también se refiere a un método para la producción de estireno que comprende: (a) poner en contacto una célula hospedadora con un sustrato fermentable (preferentemente un sustrato de carbono o un sustrato de nitrógeno), donde la célula hospedadora comprende: (i) una fenilalanina amonio liasa según se describe en la presente memoria; y (ii) una cinamato descarboxilasa según se describe en la presente memoria; y (b) cultivar la célula en un medio de cultivo durante el tiempo suficiente para producir estireno.

[0025] La tecnología objeto de esta invención también se refiere a una célula hospedadora que comprende: (a) una fenilalanina amonio liasa expresada recombinantemente según se describe en la presente memoria; (b) una cinamato descarboxilasa expresada recombinantemente según se describe en la presente memoria; y (c) un transportador unido a la membrana expresado recombinantemente según se describe en la presente memoria. La fenilalanina amonio liasa y la cinamato descarboxilasa se pueden expresar como dos proteínas independientes, o pueden estar enlazadas covalentemente mediante un conector según se describe en la presente memoria.

[0026] En un ejemplo, el transportador unido a la membrana es un transportador con dominio de unión a ATP (transportador ABC). Por ejemplo, el transportador ABC es un transportador ABC bacteriano, como uno que procede de *Pseudomonas putida*. Preferentemente, el transportador ABC es una bomba de salida de resistencia a disolventes, como la bomba SrpABC derivada de *Pseudomonas putida* S12.

[0027] La tecnología objeto de esta invención también proporciona un método para cribar proteínas candidatas mutadas atendiendo a la actividad de cinamato descarboxilasa, donde el método comprende: (a) proporcionar una muestra de proteína que comprende una proteína candidata, y un sustrato seleccionado del grupo que consiste en fenilalanina, ácido *trans*-cinámico, tirosina, ácido cumárico y combinaciones de los mismos; (b) combinar la muestra de proteína y la muestra del sustrato para formar una mezcla, e incubar la mezcla en condiciones que permitan que una cinamato descarboxilasa mutada transforme el sustrato en un producto seleccionado del grupo que consiste en estireno, 4-hidroxiestireno y una combinación de los mismos; y (c) exponer la mezcla a un material de detección que comprende una resina polimérica que absorbe el vapor del producto. En un ejemplo, la cinamato descarboxilasa mutada puede transformar el ácido *trans*-cinámico en estireno a una velocidad que es similar o superior a la de la enzima nativa (por ejemplo, FDC1). En otro ejemplo, la cinamato descarboxilasa mutada puede transformar el ácido cumárico en 4-hidroxiestireno a una velocidad que es similar o superior a la de la enzima nativa. En determinados ejemplos, la proteína candidata comprende una proteína de fusión que comprende una cinamato descarboxilasa mutada.

[0028] En un ejemplo, la tecnología objeto de esta invención también se refiere a un método para cribar proteínas candidatas atendiendo a la actividad de cinamato descarboxilasa que comprende: (a) proporcionar una muestra de proteína que comprende la proteína candidata y una muestra de sustrato que comprende el ácido *trans*-cinámico; (b) combinar la muestra de proteína y la muestra de sustrato para formar una mezcla, e incubar la mezcla en condiciones que permitan que la cinamato descarboxilasa transforme el ácido *trans*-cinámico en estireno; y (c) exponer la mezcla a un material de detección que comprende una resina polimérica que absorbe el vapor de estireno.

[0029] En determinados ejemplos, el material de detección también comprende un marcador detectable que produce un cambio de color en presencia de estireno. Por ejemplo, el marcador detectable puede ser 4-nitrobencilpiridina. El material de detección puede estar unido a un soporte sólido. En determinadas realizaciones, la resina polimérica comprende un grupo funcional aromático.

[0030] En determinados ejemplos, el cambio de color se puede detectar mediante espectrofotometría. Por ejemplo, el cambio de color se puede detectar mediante mediciones de la absorbancia de la muestra a una longitud de onda de aproximadamente 600 nm.

[0031] El método de cribado descrito en la presente memoria, pero no reivindicado, se puede utilizar para cribar simultáneamente una pluralidad de proteínas candidatas.

[0032] La tecnología objeto de esta invención también se refiere a un método para aislar a una cinamato descarboxilasa producida recombinantemente que comprende (a) proporcionar un hospedador bacteriano que comprende un ácido nucleico que codifica una cinamato descarboxilasa operativamente enlazada a una secuencia promotora; (b) cultivar el hospedador bacteriano en un medio de cultivo para expresar la cinamato descarboxilasa en la célula hospedadora, en el que la célula hospedadora se cultiva a una temperatura de aproximadamente 10 °C a aproximadamente 25 °C; y (c) aislar la cinamato descarboxilasa de la célula hospedadora, donde el aislamiento se lleva a cabo en un ambiente anaeróbico.

[0033] En determinados ejemplos, una o más soluciones amortiguadoras utilizadas para aislar la cinamato descarboxilasa comprenden un agente reductor, como tris(2-carboxietil)fosfina (TCEP), β -mercaptoetanol y una combinación de los mismos.

[0034] La tecnología objeto de esta invención también se refiere a un método para cristalizar una cinamato descarboxilasa, donde el método comprende: (a) proporcionar una solución de cinamato descarboxilasa a una concentración de aproximadamente 1 mg/ml a aproximadamente 50 mg/ml; (b) mezclar la solución de cinamato descarboxilasa con una solución de depósito en una relación de volúmenes de aproximadamente 1:10 a aproximadamente 10:1; y (c) mantener la mezcla de la solución de cinamato descarboxilasa y la solución de depósito a una temperatura adecuada para la formación de cristales de la cinamato descarboxilasa.

[0035] La tecnología objeto de esta invención también se refiere a un cristal de cinamato descarboxilasa, donde la cinamato descarboxilasa se encuentra formando un complejo con el ácido 3-hidroxicinámico.

[0036] La tecnología objeto de esta invención también se refiere a un método para producir estireno, donde el método comprende: (a) poner en contacto una célula hospedadora con un sustrato de carbono fermentable, donde la célula hospedadora que comprende (i) una fenilalanina amonio liasa; y (ii) una cinamato descarboxilasa; y (b) cultivar la célula hospedadora en un medio de cultivo durante el tiempo suficiente para producir estireno, donde un material absorbente absorbe el vapor del estireno producido.

[0037] En la presente memoria también se describe un método para detectar simultáneamente la actividad de fenilalanina amonio liasa y cinamato descarboxilasa, donde el método comprende: (a) proporcionar una proteína de fusión que comprende: (i) un primer dominio que comprende una fenilalanina amonio liasa, y (ii) un segundo dominio que comprende una cinamato descarboxilasa; (b) mezclar la proteína de fusión con un sustrato en condiciones que permiten que la proteína de fusión transforme el sustrato en un producto; y (c) detectar la cantidad de sustrato restante, o la cantidad de producto, o una combinación de los mismos. En determinados ejemplos, el cribado comprende detectar la pérdida del sustrato, o la cantidad de un producto, como ácido *trans*-cinámico, estireno y derivados posteriores del estireno. Por ejemplo, el cribado puede comprender detectar la cantidad de ácido *trans*-cinámico o estireno, o una combinación de los mismos.

[0038] La tecnología objeto de esta invención se ilustra, por ejemplo, según diversos aspectos descritos más abajo. Diversos ejemplos de aspectos de la tecnología objeto de esta invención están descritos en cláusulas numeradas (1, 2, 3, etc.) para facilitar la lectura. Estas se proporcionan como ejemplos y no limitan la tecnología objeto de esta invención. Cabe señalar que cualesquiera de las cláusulas dependientes se pueden combinar en cualquier combinación, y poner en la cláusula respectiva independiente, por ejemplo, la cláusula 1 o la cláusula 17. Las otras cláusulas se pueden presentar de una manera similar.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0039] La invención se define en las reivindicaciones y cualesquiera de las siguientes figuras que estén fuera del alcance de las reivindicaciones no forman parte de la invención, y se proporcionan para fines de comparación.

La Figura 1A es una ilustración esquemática de la estructura predicha de la FDC1 de levadura. La Figura 1B muestra cuatro posibles sitios de unión del tCA en la FDC1 (sitios A-D). La Figura 1C muestra el sitio de unión predicho del sustrato (tCA) en la FDC1 de levadura.

La Figura 2A muestra la expresión de la FDC1 en *E. coli* a 37 °C, utilizando tres soluciones amortiguadoras distintas. La FDC1 producida recombinantemente solo se acumula en las fracciones insolubles de los sedimentos (cuerpos de inclusión). La Figura 2B muestra la expresión de la FDC1 en *E. coli* a 16 °C. La FDC1 soluble se detectó en el sobrenadante.

La Figura 3 muestra el análisis por SDS-PAGE de la FDC1 purificada.

La Figura 4 muestra el efecto del pH en la actividad de la FDC1. La enzima mostró la máxima actividad a pH 6,5, y se consideró que tenía una actividad relativa del 100%. El eje de las x representa los valores de pH de las distintas soluciones amortiguadoras de reacción.

La Figura 5 muestra la estabilidad de la FDC1 con el pH. La enzima mostró una buena estabilidad con el pH. No perdió ninguna actividad cuando se incubó durante 30 minutos a pH 6,0 antes de la reacción, y se consideró que tenía una actividad relativa del 100%. El eje de las x representa los valores de pH a los que la proteína se incubó durante 30 minutos antes de la reacción enzimática.

La Figura 6 muestra el efecto de la temperatura en la actividad de la FDC1. La enzima mostró la máxima actividad cuando la reacción se llevó a cabo a 50 °C, y se consideró que tenía una actividad relativa del 100%. El eje de las x representa las temperaturas a las que se llevaron a cabo las reacciones enzimáticas.

La Figura 7 muestra la estabilidad de la FDC1 con la temperatura. La enzima mostró la máxima actividad a 50 °C, y no perdió ninguna actividad cuando se incubó a 50 °C durante 30 minutos antes de la reacción (lo que se considera como una actividad relativa del 100%). El eje de las x representa las temperaturas a las que la enzima se incubó antes de la reacción enzimática.

La Figura 8 muestra la estabilidad de la FDC1 con la temperatura a 50 °C durante diversos intervalos de tiempo. La enzima mostró la máxima actividad cuando se incubó a 50 °C durante 30 minutos antes de la reacción, y se consideró que tenía una actividad relativa del 100%. El eje de las x representa la duración (en minutos) de la incubación de la enzima a 50 °C antes de la reacción enzimática.

La Figura 9 muestra el efecto de los cofactores en la actividad de la FDC1. La actividad específica se muestra como nmol de estireno producido por mg de enzima por minuto.

5 La Figura 10A muestra el efecto de los iones metálicos en la actividad de la FDC1. La Figura 10B muestra el efecto de Zn^{2+} y Fe^{3+} en la actividad de la FDC1. La actividad específica se muestra en nmol de estireno producido por mg de enzima por minuto.

10 La Figura 11 muestra la especificidad de la FDC1 por el sustrato. La enzima mostró la máxima actividad cuando el ácido *trans*-cinámico se utilizó como sustrato, y se consideró que tenía una actividad relativa del 100%.

La Figuras 12A y 12B representan el análisis cinético de la FDC1. El eje de las x representa la reacción enzimática a diversas concentraciones de sustrato. La actividad específica se muestra como nmol de estireno producido por mg de enzima por minuto.

15 La Figura 13 ilustra el proceso de mutagénesis aleatoria para crear FDC mutadas, y el cribado de alto rendimiento de las FDC1 mutadas realizado por colorimetría.

20 La Figura 14 muestra la producción de estireno por FDC1 mutadas. El eje de las y representa la cantidad de estireno producido por las FDC1 mutadas. El eje de las x representa los aminoácidos que se han cambiado en las FDC1 mutadas.

La Figura 15 muestra la actividad de las proteínas de fusión FDC-PAL. El eje de las x muestra la longitud de los conectores.

25 La Figura 16 es una ilustración esquemática de la estructura predicha de la proteína de fusión FDC-PAL.

La Figura 17 muestra que la producción de estireno aumentó significativamente con la coexpresión de un transportador ABC.

30 La Figura 18A muestra la producción de estireno y tCA a partir de glucosa.

La Figura 18B muestra la producción de estireno a partir de ácido *trans*-cinámico o fenilalanina.

35 La Figura 19 muestra el uso de la resina STRATA-X® encima de un tubo de cultivo (izquierda) o sobre una placa de cultivo de 96 pocillos (derecha).

La Figura 20 muestra el análisis por SDS-PAGE de la FDC1 purificada.

40 La Figura 21A muestra un mapa de densidad electrónica común con un modelo inicial de la estructura cristalina de la FDC(K190E).

La Figura 21B muestra una unidad asimétrica de la estructura cristalina de la FDC(K190E).

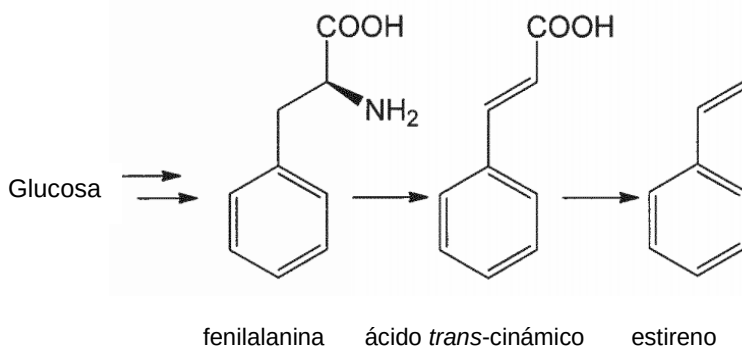
DESCRIPCIÓN DETALLADA

45

A. INTRODUCCIÓN

50 [0040] La tecnología objeto de esta invención por lo general se refiere a la biosíntesis del estireno. Determinadas realizaciones de la tecnología objeto de esta invención están basadas, en parte, en el entendimiento de que la fenilalanina se puede transformar en estireno por medio de una ruta de dos pasos de desaminación y descarboxilación, utilizando el ácido *trans*-cinámico (tCA) como intermedio. En este proceso participan de manera directa dos tipos de enzimas: la fenilalanina amonio liasa (PAL), que transforma la fenilalanina en tCA, y la cinamato descarboxilasa, que transforma el tCA en estireno. Las células hospedadoras que expresan estos dos tipos de enzimas se pueden cultivar en un biorreactor para producir estireno a partir de sustratos renovables, como la glucosa.

55



[0041] En concreto, se han adoptado diversos enfoques para aumentar la bioproducción de estireno. Por ejemplo, como se describe y ejemplifica en la presente memoria, se han diseñado proteínas de fusión en las que la fenilalanina amonio liasa y la cinamato descarboxilasa están covalentemente enlazadas mediante un conector. La proteína de fusión saca partido del fenómeno de «canalización del sustrato». La canalización del sustrato se refiere a un fenómeno por el cual los sustratos se suministran eficientemente de una enzima a otra sin equilibrarse con otras poblaciones de los mismos sustratos. En efecto, esto crea poblaciones locales de metabolitos a elevadas concentraciones con respecto a las encontradas en otras regiones de la célula. Debido a que el producto de la fenilalanina amonio liasa (ácido *trans*-cinámico) es consumido por la cinamato descarboxilasa, la proximidad entre estas dos enzimas (mediante la unión covalente de las dos enzimas en la forma de una proteína de fusión) proporcionaría un uso más eficiente del sustrato. Como tal, las proteínas de fusión en las que estas dos enzimas están unidas se benefician del fenómeno de la canalización del sustrato, y pueden reducir los costes de producción y aumentar el número de reacciones enzimáticas que pueden tener lugar durante un período de tiempo determinado. Como otra posibilidad, también se puede utilizar un complejo proteico en el que la fenilalanina amonio liasa y la cinamato descarboxilasa forman un complejo proteico mediante una interacción no covalente.

[0042] Por lo tanto, la tecnología objeto de esta invención proporciona una proteína de fusión que comprende: (i) un primer dominio que comprende una fenilalanina amonio liasa y (ii) un segundo dominio que comprende una cinamato descarboxilasa. También se proporcionan células hospedadoras que comprenden las proteínas de fusión descritas en la presente memoria, así como un método para utilizar la proteína de fusión para la producción de estireno.

[0043] En otro enfoque, se ha diseñado una biblioteca de cinamato descarboxilasas mutadas, y se han cribado atendiendo a su actividad. Se identificaron cinamato descarboxilasas mutadas que mostraban actividades catalíticas más elevadas que la forma nativa. Algunas de las cinamato descarboxilasas mutadas descritas en la presente memoria aumentaron la producción de estireno al doble o al triple. Estas cinamato descarboxilasas mutadas se pueden introducir en células hospedadoras para potenciar la bioproducción de estireno.

[0044] En otro ejemplo, un método de cribado colorimétrico de alto rendimiento puede permitir el cribado a gran escala de cinamato descarboxilasas mutadas o proteínas de fusión. El cribado colorimétrico descrito en la presente memoria es sumamente reproducible y con este se pueden cribar aproximadamente 1000 cinamato descarboxilasas mutadas y proteínas de fusión en un solo día. Esto puede proporcionar un cribado rápido y recursivo de los enzimas que participan en la biosíntesis del estireno y compuestos relacionados, lo que proporciona un método importante para mejorar la biosíntesis del estireno.

[0045] Por lo tanto, la tecnología objeto de esta invención proporciona cinamato descarboxilasas mutadas, células hospedadoras que comprenden una cinamato descarboxilasa mutada, así como un método para utilizar las cinamato descarboxilasas mutadas en la producción de estireno. En la presente memoria también se proporcionan, aunque no se reivindican, bibliotecas de cinamato descarboxilasas mutadas y un método de cribado de proteínas candidatas atendiendo a la actividad de cinamato descarboxilasa.

[0046] Otro problema que limita la bioproducción de estireno es la toxicidad del estireno para las células hospedadoras. Se sabe que la acumulación de compuestos aromáticos hidrofóbicos dentro de la membrana citoplasmática perturba su integridad. A fin de reducir la toxicidad del estireno y aumentar su producción, en la célula hospedadora se introdujo un transportador ABC (una bomba de salida que elimina el disolvente orgánico de la célula). Como se describe y se ejemplifica en la presente memoria, la producción de estireno por células de *E. coli* que expresan el transportador ABC fue casi cuatro veces mayor.

[0047] Por lo tanto, en un aspecto, la tecnología objeto de esta invención proporciona una célula hospedadora que comprende: (a) una fenilalanina amonio liasa expresada recombinantemente; (b) una cinamato descarboxilasa expresada recombinantemente; y (c) un transportador ABC expresado recombinantemente. También se proporciona un método para utilizar el transportador ABC para la producción de estireno.

B. DEFINICIONES

[0048] Como se utilizan en la presente memoria, las formas en singular «un», «uno», «una», «el» y «la» incluyen las referencias al plural a menos que el contenido claramente indique lo contrario.

[0049] Los términos «ácido cinámico» y «cinamato» se utilizan de manera intercambiable en la memoria descriptiva, y se abrevian como «AC». El término «ácido *trans*-cinámico» se abrevia como tCA.

[0050] El término «cinamato descarboxilasa» se refiere a una enzima que cataliza la conversión del ácido *trans*-cinámico en estireno. El término abarca la cinamato descarboxilasa nativa o de origen natural, así como fragmentos funcionales o variantes de una cinamato descarboxilasa nativa. La cinamato descarboxilasa de *Saccharomyces cerevisiae* descrita en la presente memoria también se denomina ferulato descarboxilasa (FDC o FDC1). El ácido ferúlico es también un ácido fenilacrílico.

[0051] El término «control», como se utiliza en la presente memoria, se refiere a una muestra que proporciona un medio para realizar una comparación. Por ejemplo, en un ensayo de cribado para determinar la actividad de cinamato descarboxilasa, un «control» puede ser una muestra paralela que comprende una cinamato descarboxilasa cuya actividad se ha caracterizado (por ejemplo, una cinamato descarboxilasa nativa, una cierta cinamato descarboxilasa mutada, etc.). Como otra posibilidad, un control puede ser un valor límite predeterminado, o un valor que está presente en una base de datos (por ejemplo, un cuadro, una base de datos electrónica, una hoja de cálculo, etc.).

[0052] Una posición aminoacídica «que corresponde a» una posición de referencia es una posición que se alinea con una secuencia de referencia, según se identifica al alinear las secuencias aminoacídicas. Dichas alineaciones se pueden realizar manualmente o mediante programas de alineación de secuencias bien conocidos, como ClustalW2, Blast 2, etc.

[0053] Una proteína «procede» de un organismo cuando la proteína se aísla a partir de ese organismo, o se modifica o se genera (por ejemplo, sintetizada químicamente o producida recombinantemente) utilizando información de la proteína de ese organismo.

[0054] El término «marcador detectable» se refiere a un compuesto químico que se añade a (o se deposita sobre) un material que absorbe el estireno, en una cantidad efectiva para detectar vapor de estireno. Los marcadores detectables preferidos son compuestos químicos que en presencia de estireno sufren una reacción química, y producen una sustancia que se puede detectar por colorimetría. La respuesta química del marcador detectable es preferiblemente dependiente de la concentración.

[0055] El término «sustrato de carbono fermentable» se refiere a una fuente de carbono que las células hospedadoras descritas en la presente memoria pueden metabolizar, y en particular, a las fuentes de carbono seleccionadas del grupo que consiste en monosacáridos, oligosacáridos, polisacáridos, sustratos de un carbono y una combinación de los mismos.

[0056] El término «fragmento funcional» de proteína se refiere a un fragmento peptídico que es una parte de la proteína completa, y que posee sustancialmente la misma actividad biológica, o desempeña sustancialmente la misma función que la proteína completa (por ejemplo, lleva a cabo la misma reacción enzimática). Por ejemplo, un fragmento funcional de una PAL puede catalizar la conversión de la fenilalanina en ácido cinámico, y un fragmento funcional de una FDC puede catalizar la conversión del ácido cinámico en estireno.

[0057] El término «variante funcional» de una proteína se refiere a una proteína en la que uno o más residuos aminoacídicos se han cambiado sin alterar la función ni la conformación global de la proteína de referencia. Entre las variantes funcionales se incluyen, por ejemplo, aquellas en las que un aminoácido se ha reemplazado con otro que tiene propiedades similares (como, por ejemplo, polaridad, capacidad para formar puentes de hidrógeno, acidez, basicidad, hidrofobicidad, aromaticidad, etc.). Los aminoácidos con propiedades similares son bien conocidos en la técnica. Por ejemplo, la arginina, la histidina y la lisina son aminoácidos hidrofílicos-básicos y se puede intercambiar. De manera similar, la isoleucina, un aminoácido hidrofóbico, se puede reemplazar con leucina, metionina o valina. Se prevé que dichos cambios tengan un efecto escaso o nulo en el peso molecular aparente o el punto isoeléctrico de la proteína o el polipéptido.

[0058] El término «homólogo/a» en todas sus formas gramaticales y variantes ortográficas se refiere a la relación entre polinucleótidos o proteínas que poseen un «origen evolutivo común», incluidos los polinucleótidos o proteínas de superfamilias, y las proteínas o polinucleótidos o homólogos de especies distintas (Reeck *et al.*, Cell 50:667, 1987). Dichos polinucleótidos o proteínas presentan homología de secuencia, según queda reflejado por la similitud entre sus secuencias, ya sea en términos del porcentaje de identidad o de la presencia de motivos o aminoácidos específicos en posiciones conservadas. Por ejemplo, dos proteínas homólogas pueden tener secuencias aminoacídicas que son idénticas en aproximadamente un 80%, en aproximadamente un 85%, en aproximadamente un 90%, en

aproximadamente un 95%, en aproximadamente un 96%, en aproximadamente un 97%, en aproximadamente un 98%, en aproximadamente un 99% o incluso en un 100%.

[0059] En este sentido, las técnicas para determinar la «similitud» entre secuencias aminoacídicas son bien conocidas en la técnica. En general, «similitud» significa la comparación exacta de aminoácido a aminoácido de dos o más polipéptidos en el lugar adecuado, donde los aminoácidos son idénticos o poseen propiedades químicas y/o físicas, como la carga o la hidrofobicidad, similares. A continuación, se puede determinar lo que se denomina el «porcentaje de similitud» entre las secuencias polipeptídicas comparadas. Las técnicas para determinar la identidad de las secuencias de los ácidos nucleicos y de aminoácidos también son bien conocidas en la técnica e incluyen determinar la secuencia nucleotídica del ARNm para ese gen (normalmente por medio de un ADNc intermedio) y determinar la secuencia aminoacídica codificada por este, y compararla a una segunda secuencia aminoacídica. En general, «identidad» se refiere a una correspondencia exacta de nucleótido a nucleótido, o de aminoácido a aminoácido, de dos secuencias polinucleotídicas o polipeptídicas, respectivamente. Dos o más secuencias polinucleotídicas se pueden comparar determinando su «porcentaje de identidad», del mismo modo que dos o más secuencias aminoacídicas. Los programas disponibles en la versión 8 del Paquete Wisconsin para análisis de secuencias (que se pueden obtener del Grupo de computación genética (Genetics Computer Group), Madison, Wisconsin, EE.UU.), como, por ejemplo, el programa GAP, pueden calcular tanto la identidad entre dos polinucleótidos como la identidad y similitud entre dos secuencias polipeptídicas, respectivamente. Los expertos en la materia conocen otros programas para calcular la identidad o similitud entre secuencias.

[0060] Los términos «mutación» o «mutada» como se utilizan en la presente memoria, se refieren a una eliminación, una inserción o una sustitución de un nucleótido o un residuo aminoacídico de una secuencia nativa. Una «secuencia nativa» hace referencia a la secuencia que se encuentra con más frecuencia en la naturaleza, y a partir de la cual se definen las formas mutadas.

[0061] El término «operativamente enlazado/a» se refiere a la asociación de secuencias del ácido nucleico a un único fragmento de ácido nucleico de manera que la función de uno está afectada por el otro. Por ejemplo, un promotor está operativamente enlazado a una secuencia codificante cuando puede afectar a la expresión de esa secuencia codificante (es decir, la secuencia codificante se encuentra bajo el control transcripcional del promotor). Las secuencias codificantes pueden estar operativamente enlazadas a secuencias reguladoras en orientación sentido y antisentido.

[0062] El término «fenilalanina amonio liasa», abreviado como PAL, se refiere a una enzima que cataliza la conversión de la fenilalanina en ácido *trans*-cinámico. El término abarca la fenilalanina amonio liasa nativa o de origen natural, así como fragmentos funcionales o variantes de una fenilalanina amonio liasa nativa.

[0063] El término «cinamato descarboxilasa mutada» se refiere a una enzima que cataliza la conversión del ácido *trans*-cinámico en estireno, o la conversión del ácido cumárico en 4-hidroxiestireno.

[0064] El término «recombinante» se refiere a una biomolécula, por ejemplo, un gen o una proteína, que (1) se ha suprimido de su entorno natural, (2) no está asociada a todo el polinucleótido, ni a una parte de este, en que el gen se encuentra de manera natural, (3) está enlazada operativamente a un polinucleótido al que no está enlazada de manera natural, o (4) no se encuentra en la naturaleza. El término «recombinante» se puede utilizar en referencia a ADN clonados aislados, análogos polinucleotídicos sintetizados químicamente, o análogos polinucleotídicos sintetizados biológicamente mediante sistemas heterólogos, así como proteínas y/o ARNm codificados por dichos ácidos nucleicos.

[0065] Una referencia a un elemento en singular no tiene por objeto significar «uno/a y solo uno/a» a menos que se indique específicamente, sino «uno o más». El término «algunos» se refiere a uno/a o más. Los encabezados y subencabezados subrayados, en cursiva y/o en negrita se utilizan únicamente para facilitar la lectura, no limitan la tecnología objeto de esta invención y no se hace referencia a ellos en relación con la interpretación de la memoria descriptiva de la tecnología objeto de esta invención. Se pretende que todos los equivalentes estructurales y funcionales de los elementos de las diversas configuraciones descritas en toda esta descripción que los expertos en la materia conozcan, o que posteriormente pasen a conocer, estén abarcados por la tecnología objeto de esta invención. Además, nada de los que está descrito en la presente memoria tiene por objeto estar dedicado al público, independientemente de si dicha descripción figura de manera explícita en la anterior memoria descriptiva.

[0066] A menos que se indique lo contrario, el porcentaje de identidad de dos polipéptidos o secuencias polinucleotídicas se refiere al porcentaje de residuos aminoacídicos o nucleótidos idénticos en toda la longitud de la secuencia más corta de las dos.

C. FENILALANINA AMONIO LIASA

[0067] La fenilalanina amonio liasa (EC 4,3,1,24; formalmente EC 4,3,1,5) es una enzima que cataliza la reacción química:

L-fenilalanina $\rightleftharpoons$ *trans*-cinamato + NH₃

Entre otros nombres que se suelen utilizar para esta enzima se incluyen tirasa, fenilalanina desaminasa, tirosina amonio liasa, *L*-tirosina amonio liasa, fenilalanina amonio liasa, PAL y *L*-fenilalanina amonio liasa. La PAL es una enzima procedente de organismos no mamíferos y ampliamente distribuida en plantas y levaduras.

[0068] En una lista representativa de PAL (identificadas por el número de acceso del Genbank y la especie) se incluyen: Q9ATN7 *Agastache rugosa*; 093967 *Amanita muscaria* (mosca agárica); P35510, P45724, P45725, Q9SS45, Q8RWP4 *Arabidopsis thaliana* (berro común); Q6ST23 *Bambusa oldhamii* (bambú gigante); Q42609 *Bromheadia finlaysoniana* (orquídea); P45726 *Camellia sinensis* (té); Q9MAX1 *Catharanthus roseus* (vincapervinca rosada) (vincapervinca de Madagascar); Q9SMK9 *Cicer arietinum* (garbanzo); Q9XFX5, Q9XFX6 *Citrus Clementina* x *Citrus reticulata*; Q42667 *Citrus Ziman* (limón); Q8H6V9, Q8H6W0 *Coffea canefora* (café Robusta); Q852S1 *Daucus carota* (zanahoria); O23924 *Digitalis lanata* (dedalera); O23865 *Daucus carota* (zanahoria); P27991 *Glycine max* (soja); O04058 *Helianthus annuus* (girasol común); P14166, (Q42858) *Ipomoea batatas* (boniato); Q8GZR8, Q8W2E4 *Lactuca sativa* (lechuga de campo); O49835, O49836 *Lithospermum erythrorhizon*; P35511, P26600 *Lycopersicon esculentum* (tomate); P35512 *Malus domestica* (manzana) (*Malus sylvestris*); Q94C45, Q94F89 *Manihot esculenta* (tapioca) (mandioca); P27990 *Medicago sativa* (alfalfa); P25872, P35513, P45733 *Nicotiana tabacum* (tabaco común); Q6T1C9 *Quercus suber* (alcornoque); P14717, P53443, Q7M1Q5, Q84VE0, Q84VE0 *Oryza sativa* (arroz); P45727 *Persea americana* (aguacate); Q9AX15 *Pharbitis nil* (violeta) (campanilla japonesa); P52777 *Pinus taeda* (pino taeda); Q01861, Q04593 *Pisum sativum* (guisante); P24481, P45728, P45729 *Petroselinum crispum* (perejil) (*Petroselinum hortense*); cultivar del híbrido Q84L12 de *Phalaenopsis* x *Doritaenopsis*; P07218, P19142, P19143 *Phaseolus vulgaris* (fríjol) (judía verde); Q7XJC3, Q7XJC4 *Pinus pinaster* (pino marítimo); Q6UD65 *Populus balsamifera* subsp. *trichocarpa* x *Populus deltoides*; P45731, Q43052, Q24266 *Populus kitakamiensis* (álamo); Q8H6V5, Q8H6V6 *Populus tremuloides* (álamo temblón); P45730 *Populus trichocarpa* (álamo negro); O64963 *Prunus avium* (cereza); Q94ENO *Rehmannia glutinosa*; P11544 *Rhodospiridium toruloides* (levadura) (*Rhodotorula gracilis*); P10248 *Rhodotorula rubra* (levadura) (*Rhodotorula mucilaginosa*); Q9M568, Q9M567 *Rubus idaeus* (frambuesa); P31425, P31426 *Solanum tuberosum* (patata); Q6SPE8 *Stellaria longipes* (estelaria de tallo largo); P45732 *Stylosanthes humilis* (stylo); P45734 *Trifolium subterraneum* (trébol subterráneo); Q43210, Q43664 *Triticum aestivum* (trigo); Q96V77 *Ustilago maydis* (hongo del fruto del maíz); P45735 *Vitis vinifera* (uva); y Q8VXG7 *Zea mays* (maíz).

[0069] También se conocen las estructuras cristalinas de diversas PAL (identificadas por el código de acceso del PDB y la especie): 1T6J (*Rhodospiridium toruloides*); 1T6P (*Rhodospiridium toruloides*); 1W27 (*Petroselinum crispum*); 1Y2M (*Rhodospiridium toruloides*); 2NYF (*Nostoc punctiforme*); y 3nz4 (*Taxus canadensis*). La PAL de *Rhodospiridium toruloides* (también denominada *Rhodotorula glutinis*) es un homotetrámero, en el que cada subunidad tiene una forma de «caballito de mar» que se engrana cabeza con cola con otras dos subunidades, de manera que se maximizan las interacciones con las subunidades adyacentes y se obtiene un tetrámero estrechamente encajado. El ensamblaje del tetrámero origina una agrupación de cuatro cisteínas vecinales (residuos 140, 455, 467 y 530), con los átomos de azufre de la Cys467 y la Cys530 separados por una distancia de 3,62 Å. La estructura del cuerpo principal de la PAL tiene un núcleo central de hélices α casi paralelas de distintas longitudes. Solo una sección de lámina β es más larga que tres residuos (hebras de residuos 231-237 y 240-246); reside en la región de embudo que desemboca en el sitio activo. MacDonald, *et al.*, *A modern view of phenylalanine ammonia lyase*, Biochem Cell Biol. 2007; 85(3):273-82.

[0070] En la técnica se conocen genes o polinucleótidos que codifican la PAL procedentes de fuentes vegetales y microbianas. Véase, por ejemplo, la patente europea n.º 321.488 (*R. toruloides*); la patente WO n.º 9.811.205 (*Eucalyptus grandis* y *Pinus radiata*); la patente WO n.º 9.732.023 (*Petunia*); la patente japonesa n.º 05.153.978 (*Pisum sativum*); la patente WO n.º 9.307.279 (patata, arroz). Las secuencias de diversos genes de la PAL se pueden determinar fácilmente a partir de la bibliografía, las bases de datos públicas (véase por ejemplo GENBANK® n.º de acceso AJ010143 y X75967) u otras fuentes comerciales. Por ejemplo, los ADNc completos de la PAL1 (At2g37040), PAL2 (At3g53260) y PAL4 (At3g10340) de *Arabidopsis* se pueden adquirir del Centro de recursos biológicos para *Arabidopsis* (Arabidopsis Biological Resource Center, ABRC) de la Universidad del estado de Ohio (EE.UU.) con los números de inventario G10120 (en el vector pENTR223.1), G12256 (en el vector pENTR223.1) y U24927 (en el vector pENTR/D-TOPO), respectivamente.

[0071] En determinadas realizaciones, la fenilalanina amonio liasa de la tecnología objeto de esta invención se obtiene de un organismo seleccionado del grupo que consiste en: una *Arabidopsis*, una *Anabaena*, una *Nostoc* y una *Saccharomyces*. Entre otros organismos fuente preferidos se incluyen, por ejemplo, las levaduras como *Rhodotorula*, *Rhodospiridium* y *Sporobolomyces*; las bacterias como *Streptomyces*; y las plantas como el guisante, la patata, el arroz, el eucalipto, el pino, el maíz, la petunia, *Arabidopsis*, el tabaco y el perejil.

[0072] La tecnología objeto de esta invención también abarca homólogos (incluidos ortólogos), fragmentos funcionales o variantes funcionales de los ejemplos de PAL descritas en la presente memoria. En la técnica son bien conocidos los métodos para obtener homólogos y variantes de PAL, incluidos, por ejemplo, los protocolos dependientes de secuencias. Entre los ejemplos de protocolos dependientes de secuencias se incluyen, por ejemplo,

la hibridación de ácidos nucleicos, la amplificación de ADN y ARN (por ejemplo, la reacción en cadena de la polimerasa (PCR), la reacción en cadena de la ligasa (LCR)), etc.

[0073] Por ejemplo, los genes que codifican homólogos de una PAL se pueden aislar directamente utilizando la totalidad o una parte de las secuencias conocidas como sondas de ADN para hibridación con el fin de realizar el cribado de bibliotecas obtenidas de cualquier planta, hongo, levadura o bacteria deseados mediante técnicas bien conocidas por los expertos en la materia. Se pueden diseñar sondas oligonucleotídicas específicas basadas en las secuencias de ácidos nucleicos descritas en la bibliografía y se pueden sintetizar mediante métodos conocidos en la técnica (Maniatis, *infra*). Además, las secuencias completas se pueden utilizar directamente para sintetizar sondas de ADN mediante métodos conocidos por los expertos en la técnica, tales como el marcado del ADN mediante cebadores aleatorios, el desplazamiento de mella, o las técnicas de marcado de los extremos, o sondas de ARN que utilizan los sistemas de transcripción *in vitro* disponibles. Además, se pueden diseñar y utilizar cebadores específicos para amplificar las secuencias completas de interés o una porción de estas. Los productos de amplificación resultantes se pueden marcar directamente durante las reacciones de amplificación o se pueden marcar después de las reacciones de amplificación, y se pueden utilizar como sondas para aislar ADNc completo o fragmentos genómicos en las condiciones de rigor adecuadas.

[0074] Además, en los protocolos de reacción en cadena de la polimerasa se pueden utilizar dos segmentos cortos de las secuencias descritas en la bibliografía para amplificar fragmentos más largos de ácidos nucleicos que codifican genes homólogos de ADN o ARN. La reacción en cadena de la polimerasa también se puede realizar en una biblioteca de fragmentos de ácidos nucleicos clonados en donde la secuencia de un cebador procede de las secuencias descritas en la bibliografía, y la secuencia del otro cebador se aprovecha de la presencia de tramos de ácido poliadenílico en el extremo 3' del precursor del ARNm que codifica los genes bacterianos. Como otra posibilidad, la secuencia del segundo cebador puede estar basada en secuencias derivadas del vector de clonación. Por ejemplo, un experto en la materia puede seguir el protocolo RACE (Frohman *et al.*, PNAS EE.UU. 85:8998 (1988)) para generar los ADNc mediante PCR con el fin de amplificar las copias de la región comprendida entre un punto único en el transcrito y el extremo 3' o 5'. Se pueden diseñar cebadores orientados en las direcciones 3' y 5' a partir de secuencias descritas en la bibliografía. Mediante el uso de sistemas 3'-RACE o 5'-RACE disponibles en el mercado, se pueden aislar fragmentos específicos en 3' o 5' del ADNc (Ohara *et al.*, Proc. Natl. Acad. Sci. USA 86:5673 (1989); Loh *et al.*, Science 243:217 (1989)).

[0075] Las secuencias del ácido nucleico y la proteína de un homólogo o una variante también se pueden identificar utilizando una «secuencia de consulta» para realizar una búsqueda en bases de datos públicas a fin de identificar otros miembros de la familia o secuencias relacionadas.

[0076] Por lo tanto, en un ejemplo de realización, la fenilalanina amonio liasa de la tecnología objeto de esta invención comprende una secuencia aminoacídica seleccionada del grupo que consiste en: SEC. N.º ID: 2, SEC. N.º ID: 4, SEC. N.º ID: 6 y un fragmento funcional o variante de las mismas. Las SEC. N.º ID: 2, SEC. N.º ID: 4 y SEC. N.º ID: 6 son las secuencias aminoacídicas de la PAL1, PAL2 y PAL4 de *Arabidopsis*, respectivamente. Las variantes de la PAL pueden incluir a aquellas secuencias polipeptídicas que comprenden una secuencia aminoacídica que es idéntica en aproximadamente un 60%, en aproximadamente un 70%, en aproximadamente un 75%, en aproximadamente un 80%, en aproximadamente un 85%, en aproximadamente un 90%, en aproximadamente un 95%, en aproximadamente un 96%, en aproximadamente un 97%, en aproximadamente un 98%, en aproximadamente un 99% o incluso en un 100% a cualquiera de las SEC. N.º ID: 2, SEC. N.º ID: 4, SEC. N.º ID: 6 o un fragmento de las mismas. El fragmento puede comprender aproximadamente 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700 o 750 residuos aminoacídicos de la PAL completa.

[0077] En otro ejemplo, la fenilalanina amonio liasa de la tecnología objeto de esta invención comprende una secuencia aminoacídica codificada por una secuencia polinucleotídica seleccionada del grupo que consiste en: SEC. N.º ID: 1, SEC. N.º ID: 3, SEC. N.º ID: 5, y un fragmento o variante de las mismas. Las SEC. N.º ID: 1, SEC. N.º ID: 3, y SEC. N.º ID: 5 son las secuencias nucleotídicas que codifican la PAL1, PAL2 y PAL4 de *Arabidopsis*, respectivamente. Algunas variantes de la secuencia codificante de la PAL pueden incluir aquellas secuencias polinucleotídicas que son idénticas en aproximadamente un 60%, en aproximadamente un 70%, en aproximadamente un 75%, en aproximadamente un 80%, en aproximadamente un 85%, en aproximadamente un 90%, en aproximadamente un 95%, en aproximadamente un 96%, en aproximadamente un 97%, en aproximadamente un 98%, en aproximadamente un 99% o incluso en un 100% a cualquiera de las SEC. N.º ID: 1, SEC. N.º ID: 3, SEC. N.º ID: 5 o un fragmento de las mismas. El fragmento puede comprender aproximadamente 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300 1400, 1500, 1600, 1700, 1800, 1900, 2000 o 2100 nucleótidos de la secuencia codificante de la PAL completa.

[0078] Para determinar el porcentaje de identidad de dos secuencias aminoacídicas o de dos secuencias polinucleotídicas, las secuencias se alinean para poder compararlas de manera óptima (por ejemplo, se pueden introducir brechas en la secuencia de una primera secuencia aminoacídica o del ácido nucleico para lograr un alineamiento óptimo con una segunda secuencia aminoacídica o del ácido nucleico, y las secuencias no homólogas se pueden descartar para fines de comparación). A menos que se indique lo contrario, un alineamiento es un

alineamiento global, (es decir, el porcentaje de residuos aminoácidos o nucleótidos idénticos en toda la longitud de la secuencia más corta de las dos). Si se desea realizar un alineamiento local, preferentemente, la longitud de una secuencia alineada para su comparación es de aproximadamente un 30%, de aproximadamente un 40%, de aproximadamente un 50%, de aproximadamente un 60%, de aproximadamente un 70%, de aproximadamente un 80% o de aproximadamente un 90% de la longitud de la secuencia más corta de las dos.

D. CINAMATO DESCARBOXILASA

[0079] La cinamato descarboxilasa es una enzima que cataliza la conversión del ácido *trans*-cinámico en estireno.

[0080] En determinadas realizaciones, la cinamato descarboxilasa de la tecnología objeto de esta invención se obtiene a partir de un organismo seleccionado del grupo que consiste en: una *Arabidopsis*, una *Anabaena*, una *Nostoc* y una *Saccharomyces*.

1. FDC1 y OHBA1

[0081] En un ejemplo de realización, la cinamato descarboxilasa de la tecnología objeto de esta invención se obtiene a partir de *Saccharomyces cerevisiae*. La cinamato descarboxilasa de *Saccharomyces cerevisiae* descrita en la presente memoria también se denomina ferulato descarboxilasa (FDC o FDC1). La secuencia de la FDC/FDC1 se da a conocer en las patentes de EE.UU. n.º 5.955.137 y 6.468.566.

[0082] Como se describe y se ejemplifica en la presente memoria, la cinamato descarboxilasa completa de *Saccharomyces cerevisiae* comprende 503 aminoácidos (SEC. N.º ID: 8). En la Figura 1A se muestra un modelo informático de la FDC1 completa; y en la Figura 2B se muestran cuatro posibles sitios de unión del tCA, siendo el sitio C el más probable. Los análisis adicionales del sitio C muestran que 1173, A174, R175, V188, 1189, K190 (no se muestra por claridad de la figura), I194, E280, M286, F291 y F440 participan en la unión del sustrato (Figura 1C). La secuencia nucleotídica que codifica la FDC1 se muestra como SEC. N.º ID: 7.

[0083] Los estudios cinéticos muestran que la Km para la FDC nativa es aproximadamente 688 µM y la V_{max} es aproximadamente 6,17 nmol·mg⁻¹·min⁻¹. Además de al tCA, la FDC1 también se une a otros sustratos, incluidos el ácido ferúlico, el ácido 2-metilcinámico y el ácido 4-hidroxicinámico.

[0084] Otro ejemplo de cinamato descarboxilasa es la 3-octaprenil-4-hidroxibenzoato carboxilasa de *Aspergillus niger* (OHBA1; SEC. N.º ID: 10). La secuencia nucleotídica que codifica la OHBA1 se muestra como SEC. N.º ID: 9.

[0085] La tecnología objeto de esta invención también abarca homólogos (incluidos ortólogos), fragmentos funcionales o variantes funcionales de los ejemplos de cinamato descarboxilasas descritas en la presente memoria. Los homólogos y las variantes de cinamato descarboxilasas se pueden obtener mediante los métodos descritos más arriba.

[0086] Por ejemplo, una variante funcional puede ser una secuencia que (i) es idéntica en aproximadamente un 50%, en aproximadamente un 60%, en aproximadamente un 70%, en aproximadamente un 80%, en aproximadamente un 85%, en aproximadamente un 90%, en aproximadamente un 95%, en aproximadamente un 96%, en aproximadamente un 97%, en aproximadamente un 98%, en aproximadamente un 99% o incluso en un 100% a la SEC. N.º ID: 8, (ii) comprende al menos 1, al menos 2, al menos 3, al menos 4, al menos 5, al menos 6, al menos 7, al menos 8, al menos 9 o al menos 10 de los siguiente residuos: I en una posición que corresponde al residuo 173 de la SEC. N.º ID: 8, A en una posición que corresponde al residuo 174 de la SEC. N.º ID: 8, R en una posición que corresponde al residuo 175 de la SEC. N.º ID: 8, V en una posición que corresponde al residuo 188 de la SEC. N.º ID: 8, I en una posición que corresponde al residuo 189 de la SEC. N.º ID: 8, K en una posición que corresponde al residuo 190 de la SEC. N.º ID: 8, L en una posición que corresponde al residuo 194 de la SEC. N.º ID: 8, E en una posición que corresponde al residuo 280 de la SEC. N.º ID: 8, M en una posición que corresponde al residuo 286 de la SEC. N.º ID: 8, F en una posición que corresponde al residuo 291 de la SEC. N.º ID: 8, y F en una posición que corresponde al residuo 440 de la SEC. N.º ID: 8; y una combinación de (i) y (ii).

[0087] Los residuos aminoácidos que «corresponden a» una posición determinada de la SEC. N.º ID: 8 se pueden identificar mediante alineamiento de la secuencia de interés con la SEC. N.º ID: 8. Como ejemplo, el alineamiento de la FDC1 y la OHBA1 se muestra más abajo (la FDC1 es la secuencia de consulta; la OHBA1 es la secuencia de referencia).

Consulta 8	LEFRDFIQVLKDEDDLEITEEIDPNLEVGAIMRKAYESHLPAPLFKNLKGASKDLFSIL	67
Refer. 8	L FR F++ LK ++DL+EI IDPNLE AI R+ E++ APLF NL G LF IL LCFRSFVEALKVDNDLVEINTPIDPNLEAAAITRRVCETNDKAPLFNNLIGMKNGLFRIL	67
Consulta 68	GCPAGLRSEKKGHDHGRIAHHLGLDPKTTIKEIIDYLLECKEKEPLPPITVPVSSAPCKTH	127
Refer. 68	G P LR +GR+A HL L P +++EI+D +L + P+ V + PCK + GAPGSLRKSSADRYGRLARHLALPPTASMREILDKMLSASDMFPI--PPTIVPTGPCKEN	125
Consulta 128	ILSEEKIHLSLPTPYLHVSDGGKYLQTYGMWILQTPDKKWTNWSIARGMVVDDKHITGL	187
Refer. 126	L + + L LP P +H SDGGKY+QTYGM I+Q+PD WTNWSIAR MV D H+TGL SLDDSEFDLTLPVPLIHKSDGGKYI QTYGMHIVQSPDGTWTNWSIARAMVHDKNHLTGL	185
Consulta 188	VIKQHIRQIADSWAAIGKANEIPFALCFGVPPAAILVSSMPIPEGVSESDYVGAILGES	247
Refer. 186	VI PQHI QI W G++ ++P+AL FGVPPAAI+ SSMPI+GV+E+ YVGA+ G S VIPPQHIWQIHQMWKKEGRS-DVPWALAFGVPPAIASSMPIPDGVTEAGYVVGAMTGSS	244
Consulta 248	VPVVKCETNDLMVPATSEMVFEGTSLTDTHLEGPFGEHMGYVFKSQGHPCPLYTVKAMS	307
Refer. 245	+ +VKC+TNDL VPATSE+V EGTLS+++T EGPFGEHMGY+F H Y V ++ LELVKCDTNDLYVPATSEIVLEGTLSISETGPEGPFGEHMGYIFPGDTHLGAKYKVNRI	304
Consulta 308	YRDNAILPVSNGCLCTDETHTLIGSLVATEAKELAIESGLPILDAFMPYEAQALWLILKV	367
Refer. 305	YR+NAI+P+S+ G TDETH+IGSL A E ++L ++ LPI DAF P+E+Q W+ L+V YRNNAIMPMSSCGRLTDETHTMIGSLAAAEIRKLCQQNDLPITDAFAPFESQVTVVALRV	364
Consulta 368	DLKGLQALKTTPEEFCKKVGDIYFRTKVGFIVHEIILVADDIDIFNFKEVIWAYVTRHTP	427
Refer. 365	D + L+A+KTT E F K+VGD+ F K G+ +H ++LV DDID++ K+V+WA+ TR P DTEKLRAMKTTSEGFPRKRVGDVVFNHKAGYTIHRLVIVGDDIDVYEGKDVWAFSTRCP	424
Consulta 428	VADQMAFDDVTSFPLAPFVSQSSRSKTMKGKCVTNCFRQQYERSFDYITCNFEKGYPK	487
Refer. 425	D+ F+DV FPL P++ + +GGK V++ + +Y ++ +F + YP+ GMDETFLFEDVRGFPLIPYMGHGN-GFAHRGGKVVS DALMPTEYTTGRNWEAADFNQSYPE	483
Consulta 488	GLVDKVNENWKRYGY	502
Refer. 484	L KV +NW + G+ DLKQKVLNDNWTMKG	498

[0088] En un ejemplo de realización, la cinamato descarboxilasa de la tecnología objeto de esta invención comprende una secuencia aminoacídica seleccionada del grupo que consiste en: SEC. N.º ID: 8, SEC. N.º ID: 10, y un fragmento funcional o variante de las mismas. Entre las variantes de la cinamato descarboxilasa se pueden incluir aquellas secuencias polipeptídicas que comprenden una secuencia aminoacídica que es idéntica en aproximadamente un 50%, en aproximadamente un 60%, en aproximadamente un 70%, en aproximadamente un 75%, en aproximadamente un 80%, en aproximadamente un 85%, en aproximadamente un 90%, en aproximadamente un 95%, en aproximadamente un 96%, en aproximadamente un 97%, en aproximadamente un 98%, en aproximadamente un 99% o incluso en un 100% a cualquiera de las SEC. N.º ID: 8, SEC. N.º ID: 10, o un fragmento de las mismas. El fragmento puede comprender aproximadamente 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700 o 750 residuos aminoacídicos de la cinamato descarboxilasa completa.

[0089] En otro ejemplo de realización, la cinamato descarboxilasa de la tecnología objeto de esta invención comprende una secuencia aminoacídica codificada por una secuencia polinucleotídica seleccionados del grupo que consiste en: SEC. N.º ID: 7, SEC. N.º ID: 9, y un fragmento o variante de las mismas. Algunas variantes de la secuencia codificante de la cinamato descarboxilasa pueden incluir aquellas secuencias polinucleotídicas que son idénticas en aproximadamente un 60%, en aproximadamente un 70%, en aproximadamente un 75%, en aproximadamente un 80%, en aproximadamente un 85%, en aproximadamente un 90%, en aproximadamente un 95%, en aproximadamente un 96%, en aproximadamente un 97%, en aproximadamente un 98%, en aproximadamente un 99% o incluso en un 100% a cualquiera de las SEC. N.º ID: 7, SEC. N.º ID: 9 o un fragmento de las mismas. El fragmento puede comprender aproximadamente 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300 1400, 1500, 1600, 1700, 1800, 1900, 2000 o 2100 nucleótidos de la secuencia codificante completa.

2. Ejemplos de FDC1 mutadas

[0090] Con base en el análisis estructural de la FDC1 de levadura, los residuos 155-156, 159, 162-164, 172-175, 187-196, 226-227, 280, 285-287, 291, 326, 331, 360-361, 395-396, 398 y 440-441 de la FDC1 (SEC. N.º ID: 8) fueron identificados como sitios potenciales para la mutagénesis. Por lo tanto, en otro aspecto, la tecnología objeto de esta invención proporciona una cinamato descarboxilasa mutada, donde se introduce una mutación en una posición de residuo aminoacídico que corresponde a una de las siguientes posiciones: 155, 156, 159, 162, 163, 164, 172, 173,

174, 175, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 226, 227, 280, 285, 286, 287, 291, 326, 331, 360, 361, 395, 396, 398, 440, 441 de la SEC. N.º ID: 3313318, y una combinación de las mismas. La mutación puede ser una mutación por eliminación, inserción o sustitución.

[0091] En un ejemplo de realización, la cinamato descarboxilasa mutada comprende una mutación en una posición de residuo aminoacídico que corresponde al residuo 190 de la SEC. N.º ID: 8. Preferentemente, el residuo que corresponde a la posición 190 se reemplaza con uno de los siguientes: E, C, D, V, N, L, H. En otro ejemplo de realización, la cinamato descarboxilasa mutada comprende una mutación en una posición de residuo aminoacídico que corresponde al residuo 175 de la SEC. N.º ID: 8. Preferentemente, el residuo que corresponde a la posición 175 se reemplaza con I.

[0092] En determinadas realizaciones, la cinamato descarboxilasa mutada es una FDC1 mutada, donde una mutación se introduce en una de las siguientes posiciones en la SEC. N.º ID: 8: 155, 156, 159, 162, 163, 164, 172, 173, 174, 175, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 226, 227, 280, 285, 286, 287, 291, 326, 331, 360, 361, 395, 396, 398, 440, 441 y una combinación de las mismas.

[0093] En determinadas realizaciones, la cinamato descarboxilasa mutada es una FDC1 mutada, donde una mutación se introduce en una de las siguientes posiciones en la SEC. N.º ID: 8: 175, 190, 193 y una combinación de las mismas.

[0094] En determinadas realizaciones, la cinamato descarboxilasa mutada es una forma mutada de la FDC1 que comprende una de las mutaciones en la SEC. N.º ID: 8: K190E, K190C, K190H, K190P, K190L, K190R, K190D, K190V, K190S, K190N, R175I, H193P y una combinación de las mismas.

[0095] En determinados ejemplos, la cinamato descarboxilasa mutada comprende cualesquiera de las secuencias seleccionadas del grupo que consiste en la SEC. N.º ID: 16; SEC. N.º ID: 18; SEC. N.º ID: 20; SEC. N.º ID: 22; SEC. N.º ID: 24; SEC. N.º ID: 26; SEC. N.º ID: 28; SEC. N.º ID: 30; SEC. N.º ID: 32; SEC. N.º ID: 34; SEC. N.º ID: 36; y SEC. N.º ID: 38.

[0096] La cinamato descarboxilasa de la tecnología objeto de esta invención también puede comprender una variante o fragmento funcional de la cinamato descarboxilasa mutada descrita en la presente memoria. Entre las variantes de cinamato descarboxilasa se pueden incluir aquellas secuencias polipeptídicas que comprenden una secuencia aminoacídica que es idéntica en aproximadamente un 50%, en aproximadamente un 60%, en aproximadamente un 70%, en aproximadamente un 75%, en aproximadamente un 80%, en aproximadamente un 85%, en aproximadamente un 90%, en aproximadamente un 95%, en aproximadamente un 96%, en aproximadamente un 97%, en aproximadamente un 98%, en aproximadamente un 99% o incluso en un 100% a cualquiera de las SEC. N.º ID: 16; SEC. N.º ID: 18; SEC. N.º ID: 20; SEC. N.º ID: 22; SEC. N.º ID: 24; SEC. N.º ID: 26; SEC. N.º ID: 28; SEC. N.º ID: 30; SEC. N.º ID: 32; SEC. N.º ID: 34; SEC. N.º ID: 36; SEC. N.º ID: 38; y un fragmento de las mismas. El fragmento puede comprender aproximadamente 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700 o 750 residuos aminoacídicos de la cinamato descarboxilasa completa.

[0097] Como otra posibilidad, o de manera adicional, las variantes pueden comprender al menos 1, al menos 2, al menos 3, al menos 4, al menos 5, al menos 6, al menos 7, al menos 8, al menos 9 o al menos 10 de los siguientes residuos: I en una posición que corresponde al residuo 173 de la SEC. N.º ID: 8, A en una posición que corresponde al residuo 174 de la SEC. N.º ID: 8, R en una posición que corresponde al residuo 175 de la SEC. N.º ID: 8, V en una posición que corresponde al residuo 188 de la SEC. N.º ID: 8, I en una posición que corresponde al residuo 189 de la SEC. N.º ID: 8, K en una posición que corresponde al residuo 190 de la SEC. N.º ID: 8, L en una posición que corresponde al residuo 194 de la SEC. N.º ID: 8, E en una posición que corresponde al residuo 280 de la SEC. N.º ID: 8, M en una posición que corresponde al residuo 286 de la SEC. N.º ID: 8, F en una posición que corresponde al residuo 291 de la SEC. N.º ID: 8, y F en una posición que corresponde al residuo 440 de la SEC. N.º ID: 8.

[0098] También se proporcionan bibliotecas de cinamato descarboxilasas mutadas. La biblioteca comprende una pluralidad de cinamato descarboxilasas mutadas. Por ejemplo, la biblioteca puede comprender cinamato descarboxilasas mutadas que representan aproximadamente 15, aproximadamente 16, aproximadamente 17, aproximadamente 18 o aproximadamente 19 mutaciones distintas en una posición de interés determinada. Por ejemplo, la biblioteca puede comprender aproximadamente 19 formas mutadas distintas en las que el K190 de la FDC1 (SEC. N.º ID: 8) está reemplazado con cada uno de los otros 19 aminoácidos. Como otra posibilidad, la biblioteca puede comprender cinamato descarboxilasas mutadas que representan aproximadamente 5, aproximadamente 6, aproximadamente 7, aproximadamente 8, aproximadamente 9, aproximadamente 10, aproximadamente 11, aproximadamente 12, aproximadamente 13, aproximadamente 14, aproximadamente 15, aproximadamente 16, aproximadamente 17, aproximadamente 18 o aproximadamente 19 mutaciones en posiciones distintas. Por ejemplo, se puede realizar un barrido de alanina para crear bibliotecas de formas mutadas en las que los residuos aminoacídicos en distintas posiciones se han reemplazado con alanina.

[0099] Los cambios en la secuencia aminoacídica se pueden generar realizando cambios en la secuencia del ácido nucleico que codifica la secuencia aminoacídica. Una secuencia del ácido nucleico que codifica una cinamato descarboxilasa mutada se puede preparar mediante métodos conocidos en la técnica, basándose en la presente memoria descriptiva para secuencias determinadas. Entre estos métodos se incluyen, entre otros, la preparación mediante mutagénesis dirigida (o mediada por oligonucleótidos) (Coombs *et al.*, *Proteins* (1998), 259-311, 1 *plate*. Editor(es): Angeletti, Ruth Hogue. Editorial: Academic, San Diego, California, EE.UU.); mutagénesis por PCR o por PCR propensa a error (Melnikov *et al.*, *Nucleic Acids Research*, (Feb. 15, 1999) Vol. 27, No. 4, págs. 1056-1062); mutagénesis por inserción de un casete de una secuencia del ácido nucleico preparada con anterioridad; «reordenamiento aleatorio de genes» (patente de EE.UU. n.º 5.605.793; patente de EE.UU. n.º 5.811.238; patente de EE.UU. n.º 5.830.721; y patente de EE.UU. n.º 5.837.458); todas estos métodos son bien conocidos en la técnica.

[0100] Como otra posibilidad, se puede utilizar la mutagénesis *in vivo* con materiales disponibles en el mercado, como la cepa XL 1-Red de *E. coli* y la cepa mutada XL 1-Red XL de *Epicurian coli* de STRATAgene (STRATAgene, La Jolla, California, EE.UU. Greener and Callahan, *Strategies* 7:32-34) (1994)). Esta cepa es deficiente en tres de las principales rutas de reparación del ADN (mutS, mutD y mutT), lo que produce una tasa de mutaciones 5000 veces más elevada que en la forma nativa.

3. Método de cribado

[0101] La tecnología objeto de esta invención proporciona un método para cribar proteínas candidatas atendiendo a la actividad de cinamato descarboxilasa, que comprende (a) proporcionar una muestra de proteína que comprende la proteína candidata y una muestra de sustrato que comprende ácido *trans*-cinámico; (b) combinar la muestra de proteína y la muestra de sustrato par formar una mezcla, e incubar la mezcla en unas condiciones que permitan que una cinamato descarboxilasa transforme el ácido *trans*-cinámico en estireno; y (c) exponer la mezcla a un material de detección que comprende (i) una resina polimérica que absorbe el vapor de estireno; y (ii) un marcador detectable que produce un cambio de color en presencia de estireno. Un cambio en el color del material de detección indica que la proteína candidata tiene actividad de cinamato descarboxilasa.

[0102] Para el cribado de FDC mutadas, se ha desarrollado un ensayo de cribado de alto rendimiento. El cribado de una gran población de la biblioteca de proteínas constituye el cuello de botella para la evolución molecular de la proteína. La caracterización funcional de la descarboxilasa se suele realizar con instrumentos analíticos, como HPLC o LC-MS. Aunque el HPLC es muy sensible, requiere mucho tiempo, es caro y genera residuos como metanol o acetonitrilo, y no es adecuado para aplicaciones que requieran una elevada capacidad.

[0103] Para sortear estos escollos técnicos, la tecnología objeto de esta invención proporciona un método de ensayo colorimétrico basado en espectroscopía. En primer lugar, una proteína candidata y el tCA se incuban en unas condiciones que permiten que una cinamato descarboxilasa transforme el ácido *trans*-cinámico en estireno. Normalmente, las condiciones de incubación deben permitir a una cinamato descarboxilasa mostrar una actividad enzimática óptima (por ejemplo, a una temperatura de aproximadamente 10 °C a aproximadamente 65 °C, de aproximadamente 25 °C a aproximadamente 55 °C, de aproximadamente 35 °C a aproximadamente 60 °C, o de aproximadamente 35 °C a aproximadamente 55 °C; a un pH de entre 6 y 10, de entre 6 y 8, o de entre 5,5 y 7,5; en presencia de iones metálicos a concentraciones de aproximadamente 1 mM a aproximadamente 20 mM; etc.). La proteína candidata y el tCA se pueden incubar durante aproximadamente 30 minutos, aproximadamente 1 hora; aproximadamente 90 minutos, aproximadamente 2 horas, aproximadamente 2,5 horas, aproximadamente 3 horas; aproximadamente 3,5 horas, aproximadamente 4 horas, aproximadamente 4,5 horas, aproximadamente 5 horas, aproximadamente 6 horas, aproximadamente 7 horas o aproximadamente 8 horas, para permitir que se produzca la cantidad suficiente de estireno.

[0104] En segundo lugar, la mezcla de reacción se expone a un material de detección que comprende (i) una resina polimérica que absorbe el vapor de estireno; y (ii) un marcador detectable que produce un cambio de color en presencia de estireno. Debido a que el estireno se evapora, el material de detección se puede colocar encima de la mezcla de reacción para absorber el vapor de estireno (sin tener que sumergirlo en la mezcla de reacción). Por lo tanto, el material de detección se puede unir a un soporte sólido, y se puede invertir durante el proceso de detección para absorber el vapor de estireno. Véase, por ejemplo, la Figura 19.

[0105] Entre las resinas poliméricas preferidas para la absorción del vapor de estireno se incluyen, por ejemplo, las resinas hidrofóbicas de fase inversa, que tienen un grupo funcional hidrocarbonado o aromático (por ejemplo, un anillo aromático de benceno) que se puede unir a compuestos polares y apolares. Más preferentemente, las resinas hidrofóbicas tienen un grupo funcional hidrocarbonado o aromático, y no tienen ningún grupo polar. Entre los ejemplos de resinas hidrofóbicas de fase inversa se incluyen las resinas adsorbentes de C18, C8, fenilo, SDB-L y combinaciones de las mismas.

[0106] Ejemplos de resinas poliméricas que se pueden utilizar para absorber compuestos orgánicos volátiles (como el vapor de estireno) incluyen, por ejemplo, las resinas poliméricas STRATA-X® (Phenomenex), las resinas poliméricas Amberlite (Sigma-Aldrich), las resinas poliméricas DOWEX™ o las resinas poliméricas AMBERJET™ (The Dow

Chemical Compani), las resinas poliméricas Macronet™, las resinas poliméricas PuroSorb™ o las resinas Chromalite® (PuroLite), etc. El área superficial y la distribución del tamaño de poro de la resina deben ser adecuados para la absorción de compuestos orgánicos volátiles. Debido a que una resina polimérica tiene pocos grupos funcionales (a menudo el grupo funcional único dominante es el que confiere a la superficie sus características de adsorción), se puede determinar la afinidad (por ejemplo, el parámetro de solubilidad de Hansen) y predecir la capacidad adsorbente con base en parámetros termodinámicos. Diversos estudios han mostrado que se puede establecer una correlación entre la capacidad adsorbente de diversos solutos y resinas poliméricas, y los parámetros de solubilidad.

[0107] Otros materiales adecuados para absorber el vapor de estireno incluyen el carbón activado y los materiales celulósicos. Por ejemplo, las rebabas de algodón dadas a conocer en la patente de EE.UU. n.º 2002/0.151.622 como materiales celulósicos se pueden utilizar en la presente invención para absorber compuestos orgánicos volátiles.

[0108] El material de detección también comprende un marcador detectable que produce un cambio de color en presencia de estireno. Un ejemplo de marcador detectable es la 4-nitrobencilpiridina (NBP). El electrón no apareado de la NBP reacciona con el anillo de oxirano del óxido de estireno para dar un cromóforo azul.

[0109] Si se desea, el cambio de color del material detectable se puede determinar mediante un ensayo cuantitativo, como mediante espectrofotometría. El cambio de color también se puede determinar cualitativamente. A veces, el cambio de color resultaría evidente a un observador.

[0110] Si se desea, la actividad de la proteína candidata se puede comparar con un control. Un control puede ser una muestra paralela que comprende una cinamato descarboxilasa cuya actividad se ha caracterizado (por ejemplo, una cinamato descarboxilasa nativa, o una cierta cinamato descarboxilasa mutada que sirve como la secuencia base para mutaciones posteriores). Como otra posibilidad, un control puede ser un valor límite predeterminado, o un valor que está presente en una base de datos (por ejemplo, una tabla, una base de datos electrónica, una hoja de cálculo, etc.).

[0111] Como se muestra en la Figura 13, el ensayo de cribado se puede llevar a cabo en más de un ciclo, y se puede combinar con modelado por ordenador para el diseño racional y para la mejora de la actividad de las cinamato descarboxilasas.

[0112] En determinados ejemplos, el cribado de las cinamato descarboxilasas mutadas que pueden transformar el ácido *trans*-cinámico en estireno, o transformar el ácido cumárico en 4-hidroxiestireno, se puede realizar mediante el método descrito en la presente memoria. En determinados ejemplos, la proteína candidata que se está cribando puede ser una proteína de fusión que comprende una cinamato descarboxilasa mutada.

4. Producción recombinante y purificación de la cinamato descarboxilasa

[0113] La tecnología objeto de esta invención también proporciona un método de producción recombinante y purificación de la cinamato descarboxilasa, como la FDC. En concreto, la tecnología objeto de esta invención proporciona un método para aislar una cinamato descarboxilasa producida recombinantemente, que comprende; (i) proporcionar un hospedador bacteriano que comprende un ácido nucleico que codifica una cinamato descarboxilasa operativamente enlazada a una secuencia promotora; (ii) cultivar el hospedador bacteriano en un medio de cultivo para expresar la cinamato descarboxilasa en la célula hospedadora, en el que la célula hospedadora se cultiva a una temperatura de aproximadamente 10 °C a aproximadamente 25 °C; y (iii) aislar la cinamato descarboxilasa de la célula hospedadora, donde el aislamiento se lleva a cabo en un ambiente anaeróbico.

[0114] En algunos ejemplos, cultivar la célula hospedadora bacteriana (por ejemplo, *E. coli*) a una temperatura más baja puede promover significativamente el plegamiento correcto de la cinamato descarboxilasa producida recombinantemente, como la FDC1. Disminuir la temperatura de cultivo también facilita la conversión de una proteína agregada o plegada incorrectamente a una forma funcionalmente soluble. En determinados ejemplos, el cultivo celular crece a una temperatura de aproximadamente 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 o 25 °C.

[0115] En el método de producción recombinante de la cinamato descarboxilasa, como la FDC1, también se pueden utilizar agentes reductores, como tris(2-carboxietil)fosfina (TCEP) o β-mercaptoetanol. Por lo tanto, para aislar la cinamato descarboxilasa se utilizan una o más soluciones amortiguadoras que pueden comprender un agente reductor, como TCEP a una concentración de aproximadamente 5 mM, 10 mM, 15 mM, 20 mM, 25 mM, 30 mM, 35 mM, 40 mM, 45 mM, 50 mM o 100 mM; o β-mercaptoetanol a una concentración de aproximadamente 5 mM, 10 mM, 15 mM, 20 mM, 25 mM, 30 mM, 35 mM, 40 mM, 45 mM, 50 mM o 100 mM; y una combinación de los mismos.

[0116] Las muestras y los extractos celulares que contienen la cinamato descarboxilasa recombinante (como la FDC1) mantenidos en un ambiente anaeróbico también presentan actividad de cinamato descarboxilasa.

[0117] La cinamato descarboxilasa producida recombinantemente (como la FDC1) mantiene la actividad de cinamato descarboxilasa en presencia de iones metálicos. Entre los metales iónicos adecuados se incluyen, por

ejemplo, los iones de calcio, zinc, magnesio, hierro o manganeso, o una combinación de los mismos. El ión metálico puede estar presente en una concentración de aproximadamente 0,1 mM a aproximadamente 100 mM, como aproximadamente 0,1 mM, aproximadamente 0,2 mM, aproximadamente 0,5 mM, aproximadamente 1 mM, aproximadamente 2 mM, aproximadamente 5 mM, aproximadamente 10 mM, aproximadamente 15 mM, aproximadamente 20 mM, aproximadamente 25 mM, aproximadamente 30 mM, aproximadamente 40 mM, aproximadamente 50 mM, aproximadamente 60 mM, aproximadamente 70 mM, aproximadamente 80 mM, aproximadamente 90 mM, aproximadamente 100 mM; de aproximadamente 0,1 mM a aproximadamente 50 mM, de aproximadamente 0,1 mM a aproximadamente 25 mM, de aproximadamente 0,1 mM a aproximadamente 20 mM, de aproximadamente 0,1 mM a aproximadamente 10 mM, de aproximadamente 1 mM a aproximadamente 50 mM, de aproximadamente 1 mM a aproximadamente 25 mM, de aproximadamente 1 mM a aproximadamente 20 mM, de aproximadamente 1 mM a aproximadamente 15 mM, o de aproximadamente 1 mM a aproximadamente 10 mM. Preferentemente, el ión metálico es un ión de un metal alcalinotérreo o un ión de un metal de transición.

[0118] La cinamato descarboxilasa purificada descrita en la presente memoria es adecuada para la cristalización.

5. Preparación de la forma cristalina de la cinamato descarboxilasa

[0119] A fin de adquirir más conocimientos sobre la estructura de la cinamato descarboxilasa, se desarrolló un método para cristalizar esta proteína. Tradicionalmente, para la cristalización se utilizan los métodos de difusión de vapor de la gota colgante y la gota sentada. Ambos métodos requieren un sistema cerrado, es decir, el sistema se debe sellar para aislarlo del exterior utilizando un recipiente hermético o aplicando grasa de alto vacío entre las superficies de vidrio.

[0120] La tecnología objeto de esta invención proporciona un método para cristalizar una cinamato descarboxilasa, donde el método comprende (a) proporcionar una solución de cinamato descarboxilasa a una concentración de aproximadamente 1 mg/ml a aproximadamente 50 mg/ml; (b) mezclar la solución de cinamato descarboxilasa con una solución de depósito en una relación de volúmenes de aproximadamente 1:10 a aproximadamente 10:1; y (c) mantener la mezcla del paso (b) a una temperatura adecuada para la formación de cristales de cinamato descarboxilasa.

[0121] Preferentemente, la concentración de proteína en el paso (a) es de aproximadamente 2 mg/ml a aproximadamente 20 mg/ml, y más preferentemente de aproximadamente 5 mg/ml a aproximadamente 10 mg/ml. La relación de volúmenes de la solución de cinamato descarboxilasa y la solución de depósito es preferentemente de aproximadamente 1:5 a aproximadamente 5:1, y más preferentemente de aproximadamente 1:2 a aproximadamente 2:1. Normalmente, se mezclan de 1 a 2 μ l de la solución de cinamato descarboxilasa (concentración de proteína de aproximadamente 5,0 a 10,0 mg/ml) con 1 a 2 μ l de una solución de depósito en un recipiente. La mezcla se pone boca abajo en un pocillo, que contiene la solución de depósito (por lo general aproximadamente 500 μ l) y el sistema se incuba a una temperatura de aproximadamente 1 °C a aproximadamente 20 °C, preferentemente de aproximadamente 4 °C a aproximadamente 12 °C. Inicialmente, la gota de solución de proteína contiene una concentración de precipitante insuficiente para la cristalización, pero a medida que al agua se evapora de la gota y se transfiere al depósito, la concentración de precipitante aumenta hasta el nivel necesario para la cristalización. Puesto que el sistema está en equilibrio, estas condiciones se mantienen hasta que la cristalización ha finalizado.

[0122] La cristalización de la cinamato descarboxilasa se puede mejorar si la proteína forma un complejo con una molécula pequeña. Por ejemplo, la molécula pequeña se puede añadir a la solución de la proteína en el paso (a) o a la mezcla de proteína/depósito en el paso (b) para permitir la formación de un complejo proteína-molécula pequeña. Normalmente, la concentración final de la molécula pequeña en la mezcla del paso (b) es de aproximadamente 0,01% (p/v) a aproximadamente 1% (p/v), preferentemente de aproximadamente 0,05% (p/v) a aproximadamente 0,5% (p/v), y más preferentemente de aproximadamente 0,08% (p/v) a aproximadamente 0,2% (p/v).

[0123] Entre las moléculas pequeñas adecuadas se incluyen los sustratos de la cinamato descarboxilasa (por ejemplo, el ácido *trans*-cinámico) y sus análogos, como el ácido 3-hidroxicinámico, el ácido ferúlico, el ácido 2-metilcinámico, el ácido 4-hidroxicinámico, el ácido 3,4-dimetoxicinámico, el ácido 2,5-dimetoxicinámico y combinaciones de los mismos.

[0124] El método de cristalización se puede mejorar aún más mediante la inclusión de uno o más aditivos en la solución, como $MnCl_2$ a una concentración de aproximadamente 0,001 M a aproximadamente 0,1 M, y preferentemente de aproximadamente 0,005 M a aproximadamente 0,02 M; polivinilpirrolidona K15 a una concentración de aproximadamente un 0,1% (p/v) a aproximadamente un 2,5% (p/v), y preferentemente de aproximadamente un 0,25% (p/v) a aproximadamente un 1% (p/v); sulfobetaína no detergente 201 (NDSB-201, $C_8H_{11}NO_3S$) a una concentración de aproximadamente 0,02 M a aproximadamente 1 M, y preferentemente de aproximadamente 0,1 M a aproximadamente 0,3 M; clorhidrato de benzamida a una concentración de aproximadamente un 0,2% (p/v) a aproximadamente un 10% (p/v), y preferentemente de aproximadamente un 1% (p/v) a aproximadamente un 3% (p/v); con todas las concentraciones determinadas en la mezcla de proteína-depósito del paso (b).

E. PROTEÍNAS DE FUSIÓN Y COMPLEJOS PROTEICOS

[0125] En otro aspecto, la tecnología objeto de esta invención proporciona una proteína de fusión que comprende: (i) un primer dominio que comprende una fenilalanina amonio liasa, y (ii) un segundo dominio que comprende una cinamato descarboxilasa. Como otra posibilidad, una fenilalanina amonio liasa y una cinamato descarboxilasa pueden formar un complejo proteico por medio de una interacción o interacciones no covalentes.

[0126] La proteína de fusión o el complejo proteico descritos en la presente memoria se benefician del fenómeno de la «canalización de sustratos». La canalización del sustrato se refiere a un fenómeno por el cual los sustratos se suministran eficientemente de una enzima a otra sin equilibrarse con otras poblaciones de los mismos sustratos. La proteína de fusión o el complejo proteico pueden canalizar los intermedios entre enzimas secuenciales, y controlar el flujo de sustratos a ramas competidoras de la ruta. En efecto, esto crea poblaciones locales de metabolitos a elevadas concentraciones con respecto a las encontradas en otras regiones de la célula.

[0127] Cualquiera de las fenilalanina amonio liasas o cinamato descarboxilasas descritas en la presente memoria se pueden utilizar para producir una proteína de fusión. En determinadas realizaciones, la proteína de fusión también comprende un conector que enlaza covalentemente el primer dominio (PAL) y el segundo dominio (cinamato descarboxilasa). El conector preferentemente comprende uno o más residuos aminoácidos (por ejemplo, un conector aminoácido o un conector peptídico). Los residuos aminoácidos del conector pueden comprender *L*-aminoácido(s), *D*-aminoácido(s), análogos de aminoácidos o una combinación de los mismos. Otros posibles conectores incluyen, por ejemplo, un enlace covalente, un alquilo de C₁-C₆, un cicloalquilo como un ciclopentilo o un ciclohexilo, un cicloalqueno, un arilo o un heteroarilo. Un conector también puede comprender una combinación de uno o más aminoácido(s) con otra parte conectora (como un alquilo de C₁-C₆, un cicloalquilo-(C₅, C₆), un arilo o un heteroarilo).

[0128] En una realización determinada, el conector es un conector peptídico. El péptido conector puede tener una longitud de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35 o 40 aminoácidos. Preferentemente, la longitud del conector es de 2 a 15 aminoácidos.

[0129] En determinadas realizaciones, el conector es un conector de glicina/serina, es decir, un péptido conector que consiste fundamentalmente en glicina y serina. En un ejemplo de realización, el conector comprende GS o GSG. En otro ejemplo de realización, el conector comprende el motivo Gly-Ser-Gly (GSG), como GGSG (SEC. N.º ID: 39), (GS)×3 (SEC. N.º ID: 40), (GSG)×2 (SEC. N.º ID: 41), SGGSGGSGG (SEC. N.º ID: 42), GSGGGSGGGSG (SEC. N.º ID: 43), (GGGS)×3 (SEC. N.º ID: 44), como se describe a continuación en el Cuadro 1.

Cuadro 1 Conectores de glicina/serina

Conector	Secuencia aminoácida
Conector de GS	GS
Conector de GSG	GSG
SEC. N.º ID: 39	GGSG
SEC. N.º ID: 40	GSGSGS
SEC. N.º ID: 41	GGSGGGSG
SEC. N.º ID: 42	SGGSGGSGG
SEC. N.º ID: 43	GSGGGSGGGSG
SEC. N.º ID: 44	GGGSGGGSGGGGS

[0130] Las proteínas de fusión descritas en la presente memoria se pueden producir mediante técnicas bien conocidas en la técnica. Por ejemplo, cuando el conector es un conector peptídico, la proteína de fusión se puede producir mediante tecnología de ADN recombinante convencional. Se pueden acoplar otros tipos de conectores mediante, por ejemplo, técnicas convencionales de conjugación. Véase, por ejemplo, Hermanson *et al.*, *Bioconjugate Techniques*, 2.ª Ed., 2008; Academic Press.

[0131] La proteína de fusión puede comprender múltiples unidades de fenilalanina amonio liasa y cinamato descarboxilasa. En determinadas realizaciones, la proteína de fusión puede ser PAL-FDC, PAL-FDC-PAL-FDC o PAL-FDC-FDC-PAL, etc. Por ejemplo, el FDC puede estar en el extremo 5' del ADN que codifica la proteína de fusión, y el PAL puede estar en el extremo 3' del ADN que codifica la proteína de fusión. En otro ejemplo, el PAL puede estar en el extremo 5' del ADN que codifica la proteína de fusión, y el FDC puede estar en el extremo 3' del ADN que codifica la proteína de fusión. Además, la unidad de cinamato descarboxilasa de la proteína de fusión puede ser nativa (WT)

o una forma mutada, como cualquiera de las proteínas FDC mutadas descritas más arriba. Por ejemplo, la proteína de fusión puede ser PAL-FDC(WT) o PAL-FDC(K190E).

[0132] Como otra posibilidad o de manera adicional, la fenilalanina amonio liasa y la cinamato descarboxilasa pueden formar un complejo proteico por medio de una interacción o interacciones no covalentes. Un «complejo proteico» se refiere a una asociación de más de una proteína. Las proteínas del complejo se pueden asociar mediante, por ejemplo, asociación funcional, estereoquímica, conformacional, bioquímica o electrostática. Se pretende que el término abarque asociaciones de cualquier número de proteínas.

[0133] Como otra posibilidad, o de manera adicional, la fenilalanina amonio liasa y la cinamato descarboxilasa se pueden modificar para incluir una «parte que interactúa». Por ejemplo, una enzima puede comprender un anticuerpo, y la otra puede comprender un antígeno afín; o una enzima puede comprender un ligando, y la otra puede comprender un receptor afín; una enzima puede comprender biotina, y la otra puede comprender avidina, etc. Las partes que interactúan pueden interaccionar entre sí, de manera que las dos enzimas se aproximan la una a la otra. Se puede utilizar cualquier par de partes que interactúan.

F. CÉLULAS HOSPEDADORAS Y CULTIVOS CELULARES

[0134] En los métodos descritos en la presente memoria se utilizan células hospedadoras para producir estireno. Una célula hospedadora puede proceder de una bacteria, hongo (por ejemplo, levadura), protista (por ejemplo, algas), planta, insecto, anfibio, pez, reptil, ave, mamífero (incluidas personas), o puede ser una célula de hibridoma. Las células hospedadoras pueden ser células sin modificar o líneas celulares, o líneas celulares modificadas genéticamente (por ejemplo, para facilitar la producción de estireno). En algunas realizaciones, la célula hospedadora es una línea celular que se ha modificado para permitir el crecimiento en las condiciones deseadas, como a una temperatura más baja.

[0135] Como se describe en la presente memoria, las células hospedadoras adecuadas que expresan fenilalanina amonio liasa o cinamato descarboxilasa se pueden cultivar, a veces a gran escala (es decir, aproximadamente 1 litro, aproximadamente 10 litros, al menos 100 litros, etc.), para producir cantidades útiles desde el punto de vista comercial de estireno u otros compuestos obtenidos posteriormente a partir del estireno (compuestos que se producen por medio del tratamiento adicional del estireno en estos microorganismos a través de rutas enzimáticas o biológicas).

[0136] Los métodos descritos en la presente memoria se pueden aplicar a cualquier tamaño de biorreactor y/o matraz de cultivo celular. Por ejemplo, los métodos se pueden aplicar en biorreactores o cultivos celulares de 1 L, 10 L, 30 L, 50 L, 100 L, 150 L, 200 L, 300 L, 500 L, 1000 L, 2000 L, 3000 L, 4000 L, 5000 L, 10.000 L o de mayor capacidad.

[0137] El pH del medio líquido se puede mantener constante, es decir, regulado durante el período de cultivo, o no. Los cultivos pueden crecer en discontinuo, en semicontinuo o en continuo. Los nutrientes se pueden proporcionar al inicio de la fermentación, o se pueden proporcionar en semicontinuo o en continuo.

[0138] Entre los componentes del medio se incluyen, por ejemplo, la solución amortiguadora, el contenido en aminoácidos, el contenido en vitaminas, el contenido en sales, el contenido en minerales, el contenido en suero, el contenido en la fuente de carbono, el contenido en lípidos, el contenido en ácidos nucleicos, el contenido en hormonas, el contenido en oligoelementos, el contenido en amoníaco, el contenido en cofactores, el contenido en indicadores, el contenido en moléculas pequeñas, el contenido en hidrolizados y el contenido en moduladores de la enzima.

[0139] El medio de cultivo a utilizar debe cumplir adecuadamente los requisitos de las cepas en cuestión. Se pueden encontrar descripciones de medios de cultivo para diversos microorganismos en el libro de texto *Manual of Methods for General Bacteriology* de la Sociedad norteamericana de bacteriología (American Society for Bacteriology) (Washington D.C., EE.UU., 1981). Estos medios que se pueden utilizar conforme a la tecnología objeto de esta invención normalmente comprenden una o más fuentes de carbono, fuentes de nitrógeno, sales inorgánicas, vitaminas y/o oligoelementos.

[0140] Las fuentes de carbono para uso en el medio de cultivo de las células hospedadoras comprenden glúcidos, como mono-, di- o polisacáridos. Algunos ejemplos de fuentes de carbono son la glucosa, la fructosa, la manosa, la galactosa, la ribosa, la sorbosa, la ribulosa, la lactosa, la maltosa, la sucrosa, la rafinosa, el almidón o la celulosa. Los glúcidos también se pueden añadir al medio a través de compuestos complejos como la melaza u otros subproductos procedentes del refinado del azúcar. La adición de mezclas de diversas fuentes de carbono también puede presentar ventajas. Otras posibles fuentes de carbono son los aceites y las grasas como, por ejemplo, el aceite de soja, el aceite de girasol, el aceite de cacahuete y/o la grasa de coco, los ácidos grasos como, por ejemplo, el ácido palmítico, el ácido esteárico y/o el ácido linoleico, los alcoholes y/o polialcoholes como, por ejemplo, el glicerol, el metanol y/o el etanol, y/o los ácidos orgánicos como, por ejemplo, el ácido acético y/o el ácido láctico.

[0141] Las fuentes de nitrógeno suelen ser compuestos nitrogenados orgánicos o inorgánicos, o materiales que comprenden estos compuestos. Ejemplos de fuentes de nitrógeno comprenden el amoníaco en forma líquida o gas, o sales de amonio como el sulfato de amonio, el cloruro de amonio, el fosfato de amonio, el carbonato de amonio o el nitrato de amonio, nitratos, la urea, aminoácidos o fuentes nitrogenadas complejas como el licor de maíz macerado, la harina de soja, la proteína de soja, el extracto de levadura, el extracto de carne, etc. Las fuentes de nitrógeno se pueden utilizar de manera individual o como una mezcla.

[0142] En una realización, las sales inorgánicas presentes en el medio de cultivo comprenden aproximadamente una de las sales de cloruro, fósforo y sulfato con calcio, magnesio, sodio, cobalto, molibdeno, potasio, manganeso, zinc, cobre o hierro.

[0143] Como fuentes de azufre para la producción de derivados del estireno que contienen azufre se pueden utilizar compuestos inorgánicos que contienen azufre, como, por ejemplo, sulfatos, sulfitos, ditionitas, tetratiónatos, tiosulfatos, sulfuros o compuestos orgánicos de azufre como mercaptanos y tioles.

[0144] Como fuente de fósforo se puede utilizar el ácido fosfórico, el dihidrógenofosfato de potasio o el hidrógenofosfato de dipotasio, o las sales correspondientes de sodio.

[0145] Otros componentes que se pueden añadir al medio de cultivo para mantener los iones metálicos en solución comprenden dihidroxifenoles como el catecol o el protocatecolato, y ácidos orgánicos como el ácido cítrico.

[0146] El medio de cultivo puede comprender uno o más iones metálicos, como el calcio, el zinc, el magnesio, el hierro, el manganeso y una combinación de los mismos. El ión metálico puede estar presente en una concentración de aproximadamente 0,1 mM a aproximadamente 100 mM, como aproximadamente 0,1 mM, aproximadamente 0,2 mM, aproximadamente 0,5 mM, aproximadamente 1 mM, aproximadamente 2 mM, aproximadamente 5 mM, aproximadamente 10 mM, aproximadamente 15 mM, aproximadamente 20 mM, aproximadamente 25 mM, aproximadamente 30 mM, aproximadamente 40 mM, aproximadamente 50 mM, aproximadamente 60 mM, aproximadamente 70 mM, aproximadamente 80 mM, aproximadamente 90 mM, aproximadamente 100 mM; de aproximadamente 0,1 mM a aproximadamente 50 mM, de aproximadamente 0,1 mM a aproximadamente 25 mM, de aproximadamente 0,1 mM a aproximadamente 20 mM, de aproximadamente 0,1 mM a aproximadamente 10 mM, de aproximadamente 1 mM a aproximadamente 50 mM, de aproximadamente 1 mM a aproximadamente 25 mM, de aproximadamente 1 mM a aproximadamente 20 mM, de aproximadamente 1 mM a aproximadamente 15 mM, o de aproximadamente 1 mM a aproximadamente 10 mM. Preferentemente, el ión metálico es un ión de un metal alcalinotérreo o un ión de un metal de transición.

[0147] Los medios de fermentación utilizados conforme a la tecnología objeto de esta invención para cultivar las células hospedadoras de la tecnología objeto de esta invención normalmente también comprenden otros factores de crecimiento como vitaminas o potenciadores del crecimiento, que incluyen, por ejemplo, la biotina, la riboflavina, la tiamina, el ácido fólico, el ácido nicotínico, el pantotenato y la piridoxina. Los factores de crecimiento y las sales a menudo proceden de componentes complejos del medio, como el extracto de levadura, la melaza, el licor de maíz macerado, etc. Además, es posible añadir precursores adecuados al medio de cultivo. La composición exacta de los compuestos del medio depende en gran medida del experimento en particular y se decide de manera individual para cada caso específico. Se puede encontrar información sobre el cultivo del medio en el libro de texto *Applied Microbial Physiology, A Practical Approach* (Editores: P. M. Rhodes, P. F. Stanbury, IRL Press (1997) págs. 53-73). El medio de crecimiento también se puede obtener de proveedores comerciales, como, por ejemplo, Standard 1 (Merck), BHI (caldo de infusión cerebro-corazón, DIFCO), etc.

[0148] Todos los componentes del medio se esterilizan, o bien con calor (20 min a 1,5 bar y 121° C), o bien mediante esterilización por filtración. Los componentes se pueden esterilizar juntos o, si fuera necesario, por separado. Todos los componentes del medio pueden estar presentes al inicio del cultivo o se pueden añadir en continuo o en discontinuo, según se prefiera.

[0149] La temperatura del cultivo celular se seleccionará principalmente con base en el intervalo de temperaturas a las que el cultivo celular permanece viable, a las que se produce un elevado nivel de polipéptido, a las que el plegamiento incorrecto y/o la agregación del polipéptido son menores, a las que el polipéptido muestra una modificación postraducciona (por ejemplo, glicosilación, fosforilación, etc.) más extensa o más conveniente, o cualquier combinación de estos u otros factores que el profesional considere importantes. En general, la mayoría de células hospedadoras crecen bien y pueden producir elevados niveles de proteína o polipéptido dentro de un intervalo de aproximadamente 15 °C a 45 °C. En determinadas realizaciones, el cultivo celular crece a una temperatura de aproximadamente 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44 o 45 °C en uno o más puntos temporales durante el proceso de cultivo celular. Los expertos en la materia tendrán la habilidad de seleccionar la temperatura o temperaturas adecuadas para el crecimiento de las células, dependiendo de las necesidades de las células y de las necesidades de producción.

[0150] Además, el cultivo se puede someter a uno o más cambios de temperatura durante el transcurso del cultivo. Cuando se modifica la temperatura del cultivo, el cambio de temperatura puede ser relativamente gradual. Por ejemplo, completar el cambio de temperatura puede llevar varias horas o días. Como otra posibilidad, el cambio de temperatura puede ser relativamente abrupto. La temperatura se puede aumentar o disminuir constantemente durante el proceso de cultivo. De manera adicional, o como otra posibilidad, la temperatura se puede aumentar o disminuir en cantidades discretas en distintos puntos temporales durante el proceso de cultivo. La(s) temperatura(s) o intervalo(s) de temperatura(s) posteriores pueden ser más bajos o más elevados que la(s) temperatura(s) o intervalo(s) de temperatura(s) previos o iniciales. Un experto en la materia sabrá que la tecnología objeto de esta invención abarca múltiples cambios de temperatura. Por ejemplo, la temperatura se puede modificar una vez (a una temperatura o intervalo de temperaturas más alto o más bajo), las células se mantienen a esta temperatura o intervalo de temperaturas durante un determinado intervalo de tiempo, tras el cual la temperatura se puede modificar de nuevo a otra temperatura o intervalo de temperaturas, que puede ser más alto o más bajo que la temperatura o intervalo de temperaturas de la temperatura o intervalo de temperaturas anterior. La temperatura del cultivo después de cada cambio discreto puede ser constante o se puede mantener dentro de un intervalo de temperaturas determinado.

[0151] Al igual que con la temperatura o intervalo de temperaturas iniciales, la temperatura o intervalo de temperaturas del cultivo celular después del cambio o cambios de temperatura por lo general se selecciona principalmente con base en la temperatura o temperaturas a las que el cultivo celular permanece viable, el intervalo en el que se produce una cantidad elevada de proteína. Por ejemplo, un cultivo celular bacteriano puede crecer a 37 °C tras la siembra para favorecer la proliferación celular; una vez las células alcanzan la densidad deseada, se induce la expresión de una proteína recombinante, y la temperatura se modifica a 25 °C para reducir el plegamiento incorrecto o la agregación de la proteína producida recombinantemente.

[0152] También se contemplan las condiciones anaeróbicas. Entre los ejemplos de hospedadores bacterianos anaeróbicos se incluyen, por ejemplo, *Bacteroids*, *Fusobacterium*, *Clostridium*, *Propionibacterium*, *Lactobacillus*, *Peptococcus*, *Peptostreptococcus* y *Veillonella*.

[0153] Si se desea, el estireno producido en las células hospedadoras puede además sufrir una reacción enzimática y transformarse en un derivado posterior del estireno, como el tolueno, el xileno, el poliestireno, el ABS, el caucho de estireno-butadieno (SBR), el látex de estireno-butadieno, el estireno-isopreno-estireno (SIS), el estireno-etileno/butileno-estireno (S-EB-S), el estireno-divinilbenceno (S-DVB), la resina de estireno-acrilonitrilo (SAN), poliésteres insaturados, etc. Por lo tanto, en algunas realizaciones, las células hospedadoras de la tecnología objeto de esta invención pueden transformar un determinado porcentaje del estireno producido en un derivado posterior del estireno. Por ejemplo, la célula hospedadora puede transformar aproximadamente un 5% del estireno en un derivado posterior, o de aproximadamente un 5% a aproximadamente un 15% del estireno en un derivado posterior, o de aproximadamente un 10% a aproximadamente un 25% del estireno en un derivado posterior, o de aproximadamente un 20% a aproximadamente un 35% del estireno en un derivado posterior, o de aproximadamente un 30% a aproximadamente un 45% del estireno en un derivado posterior, o de aproximadamente un 40% a aproximadamente un 55% del estireno en un derivado posterior, o de aproximadamente un 50% a aproximadamente un 65% del estireno en un derivado posterior, o de aproximadamente un 60% a aproximadamente un 75% del estireno en un derivado posterior, o de aproximadamente un 70% a aproximadamente un 85% del estireno en un derivado posterior, o de aproximadamente un 80% a aproximadamente un 95% del estireno en un derivado posterior.

1. Hospedadores microbianos

[0154] Entre los microorganismos útiles para la producción de estireno se pueden incluir las bacterias, como las bacterias entéricas (por ejemplo, *Escherichia* y *Salmonella*), *Bacillus*, *Acinetobacter*, actinomicetos (como *Streptomyces*), *Corynebacterium*, metanotrofos (como *Metilosinus*), *Methylomonas*, *Rhodococcus*, *Pseudomona*, cianobacterias (como *Rhodobacterand*, *Synechocistis*), *Klebsiella*, *Pantoea*, *Corynebacterium*, *Clostridium*, etc.; levaduras, como *Saccharomyces*, *Zygosaccharomyces*, *Kluyveromyces*, *Candida*, *Hansenula*, *Debaryomyces*, *Mucor*, *Pichia* y *Torulopsis*; hongos filamentosos como *Aspergillus*, *Arthrobotrys*; algas, etc.

[0155] Aunque cualesquiera de los microorganismos mencionados arriba serían útiles para la producción de estireno, se prefieren las cepas mutadas de bacterias que sobreproducen fenilalanina. La cepa «que sobreproduce fenilalanina» es una cepa microbiana mutada que produce una cantidad más elevada de fenilalanina con respecto a la de la cepa nativa que no contiene la mutación. La fenilalanina se encuentra presente de manera natural en los microorganismos. Sin embargo, para una síntesis óptima del estireno, la célula hospedadora preferentemente sobreproduce fenilalanina de manera que la cantidad de sustrato no limita la producción de estireno por la célula hospedadora. En la técnica se conocen métodos para aumentar la síntesis de aminoácidos aromáticos en un microorganismo.

[0156] Un ejemplo concreto de una *E. coli* que sobreproduce fenilalanina es la cepa NST74 de *E. coli* (patente de EE.UU. n.º 4.681.852). Entre otras cepas que sobreproducen fenilalanina se incluyen, por ejemplo, *Corynebacterium glutamicum* (Ikeda, M. and Katsumata, R. *Metabolic engineering to produce tyrosine or phenylalanine in a tryptophan-producing Corynebacterium glutamicum* strain, Appl. Environ. Microbial. (1992), 58(3), págs. 781-785); *Microbacterium*

ammoniafilum (ATCC, 10155); *Corynebacterium lillium* (NRRL-B-2243), *Brevibacterium divaricatum* (NRRL-B-2311); *Arthrobacter citreus* (ATCC, 11624). Otras cepas adecuadas que sobreproducen fenilalanina se pueden encontrar, por ejemplo, en Maiti *et al*, *Supra and Metabolic Engineering for Microbial Production of Aromatic Amino Acids and Derived Compounds*, J. Bongaertes *et al*, *Metabolic Engineering*, vol. 3, 289-300), 2001.

[0157] La célula hospedadora también se puede seleccionar para que tenga una mayor resistencia frente a un análogo tóxico de un aminoácido aromático (por ejemplo, la fenilalanina). Por ejemplo, se pueden seleccionar microorganismos mutados para que sean resistentes frente a análogos tóxicos (*m*-fluoro-) de la fenilalanina. Estas formas mutadas insensibles a menudo producen elevadas cantidades de fenilalanina y tirosina (patente británica n.º 1.071.935; patente de EE.UU. n.º 3.709.785).

[0158] También es posible obtener una célula hospedadora recombinante que produzca más fenilalanina mediante la sobreexpresión de uno o más genes clave en la biosíntesis de la fenilalanina (Ikeda 2003. *Amino acid production processes*, págs. 1-35 en T. Scheper (Ed.), *Advances in Biochemical Engineering/Biotechnology*, vol. 79. Springer-Verlag, Berlin Heidelberg).

[0159] Se pueden utilizar metodologías de ADN recombinante convencionales para obtener un ácido nucleico que codifique una proteína descrita en la presente memoria (por ejemplo, fenilalanina amonio liasa, cinamato descarboxilasa o una proteína de fusión), incorporar el ácido nucleico en un vector de expresión, e introducir el vector en una célula hospedadora, como las descritas en Sambrook, *et al.* (eds), *Molecular Cloning; A Laboratory Manual*, 3.ª edición, Cold Spring Harbor, (2001); Ausubel, F. M. *et al.* (eds.) *Current Protocols in Molecular Biology*, John Wiley & Sons (1995). Un ácido nucleico que codifica una proteína se puede insertar en un vector o vectores de expresión de manera que los ácidos nucleicos estén operativamente enlazados a las secuencias de control transcripcional y traduccional (como una secuencia promotora, una secuencia de terminación de la transcripción, etc.). El vector de expresión y las secuencias de control de expresión por lo general se eligen para que sean compatibles con la célula hospedadora utilizada para la expresión.

[0160] La expresión de proteínas en un hospedador microbiano descrito en la presente memoria se puede mejorar aún más mediante la optimización de codones. Por ejemplo, modificar un codón menos común con un codón más común puede afectar a la semivida del ARNm o alterar su estructura al introducir una estructura secundaria que interfiere con la traducción del mensaje. Se puede optimizar la totalidad o una parte de una región codificante. En algunos casos, la modulación deseada de la expresión se logra mediante la optimización de prácticamente todo el gen. En otros casos, la modulación deseada se logra mediante la optimización de una parte de la secuencia completa del gen, y no de la secuencia completa.

[0161] El uso de codones de cualquier secuencia codificante se puede ajustar para lograr una propiedad deseada, por ejemplo elevados niveles de expresión en un tipo determinado de célula. El punto de partida para dicha optimización puede ser una secuencia codificante con un 100% de codones comunes, o una secuencia codificante que contiene una mezcla de codones comunes y no comunes.

[0162] Se pueden generar dos o más secuencias candidatas que difieran en el uso de codones, y se pueden analizar para determinar si poseen las propiedades deseadas. Las secuencias candidatas se pueden evaluar por ordenador para investigar la presencia de elementos reguladores, como silenciadores o potenciadores, y para investigar la presencia de regiones de la secuencia codificante que se podrían transformar en dichos elementos reguladores mediante una alteración del uso de codones. Entre los criterios adicionales se puede incluir el enriquecimiento en determinados nucleótidos, por ejemplo, A, C, G o U, la preferencia codónica por un aminoácido en particular, o la presencia o ausencia de una estructura secundaria o terciaria de ARNm en particular. El ajuste a la secuencia candidata se puede realizar con base en varios de dichos criterios.

[0163] En determinadas realizaciones, la secuencia del ácido nucleico para la que se han optimizado los codones puede expresar su proteína a un nivel de aproximadamente un 110%, de aproximadamente un 150%, de aproximadamente un 200%, de aproximadamente un 250%, de aproximadamente un 300%, de aproximadamente un 350%, de aproximadamente un 400%, de aproximadamente un 450% o de aproximadamente un 500% con respecto al expresado por la secuencia del ácido nucleico para la que no se han optimizado los codones.

[0164] Además del ácido nucleico que codifica la proteína, el vector de expresión también puede contener secuencias reguladoras que controlan la expresión de la proteína en una célula hospedadora, como promotores, potenciadores u otros elementos de control de la expresión que controlan la transcripción o la traducción del ácido nucleico o ácidos nucleicos. Dichas secuencias reguladoras se conocen en la técnica (véase, por ejemplo, Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press (1990)). Los expertos en la técnica apreciarán que el diseño del vector de expresión, incluida la selección de secuencias reguladoras puede depender de factores tales como la elección de la célula hospedadora a transformar, el nivel de expresión deseado de proteína, etc.

[0165] Además de las secuencias que codifican la proteína y las secuencias reguladoras, los vectores de expresión recombinantes de la tecnología objeto de la presente invención pueden contener otras secuencias, como secuencias que regulan la replicación del vector en las células hospedadoras (por ejemplo, orígenes de la replicación) y genes marcadores seleccionables.

[0166] El vector o vectores de expresión que codifican la proteína se pueden introducir por transformación o por transfección en una célula hospedadora mediante técnicas convencionales, como electroporación, precipitación con fosfato de calcio o transfección con DEAE-dextrano.

[0167] Para la producción comercial de estireno, se pueden aplicar diversos métodos de fermentación. Por ejemplo, la producción a gran escala se puede realizar mediante fermentación en discontinuo o en continuo.

[0168] Una fermentación en discontinuo clásica consiste en un sistema cerrado en el que la composición del medio se fija al inicio de la fermentación y no se somete a alteraciones artificiales durante la fermentación. Por lo tanto, al inicio de la fermentación, el medio se inocula con el microorganismo o microorganismos deseados, y se deja que la fermentación transcurra sin añadir nada al sistema. Sin embargo, normalmente la concentración de la fuente de carbono en una fermentación «en discontinuo» es limitada y con frecuencia se intenta controlar factores como el pH y la concentración de oxígeno. En los sistemas en discontinuo, la composición en metabolitos y biomasa del sistema cambia constantemente hasta el momento en que la fermentación se para. En los cultivos en discontinuo, las células pasan de una fase de latencia estática a una fase de elevado crecimiento logarítmico y, finalmente, a una fase estacionaria en la que la tasa de crecimiento es inferior o nula. Si no se tratan, las células en fase estacionaria con el tiempo mueren. Las células en fase logarítmica por lo general son responsables de la mayor parte de la producción de producto final o intermedio.

[0169] Una variación del sistema convencional en discontinuo es el sistema en semicontinuo. Los procesos de fermentación en semicontinuo también son adecuados para la tecnología objeto de esta invención y comprenden un sistema en discontinuo convencional, con la excepción de que el sustrato se añade en incrementos a medida que la fermentación avanza. Los sistemas en semicontinuo son útiles cuando la represión por catabolito puede inhibir el metabolismo de las células y cuando se desea tener cantidades limitadas de sustrato en el medio. En los sistemas en semicontinuo es difícil medir la concentración real de sustrato y, por lo tanto, la concentración se estima sobre la base de los cambios en factores mensurables, como el pH, el oxígeno disuelto y la presión parcial de los gases residuales, como el CO₂. Las fermentaciones en discontinuo y en semicontinuo son habituales y bien conocidas en la técnica, y se pueden encontrar ejemplos en Brock, T. D.; *Biotechnology: A Textbook of Industrial Microbiology*, 2.^a ed.; Sinauer Associates: Sunderland, Mass., 1989; o Deshpande, M. V. *Appl. Biochem. Biotechnol.* 36:227, (1992).

[0170] La producción comercial de estireno también se puede llevar a cabo mediante fermentación en continuo. La fermentación en continuo consiste en un sistema abierto donde un medio de fermentación definido se añade continuamente a un biorreactor y la misma cantidad de medio acondicionado se extrae simultáneamente para su tratamiento. La fermentación en continuo por lo general mantiene los cultivos a una densidad elevada y constante, y las células se encuentran principalmente en su fase de crecimiento logarítmico.

[0171] La fermentación en continuo permite la modulación de cualquier número de factores que afecten al crecimiento celular o a la concentración de producto final. Por ejemplo, un método mantendrá un nutriente limitante, como la fuente de carbono o de nitrógeno, a una velocidad fija y permitirá que se moderen todos los otros parámetros. En otros sistemas, varios factores que afectan al crecimiento se pueden alterar de manera continua, mientras que la concentración celular, determinada a partir de la turbidez del medio, se mantiene constante. Los sistemas en continuo tratan de mantener condiciones de crecimiento en equilibrio y, por lo tanto, la pérdida de células debida a la eliminación del medio se debe compensar con la tasa de crecimiento celular en la fermentación. Los métodos para modular los nutrientes y los factores de crecimiento en los procesos de fermentación en continuo, así como las técnicas para aumentar al máximo la velocidad de formación del producto, son bien conocidos en la técnica de la microbiología industrial y en el Brock (*supra*) se describen diversos métodos.

2. Hospedadores vegetales

[0172] Las células vegetales también se pueden utilizar como hospedadores para la producción de estireno. Los hospedadores vegetales preferidos son cualquier variedad que promueva un nivel elevado de expresión de fenilalanina amonio liasa y/o cinamato descarboxilasa. Entre las plantas verdes adecuadas se incluyen, por ejemplo, la soja, la colza (*Brassica napus*, *B. campestris*), el girasol (*Helianthus annuus*), el algodón (*Gossypium hirsutum*), el maíz, el tabaco (*Nicotiana tabacum*), la alfalfa (*Medicago sativa*), el trigo (*Triticum sp*), la cebada (*Hordeum vulgare*), la avena (*Avena sativa*, L), el sorgo (*Sorghum bicolor*), el arroz (*Oryza sativa*), *Arabidopsis*, las verduras crucíferas (broccoli, coliflor, col, chirivía, etc.), el melón, la zanahoria, el apio, el perejil, el tomate, la patata, la fresa, el cacahuete, la uva, el cultivo de semillas de césped, la remolacha azucarera, la caña de azúcar, la judía, el guisante, el centeno, la linaza, los árboles de madera dura, los árboles de madera blanda y los pastos forrajeros. Entre las especies de alga se incluyen, por ejemplo, hospedadores importantes desde un punto de vista comercial como la espirulina, la *Haematococcus* y la *Dunaliella*. Entre las plantas adecuadas también se incluyen el biocombustible, la biomasa y las

plantas para cultivos de bioenergía. Entre los ejemplos de plantas se incluyen la *Arabidopsis thaliana*, el arroz (*Oryza sativa*), el mijo perenne (*Panicum virgatum*), la *Brachypodium spp.*, la *Brassica spp.* y la *Crambe abyssinica*.

[0173] En algunas realizaciones, la célula vegetal es una célula vegetal de *Arabidopsis*, una célula vegetal de tabaco, una célula vegetal de petunia o una célula de un cultivo de semillas oleaginosas (incluidas, por ejemplo, una célula vegetal de soja, una célula vegetal de canola, una célula vegetal de colza, una célula vegetal de palma, una célula vegetal de girasol, una célula vegetal de algodón, una célula vegetal de maíz, una célula vegetal de cacahuete, una célula vegetal de linaza, una célula vegetal de sésamo, etc.).

[0174] Las células hospedadoras adecuadas se pueden modificar mediante técnicas de ingeniería genética para que expresen fenilalanina amonio liasa y cinamato descarboxilasa. Por ejemplo, el ácido nucleico que codifica la fenilalanina amonio liasa o la cinamato descarboxilasa se pueden enlazar operativamente a los promotores que puedan dirigir la expresión de una proteína en los tejidos deseados y en el estadio deseado de desarrollo. Se puede utilizar cualquier promotor y/o terminador adecuados que puedan inducir la expresión de una región codificante. Algunos ejemplos de promotores y terminadores adecuados incluyen a aquellos de la nopalina sintasa (nos), octopina sintasa (ocs) y los genes del virus del mosaico de la coliflor (CaMV).

[0175] Un tipo de promotor de origen vegetal eficiente que se puede utilizar es un promotor de origen vegetal de alto nivel. Entre los promotores de alto nivel origen vegetal que se pueden utilizar en la tecnología objeto de esta invención se incluyen el promotor de la subunidad pequeña (ss) de la ribulosa-1,5-bisfosfato carboxilasa, por ejemplo, de la soja (Berry-Lowe *et al.*, J. Molecular and App. Gen., 1:483-498) (1982)), y el promotor de la proteína de unión a la clorofila a/b. Se sabe que en las células vegetales estos dos promotores son inducidos por la luz (véase, por ejemplo, *Genetic Engineering of Plants, an Agricultural Perspective*, A. Cashmore, Plenum, N.Y. (1983), págs. 29-38); Coruzzi, G. *et al.*, The Journal of Biological Chemistry, 258: 1399 (1983), y Dunsmuir, P. *et al.*, Journal of Molecular and Applied Genetics, 2:285 (1983)).

[0176] Se pueden utilizar metodologías convencionales de ADN recombinante para obtener un ácido nucleico que codifique una proteína descrita en la presente memoria, incorporar el ácido nucleico en un vector de expresión e introducir el vector en una célula hospedadora. La elección de vector depende del método utilizado para transformar los hospedadores vegetales. Un experto en la materia conocerá bien los elementos genéticos que deben estar presentes en el vector plasmídico a fin de transformar, seleccionar y propagar con éxito las células hospedadoras que contienen el vector. Un experto en la materia también entenderá que distintos acontecimientos independientes de transformación producirán distintos niveles y pautas de expresión (Jones *et al.*, EMBO J. 4:2411-2418) (1985); De Almeida *et al.*, Mol. Gen. Genetics 218:78-86) (1989)), y, por lo tanto, que se deben cribar múltiples acontecimientos a fin de obtener líneas que muestren los patrones y el nivel de expresión deseados. Dicho cribado se puede llevar a cabo mediante transferencia de Southern de ADN (Southern, J. Mol. Biol. 98, 503, (1975)). Northern analysis of ARNm expression (Kroczeck, J. Chromatogr. Biomed. Appl., 618 (12): 133-145) (1993)), análisis de la expresión de proteínas mediante transferencia de Western o análisis fenotípicos.

[0177] La expresión de proteínas en un hospedador vegetal descrita en la presente memoria se puede mejorar más mediante la optimización de codones, como se describe más arriba. En determinadas realizaciones, la secuencia del ácido nucleico para la que se han optimizado los codones puede expresar su proteína a un nivel de aproximadamente un 110%, de aproximadamente un 150%, de aproximadamente un 200%, de aproximadamente un 250%, de aproximadamente un 300%, de aproximadamente un 350%, de aproximadamente un 400%, de aproximadamente un 450% o de aproximadamente un 500% del expresado por la secuencia del ácido nucleico para la que no se han optimizado los codones.

[0178] La tecnología objeto de esta invención también proporciona células hospedadoras transgénicas o células hospedadoras que se han transformado con uno o más ácidos nucleicos dados a conocer en la presente memoria. La molécula de ácido nucleico puede estar integrada de manera estable en el genoma de la célula, o la molécula de ácido nucleico también puede estar presente como una molécula extracromosómica. Dicha molécula extracromosómica puede ser autorreplicante. Se entiende que las células, tejidos o sujetos transformados abarcan no solo el producto final de un proceso de transformación, sino también la progenie transgénica de los mismos.

[0179] La introducción de un ácido nucleico de la tecnología objeto de esta invención en una célula vegetal se puede llevar a cabo mediante diversos métodos conocidos por los expertos en la materia, incluidos, entre otros, la inserción de una secuencia del ácido nucleico de interés en un plásmido Ri de *Agrobacterium rhizogenes* o un plásmido Ti de *Agrobacterium tumefaciens*, la microinyección, la electroporación o la precipitación directa. A modo de ejemplo, en algunas realizaciones, la expresión transitoria de una secuencia del ácido nucleico o gen de interés se puede llevar a cabo mediante métodos de agroinfiltración. A este respecto, una suspensión de *Agrobacterium tumefaciens* que contiene la secuencia del ácido nucleico o el gen de interés puede crecer en un cultivo y a continuación, se puede inyectar a una planta colocando la punta de una jeringa contra el reverso de una hoja mientras se aplica una contrapresión suave en el otro lado de la hoja. A continuación, la solución de *Agrobacterium* se inyecta en los espacios con aire dentro de la hoja a través de los estomas. Una vez dentro de la hoja, *Agrobacterium* transforma una parte de las células vegetales con el gen de interés, en donde dicho gen a continuación se expresa de forma transitoria.

[0180] Como otro ejemplo, un vector o ácido nucleico de interés se pueden introducir por transformación en una célula vegetal mediante técnicas de bombardeo con partículas. A este respecto, una suspensión de embriones vegetales puede crecer en un cultivo líquido y a continuación, se puede bombardear con plásmidos o ácidos nucleicos que están acoplados a partículas de oro, donde las partículas de oro unidas al plásmido o ácido nucleico de interés se pueden propulsar a través de las membranas de los tejidos vegetales, como el tejido embrionario. Tras el bombardeo, los embriones transformados se pueden seleccionar utilizando un antibiótico adecuado para generar nuevos cultivos embrionarios en suspensión transformados y propagados mediante clonación.

[0181] Para obtener más información sobre métodos para transformar y producir células vegetales transgénicas, véanse las patentes de EE.UU. n.º 4.459.355; 4.536.475; 5.464.763; 5.177.010; 5.187.073; 4.945.050; 5.036.006; 5.100.792; 5.371.014; 5.478.744; 5.179.022; 5.565.346; 5.484.956; 5.508.468; 5.538.877; 5.554.798; 5.489.520; 5.510.318; 5.204.253; 5.405.765; las patentes europeas n.º 267.159; 604.662; 672.752; 442.174; 486.233; 486.234; 539.563; 674.725; y las patentes mundiales n.º WO 91/02071 y WO 95/06128.

3. Reducción de la toxicidad del estireno

[0182] En otro aspecto, la tecnología objeto de esta invención proporciona una célula hospedadora que comprende: (a) una fenilalanina amonio liasa expresada recombinantemente; (b) una cinamato descarboxilasa expresada recombinantemente; y (c) un transportador unido a la membrana expresado recombinantemente.

[0183] Un problema significativo que limita la bioproducción de estireno es la toxicidad del estireno para las células hospedadoras. Se sabe que la acumulación de compuestos aromáticos hidrofóbicos dentro de la membrana citoplasmática perturba su integridad. Para reducir la toxicidad del estireno y aumentar su producción, se puede introducir un transportador unido a la membrana (por ejemplo, una bomba de salida) en la célula hospedadora para eliminar el disolvente orgánico de la célula. Por lo tanto, la célula hospedadora muestra un fenotipo tolerante hacia los disolventes hidrofóbicos.

[0184] En una realización, el transportador unido a la membrana puede ser un transportador ABC (transportador con dominio de unión a ATP), que son proteínas transmembranarias que aprovechan la energía procedente de la hidrólisis de la adenosina trifosfato (ATP) para llevar a cabo determinados procesos biológicos, incluidos la translocación de diversos sustratos a través de las membranas y procesos no relacionados con el transporte, como la traducción de ARN y la reparación de ADN. Transportan una amplia variedad de sustratos a través de las membranas extracelulares e intracelulares, incluidos productos del metabolismo, lípidos y esteroides, y fármacos. Las proteínas se clasifican en transportadores ABC con base en la secuencia y la organización de su dominio o dominios de unión a ATP (ABC, por sus siglas en inglés).

[0185] En la presente memoria se proporcionan células hospedadoras que expresan un transportador ABC que permite que las células hospedadoras secreten estireno en el medio de cultivo. En concreto, el transportador ABC es una bomba de resistencia a disolventes. Se ha demostrado que las bombas de resistencia a disolventes que confieren resistencia o tolerancia frente a los disolventes orgánicos poseen una especificidad muy amplia, y absorben como sustratos compuestos orgánicos, como compuestos aromáticos, alcoholes, alcanos, etc., atendiendo a sus características físicoquímicas (por ejemplo, acumulación en la membrana bacteriana) (Kieboom *et al.* 1998. J. Biol. Chem. 273:85-91). Los compuestos aromáticos también se reparten eficazmente en la membrana celular, donde actúan como sustratos para las bombas de resistencia a disolventes.

[0186] En una realización de la tecnología objeto de esta invención, una célula hospedadora comprende un miembro de la familia de bombas de protones de tipo resistencia/nodulación/división celular (RND). Las bombas de salida tipo RND pertenecen a las bombas de multiresistencia a fármacos (MDR). Estas tienen una especificidad frente a sustratos extremadamente amplia y protegen a las células bacterianas de las acciones de los antibióticos en ambos lados de la membrana citoplasmática. Se ha demostrado que los miembros de esta familia están implicados en la expulsión de los antibióticos, metales y oligosacáridos que participan en la transducción de señales de la nodulación. Las bombas de salida tipo RND suelen funcionar como ensamblajes de tres componentes que se extienden por las membranas exterior y citoplasmática, así como por el espacio periplásmico, de las bacterias gram-negativas. Ejemplos de bombas de salida tipo RND adecuadas para su uso en un método de la tecnología objeto de esta invención se pueden encontrar en Tseng, T. T., Gratwick, K. S., Kollman, J., Park., D., Nies, D. H., Goffeau, A., & Saier Jr., M. H. (1999), J. Mol. Microbial. Biotechnol. 1: 107-125.

[0187] En una realización, la célula hospedadora comprende la bomba *srpABC* de resistencia a disolventes de *P. putida* S12 (Isken *et al.* 1996 J. Bacterial. 178:6056; Kieboom *et al.* 1998. J. Biol. Chem. 273:85-91). Las secuencias aminoácidas deducidas de las proteínas codificadas por los genes *srpABC* tienen una amplia homología con las de la familia RND de bombas de salida. Consta de tres componentes proteicos que juntos se extienden por las membranas internas y externas de las bacterias gram-negativas: un transportador unido a la membrana interna (análogos de la *SrpB*), un canal en la membrana externa (análogos de la *SrpC*) y una proteína conectora periplásmica (análogos de la *SrpA*). En Kieboom *et al.* (1998 J. Biol. Chem. 273:85-91) se muestran dendrogramas con la relación

filogenética de la SrpA, la SrpB y la SrpC con otras proteínas que participan en la multirresistencia a fármacos. Las proteínas codificadas por *srpABC* son las que muestran más homología con las de la bomba de multirresistencia a fármacos codificada por *mexAB/oprM* y encontrada en *Pseudomonas aeruginosa*. La SrpA, la SrpB y la SrpC son idénticas en un 57,8, 64,4 y 58,5% a la MexA, la MexB y la OprM, respectivamente. En una realización, una célula hospedadora comprende una bomba de salida que consiste en un transportador unido a la membrana interna, un canal en la membrana externa y una proteína conectora periplásmica que pertenece a la familia RND de bombas de salida, donde las proteínas muestran una identidad de secuencia (homología) de aproximadamente un 50%, de aproximadamente un 55%, de aproximadamente un 60%, de aproximadamente un 70%, de aproximadamente un 80%, de aproximadamente un 85%, de aproximadamente un 95%, de aproximadamente un 98%, de aproximadamente un 99% o incluso del 100% con respecto a las proteínas SrpA, SrpB o SrpC de *P. putida* S12. Se puede utilizar cualquier equivalente funcional de bombas de salida de disolventes conocidas que puedan utilizar un compuesto aromático como sustrato.

[0188] En una realización, la célula hospedadora puede transformar el sustrato de carbono fermentable en un aminoácido aromático (por ejemplo, fenilalanina), que a continuación se transforma en estireno. Una vez producido, el estireno se transporta activamente fuera de la célula hospedadora mediante una bomba de salida, preferentemente mediante un miembro de la familia de bombas de protones de tipo resistencia/nodulación/división celular (RND), como la *srpABC*.

[0189] La bacteria *P. putida* S12 también se ha modificado mediante técnicas de ingeniería genética como sistema de tolerancia a disolventes para la biosíntesis del *p*-hidroxibenzoato y del *p*-hidroxiestireno (Verhoef *et al.*, 2007, *Bioproduction of p-hydroxybenzoate from renewable feedstock by solvent-tolerant Pseudomonas putida* S12. *Journal of Biotechnology* 132, 49-56; Verhoef *et al.*, 2009, *Bioproduction of p-hydroxystyrene from glucose by the solvent-tolerant bacterium Pseudomonas putida* S12 in a two-phase water-decanol fermentation. *Applied and Environmental Microbiology* 75, 931-936). La cepa obtenida mediante técnicas de ingeniería genética se puede utilizar para la bioproducción de estireno. Se pueden obtener otros huéspedes mediante técnicas de ingeniería genética de una manera similar para aumentar la tolerancia a los compuestos orgánicos.

G. RECOLECCIÓN DEL ESTIRENO A PARTIR DEL CULTIVO CELULAR

[0190] El estireno se puede recolectar del cultivo celular mediante métodos convencionales. Por ejemplo, las células hospedadoras se pueden eliminar mediante filtración o centrifugación. Las fases acuosa y oleosa se pueden separar por centrifugación o por cromatografía. Se pueden utilizar técnicas adicionales de adsorción, destilación y microfiltración para purificar más el estireno. Un esquema general de purificación conlleva la eliminación de los inhibidores de la polimerización con una solución acuosa alcalina (normalmente NaOH al 5-10%), el lavado en agua destilada, el secado y la destilación fraccionada a presión reducida, la microfiltración de estireno monomérico en fase gas, y una combinación de los mismos.

[0191] Por lo tanto, la tecnología objeto de esta invención proporciona una ruta biológica de bajo coste para la producción de estireno, el cual es de utilidad en un abanico de materiales comerciales, incluido el poliestireno, el ABS, el caucho de estireno-butadieno (SBR), el látex de estireno-butadieno, el SIS (estireno-isopreno-estireno), el S-EB-S (estireno-etileno/butileno-estireno), el estireno-divinilbenceno (S-DVB), la resina de estireno-acrilonitrilo (SAN) y los poliésteres insaturados. Estos materiales se utilizan en el caucho, el plástico, el aislamiento, la fibra de vidrio, tuberías, automóviles y piezas de buque, recipientes para alimentos y refuerzo de moqueta.

H. BIOSÍNTESIS DEL ESTIRENO

[0192] Normalmente, el estireno se puede producir mediante la incubación de un sustrato descrito en la presente memoria (como la glucosa, la fenilalanina o el ácido *trans*-cinámico) con la célula hospedadora que comprende cualquier cinamato descarboxilasa descrita en la presente memoria o cualquier proteína de fusión descrita en la presente memoria. La concentración de glucosa durante la incubación normalmente es de aproximadamente un 0,02% a aproximadamente un 3%, preferentemente de aproximadamente un 0,05% a aproximadamente un 1,5%, y más preferentemente de aproximadamente un 0,1% a aproximadamente un 1%. La concentración de ácido *trans*-cinámico durante la incubación normalmente es de aproximadamente un 0,02% a aproximadamente un 0,5%, y preferentemente de aproximadamente un 0,05% a aproximadamente un 0,2%. La concentración de fenilalanina durante la incubación normalmente es de aproximadamente un 0,02% a aproximadamente un 0,5%, y preferentemente de aproximadamente un 0,05% a aproximadamente un 0,2%. En una realización, tras la adición del sustrato, las células hospedadoras se cultivan continuamente durante aproximadamente 10 a aproximadamente 72 horas, y preferentemente durante aproximadamente 15 a aproximadamente 36 horas, a una temperatura de aproximadamente 16°C a aproximadamente 37 °C, y preferentemente de 22 °C a aproximadamente 30 °C.

[0193] Con base en lo anterior, la tecnología objeto de esta invención proporciona un método para producir estireno, donde el método comprende (a) poner en contacto una célula hospedadora con un sustrato de carbono fermentable, donde la célula hospedadora comprende una proteína de fusión como las descritas más arriba; y (b) cultivar la célula hospedadora en un medio de cultivo durante un tiempo suficiente para producir estireno.

[0194] La tecnología objeto de esta invención también proporciona un método para producir estireno, donde el método comprende (a) poner en contacto una célula hospedadora con un sustrato de carbono fermentable, donde la célula hospedadora comprende (i) una fenilalanina amonio liasa; y (ii) una cinamato descarboxilasa mutada como las descritas más arriba; y (b) cultivar la célula hospedadora en un medio de cultivo durante un tiempo suficiente para producir estireno.

[0195] La tecnología objeto de esta invención también proporciona una célula hospedadora que comprende: (a) una fenilalanina amonio liasa expresada recombinantemente; (b) una cinamato descarboxilasa expresada recombinantemente; y (c) un transportador unido a la membrana expresado recombinantemente. Por lo tanto, la tecnología objeto de esta invención también proporciona un método para producir estireno, donde el método comprende (a) poner en contacto la célula hospedadora de arriba con un sustrato de carbono fermentable; y (b) cultivar la célula hospedadora en un medio de cultivo durante un tiempo suficiente para producir estireno.

[0196] En los métodos convencionales, la biosíntesis del estireno se realiza en recipientes cerrados para evitar que el estireno se evapore del recipiente. Este sistema cerrado no puede mantener las condiciones opcionales anaeróbicas (condiciones óptimas) para que un sistema biológico produzca estireno. Por lo tanto, con el método convencional, el estireno producido se acumula en el recipiente cerrado y pasado un tiempo, la biosíntesis del estireno se para debido al efecto de toxicidad del estireno en el sistema de biosíntesis (por ejemplo, una célula hospedadora). Sorprendentemente, la inclusión de un material absorbente para eliminar el vapor de estireno del proceso de biosíntesis no solo permite una detección fiable del estireno producido (por ejemplo, por el método de detección de la actividad de las FDC mutadas descrito más arriba), sino que también mejora el rendimiento de la biosíntesis del estireno.

[0197] Por lo tanto, la tecnología objeto de esta invención también proporciona un método para producir estireno, donde el método comprende: (a) poner en contacto una célula hospedadora con un sustrato de carbono fermentable, donde la célula hospedadora comprende (i) una fenilalanina amonio liasa; y (ii) una cinamato descarboxilasa; y (b) cultivar la célula hospedadora en un medio de cultivo durante un tiempo suficiente para producir estireno, donde un material absorbente absorbe el vapor de estireno producido.

[0198] En este ejemplo, la fenilalanina amonio liasa y la cinamato descarboxilasa pueden incluir tanto proteínas nativas como proteínas mutadas. Por ejemplo, la cinamato descarboxilasa también puede ser una proteína FDC mutada como la descrita más arriba. Además, la fenilalanina amonio liasa y la cinamato descarboxilasa pueden estar presentes como proteínas independientes o como una proteína de fusión, como las proteínas de fusión PAL-FDC(WT) o PAL-FDC(K190E) descritas más arriba.

[0199] En un ejemplo, una resina polimérica que puede absorber moléculas orgánicas absorbe el vapor de estireno, mientras que se permite que el aire circule libremente por el sistema de biosíntesis (por ejemplo, una célula hospedadora). Cualquier material absorbente utilizado en el proceso descrito más arriba de cribado de alto rendimiento según la actividad de cinamato descarboxilasa también se puede emplear en el proceso de biosíntesis del estireno con dispositivos similares (Figura 19). Por ejemplo, la columna STRATA-X® (que contiene resina polimérica de fase inversa) se puede utilizar para absorber el vapor de estireno producido a partir de una célula hospedadora (por lo tanto, reduciendo la toxicidad para la célula hospedadora), a la vez que se permite que el oxígeno pase a través de la resina para mantener condiciones anaeróbicas opcionales para el crecimiento celular. Normalmente, este método permite producir estireno en una concentración superior a 1 g/L.

[0200] Todos los métodos descritos más arriba para la producción de estireno, incluidos el método en el que se utiliza una proteína de fusión, el método en el que se utiliza una cinamato descarboxilasa mutada, el método en el que se utiliza un transportador unido a la membrana y el método en el que se elimina el vapor de estireno con un material absorbente, también se pueden utilizar para la producción de 4-hidroxiestireno. En un ejemplo, se puede utilizar tirosina o ácido cumárico, o ambos, como sustratos en los métodos descritos más arriba para producir 4-hidroxiestireno.

[0201] La capacidad para construir una proteína de fusión y para analizar rápidamente la actividad de cinamato descarboxilasa en un gran número de muestras permite la detección simultánea de la actividad de las unidades de PAL y cinamato descarboxilasa en la proteína de fusión, de manera que se pueda lograr un mayor rendimiento global de estireno. Este proceso resulta ventajoso con respecto al proceso de determinar la actividad de PAL y cinamato descarboxilasa por separado antes de construir una proteína de fusión.

[0202] Por lo tanto, la tecnología objeto de la presente invención también proporciona un método para detectar simultáneamente la actividad de fenilalanina amonio liasa y cinamato descarboxilasa, donde el método comprende: (a) proporcionar una proteína de fusión que comprende: (i) un primer dominio que comprende una fenilalanina amonio liasa, y (ii) un segundo dominio que comprende una cinamato descarboxilasa; (b) mezclar la proteína de fusión con un sustrato en condiciones que permiten que la proteína de fusión transforme el sustrato en un producto; y (c) detectar la cantidad restante de sustrato, o la cantidad de producto, o ambas.

[0203] En un ejemplo, la tecnología objeto de la presente invención también proporciona un método para detectar simultáneamente la actividad de fenilalanina amonio liasa y cinamato descarboxilasa, donde el método comprende: (a) proporcionar una proteína de fusión que comprende: (i) un primer dominio que comprende una fenilalanina amonio liasa, y (ii) un segundo dominio que comprende una cinamato descarboxilasa; (b) proporcionar un sustrato seleccionado del grupo que consiste en fenilalanina, ácido *trans*-cinámico, tirosina, ácido cumárico y combinaciones de los mismos; (c) incubar la proteína de fusión y el sustrato en condiciones que permitan que la proteína de fusión transforme el sustrato en un producto seleccionado del grupo que consiste en estireno, 4-hidroxiestireno y una combinación de los mismos; y d) detectar la cantidad restante de sustrato, o la cantidad de producto, o ambas.

[0204] La proteína de fusión, como la PAL-FDC o la PAL-FDC(K190E), se puede preparar utilizando cualesquiera de las PAL y cinamato descarboxilasas descritas más arriba. Normalmente, se mezcla un sustrato (como fenilalanina o ácido *trans*-cinámico) con la proteína de fusión, y se detecta simultáneamente la actividad de PAL y de cinamato descarboxilasa en la proteína de fusión mediante determinaciones de la cantidad restante de sustrato, o la cantidad de producto, o ambas. Por ejemplo, cuando la fenilalanina como sustrato se mezcla con la proteína de fusión, se puede determinar la cantidad de ácido *trans*-cinámico y estireno como productos. En este ejemplo, la concentración de ácido cinámico refleja tanto su producción a partir de fenilalanina (por la PAL) como su conversión en estireno (por la cinamato descarboxilasas), y la concentración de estireno refleja la producción global de estireno por la proteína de fusión a partir de cualquier sustrato (Figura 18). De manera similar, cuando el ácido *trans*-cinámico se mezcla con la proteína de fusión para su conversión en estireno, se puede determinar la cantidad restante de ácido *trans*-cinámico como sustrato y la cantidad de estireno producido.

[0205] En un ejemplo, la detección simultánea de la actividad de fenilalanina amonio liasa y cinamato descarboxilasa se puede realizar en condiciones similares a las utilizadas en los ensayos de cribado de alto rendimiento descritos en la presente memoria para detectar la actividad de cinamato descarboxilasa. En un ejemplo de realización, la detección del estireno comprende exponer la mezcla de la proteína de fusión y el sustrato a una resina polimérica que absorbe el vapor de estireno. Entre las resinas adecuadas se incluyen las resinas hidrofóbicas, como las resinas sorbentes de C18, C8, fenilo, SDB-L y combinaciones de las mismas.

EJEMPLOS

[0206] La invención se define en las reivindicaciones, y cualesquiera de los siguientes ejemplos que queden fuera del alcance de las reivindicaciones no forman parte de la invención, pero se proporcionan para fines de comparación.

[0207] La tecnología objeto de esta invención también se define en los siguientes ejemplos. Se debe entender que estos ejemplos, aunque indican realizaciones preferidas de la tecnología objeto de esta invención, se facilitan únicamente a título ilustrativo. Con base en la explicación anterior y estos ejemplos, un experto en la materia puede determinar las características fundamentales de la tecnología objeto de esta invención y, sin alejarse del alcance de esta, puede realizar diversos cambios y modificaciones a la tecnología objeto de esta invención para adaptarla a diversos usos y condiciones.

EJEMPLO I. PRODUCCIÓN POR RECOMBINACIÓN, PURIFICACIÓN Y CARACTERIZACIÓN DE LA FDC1 DE LEVADURA.

1. Expresión de la FDC1 de levadura en *E. coli*

[0208] La descarboxilación del ácido *trans*-cinámico (tCA) por la ferulato descarboxilasa (FDC1) es el último paso en la ruta biosintética del estireno para producir estireno. Aunque en una publicación se afirmó que la expresión de la FDC1 por sí sola era insuficiente para la conversión del tCA en estireno (Jiang *et al.*, 2005, *Applied and Environmental Microbiology*, 71: 2962-2969; Clausen *et al.*, 1994, *Gene*, 142: 107-112), más recientemente, se publicó que la expresión de la FDC1 por sí sola era suficiente para la conversión *in vivo* del tCA en estireno (McKenna *et al.*, 2011, *Metabolic Engineering*, 13 (5): 544-554). McKenna especuló que otras proteínas, como la Ubix de *E. coli*, que se asocian *in vivo* a la FDC1, permiten la conversión del tCA en estireno. No existe ninguna publicación que muestre claramente que la FDC1 por sí sola, sin asociarse a ninguna otra proteína, sea suficiente para transformar el tCA en estireno.

[0209] Para examinar la actividad *in vitro* de la FDC1 purificada, se construyó un vector de expresión y la FDC1 se expresó en *E. coli* en diversas condiciones. El plásmido pDEST1 7-FDC1 se introdujo por transformación en células competentes de la cepa BL21 (DE3) de *E. coli*. El medio de caldo Terrific suplementado con ampicilina (100 µg/ml) se inoculó con un cultivo crecido durante toda la noche y diluido 1000 veces. Las bacterias se cultivaron a 37 °C hasta alcanzar una OD₆₀₀ de 0,8 a 1,0, en cuyo momento se añadió isopropil β-D-1-tiogalactopiranosido (IPTG) a una concentración final de 0,2 mM, y el cultivo se continuó durante 3 horas a la misma temperatura. Se utilizaron tres soluciones amortiguadoras distintas para comprobar la expresión de la FDC1. La solución amortiguadora A contenía solución amortiguadora de fosfato de potasio 25 mM a pH 7,5, cloruro de sodio 500 mM, imidazol 10 mM, βME 20 mM, PMSF 0,5 mM, MgCl₂ 10 mM, ADNasa 2 µg/ml, ARNasa 2 µg/ml y lisozima a 4 µg/ml. La solución amortiguadora B

contenía solución amortiguadora de fosfato de potasio 25 mM a pH 7,0, DTT 1 mM, tiosulfato de sodio 50 mM, glicerol al 20%, NaCl 500 mM, MgCl₂ 10 mM, imidazol 20 mM, ADNasa 2 µg/ml, ARNasa 2 µg/ml y lisozima a 4 µg/ml. La solución amortiguadora C estaba compuesta por la solución amortiguadora A que contenía tiosulfato de sodio 50 mM y glicerol al 20%.

[0210] A 37 °C, la FDC1 recombinante no se expresó como una forma soluble en *E. coli*, y formó cuerpos de inclusión, como se muestra en la Figura 2A. Los protocolos convencionales para la expresión de proteínas recombinantes no se pueden aplicar a la FDC1.

[0211] Para producir la FDC1 funcional en *E. coli*, la expresión de la FDC1 se indujo a diversas temperaturas (37 °C, 30 °C, 25 °C, 18 °C y 16 °C) en presencia de IPTG a diversas concentraciones (0,2 mM, 0,5 mM). La concentración de IPTG no mostró un efecto significativo en la expresión; sin embargo, a temperaturas más bajas la expresión y la solubilidad de esta proteína mejoraron. La expresión óptima de la FDC1 se logró tras la inducción con IPTG 0,2 mM a 16 °C durante 16 horas (Figura 2B). Las células se recolectaron mediante centrifugación a 4500 rpm a 4 °C, se lavaron una vez con solución amortiguadora de IX PBS y, a continuación, se almacenaron a -80 °C.

[0212] La expresión de la FDC1 se estudió mediante SDS-PAGE (gel de acrilamida al 10% en placa con un grosor de 0,75 mm) utilizando el protocolo de Laemmli. Para la tinción de la banda de la proteína se utilizó azul brillante de Coomassie R-250.

[0213] Nuestros resultados indicaron que la expresión de la FDC1 a una temperatura más baja y a una concentración más baja de IPTG contribuyó al correcto plegamiento de la enzima, y a la conversión de formas agregadas o plegadas incorrectamente a una forma soluble y funcional.

2. Purificación de la FDC1 de levadura

[0214] No existen publicaciones sobre la purificación y caracterización de la FDC1 de levadura. Para estudiar la actividad y las propiedades de la FDC1, se purificó la FDC1 funcional que se produjo recombinantemente a partir de *E. coli*.

[0215] La purificación se llevó a cabo en un ambiente anaeróbico a 4 °C. Las células se suspendieron en solución amortiguadora A (compuesta por Tris-HCl 50 mM a pH 8,0 que contenía Na₂S₂O₃ 50 mM, TCEP 25 mM, NaCl 500 mM, PMSF 0,5 mM, β-mercaptoetanol 20 mM, glicerol al 20% e imidazol 10 mM) que contenía MgCl₂ 10 mM, Triton X-100 al 0,2%, ADNasa 2 µg/ml, ARNasa 2 µg/ml y lisozima a 4 µg/ml. Las células se rompieron mediante ultrasonificación y el sobrenadante se recuperó mediante centrifugación a 15.000 g durante 20 min a 4 °C. El sobrenadante resultante se filtró a través de un filtro de PES de 0,45 µm y a continuación, se depositó en una columna de afinidad de Ni²⁺-agarosa (GE Healthcare) equilibrada con la solución amortiguadora A. La columna se lavó bien con solución amortiguadora A hasta que todas las proteínas que se unen de manera no específica se eluyeron de la columna. La proteína se eluyó con solución amortiguadora A que contenía imidazol 250 mM. El contenido de proteína se determinó mediante espectrofotometría a 280 nm. La proteína estaba bastante pura (Figura 3).

[0216] Se purificó la FDC1 recombinante que podía transformar el tCA en estireno. Sin embargo, la proteína perdió su actividad muy rápidamente durante la purificación. La actividad de la proteína se puede mantener cuando la purificación se realiza en condiciones anaeróbicas, como al añadir cantidades más elevadas de agentes reductores (Na₂S₂O₃ 50 mM, TCEP 25 mM y β-mercaptoetanol 20 mM) a las soluciones amortiguadoras.

3. Ensayos para determinar la actividad de las FDC1 producidas recombinantemente

[0217] A fin de estudiar la función de la FDC1 recombinante *in vitro*, se investigaron los efectos de diversas condiciones (por ejemplo, soluciones amortiguadoras a distintos pH, diversas concentraciones de sustrato y diversos disolventes orgánicos para la extracción del producto) en la actividad enzimática de la FDC1.

[0218] Primeramente, se desarrolló un ensayo enzimático muy exacto y reproducible, con el que se podía determinar la actividad de la FDC1 a una concentración baja de sustrato.

[0219] La mezcla de reacción estándar para la descarboxilación consistía en solución amortiguadora de fosfato de potasio 25 mM (pH 6,5) que contenía ditiotreitol 5 mM, ácido *trans*-cinámico 1,4 mM y enzima (0,50 mg) en un volumen final de 1,0 ml. La reacción se inició mediante la adición de la enzima y se incubó a 30 °C durante 5 min. La reacción se paró mediante la adición de 24 µl de ácido acético glacial (17,4 N), tras lo cual se añadió 2-propanol en un volumen igual a la mezcla de reacción para solubilizar el producto. La cantidad de estireno producida se determinó mediante HPLC utilizando un UHPLC Dionex Ultimate3000 equipado con un cargador de muestras automático, un detector UV/Vis de diodos en serie y una columna de fase inversa Acclaim 120 C18 (2,1 x 150 mM, Dionex, EE.UU.). Las muestras se resolvieron a un caudal de 1 ml/min en ácido acético al 0,15% (A) con un gradiente de concentración en aumento de ácido acético al 0,15% (B): 0 a 4 min, 5%; 4 a 5 minutos, 5 a 40 %; 5 a 7 min, 40 a 45%; 7 a 8 min, 45 a

85%; 8 a 12 min, 85 a 95%; 12 a 14 minutos, 95 a 5%. La actividad específica se expresó en U (nmol de estireno)·mg⁻¹·min⁻¹.

4. Efecto del pH en la actividad de la FDC1

[0220] El pH de la solución amortiguadora es uno de los principales factores que pueden afectar a la reacción enzimática. Se llevaron a cabo reacciones respectivas en soluciones amortiguadoras a pH 6,0, 6,5, 7,0, 7,5, 8,0 y 8,5 para evaluar el pH óptimo para la actividad de la FDC1. Para los valores de pH de 6,0 a 7,5 se utilizó una solución amortiguadora de fosfato de potasio, y para los valores de pH de 8,0 y 8,5 se utilizó una solución amortiguadora de Tris-HCl. La mezcla de reacción estaba compuesta por ácido cinámico a 200 µM, DTT 5 mM, soluciones amortiguadoras a varios pH a 100 mM y 150 µl de extracto crudo de FDC1 en un volumen final de 1 ml. Las mezclas de reacción se incubaron durante 30 minutos a 30 °C. El estireno se extrajo con 100 µl de butanol. Para el control experimental negativo se utilizaron las mismas condiciones, con la excepción de que la mezcla de reacción no contenía sustrato. Se encontró que el pH óptimo para la descarboxilación en el extracto crudo de FDC1 era aproximadamente 6,5. Nuestros resultados muestran que la solución amortiguadora tiene un efecto significativo en la actividad de la FDC1, y la FDC1 mostró una actividad más elevada a un pH de 6,5 (Figura 4). Esta información es útil para la producción industrial de estireno.

5. Estabilidad de la FDC1 con el pH

[0221] La solución amortiguadora también afecta a la estabilidad de la FDC1 y a la velocidad de la reacción. Para estudiar la estabilidad de la FDC1, la proteína se incubó en soluciones amortiguadoras a distintos pH, de 5,0 a 11,0.

[0222] La proteína (extracto crudo, 586 µl) se añadió a 112 µl de soluciones amortiguadoras a 500 mM a un pH de 5, 6, 7, 8, 9, 10 y 11, y se incubó a 30 °C durante 30 minutos. Después del tratamiento en soluciones amortiguadoras a distintos pH, se añadió una alícuota de proteína (42 mg) a una solución amortiguadora de fosfato de potasio 100 mM (concentración final) a pH 6,5, DTT 5 mM y ácido cinámico 1,4 mM, y se enrasó a un volumen final de 2 ml. A continuación, la mezcla de reacción se incubó durante 8 minutos a 30 °C. El estireno se extrajo mediante la adición de 250 µl de butanol. Para el control experimental negativo se utilizaron las mismas condiciones, con la excepción de que la mezcla de reacción no contenía sustrato. Se encontró que el extracto crudo de FDC1 era estable en el intervalo de pH de 6 a 10 (Figura 5). Esta proteína es estable y activa en un amplio intervalo de pH y se puede utilizar para la producción industrial de estireno.

6. Efecto de la temperatura en la actividad de la FDC1

[0223] La temperatura puede afectar a la estabilidad de la FDC1 y a la velocidad de la reacción. El efecto de temperatura en la velocidad de la reacción se describe por la ecuación de Arrhenius. Como regla general, la velocidad de reacción de muchas reacciones se duplica o triplica por cada aumento de 10 °C en la temperatura, aunque el efecto de la temperatura puede ser mucho mayor o menor.

[0224] Para evaluar la temperatura óptima para la FDC1, se llevaron a cabo reacciones respectivas a temperaturas de 25 °C, 30 °C, 32 °C, 35 °C, 40 °C, 50 °C y 60 °C durante 30 minutos. La proteína (extracto crudo de FDC1, 13 mg) se añadió a mezclas de reacción que contenían fosfato de potasio 25 mM a pH 6,5, DTT 5 mM y ácido cinámico 1,4 mM en un volumen de 1 ml. Tras la reacción, se añadió 1 ml de propanol a las mezclas de reacción para disolver el estireno. Para el control experimental negativo se utilizaron las mismas condiciones, con la excepción de que la mezcla de reacción no contenía sustrato. La temperatura óptima para la actividad de la FDC1 fue de aproximadamente 50 °C (Figura 6).

[0225] De manera inesperada, la enzima mostró una actividad máxima a una temperatura más elevada que otras enzimas de levadura. Esta información es de utilidad para la producción industrial de estireno, ya que una cantidad considerable del coste asociado a la fermentación se debe al enfriamiento del sistema de fermentación. Dado que la enzima FDC es activa a una temperatura de reacción más elevada, el intervalo de temperaturas para la fermentación se puede controlar para aumentar el rendimiento y reducir el coste de enfriamiento.

7. Estabilidad de la FDC1 con la temperatura

[0226] Para estudiar la estabilidad de la FDC1 con la temperatura, se llevaron a cabo reacciones respectivas a temperaturas de 30 °C, 50 °C, 60 °C y 70 °C durante 30 minutos. Tras la incubación a diversas temperaturas, se añadió una alícuota de proteína (42 mg) a una solución amortiguadora de fosfato de potasio 25 mM a pH 6,5, DTT 5 mM y tCA 1,4 mM en un volumen final de 1 ml. Las mezclas de reacción se incubaron durante 8 minutos a 30 °C. El estireno se extrajo con 250 µl de butanol y se cuantificó mediante HPLC con una columna C18. Para el control experimental negativo se utilizaron las mismas condiciones, con la excepción de que la mezcla de reacción no contenía sustrato. La enzima se mantuvo estable a 50 °C (Figura 7).

[0227] La estabilidad del extracto crudo de FDC1 a 50 °C también se evaluó mediante la incubación por separado del extracto crudo a 50 °C durante 10, 30, 60 y 120 minutos. Tras la incubación de la proteína durante varios tiempos, se añadió una alícuota (42 mg) a solución amortiguadora de fosfato de potasio 25 mM a pH 6,5, DTT 5 mM y tCA 1,4 mM en un volumen final de 1 ml. A continuación, las mezclas de reacción se incubaron durante 5 minutos a 30°C. El estireno se extrajo tras la adición de 250 µl de butanol a las mezclas de reacción y se cuantificó mediante HPLC. Para el control experimental negativo se utilizaron las mismas condiciones, con la excepción de que la mezcla de reacción no contenía sustrato. La proteína se mantuvo estable a 50 °C durante 2 h (Figura 8). Este resultado demuestra que la proteína retuvo un 100% de su actividad incluso después de incubarla durante aproximadamente 2 horas. Esta información es de utilidad para la producción industrial de estireno, debido a la termoestabilidad de la FDC1.

8. Efecto de los cofactores en la actividad de la FDC1

[0228] Para examinar el efecto de los cofactores en la actividad enzimática de la FDC1, se llevaron a cabo reacciones con diversos cofactores. Se evaluaron los efectos de varios cofactores, incluidos la tiamina pirofosfato (TPP), la biotina y el piridoxal fosfato (PLP), en la actividad de la FDC1. Las mezclas de reacción que contenían solución amortiguadora de fosfato de potasio 25 mM a pH 6,5, DTT 5 mM, tCA 0,5 mM, cofactor a 0,5 mM y 0,5 mg de FDC1 purificada en un volumen de 1 ml se incubaron durante 30 minutos a 30 °C. El estireno se extrajo con 250 µl de butanol. Para el control experimental se utilizaron las mismas condiciones, con la excepción de que la mezcla de reacción no contenía cofactores. Para el control experimental negativo se utilizaron las mismas condiciones, con la excepción de que la mezcla de reacción no contenía sustrato.

[0229] Ninguno de los cofactores aumentó la actividad de la FDC1. De hecho, se observó que disminuían la actividad de la FDC1 (Figura 9). Este resultado demostró que esta enzima no necesita ninguno de los cofactores comúnmente conocidos para mostrar actividad.

9. Efecto de los iones metálicos en la actividad de la FDC1

[0230] Para estudiar el efecto de los iones metálicos en la actividad de la FDC1, se llevaron a cabo reacciones con diversos iones metálicos, incluidos ZnSO₄, FeCl₃, MnCl₂, MgSO₄, CaCl₂, MnSO₄ y FeSO₄. Las mezclas de reacción que contenían solución amortiguadora de fosfato de potasio 25 mM a pH 6,5, tCA 1,4 mM, DTT 5 mM, ión metálico a 10 mM y 0,5 mg de FDC1 purificada en un volumen final de 1 ml se incubaron durante 5 minutos a 30 °C. Tras la reacción, se añadió 1 ml de propanol a las mezclas de reacción para solubilizar el estireno. Para el control experimental se utilizaron las mismas condiciones, con la excepción de que la mezcla de reacción no contenía iones metálicos. Para el control experimental negativo se utilizaron las mismas condiciones, con la excepción de que la mezcla de reacción no contenía sustrato. Los iones Ca²⁺, Mg²⁺, Zn²⁺ y Fe³⁺ aumentaron la actividad de la FDC1 (Figura 10A).

[0231] Los efectos del Zn²⁺ y Fe³⁺ en la actividad de la FDC1 se investigaron más en profundidad utilizando EDTA para contrarrestar los efectos de los iones metálicos. Las mezclas de reacción que contenían solución amortiguadora de fosfato de potasio 25 mM a pH 6,5, tCA 1,4 mM, DTT 5 mM, ión metálico a 10 mM, EDTA 10 mM y 0,5 mg de FDC1 purificada en un volumen final de 1 ml se incubaron durante 5 minutos a 30°C. Tras la reacción, se añadió 1 ml de propanol a las mezclas de reacción para solubilizar el estireno. Para el control experimental se utilizaron las mismas condiciones, con la excepción de que la mezcla de reacción no contenía iones metálicos ni EDTA. Para el control experimental negativo se utilizaron las mismas condiciones, con la excepción de que la mezcla de reacción no contenía sustrato. Las reacciones que contenían el ión metálico y EDTA mostraron una actividad similar a la del control (Figura 10B), lo que indica que los iones Zn²⁺ y Fe³⁺ aumentan la actividad de la FDC1. Según los resultados presentados en las Figuras 10A y 10B, la actividad enzimática de la FDC1 aumentó en presencia de metales iónicos, y no se debió a un artefacto experimental. Para la producción industrial de estireno, esta información indica que la adición de determinados iones metálicos al medio de cultivo puede favorecer la biosíntesis del estireno.

10. Especificidad frente al sustrato

[0232] A fin de examinar si la FDC1 es altamente específica frente al tCA (o si muestra promiscuidad con otros sustratos), se analizó el tCA y sus sustratos análogos, como el ácido ferúlico, el ácido 2-metilcinámico, el ácido 2-hidroxicinámico, el ácido 3-hidroxicinámico, el ácido 4-hidroxicinámico, el ácido 3,4-dimetoxicinámico y el ácido 2,5-dimetoxicinámico. Las mezclas de reacción de amortiguador de fosfato de potasio 25 mM a pH 6,5 que contenían DTT 5 mM, sustrato 0,2 mM y 0,5 mg de FDC1 purificada se ajustaron a un volumen final de 1 ml. A continuación, las mezclas de reacción se incubaron durante 5 minutos a 30 °C. Tras la reacción, se añadió 1 ml de propanol a las mezclas de reacción para solubilizar el estireno. Para el control experimental negativo se utilizaron las mismas condiciones, con la excepción de que la mezcla de reacción no contenía sustrato. Los sustratos ácido ferúlico, ácido 2-metilcinámico y ácido 4-hidroxicinámico mostraron una actividad del 57%, 35% y 68%, respectivamente, en comparación con el tCA (Figura 11). La enzima no mostró ninguna actividad para los sustratos ácido 2-hidroxicinámico, ácido 3-hidroxicinámico, ácido 3,4-dimetoxicinámico y ácido 2,5-dimetoxicinámico (Figura 11). La enzima no mostró especificidad estricta frente al tCA; en su lugar, mostró una actividad moderada para el ácido ferúlico, el ácido 2-metilcinámico y el ácido 4-hidroxicinámico.

[0233] Los análisis de la especificidad frente al sustrato contribuirán a dilucidar el sitio de unión del sustrato y el mecanismo de la actividad enzimática. Además, este abanico de sustratos indicó que se puede utilizar el mismo sistema de fermentación para producir distintos monómeros industriales. Además de convertir el ácido cinámico en estireno, ahora se puede producir 4-hidroxi-3-metoxiestireno (también conocido como, 4-vinilguayacol, 4VG) a partir del ácido ferúlico; 2-metilestireno a partir del ácido 2-metilcinámico; y 4-hidroxiestireno (también conocido como 4-vinilfenol) a partir del ácido 4-hidroxicinámico.

11. Cinética

[0234] A fin de examinar la eficiencia catalítica de esta proteína, se realizaron estudios cinéticos de la FDC1 utilizando el tCA como sustrato. Para medir las constantes cinéticas de equilibrio de la FDC1 nativa, la actividad enzimática se determinó a diferentes concentraciones de tCA (100-2000 μM). Las reacciones se realizaron en un total de 1,0 ml de mezcla de reacción estándar con 0,5 mg de proteína purificada, y se dejaron avanzar durante 5 min a 30 °C. El estireno se extrajo con 250 μl de butanol. Para el control experimental negativo se utilizaron las mismas condiciones, con la excepción de que la mezcla de reacción no contenía sustrato. Las actividades se cuantificaron mediante un método de ensayo convencional. Se realizaron ensayos duplicados y se promediaron los resultados. La V_{max} y la K_m se determinaron mediante análisis de regresión no lineal de los datos de velocidad-concentración ajustados a la ecuación de Michaelis-Menten.

[0235] La K_m para la FDC1 nativa fue 688 μM y la V_{max} fue 6,17 $\text{nmol}\cdot\text{mg}^{-1}\cdot\text{min}^{-1}$. La eficiencia catalítica de esta proteína fue 8,4 $\text{M}^{-1}\text{S}^{-1}$, que es más baja que la de otras enzimas nativas. La eficiencia catalítica más baja de la FDC1 constituye el mayor obstáculo para la producción de estireno. La ingeniería de proteínas basada en la estructura será un buen enfoque para la evolución molecular de la FDC1 con el fin de aumentar la actividad.

25 EJEMPLO II. FDC1 MUTADAS Y BIBLIOTECAS DE FORMAS MUTADAS

1. Modelos de la estructura de la FDC1

[0236] No hay estructura terciaria de la FDC1 que se pueda utilizar para analizar los sitios de unión del sustrato. Una estructura de una proteína homóloga (3-octaprenil-4-hidroxibenzoato descarboxilasa, código de PDB: 2IDB) mostró únicamente una identidad del 20% con la FDC1. Para analizar el sitio de unión del sustrato se construyó un modelo para la FDC1 truncada mediante el programa SWISMODEL y un modelo para la FDC1 completa mediante el programa I-TASSER. Estos dos modelos son fiables ya que se superponen muy bien. Debido a que la menor eficiencia catalítica de la FDC1 constituye un cuello de botella para la producción de cantidades más elevadas de estireno, aplicamos un método combinado de biología molecular y biología estructural para la evolución en el laboratorio de un modelo basado en proteínas de la secuencia completa de la FDC1 nativa (Figura 1A).

[0237] Acoplamiento simulado del sustrato con la FDC1. Para la evolución en el laboratorio de la proteína, se examinó el sitio de unión del sustrato en la FDC1. No hay publicaciones sobre el sitio de unión del sustrato en la FDC1. El acoplamiento simulado del tCA con la FDC1 se realizó con el programa informático SWISDOCK. Tras el acoplamiento simulado, se encontraron cuatro posibles sitios de unión (Figura 1B).

[0238] Tras el análisis de la estructura, el sitio C se estudió para la evolución molecular con el fin de mejorar la actividad. Sin limitar el alcance de la tecnología objeto de esta invención, se plantea la hipótesis de que I173, A174, R175, V188, I189, K190 (no se muestra por claridad en la figura), I194, E280, M286, F291 y F440 contribuyen a la unión del sustrato (Figura 1C). Por lo tanto, los residuos hidrofóbicos (A174, I194 y V188) crean un bolsillo para la unión del anillo de fenilo del tCA; y el R175, con carga positiva, forma puentes de hidrógeno con el grupo carboxilo del tCA y el E280, que tiene carga negativa.

2. Mutagénesis por saturación de la FDC1

[0239] Debido a que la mutagénesis convencional no mejoró la actividad de la FDC1, se aplicó la mutagénesis por saturación dirigida. La mutagénesis por saturación permite el cambio de un aminoácido a los otros 19 posibles residuos aminoácidos. La mutagénesis por saturación se realizó en los sitios 155-156, 159, 162-164, 172-175, 187-196, 226-227, 285-287, 291, 326, 331, 360-361, 395-396, 398 y 440-441 de la FDC1 mediante la estrategia de mutagénesis dirigida con el kit QuickChange (STRATAgene, California, EE.UU.) utilizando cebadores degenerados NNK (N representa la mezcla de A, T, G, C y K para G/T). El codón NNK está degenerado 32 veces y codifica los 20 aminoácidos sin codones raros. Los productos de la PCR realizada con el kit QuikChange se examinaron mediante electroforesis en gel de agarosa y a continuación, 15 μl de los productos de PCR se digirieron con 1 μl de DpnI (New England Biolabs) a 37 °C durante 4 horas para eliminar el plásmido molde. Alícuotas (2 μl) de los productos digeridos se introdujeron por transformación en células competentes BL21-Gold (DE3) (STRATAgene, California, EE.UU.) y se inocularon en placas de agar Luria-Bertani (LB) que contenían kanamicina. Se confirmó la calidad de la biblioteca obtenida mediante secuenciación del ADN. La biblioteca abarca un 90% de mutagénesis (17 formas mutadas de 19). Para abarcar un 100% (19 formas mutadas de 19), cribamos 150 formas mutadas para cada sitio.

3. Método colorimétrico para el cribado de alto rendimiento de la biblioteca de FDC1 mutadas

[0240] El cribado de una población de gran tamaño de la biblioteca de proteínas constituye el cuello de botella para la evolución molecular de la proteína. La caracterización funcional de la descarboxilasa se suele realizar con instrumentos analíticos, como HPLC o LC-MS. Aunque el HPLC es muy sensible, requiere mucho tiempo, es caro y genera residuos como metanol o acetonitrilo, y no es adecuado para aplicaciones que requieran una elevada capacidad. Para superar estas barreras técnicas, se desarrolló un método de ensayo colorimétrico basado en la espectroscopia, que consistía fundamentalmente en la detección del estireno producido enzimáticamente a partir del tCA. El método detallado se facilita más abajo.

[0241] Las transformadas (ejemplo II.2) se inocularon en placas de 96 pocillos (NUNC, Roskilde, Dinamarca) que contenían 100 µl de CL por pocillo, y se incubaron a 37 °C durante la noche. A continuación, los cultivos se mezclaron con una cantidad igual de glicerol al 50% y se almacenaron a -80 °C como la placa principal. Una alícuota de cultivo de 10 µl se inoculó en placas de pocillos profundos de 2 ml (USA Scientific) con 1 ml de caldo Terrific (TB) por pocillo y se incubó a 37 °C hasta que la OD₆₀₀ alcanzó un valor de 1,0. A continuación, los cultivos se indujeron con IPTG 0,2 mM y se incubaron a 18 °C durante 20 horas con agitación a 250 rpm.

[0242] Las células se recolectaron y se suspendieron con 0,25 ml de solución amortiguadora de PBS a pH 7,0 y el sustrato (tCA) se añadió a una concentración final de 1 g/L. El cultivo se incubó a 30 °C durante 4 horas y el producto volátil, el estireno, se recuperó con una placa de fase inversa STRATA-X[®] de 96 pocillos que contenía 10 mg de resina polimérica (Phenomenex) sobre la placa de cultivo (Figura 19). A fin de evitar la difusión del vapor de estireno, para sellar los pocillos se utilizaron alfombrillas de silicona con prehendadura de 96 pocillos entre la placa de cultivo de 96 pocillos y la placa de fase inversa STRATA-X[®] de 96 pocillos. El producto se recuperó con 200 µl de propanol por cada pocillo y la cantidad de producto se determinó mediante el método colorimétrico.

[0243] Se evaluaron diversas sustancias químicas como indicador para la detección del estireno, incluida la NBP (4-nitrobencilpiridina) como indicador preferido. El estireno se mezcló con citocromo P450 BM-3 en solución amortiguadora de fosfato 50 mM a pH 8,0. La reacción de oxidación del estireno se inició añadiendo NADPH 0,2 mM, y la mezcla de reacción se incubó a temperatura ambiente durante 5 minutos con agitación. La solución de NBP se añadió a la mezcla y se incubó a temperatura ambiente durante otros 5 minutos con agitación, lo que produjo un precipitado blanco. El tubo o la placa que contenía las mezclas de reacción se calentaron a 70 °C durante 30 minutos y se enfriaron sobre hielo durante 5 minutos. Tras añadir dimetilformamida (150 µl) y K₂CO₃ 1M (25 µl) la mezcla de reacción adquirió un color azul que se puede medir inmediatamente mediante espectrofotometría a 600 nm. El electrón no apareado de la NBP reacciona con el anillo de oxirano del óxido de estireno para dar un cromóforo azul cuyo color azul se puede monitorizar mediante espectrofotometría a 600 nm.

[0244] Los cultivos en placa de 96 pocillos que contenían estireno se monitorizaron directamente utilizando este método de alto rendimiento (Figura 13). Con este método se puede detectar estireno en concentraciones tan bajas como 1,0 mM. Este método colorimétrico es muy rápido y menos caro, es reproducible y puede realizar mediciones en 1000 colonias en un solo día.

[0245] Las formas mutadas que mostraron una actividad más elevada se confirmaron mediante HPLC con una columna C18 ACQUITY UPLC BEH (1,7 µm, 2,1 x 50 mm, Waters, EE.UU.). Las muestras se resolvieron a un caudal de 1 ml/min en acetonitrilo (A) al 30% con un gradiente de concentración en aumento de acetonitrilo (B): durante 1 min, 95%; entonces a 1 min, 30%. La absorción por UV se monitorizó a 254, 280 y 310 nm. Los residuos aminoácidos cambiados en las formas mutadas que mostraron la actividad más elevada se confirmaron con la secuencia de ADN.

[0246] Se identificaron varias posibles formas mutadas únicas (como K190E, K190C, K190D, K190V, K190N, K190L, K190H y R175I) que producían una cantidad de estireno significativamente más elevada en comparación con la forma nativa (Figura 14). Las formas mutadas K190E, K190C y K190V produjeron 3 veces más, y la forma mutada R175I produjo 1,5 veces más estireno, que la forma nativa. Como tal, las FDC1 mutadas produjeron una cantidad significativamente más elevada de estireno que la forma nativa. La mutagénesis en distintos sitios mostró un efecto aditivo o sinérgico en la evolución proteica de la metiltransferasa (Bhuiya *et al.*, 2010, *Journal of Biological Chemistry*, 285: 277-285). La mutagénesis en los sitios K190 y R175 de la FDC1 puede aumentar aún más la producción de estireno.

[0247] Se llevó a cabo la producción de estireno manteniendo el cultivo en un ambiente aeróbico. La FDC1 nativa produjo la mitad de estireno en condiciones anaeróbicas, en comparación con la cantidad producida en condiciones aeróbicas. Con la columna STRATA-X[®] se puede atrapar el estireno volátil y mantener las condiciones aeróbicas.

EJEMPLO III. PROTEÍNAS DE FUSIÓN FDC1-PAL

[0248] *Expresión de la proteína de fusión FDC-PAL2 en E. coli.* Se incluyeron canales artificiales en la producción de estireno. Para construir la proteína de fusión de la PAL2 y la FDC mediante la tecnología Gateway, se diseñaron cuatro cebadores para la amplificación por PCR. El cebador FDC-5' contenía la secuencia CACC, y el cebador FDC-

3' tenía un conector de 9 aminoácidos y dos sitios de restricción NcoI y PstI en el extremo 3'. Por consiguiente, para fusionar la PAL2 con la FDC, el cebador PAL2-5' tenía un sitio NcoI, y el cebador PAL2-3' tenía un sitio PstI. El gen de la FDC (~1,5 kb) y el gen de la PAL2 (2,2 kb) se amplificaron utilizando los cebadores de arriba y se clonaron en el vector Gateway® pDEST17. Los constructos de la proteína de fusión de la PAL2 y la FDC se introdujeron por transformación en la cepa BL21(DE3). También se pueden obtener proteínas de fusión de la FDC mutada y la PAL mediante el proceso descrito más arriba.

[0249] El uso de un conector para las proteínas de fusión se evaluó con base en la información sobre la ruta de biosíntesis del resveratrol. Los resultados previos muestran que la proteína de fusión produjo 15 veces más producto en comparación con la expresión de dos proteínas independientes (Yechun Wang *et al.* (2011), JACS, 133: 20684-20687). Los resultados indican que el conector desempeña un importante papel en la proteína de fusión en lo que respecta a la mejora de la ruta de biosíntesis. Se diseñaron conectores de distintas longitudes, con 2, 3, 4, 6, 8, 9, 12 y 15 aminoácidos, utilizando el motivo GSG y se examinó el efecto de los conectores en la ruta de biosíntesis del resveratrol. Encontramos que el conector con 9 aminoácido mostraba el rendimiento más elevado, en comparación con los otros conectores (Figura 15). Con base en nuestros descubrimientos, diseñamos conectores de 9 aminoácidos para nuestras proteínas de fusión PAL-FDC.

[0250] Con base en la observación anterior, se desarrolló un modelo informático para demostrar que el conector artificial realmente aumentaba la canalización de metabolitos. El conector de 9 aminoácidos conectaba a dos proteínas (Figura 16) de tal manera que se evitaba la difusión del intermedio (en este caso el ácido cinámico), lo que permitía que un segundo enzima (FDC) pudiera captar eficientemente el intermedio, lo que en última instancia aumenta la cantidad final de estireno producido. El modelo informático predijo que la distancia entre los dos centros de reacción (marcados con las dos flechas) es de 70 Å o menos, y forma un canal metabólico.

[0251] El análisis de la expresión mostró que la concentración de estireno en el medio era de aproximadamente 615 mg/L. Esta proteína de fusión produjo aproximadamente 6 veces más estireno que la FDC expresada en *E. coli* (95 mg/L en el medio). Esta cantidad de estireno también fue mucho más elevada que la obtenida con las proteínas PAL2 y FDC no fusionadas (expresadas en levadura, con una concentración de estireno de 69 mg/L). Estos análisis indican que la proteína de fusión de la PAL2 y la FDC se puede utilizar para mejorar la bioproducción de estireno.

EJEMPLO IV. EXPRESIÓN DE UN TRANSPORTADOR ABC EN *E. COLI*

[0252] Algunas bacterias han desarrollado múltiples mecanismos para adaptarse a condiciones desfavorables. Por ejemplo, las cepas de *Pseudomonas* disponen de aproximadamente tres estrategias para hacer frente a la toxicidad de los disolventes orgánicos, es decir, estructura modificada de la membrana, bombas de salida activas y detoxificación enzimática. En este trabajo, se produjo un huésped de *E. coli* recombinante que expresaba una bomba de salida (denominada bomba de resistencia a disolventes). La bomba, un miembro de la familia de transportadores ABC, se compone de tres subunidades: conector periplásmico (srpA); transportador unido a la membrana interna (srpB); y canales en la membrana externa (srpC). La secuencia completa (~6 kb) de la bomba srpABC se clonó a partir de ADN genómico de *P. putida*. En primer lugar, el clon se insertó en el vector pENTR con resistencia a la zeocina y a continuación, en el vector de expresión pCONA-2-DEST de *E. coli* con resistencia a la kanamicina.

[0253] Finalmente, la bomba se introdujo por transformación en la cepa BL21 (DE3) de *E. coli* resistente a la ampicilina que contenía la fusión de PAL2 y FDC. Tras la expresión de prueba de 5 clones en matraces, se seleccionó un clon para otro experimento con fermentador. Se inocularon 50 ml de cultivo crecido durante la noche en 1,5 L de medio LB que contenía los antibióticos adecuados. Cuando las células alcanzaron una OD₆₀₀ de aprox. 0,8, se añadió IPTG 5 mM al cultivo para la inducción. Después de 2 h, se añadió el sustrato fenilalanina a una concentración final de 5 g/L en el cultivo celular. El vapor de estireno se rastreó en una botella que contenía 250 ml de butanol. Las muestras procedentes de la fermentación nocturna (~16 horas) se tomaron y se analizaron por HPLC con una columna Acclaim RSLC C18 y detección a 280 nm. Después de 16 h de fermentación, apenas se encontró estireno en el medio, pero se encontró estireno a una concentración elevada en el butanol. En comparación con el control que solo contiene la fusión de PAL2 y FDC, el clon que contiene genes triples (fusión de PAL2 y FDC con srpABC) pudo producir casi cuatro veces más producto, es decir, 521 mg/L frente a 139 mg/L (véase la Figura 17).

EJEMPLO V. BIOSÍNTESIS DEL ESTIRENO

[0254] *Producción de estireno a partir de glucosa.* Se construyó un vector que comprendía la proteína de fusión FDC1(K190E)-PAL, y se introdujo por transformación en cepas que producían fenilalanina (ATCC 31884 y HG), que se hicieron crecer en LB que contenía ampicilina y/o kanamicina. Los cultivos crecieron a 30 °C durante 12 horas, tras lo cual los cultivos se recolectaron y una alícuota de estos se inoculó en medio M9. Los cultivos crecieron a 30 °C hasta que la OD₆₀₀ alcanzó un valor de 0,8 y entonces, se indujeron con IPTG 0,2 mM y se continuó el cultivo a 30 °C o 37 °C durante 48 h. El estireno volátil se recuperó con la columna STRATA-X®. La cantidad de fenilalanina, tCA y estireno se determinó por HPLC. La cepa con número 31884 de la ATCC produjo 177 mg/L de estireno a partir de glucosa a 30 °C; no se observó acumulación de tCA ni de fenilalanina (Figura 18A). La cepa HG produjo 13,1 mg/L de estireno y 5,0 mg/L de tCA a partir de glucosa a 30 °C; no se observó acumulación de fenilalanina (Figura 18 A).

[0255] Resulta interesante que las cepas BL-21(DE3) produjeron 182 mg/L de estireno a partir de glucosa mediante la coexpresión de la proteína de fusión FDC-PAL y el vector productor de fenilalanina (HGTM) a 37 °C. El sistema completo de producción de estireno a partir de glucosa es transferible a cualquier sistema hospedador.

[0256] Cuando la FDC-PAL produjo estireno a partir de fenilalanina, con la columna STRATA-X® se produjo una concentración de estireno 25 veces mayor (125 mg/L), en comparación con la obtenida en condiciones anaeróbicas y sin la columna STRATA-X® (5 mg/L). La columna STRATA-X® mantiene condiciones aeróbicas, atrapa el estireno volátil y lo elimina, lo que, en última instancia, aumenta la producción de estireno y elimina el efecto de toxicidad en el cultivo.

[0257] *Producción de estireno a partir de ácido trans-cinámico o L-fenilalanina.* Los nucleótidos que codifican la FDC nativa, la FDC mutada (K190E), la proteína de fusión FDC(WT)-PAL y la proteína de fusión FDC(K190E)-PAL se introdujeron por transformación en *E. coli* BL-21 (DE3). Las células crecieron en LB a 30 °C durante la noche. Las células se recolectaron y se lavaron con medio M9. Las células se inocularon a 2,0 ml de medio M9 con una OD₆₀₀ inicial de 0,3 y se cultivaron a 30 °C hasta que la OD₆₀₀ del cultivo alcanzó un valor de 0,6 a 0,8. A continuación, las células se indujeron con IPTG 0,2 mM durante 8,0 h a 30 °C. Las células transformadas con la FDC nativa y FDC(K190E) fueron alimentadas con ácido *trans*-cinámico al 0,1%. Las células transformadas con las proteínas de fusión FDC(WT)-PAL y FDC(K190E)-PAL fueron alimentadas con, o bien ácido *trans*-cinámico al 0,1%, o bien *L*-fenilalanina al 0,1%. Las células alimentadas se cultivaron en continuo durante 36 horas. La columna STRATA-X® se utilizó para recuperar el estireno. El producto se eluyó a partir de la columna con butanol y la cantidad se determinó mediante HPLC.

[0258] Como se muestra en la Figura 18B, las células transformadas con la proteína mutada FDC(K190E) produjeron una cantidad más elevada de estireno (758 mg/L) que las transformadas con la FDC nativa (175 mg/ml). En comparación, la FDC(K190E) fusionada con la PAL produjo una cantidad de estireno ligeramente más baja a partir de tanto el ácido *trans*-cinámico como de la *L*-fenilalanina que la FDC nativa fusionada con la PAL.

EJEMPLO VI. PURIFICACIÓN DE LA CINAMATO DESCARBOXILASA PARA SU CRISTALIZACIÓN

[0259] Se construyó un vector de expresión de la FDC1 que tenía 6 histidinas y SUMO en el extremo amino. El ADN que codifica la FDC1 se introdujo por transformación en células competentes de *E. coli* (Rosetta 2) y se hizo crecer en LB que contenía ampicilina. Los cultivos crecieron en LB a 37 °C durante la noche. Los cultivos se inocularon en medio de caldo Terrific y crecieron a 37 °C hasta que la OD₆₀₀ alcanzó un valor de 0,8 a 1,0, momento en el que se añadió IPTG a una concentración final de 0,2 mM, y el cultivo se continuó a 16 °C durante 16 horas. Las células se recolectaron mediante centrifugación durante 15 minutos a 4°C a 4500 rpm. El sedimento celular se suspendió en la solución amortiguadora A (solución amortiguadora de fosfato de potasio 50 mM, tiosulfato de sodio 50 mM, TCEP-HCl 50 mM, NaCl 500 mM, PMSF 0,5 mM, MgCl₂ 10 mM, imidazol 10 mM, βME 20 mM y glicerol al 20% ajustado a pH 7,5) que contenía Triton X-100 al 0,1%, RNasa 2 µg/ml, DNasa 2 µg/ml y lisozima a 4 µg/ml. Las células se rompieron mediante ultrasonificación 10 veces a una amplitud de 15, durante un tiempo de proceso de 5 segundos y con pulsos intermitentes de 1 segundo. El sobrenadante se recuperó mediante centrifugación durante 15 minutos a 15.000 rpm a 4 °C. El sobrenadante se filtró a través de un filtro de PES de 0,45 µm y se aplicó a una columna de afinidad de Ni+-agarosa (GE Healthcare). La columna se lavó con la solución amortiguadora A hasta que las proteínas que se unen de manera no específica se eluyeron de la columna. La FDC1 se eluyó con fosfato de potasio 50 mM, tiosulfato de sodio 50 mM, TCEP-HCl 50 mM, cloruro de sodio 500 mM, PMSF 0,5 mM, cloruro de magnesio 10 mM, imidazol 250 mM, βME 20 mM y glicerol al 20% ajustado a pH 8,0. Se obtuvo un total de 45,0 mg de proteína después de la columna de afinidad de Ni+-agarosa.

[0260] Se añadió hidrolasa (0,5 mg) para disociar el SUMO de la FDC1. La digestión se llevó a cabo a 4 °C mediante diálisis de la proteína durante la noche en 1 L de fosfato de potasio 25 mM, tiosulfato de sodio 50 mM, NaCl 500 mM, DTT 5 mM y glicerol al 20% ajustado a pH 7,5. La diálisis se continuó a la mañana siguiente en un litro fresco de solución amortiguadora de diálisis durante 2-4 horas.

[0261] *Purificación sustractiva.* La columna de afinidad de Ni+-agarosa se lavó con agua, se recargó con sulfato de níquel y se equilibró con la solución amortiguadora A antes de depositar la proteína en la columna. La proteína se depositó en una columna, el eluyente se recuperó y se añadieron 15 ml de solución amortiguadora A, que se recuperó como eluyente. Después de recuperar toda la FDC1, la columna se lavó con una solución amortiguadora de elución para eliminar el SUMO y la hidrolasa de la columna. El eluyente que contiene la FDC1 se combinó y se determinó la cantidad de proteína. Tras la purificación sustractiva, se encontró un total de 27,5 mg. La proteína se dializó durante la noche en la columna Q con 1 L de solución amortiguadora A (fosfato de potasio 25 mM, DTT 5 mM y tiosulfato de sodio 25 mM ajustado a pH 7,5).

[0262] *Cromatografía de intercambio aniónico.* La columna Q (GE Healthcare) de intercambio aniónico se equilibró con la solución amortiguadora A para la columna Q. La proteína se depositó en la columna y se lavó con solución amortiguadora A hasta que todas las proteínas que se unen de manera no específica se eluyeron de la columna. La

proteína se eluyó con solución amortiguadora A con cloruro de sodio 1 M para la columna Q. La elución empezó con un 27% de solución amortiguadora B para columna Q. Las fracciones que contenían FDC1 se utilizaron para comprobar la pureza de la proteína mediante SDS-PAGE. Tras la cromatografía de intercambio aniónico, se encontró un total de 5,0 mg de proteína. La proteína se dializó durante la noche en solución amortiguadora (fosfato de sodio 50 mM a pH 7,5, NaCl 150 mM, DTT 5 mM y tiosulfato de sodio 25 mM) para exclusión molecular.

[0263] *Cromatografía de exclusión molecular.* La proteína se depositó en la columna de exclusión molecular (GE Healthcare) y se eluyó con fosfato de sodio 50 mM a pH 7,5, cloruro de sodio 150 mM, DTT 5 mM y tiosulfato de sodio 25 mM. Las fracciones (tubos del 33 al 38) que contenían FDC1 se combinaron y se concentraron. Se encontró un total de 3,5 mg de proteína purificada después de la cromatografía de exclusión molecular. La pureza se comprobó mediante SDS-PAGE (Figura 20). La proteína tenía una pureza de más del 98% y estaba lista para la cristalización.

[0264] Las técnicas convencionales de clonación molecular y ADN recombinante aquí utilizadas son bien conocidas en la técnica y Sambrook, J., Fritsch, E. F. y Maniatis las describen en *T. Molecular Cloning: A Laboratory Manual*, 2.^a ed.; Cold Spring Harbor Laboratory: Cold Spring Harbor, N.Y., 1989 (en lo sucesivo, «Maniatis»); y por Silhavy, T. J., Bannan, M. L. y Enquist, L. W. *Experiments with Gene Fusions*; Cold Spring Harbor Laboratory: Cold Spring Harbor, N.Y., 1984; y por Ausubel, F. M. et al., en *Current Protocols in Molecular Biology*, publicado por Greene Publishing y Wiley-Interscience, 1987.

[0265] Las FDC mutadas, por ejemplo, FDC(K190E), que tienen una secuencia aminoacídica como la establecida en la SEC. N.º ID: 16, se pueden purificar mediante el mismo procedimiento descrito más arriba para la FDC1.

EJEMPLO VII. CRISTALIZACIÓN DE LA CINAMATO DESCARBOXILASA

[0266] Los cristales de cinamato descarboxilasa mutada FDC(K190E) que forma un complejo con el ácido 3-hidroxicinámico se dejaron crecer por el método de difusión por vapor. Para cribar las condiciones de cristalización se utilizó un kit comercial de cribado (Hampton Research, Qiagen, Emerald Biosystems). Se evaluaron distintas relaciones de volúmenes de soluciones de proteína y depósito para la cristalización. Gotas colgantes que contenían una mezcla 1:1, 1:2 o 2:1 de proteína (5-10 mg/ml) y de solución amortiguadora para cristalización (polietilenglicol (PEG) 6000 al 10% (p/v), 2-metil-2,4-pentanediol (MPD) al 5% (p/v), HEPES 0,1 M, pH 7,5 y DTT 2 mM) se mantuvieron a 4°C. En particular, los cristales bien difractados se dejaron crecer en polietilenglicol (PEG) 6000 al 7-11% (p/v), 2-metil-2,4-pentanediol (MPD) al 3% (p/v), pH 6,5-7,5 y DTT 2 mM con ácido 3-hidroxicinámico al 0,1%. Las condiciones del cristal se mejoraron aún más al añadir MnCl₂ 0,01 M, polivinilpirrolidona K15 al 0,5% (p/v), NDSB-201 0,2 M o clorhidrato de benzamidina al 2% (p/v) como aditivo.

[0267] Los cristales de FDC(K190E) crecieron en el grupo espacial C2 con 3 cadenas por unidad asimétrica. Las dimensiones de la celda unitaria para el cristal fueron a=249,51 Å, b=120,67 Å, c=158,49 Å, β=94,9°; los datos de difracción se obtuvieron de cristales sueltos montados en un cryoloop, se congelaron rápidamente en un chorro de nitrógeno a 100 K y se redujeron con la suite HKL. Los cristales se difractaron a una resolución de 2,15 Å. Se utilizó un reemplazo molecular (Phaser) en la suite CCP4i para resolver la estructura de la FDC(K190E) mutada utilizando la estructura de la 3-octaprenil-4-hidroxibenzoato descarboxilasa (21DB) como modelo. El modelo inicial de las estructuras se construyó manualmente utilizando el software COOT y el modelo se refinó con Refmac5. En la Figura 21A se aporta un mapa de densidad electrónica típico con el modelo inicial actual de la FDC. En la Figura 21B se muestra una unidad asimétrica de la estructura cristalina que contiene 3 cadenas. Como se ilustra en la Figura 21B, las moléculas A y B de la FDC(K190E) forman un dímero que es biológicamente activo, y la molécula C de la FDC(K190E) forma otro dímero en el que su compañero pertenece a otra unidad asimétrica (no se muestra).

[0268] La memoria descriptiva se entiende mejor por medio de la información dada a conocer en las referencias citadas en la memoria descriptiva. Las realizaciones de la memoria descriptiva proporcionan una ilustración de las realizaciones de la tecnología objeto de esta invención y no se deben interpretar como si limitaran el alcance de la tecnología objeto de esta invención. Un experto en la materia fácilmente entenderá que la tecnología objeto de esta invención abarca muchas otras realizaciones. La citación de cualquier referencia en la presente memoria no es una admisión de que dichas referencias constituyen estado de la técnica para la presente invención.

[0269] Los expertos en la materia entenderán, o podrán averiguar por medio de experimentos habituales, muchos equivalentes de las realizaciones específicas de la tecnología objeto de esta invención descritas en la presente memoria. Se tiene por objeto que las realizaciones abarquen dichos equivalentes.

LISTADO DE SECUENCIAS

[0270]

<110> Phytogene, Inc.
Bhuiya, Mohammad W
Chen, Hui

Cai, Xianpeng
Han, Jixiang
Yu, Xiaodan

5 <120> ENZIMAS Y MÉTODOS PARA LA SÍNTESIS DE ESTIRENO
<130> 32990-5
<150> US 61/663,500
10 <151> 2012-06-22
<160> 44
<170> PatentIn version 3.5
15 <210> 1
<211> 2178
<212> ADN
<213> Arabidopsis thaliana
20 <400> 1
atggagatta acggggcaca caagagcaac ggaggaggag tggacgctat gttatgcggc 60
ggagacatca agacaaagaa catggtgatc aacgcggagg atcctctcaa ctggggagct 120
25 gcagcggagc aaatgaaagg tagccatttg gatgaagtga agagaatggg tgctgagttt 180
aggaagccag ttgtgaatct tgggtggtgag actctgacca ttggacaagt ggctgcgatc 240
30 tcaactattg gtaacagtgt gaagggtggag ctatcggaga cagctagagc cgggtgtgaat 300
gctagtagtg attgggttat ggagagtatg aacaaaggca ctgatagtta tgggtgttact 360
actggttttg gtgctacttc tcatcggaga accaaaaacg gtgtcgcact tcagaaggaa 420
35 cttattagat tccttaacgc cggaatattc ggaagcacga aagaaacaag ccacacattg 480
ccacactccg ccacaagagc cgccatgctt gtacgaatca acactctcct ccaaggattt 540
40 tccggtatcc gatttgagat tctcgaagca attaccagtt tcctcaacaa caacatcact 600
ccatctctcc ccctccgtgg tacaatcacc gcctccggag atctcgttcc tctctcctac 660
atcgccggac ttctcaccgg tcgtcccaat tccaaagcta ctggtcccaa cgggtgaagct 720
45 ttaacagcag aggaagcttt caaattagca ggaatcagct ccggattctt tgatctccag 780
cctaaggaag gtctcgcgct agtcaatggc acggcggttg gatctggaat ggcgtcaatg 840
50 gtgttattcg aaacgaatgt tctctctggt ttggctgaga ttttgtcggc ggttttcgca 900
gaggtgatga gtggttaagc tgagttcacc gatcatctca ctacagact taaacatcat 960
cccgggtcaaa tcgaagcggc ggcgataatg gagcatatcc tcgacggaag ctcgtagatg 1020
55 aaattagctc agaagcttca cgagatggat ccgttacaga aacctaaaca agatcgttac 1080
gctcttcgta cttctcctca atggttaggt cctcaaatacg aagtgatccg ttacgcaacg 1140
60 aaatcgatcg agcgtgagat taactccgtc aacgataatc cgttgatcga tgtttcgagg 1200
aacaaggcga ttcacggtgg taacttccaa ggaacaccaa tcggagtttc aatggataac 1260
acgagattgg cgatagcagc gattggtaaa ctcatgtttg ctcaattctc agagcttgtg 1320
65 aatgatttct acaacaatgg tttaccctcg aatctaaccg cttcgaggaa tccaagtttg 1380
gattatggat tcaagggagc tgagattgca atggcttctt attgttcaga gtttcaatac 1440

ES 2 709 217 T3

ttagctaadc ctgtgactag ccatgttcaa tcagcagagc aacataacca agatgtcaac 1500
 5 tctttgggac taatctcgtc tcgcaaaact tctgaagctg ttgatattct caagcttatg 1560
 tcaacaacgt tcctcgttgc gatttgtcaa gctgtggatt tgagacattt ggaggagaat 1620
 ttgagacaga ctgtgaagaa cactgtctct caagtggcga agaaagttct tactactgga 1680
 10 gtcaatggtg agcttcatcc ttctcgcttc tgcgaaaagg atttactcaa agttgtagac 1740
 cgtgaacaag tctacacata cgcggatgat ccttgtagcg caacgtaccc gttgattcag 1800
 aagctgagac aagttattgt tgaccatgct ttgatcaatg gtgagagtga gaagaatgca 1860
 15 gtgacttcaa tcttccataa gattggagct ttcgaggagg agcttaaggc agtgctaccg 1920
 aaagaagtgg aagcagcaag agcagcctac gataacggaa catcggctat cccgaacagg 1980
 20 atcaaggaat gtaggtcgtg tccattgtat agattcgtga gggaagagct tggaacagag 2040
 cttttgaccg gagagaaagt gacgtcgctt ggagaagagt tcgacaaggt tttcacggcg 2100
 25 atttgtgaag gtaaaatcat tgatccgatg atggaatgtc tcaacgagtg gaacggagct 2160
 cccattccaa tatgttaa 2178

30 <210> 2
 <211> 725
 <212> PRT
 <213> Arabidopsis thaliana

35 <400> 2
 Met Glu Ile Asn Gly Ala His Lys Ser Asn Gly Gly Gly Val Asp Ala
 1 5 10 15
 40 Met Leu Cys Gly Gly Asp Ile Lys Thr Lys Asn Met Val Ile Asn Ala
 20 25 30
 45 Glu Asp Pro Leu Asn Trp Gly Ala Ala Ala Glu Gln Met Lys Gly Ser
 35 40 45
 50 His Leu Asp Glu Val Lys Arg Met Val Ala Glu Phe Arg Lys Pro Val
 50 55 60
 Val Asn Leu Gly Gly Glu Thr Leu Thr Ile Gly Gln Val Ala Ala Ile
 65 70 75 80
 55 Ser Thr Ile Gly Asn Ser Val Lys Val Glu Leu Ser Glu Thr Ala Arg
 85 90 95
 60 Ala Gly Val Asn Ala Ser Ser Asp Trp Val Met Glu Ser Met Asn Lys
 100 105 110
 65 Gly Thr Asp Ser Tyr Gly Val Thr Thr Gly Phe Gly Ala Thr Ser His
 115 120 125
 Arg Arg Thr Lys Asn Gly Val Ala Leu Gln Lys Glu Leu Ile Arg Phe
 40

ES 2 709 217 T3

	130					135						140				
5	Leu 145	Asn	Ala	Gly	Ile	Phe 150	Gly	Ser	Thr	Lys	Glu 155	Thr	Ser	His	Thr	Leu 160
10	Pro	His	Ser	Ala	Thr 165	Arg	Ala	Ala	Met	Leu 170	Val	Arg	Ile	Asn	Thr 175	Leu
15	Leu	Gln	Gly	Phe 180	Ser	Gly	Ile	Arg	Phe 185	Glu	Ile	Leu	Glu	Ala 190	Ile	Thr
20	Ser	Phe	Leu 195	Asn	Asn	Asn	Ile	Thr 200	Pro	Ser	Leu	Pro	Leu 205	Arg	Gly	Thr
25	Ile	Thr 210	Ala	Ser	Gly	Asp	Leu 215	Val	Pro	Leu	Ser	Tyr 220	Ile	Ala	Gly	Leu
30	Leu 225	Thr	Gly	Arg	Pro	Asn 230	Ser	Lys	Ala	Thr	Gly 235	Pro	Asn	Gly	Glu	Ala 240
35	Leu	Thr	Ala	Glu	Glu 245	Ala	Phe	Lys	Leu	Ala 250	Gly	Ile	Ser	Ser	Gly 255	Phe
40	Phe	Asp	Leu	Gln 260	Pro	Lys	Glu	Gly	Leu 265	Ala	Leu	Val	Asn	Gly 270	Thr	Ala
45	Val	Gly	Ser 275	Gly	Met	Ala	Ser	Met 280	Val	Leu	Phe	Glu	Thr 285	Asn	Val	Leu
50	Ser	Val 290	Leu	Ala	Glu	Ile	Leu 295	Ser	Ala	Val	Phe	Ala 300	Glu	Val	Met	Ser
55	Gly 305	Lys	Pro	Glu	Phe	Thr 310	Asp	His	Leu	Thr	His 315	Arg	Leu	Lys	His	His 320
60	Pro	Gly	Gln	Ile	Glu 325	Ala	Ala	Ala	Ile	Met 330	Glu	His	Ile	Leu	Asp 335	Gly
65	Ser	Ser	Tyr	Met 340	Lys	Leu	Ala	Gln	Lys 345	Leu	His	Glu	Met	Asp 350	Pro	Leu
70	Gln	Lys	Pro 355	Lys	Gln	Asp	Arg	Tyr 360	Ala	Leu	Arg	Thr	Ser 365	Pro	Gln	Trp
75	Leu	Gly 370	Pro	Gln	Ile	Glu	Val 375	Ile	Arg	Tyr	Ala	Thr 380	Lys	Ser	Ile	Glu
80	Arg 385	Glu	Ile	Asn	Ser	Val 390	Asn	Asp	Asn	Pro	Leu 395	Ile	Asp	Val	Ser	Arg 400
85	Asn	Lys	Ala	Ile	His	Gly	Gly	Asn	Phe	Gln	Gly	Thr	Pro	Ile	Gly	Val

ES 2 709 217 T3

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

ES 2 709 217 T3

	675	680	685	
5	Ser Pro Gly Glu Glu Phe Asp Lys Val Phe Thr Ala Ile Cys Glu Gly 690 695 700			
10	Lys Ile Ile Asp Pro Met Met Glu Cys Leu Asn Glu Trp Asn Gly Ala 705 710 715 720			
	Pro Ile Pro Ile Cys 725			
15	<210> 3 <211> 2154 <212> ADN <213> Arabidopsis thaliana			
20	<400> 3			
	atggatcaaa tcgaagcaat gttgtgcggc ggaggagaga agacaaaagt ggcggttact			60
25	acgaagactt tggcagatcc attgaattgg ggttttagcag cggaatcaaat gaaaggaagt			120
	catttagatg aagtgaagaa gatggctcga gagtatcgtg gaccagtcgt gaatcttggc			180
	ggagaaacac tgacgatcgg acaagttgct gccatctcca ccgtaggagg cagcgttaag			240
30	gttgagttag cggagacttc aagagccggg gtgaaagcta gcagtgattg ggttatggag			300
	agcatgaaca aaggtactga cagttacgga gtcaccaccg gctttgggtg tacttctcac			360
35	cggagaacca aaaacggcac cgcattacaa acagaactca ttagattttt gaacgccgga			420
	atattcggaa acacgaagga gacatgtcac aactgccgc aatccgccac aagagccgcc			480
	atgctcgtca gagtcaacac tcttctccaa ggatactccg ggatccgatt cgagatcctc			540
40	gaagcgatta caagtctcct caaccacaac atctctccgt cactacctct ccgtggaacc			600
	attaccgcct ccggcgatct cgttcctctc tcttacatcg ccggacttct caccggccgt			660
45	cctaattcca aagccaccgg tcccgcagg gtatcgctaa ccgcgaaaga agcttttgag			720
	aaagccggaa tcagtactgg attcttcgat ttacaacctt aggaagggtt agctctcggt			780
	aatggcacgg cggttggatc tggaaatggcg tcgatgggtt tattcgaagc gaatgtccaa			840
50	gcggtgttag cggaggtttt atcagcgatc ttcgcggagg ttatgagcgg gaaacctgag			900
	tttaccgatc atctgactca tcgtttaaaa catcatcccc gacaaatcga agcggcggcg			960
55	ataatggagc acatactcga cggaaagctc tacatgaaat tagctcaaaa gggtcacgag			1020
	atggatccat tgcagaaacc aaaacaagat cgttacgctc ttcgtacatc tcctcaatgg			1080
	ctaggtcctc aaattgaagt aatccgtcaa gctacgaaat cgatagagcg tgaaatcaac			1140
60	tccgttaacg ataatccgtt gatcgatggt tcgaggaaca aggcgattca cggtggtaac			1200
	ttccaaggaa caccaatcgg agtttctatg gataacacga gattggcgat tgctgcgatt			1260
65	gggaagctaa tgtttgctca attctctgag cttgttaatg atttctacaa caatggactt			1320
	ccttcgaatc taactgcttc gagtaatcca agtttggatt atggattcaa aggagcagag			1380
	attgctatgg cttcttattg ttctgagctt caatacttgg ctaatccagt cacaagccat			1440

ES 2 709 217 T3

gttcaatcag ctgagcaaca taatcaagat gtgaactctc ttggtttgat ctcgtctcgt 1500
 5 aaaacatctg aagctgtgga tattcttaag ctaatgtcaa caacgttcct tgtggggata 1560
 tgtcaagctg ttgatttgag acatttgag gagaatctga gacaaactgt gaagaacaca 1620
 gtttctcaag ttgctaagaa agtgtaacc actggaatca acggtgagtt acatccgtca 1680
 10 aggttttgcg agaaggactt gcttaagggtt gttgatcgtg agcaagtgtt cacgtatgtg 1740
 gatgatcctt gtagcgctac gtacccgttg atgcagagac taagacaagt tattgttgat 1800
 cacgctttgt ccaacggtga gactgagaag aatgcagtga cttcgatctt tcaaaagatt 1860
 15 ggagcttttg aagaggagct taaggctgtg cttccaaagg aagttgaagc ggctagagcg 1920
 gcttatggga atggaactgc gccgattcct aaccggatta aggaatgtag gtcgtatccg 1980
 20 ttgtataggt tcgtgagggg agagcttgga acgaagttgt tgactggaga aaaggttgtg 2040
 tctccgggag aggagtttga taaggctctc actgctatgt gtgaaggtaa acttattgat 2100
 ccgttgatgg attgtctcaa ggaatggaac ggagctccga ttccgatttg ctaa 2154
 25
 <210> 4
 <211> 717
 <212> PRT
 30 <213> Arabidopsis thaliana
 <400> 4
 35 Met Asp Gln Ile Glu Ala Met Leu Cys Gly Gly Gly Glu Lys Thr Lys
 1 5 10 15
 Val Ala Val Thr Thr Lys Thr Leu Ala Asp Pro Leu Asn Trp Gly Leu
 40 20 25 30
 Ala Ala Asp Gln Met Lys Gly Ser His Leu Asp Glu Val Lys Lys Met
 45 35 40 45
 Val Glu Glu Tyr Arg Arg Pro Val Val Asn Leu Gly Gly Glu Thr Leu
 50 50 55 60
 Thr Ile Gly Gln Val Ala Ala Ile Ser Thr Val Gly Gly Ser Val Lys
 65 65 70 75 80
 Val Glu Leu Ala Glu Thr Ser Arg Ala Gly Val Lys Ala Ser Ser Asp
 55 85 90 95
 Trp Val Met Glu Ser Met Asn Lys Gly Thr Asp Ser Tyr Gly Val Thr
 60 100 105 110
 Thr Gly Phe Gly Ala Thr Ser His Arg Arg Thr Lys Asn Gly Thr Ala
 65 115 120 125
 Leu Gln Thr Glu Leu Ile Arg Phe Leu Asn Ala Gly Ile Phe Gly Asn
 130 135 140

ES 2 709 217 T3

	Thr	Lys	Glu	Thr	Cys	His	Thr	Leu	Pro	Gln	Ser	Ala	Thr	Arg	Ala	Ala	
	145					150					155					160	
5	Met	Leu	Val	Arg	Val	Asn	Thr	Leu	Leu	Gln	Gly	Tyr	Ser	Gly	Ile	Arg	
					165					170					175		
10	Phe	Glu	Ile	Leu	Glu	Ala	Ile	Thr	Ser	Leu	Leu	Asn	His	Asn	Ile	Ser	
				180					185					190			
15	Pro	Ser	Leu	Pro	Leu	Arg	Gly	Thr	Ile	Thr	Ala	Ser	Gly	Asp	Leu	Val	
			195					200					205				
20	Pro	Leu	Ser	Tyr	Ile	Ala	Gly	Leu	Leu	Thr	Gly	Arg	Pro	Asn	Ser	Lys	
		210					215					220					
	Ala	Thr	Gly	Pro	Asp	Gly	Glu	Ser	Leu	Thr	Ala	Lys	Glu	Ala	Phe	Glu	
	225					230					235					240	
25	Lys	Ala	Gly	Ile	Ser	Thr	Gly	Phe	Phe	Asp	Leu	Gln	Pro	Lys	Glu	Gly	
					245					250					255		
30	Leu	Ala	Leu	Val	Asn	Gly	Thr	Ala	Val	Gly	Ser	Gly	Met	Ala	Ser	Met	
				260					265					270			
35	Val	Leu	Phe	Glu	Ala	Asn	Val	Gln	Ala	Val	Leu	Ala	Glu	Val	Leu	Ser	
			275					280					285				
40	Ala	Ile	Phe	Ala	Glu	Val	Met	Ser	Gly	Lys	Pro	Glu	Phe	Thr	Asp	His	
		290					295					300					
	Leu	Thr	His	Arg	Leu	Lys	His	His	Pro	Gly	Gln	Ile	Glu	Ala	Ala	Ala	
	305					310					315					320	
45	Ile	Met	Glu	His	Ile	Leu	Asp	Gly	Ser	Ser	Tyr	Met	Lys	Leu	Ala	Gln	
					325					330					335		
50	Lys	Val	His	Glu	Met	Asp	Pro	Leu	Gln	Lys	Pro	Lys	Gln	Asp	Arg	Tyr	
				340					345					350			
55	Ala	Leu	Arg	Thr	Ser	Pro	Gln	Trp	Leu	Gly	Pro	Gln	Ile	Glu	Val	Ile	
			355					360					365				
60	Arg	Gln	Ala	Thr	Lys	Ser	Ile	Glu	Arg	Glu	Ile	Asn	Ser	Val	Asn	Asp	
		370					375					380					
	Asn	Pro	Leu	Ile	Asp	Val	Ser	Arg	Asn	Lys	Ala	Ile	His	Gly	Gly	Asn	
	385					390					395					400	
65	Phe	Gln	Gly	Thr	Pro	Ile	Gly	Val	Ser	Met	Asp	Asn	Thr	Arg	Leu	Ala	
					405					410					415		

ES 2 709 217 T3

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

ES 2 709 217 T3

Val Phe Thr Ala Met Cys Glu Gly Lys Leu Ile Asp Pro Leu Met Asp
690 695 700

5
Cys Leu Lys Glu Trp Asn Gly Ala Pro Ile Pro Ile Cys
705 710 715

10 <210> 5
<211> 2124
<212> ADN
<213> Arabidopsis thaliana

15 <400> 5
atggagctat gcaatcaaaa caatcacatc accgccgtct cgggcgatcc gttgaactgg 60
aacgcgacgg ccgaagcttt gaaagggagc cacctggatg aggtgaaacg aatggtgaaa 120
20 gagtatagga aagaggcggg gaagttagga ggtgagactt tgacgattgg tcaagtagcc 180
gccgtggcta gaggaggagg aggatctacg gtggagctag cggaggaggc tcgtgccgga 240
gtcaaggcga gtagcgaatg ggtgatggag agcatgaacc gaggaacgga cagttatgga 300
25 gttaccacag ggtttggtgc aacttcccat agaagaacca aacaaggcgg tgcacttcaa 360
aatgagctta ttaggttctt gaatgccgga atatttggcc ccggcgccgg ggacacgtca 420
30 cacacgttgc caaagccgac aacaagagcg gcaatgctcg tccgtgtcaa cactctcctc 480
caaggctact ccggtatacg cttcgagatt ctcgaagcaa ttacaaagct tctcaaccac 540
gaaatcactc cgtgcctccc tctccgtggc accatcaccg cctccggtga cttgttctc 600
35 ctctcttaca tcgccggact tctcactggc cgtcccaact ccaaagccgt gggtcctct 660
ggtgagactc tactgcctc tgaggccttt aagctcgccg gagtatcgtc ctttttcgag 720
40 ctgcagccta aggaaggact agcacttggtg aacgggacag cggttggatc gggtttggcc 780
tcaacggttt tgttcgaatg aaatattttg gctgttttat cggaagtatg gtctgcatg 840
ttcgcagagg ttatgcaagg gaaaccggag ttacagatc atcttacgca taagctcaag 900
45 caccatcccg gtcagatcga agccgccgca attatggaac atatattaga cggaagctct 960
tacgttaaag aagctcaact tctccacgaa atggatcctc ttcaaaaacc taaacaagat 1020
50 cggtagcgtt tacgtacgtc accacaatgg cttggggccg agattgaagt gatcagagcg 1080
gctactaaaa tgattgagcg tgagatcaac tctgttaatg ataacccttt gatagatgtg 1140
tcgaggaaca aggcgttgca cgggtggaaat ttccaaggga caccgatcgg tgttgccatg 1200
55 gataattccc gtctagccat tgcttccatt gggaaactca tgtttgcgca gttttctgaa 1260
ctagtgaacg atttctacaa caatggtttg ctttctaatac tatctggtgg gagaaaccct 1320
60 agtcttgatt acgggtttta aggcgcggaa atagccatgg cttcttattg ctccgagctt 1380
cagttccttg ctaatcccgt gaccaaccat gtccaaagcg cagagcagca taaccaagac 1440
gttaattccc tagggctaata ctctagcagg aaaactgcag aagcagtggg tatcctcaag 1500
65 ctaatgtcca caacctactt agtcgcgctt tgccaagccg ttgatctaag acatcttgaa 1560
gagaatctga agaaggcggg taaatcagca gtgagtcagg tggcgaaacg ggtcttaacc 1620

ES 2 709 217 T3

```

gttggtgccac acggggagct acatccgtca aggttcacag aacgtgatgt cctccaagtg      1680
gttgaccgag agtacgtgtt ctcatacgca gacgatccct gcagcctcac ttacccgcta      1740
5 atgcagaaac ttagacacat tcttgttagac cacgcttttag cggatccaga acgcgaggcc      1800
aattccgcga catcggtttt ccacaaaatc ggagcttttg aagccgagct gaaactgctt      1860
10 ctccttaaag aagtagaacg cgtccgggtt gaatacgagg aaggaacatc ggctatagct      1920
aaccggatta aggaatgtcg gtcttatcca ttgtatcggg ttgtccgcga tgagctaaat      1980
actgaactgc ttactggaga gaatgttcgg tcgccaggag aggagtttga taaagtgttc      2040
15 ttagcgattt ctgatggaaa acttattgat ccgttggttg aatgtctcaa ggagtggaac      2100
ggagctccgg tttcaatctg ttga      2124
20
<210> 6
<211> 707
<212> PRT
<213> Arabidopsis thaliana
25
<400> 6
Met Glu Leu Cys Asn Gln Asn Asn His Ile Thr Ala Val Ser Gly Asp
1 5 10 15
30
Pro Leu Asn Trp Asn Ala Thr Ala Glu Ala Leu Lys Gly Ser His Leu
20 25 30
35
Asp Glu Val Lys Arg Met Val Lys Glu Tyr Arg Lys Glu Ala Val Lys
35 40 45
40
Leu Gly Gly Glu Thr Leu Thr Ile Gly Gln Val Ala Ala Val Ala Arg
50 55 60
45
Gly Gly Gly Gly Ser Thr Val Glu Leu Ala Glu Glu Ala Arg Ala Gly
65 70 75 80
50
Val Lys Ala Ser Ser Glu Trp Val Met Glu Ser Met Asn Arg Gly Thr
85 90 95
55
Asp Ser Tyr Gly Val Thr Thr Gly Phe Gly Ala Thr Ser His Arg Arg
100 105 110
60
Thr Lys Gln Gly Gly Ala Leu Gln Asn Glu Leu Ile Arg Phe Leu Asn
115 120 125
65
Ala Gly Ile Phe Gly Pro Gly Ala Gly Asp Thr Ser His Thr Leu Pro
130 135 140
Lys Pro Thr Thr Arg Ala Ala Met Leu Val Arg Val Asn Thr Leu Leu
145 150 155 160
Gln Gly Tyr Ser Gly Ile Arg Phe Glu Ile Leu Glu Ala Ile Thr Lys
48

```

ES 2 709 217 T3

	165							170							175						
5	Leu	Leu	Asn	His 180	Glu	Ile	Thr	Pro	Cys 185	Leu	Pro	Leu	Arg	Gly 190	Thr	Ile					
10	Thr	Ala	Ser 195	Gly	Asp	Leu	Val	Pro 200	Leu	Ser	Tyr	Ile	Ala 205	Gly	Leu	Leu					
15	Thr	Gly 210	Arg	Pro	Asn	Ser	Lys 215	Ala	Val	Gly	Pro	Ser 220	Gly	Glu	Thr	Leu					
20	Thr 225	Ala	Ser	Glu	Ala	Phe 230	Lys	Leu	Ala	Gly	Val 235	Ser	Ser	Phe	Phe	Glu 240					
25	Leu	Gln	Pro	Lys	Glu 245	Gly	Leu	Ala	Leu	Val 250	Asn	Gly	Thr	Ala	Val 255	Gly					
30	Ser	Gly	Leu	Ala 260	Ser	Thr	Val	Leu	Phe 265	Asp	Ala	Asn	Ile	Leu 270	Ala	Val					
35	Leu	Ser	Glu 275	Val	Met	Ser	Ala	Met 280	Phe	Ala	Glu	Val	Met 285	Gln	Gly	Lys					
40	Pro	Glu 290	Phe	Thr	Asp	His	Leu 295	Thr	His	Lys	Leu	Lys 300	His	His	Pro	Gly					
45	Gln 305	Ile	Glu	Ala	Ala	Ala 310	Ile	Met	Glu	His	Ile 315	Leu	Asp	Gly	Ser	Ser 320					
50	Tyr	Val	Lys	Glu	Ala 325	Gln	Leu	Leu	His	Glu 330	Met	Asp	Pro	Leu	Gln 335	Lys					
55	Pro	Lys	Gln	Asp 340	Arg	Tyr	Ala	Leu	Arg 345	Thr	Ser	Pro	Gln	Trp 350	Leu	Gly					
60	Pro	Gln	Ile 355	Glu	Val	Ile	Arg	Ala 360	Ala	Thr	Lys	Met	Ile 365	Glu	Arg	Glu					
65	Ile	Asn 370	Ser	Val	Asn	Asp	Asn 375	Pro	Leu	Ile	Asp	Val 380	Ser	Arg	Asn	Lys					
70	Ala 385	Leu	His	Gly	Gly	Asn 390	Phe	Gln	Gly	Thr	Pro 395	Ile	Gly	Val	Ala	Met 400					
75	Asp	Asn	Ser	Arg	Leu 405	Ala	Ile	Ala	Ser	Ile 410	Gly	Lys	Leu	Met	Phe 415	Ala					
80	Gln	Phe	Ser	Glu 420	Leu	Val	Asn	Asp	Phe 425	Tyr	Asn	Asn	Gly	Leu 430	Pro	Ser					
85	Asn	Leu	Ser	Gly	Gly	Arg	Asn	Pro	Ser	Leu	Asp	Tyr	Gly	Phe	Lys	Gly					

ES 2 709 217 T3

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

705

5	<210> 7 <211> 1512 <212> ADN <213> <i>Saccharomyces cerevisiae</i>	
10	<400> 7 atgaggaagc taaatccagc tttagaattt agagacttta tccaggctctt aaaagatgaa	60
	gatgacttaa tcgaaattac cgaagagatt gatccaaatc tcgaagtagg tgcaattatg	120
15	aggaaggcct atgaatccca cttaccagcc ccgttatatta aaaatctcaa aggtgcttcg	180
	aaggatcttt tcagcatttt aggttgccca gccggtttga gaagtaagga gaaaggagat	240
	catggtagaa ttgcccata tctggggctc gacccaaaaa caactatcaa ggaaatcata	300
20	gattatttgc tggagtgtaa ggagaaggaa cctctccccc caatcactgt tcctgtgtca	360
	tctgcacctt gtaaaacaca tatactttct gaagaaaaaa tacatctaca aagcctgcca	420
25	acaccatata tacatgtttc agacggtggc aagtacttac aaacgtacgg aatgtggatt	480
	cttcaaactc cagataaaaa atggactaat tggatcaattg ctagaggatg ggtttagat	540
	gacaagcata tcaactggtc ggtaattaaa ccacaacata ttagacaaat tgctgactct	600
30	tgggcagcaa ttggaaaagc aaatgaaatt cctttcgcgt tatgttttgg cgttcccca	660
	gcagctatct tagttagttc catgccaat cctgaagggtg tttctgaatc ggattatgtt	720
35	ggcgcaatct tgggtgagtc gggtccagta gtaaaatgtg agaccaacga tttaatgggt	780
	cctgcaacga gtgagatggt atttgagggg actttgtcct taacagatac acatctggaa	840
	ggcccatctg gtgagatgca tggatatgtt ttcaaaagcc aaggtcatcc ttgtccattg	900
40	tacactgtca aggctatgag ttacagagac aatgctattc tacctgtttc gaaccccggt	960
	ctttgtacgg atgagacaca taccttgatt gggtcactag tggctactga ggccaaggag	1020
45	ctggctattg aatctggctt gccaattctg gatgccttta tgccttatga ggctcaggct	1080
	ctttggctta tcttaaagggt ggatttgaaa gggctgcaag cattgaagac aacgcctgaa	1140
	gaattttgta agaaggtagg tgatatttac tttaggacaa aagttgggtt tatagtccat	1200
50	gaaataattt tgggtggcaga tgatatcgac atattttaact tcaaagaagt catctgggcc	1260
	tacgttacaa gacatacacc tgttgcagat cagatggctt ttgatgatgt cacttctttt	1320
55	cctttggctc cctttgtttc gcagtcattc agaagtaaga ctatgaaagg tggaaagtgc	1380
	gttactaatt gcatatttag acagcaatat gagcgcagtt ttgactacat aacttgtaat	1440
	tttgaaaagg gatatccaaa aggattagtt gacaaagtaa atgaaaattg gaaaaggtag	1500
60	ggatataaat aa	1512
65	<210> 8 <211> 503 <212> PRT <213> <i>Saccharomyces cerevisiae</i>	
	<400> 8	

ES 2 709 217 T3

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

ES 2 709 217 T3

Ser Leu Thr Asp Thr His Leu Glu Gly Pro Phe Gly Glu Met His Gly
 275 280 285
 5 Tyr Val Phe Lys Ser Gln Gly His Pro Cys Pro Leu Tyr Thr Val Lys
 290 295 300
 10 Ala Met Ser Tyr Arg Asp Asn Ala Ile Leu Pro Val Ser Asn Pro Gly
 305 310 315 320
 15 Leu Cys Thr Asp Glu Thr His Thr Leu Ile Gly Ser Leu Val Ala Thr
 325 330 335
 20 Glu Ala Lys Glu Leu Ala Ile Glu Ser Gly Leu Pro Ile Leu Asp Ala
 340 345 350
 Phe Met Pro Tyr Glu Ala Gln Ala Leu Trp Leu Ile Leu Lys Val Asp
 355 360 365
 25 Leu Lys Gly Leu Gln Ala Leu Lys Thr Thr Pro Glu Glu Phe Cys Lys
 370 375 380
 30 Lys Val Gly Asp Ile Tyr Phe Arg Thr Lys Val Gly Phe Ile Val His
 385 390 395 400
 35 Glu Ile Ile Leu Val Ala Asp Asp Ile Asp Ile Phe Asn Phe Lys Glu
 405 410 415
 Val Ile Trp Ala Tyr Val Thr Arg His Thr Pro Val Ala Asp Gln Met
 420 425 430
 40 Ala Phe Asp Asp Val Thr Ser Phe Pro Leu Ala Pro Phe Val Ser Gln
 435 440 445
 45 Ser Ser Arg Ser Lys Thr Met Lys Gly Gly Lys Cys Val Thr Asn Cys
 450 455 460
 50 Ile Phe Arg Gln Gln Tyr Glu Arg Ser Phe Asp Tyr Ile Thr Cys Asn
 465 470 475 480
 55 Phe Glu Lys Gly Tyr Pro Lys Gly Leu Val Asp Lys Val Asn Glu Asn
 485 490 495
 60 Trp Lys Arg Tyr Gly Tyr Lys
 500
 <210> 9
 <211> 1503
 <212> ADN
 65 <213> Aspergillus niger
 <400> 9
 atgtctgcgc aacctgctca cctgtgtttc cgctccttcg tcgaagccct caaggctcac

ES 2 709 217 T3

	aacgaccttg ttgaaatcaa taccccaatt gacccaatc tcgaagctgc tgctattacc	120
5	cgccgagtat gtgagaccaa cgacaaggct cttttattca acaacctcat cggcatgaaa	180
	aatggcctct tccgtatact tggggctcct ggctctctca ggaagtcgtc tgctgatcgc	240
	tacggccgcc ttgctcgtca cctagccctc ccacctacgg cctcaatgcg tgagattctc	300
10	gataagatgc tctccgccag cgatatgcct cccatccctc cgaccattgt tcccaccggg	360
	ccatgcaagg agaacagctt agatgactct gaattcgacc ttaccgaact ccccgttcct	420
	cttattcaca aatcggatgg tggtaaatac atccaaacct atggcatgca cattgtgcag	480
15	tctccggatg gaacctggac caactgggtct attgcccgtg cgatgggtcca tgacaagaac	540
	catctgaccg gcctgggttat tccccctcag cacatctggc agattcacca gatgtggaag	600
20	aaggaaggcc gcagtgcgt tccctgggct ttggcctttg gtgtcccacc cgctgccatt	660
	atggcctcta gcatgcctat tcccgatggg gtcaccgaag ctgggtacgt gggagctatg	720
	acgggatcct ccctggagct tgttaaagt gatacgaacg atctatatgt ccccgctacc	780
25	tcagaaatcg ttctcgaggg cacactctct atcagcgaga caggcccaga gggacctttc	840
	ggtgagatgc atgggttacat cttccccggg gatactcacc tcggcgccaa atacaagggt	900
30	aaccggatca cctaccgcaa caacgccatc atgcccattgt cttcttggtg ccgcttgacg	960
	gatgaaacgc acaccatgat cggctctctg gctgcggcgg agatccgtaa gctctgccag	1020
	cagaatgacc tccctatcac tgatgccttc gctcctttcg agtctcaagt tacctgggtt	1080
35	gctctgcggg tcgatactga gaagctacgt gccatgaaga caacgtctga gggattccgc	1140
	aagagagtgg gagacgtcgt cttcaaccac aaggccggat acaccattca tcgtctgggtg	1200
40	ttggtcggtg acgacattga tgtctatgaa ggaaaggatg tgctctgggc gttctccacc	1260
	cgttgccgtc ctggtatgga cgagactttg tttgaggatg ttcgtgggtt ccccttgatt	1320
	ccgtatatgg gacacgggaa tggggccgcc caccgcggcg gaaagggtgt gtccgacgct	1380
45	cttatgccga ctgagtacac cactggtcgc aactgggagg ctgctgactt caaccaatct	1440
	tatcccaggg atctgaagca gaagggtgtg gacaactgga cgaagatggg tttcagcaac	1500
50	taa	1503

	<210> 10
	<211> 500
55	<212> PRT
	<213> Aspergillus niger
	<400> 10

60	Met Ser Ala Gln Pro Ala His Leu Cys Phe Arg Ser Phe Val Glu Ala
	1 5 10 15

65	Leu Lys Val Asp Asn Asp Leu Val Glu Ile Asn Thr Pro Ile Asp Pro
	20 25 30

Asn Leu Glu Ala Ala Ala Ile Thr Arg Arg Val Cys Glu Thr Asn Asp
54

ES 2 709 217 T3

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

ES 2 709 217 T3

	305		310		315		320									
5	Asp	Glu	Thr	His	Thr 325	Met	Ile	Gly	Ser	Leu 330	Ala	Ala	Ala	Glu	Ile 335	Arg
10	Lys	Leu	Cys	Gln 340	Gln	Asn	Asp	Leu	Pro 345	Ile	Thr	Asp	Ala	Phe 350	Ala	Pro
15	Phe	Glu	Ser 355	Gln	Val	Thr	Trp	Val 360	Ala	Leu	Arg	Val	Asp 365	Thr	Glu	Lys
20	Leu	Arg 370	Ala	Met	Lys	Thr	Thr 375	Ser	Glu	Gly	Phe	Arg 380	Lys	Arg	Val	Gly
25	Asp 385	Val	Val	Phe	Asn	His 390	Lys	Ala	Gly	Tyr	Thr 395	Ile	His	Arg	Leu	Val 400
30	Leu	Val	Gly	Asp	Asp 405	Ile	Asp	Val	Tyr	Glu 410	Gly	Lys	Asp	Val	Leu	Trp 415
35	Ala	Phe	Ser	Thr 420	Arg	Cys	Arg	Pro	Gly 425	Met	Asp	Glu	Thr	Leu 430	Phe	Glu
40	Asp	Val	Arg 435	Gly	Phe	Pro	Leu	Ile 440	Pro	Tyr	Met	Gly	His 445	Gly	Asn	Gly
45	Pro	Ala 450	His	Arg	Gly	Gly	Lys 455	Val	Val	Ser	Asp	Ala 460	Leu	Met	Pro	Thr
50	Glu	Tyr	Thr	Thr	Gly	Arg 470	Asn	Trp	Glu	Ala	Ala 475	Asp	Phe	Asn	Gln	Ser 480
55	Tyr	Pro	Glu	Asp	Leu 485	Lys	Gln	Lys	Val	Leu 490	Asp	Asn	Trp	Thr	Lys 495	Met
60	Gly	Phe	Ser	Asn 500												
65	<210>	11														
	<211>	3774														
	<212>	ADN														
	<213>	Secuencia artificial														
	<220>															
	<223>	Sintética														
	<400>	11														
	atgtcgtact	accatcacca	tcaccatcac	ctcgaatcaa	caagtttgta	caaaaaagca										60
	ggctccgcgg	ccgccccctt	caccatgagg	aagctaaatc	cagctttaga	atttagagac										120
	tttatccagg	tcttaaaaga	tgaagatgac	ttaatcgaaa	ttaccgaaga	gattgatcca										180
	aatctcgaag	taggtgcaat	tatgaggaag	gcctatgaat	cccacttacc	agccccgtta										240

	tttaaaaatc	tcaaaggtgc	ttcgaaggat	cttttcagca	ttttaggttg	cccagccggt	300
	ttgagaagta	aggagaaagg	agatcatggt	agaattgccc	atcatctggg	gctcgacca	360
5	aaaacaacta	tcaaggaaat	catagattat	ttgctggagt	gtaaggagaa	ggaacctctc	420
	cccccaatca	ctgttcctgt	gtcatctgca	ccttgtaaaa	cacatatact	ttctgaagaa	480
10	aaaatacatc	tacaaagcct	gccaacacca	tatctacatg	tttcagacgg	tggcaagtac	540
	ttacaaacgt	acggaatgtg	gattcttcaa	actccagata	aaaaatggac	taattggtca	600
	attgctagag	gtatggttgt	agatgacaag	catatcactg	gtctggtaat	taaaccacaa	660
15	catattagac	aaattgctga	ctcttgggca	gcaattggaa	aagcaaata	aattcctttc	720
	gcgttatgtt	ttggcgttcc	cccagcagct	attttagtta	gttccatgcc	aattcctgaa	780
20	gggtgtttctg	aatcggatta	tgttggcgca	atcttgggtg	agtcggttcc	agtagtaaaa	840
	tgtgagacca	acgatttaat	ggttcctgca	acgagtgaga	tggatattga	gggtactttg	900
	tccttaacag	atacacatct	ggaaggccca	tttggtgaga	tgcatggata	tgttttcaaa	960
25	agccaaggtc	atccttgtcc	attgtacact	gtcaaggcta	tgagttacag	agacaatgct	1020
	attctacctg	tttcgaaccc	cggctcttgt	acggatgaga	cacatacctt	gattgggtca	1080
30	ctagtggcta	ctgaggccaa	ggagctggct	attgaatctg	gcttgccaat	tctggatgcc	1140
	tttatgcctt	atgaggctca	ggctctttgg	cttatcttaa	aggtggattt	gaaagggctg	1200
	caagcattga	agacaacgcc	tgaagaattt	tgtagaagg	taggtgatat	ttactttagg	1260
35	acaaaagttg	gttttatagt	ccatgaaata	attttgggtg	cagatgatat	cgacatattt	1320
	aacttcaaag	aagtcactct	ggcctacgtt	acaagacata	cacctgttgc	agatcagatg	1380
40	gcttttgatg	atgtcacttc	ttttcctttg	gctccctttg	tttcgcagtc	atccagaagt	1440
	aagactatga	aaggtggaaa	gtgctgtact	aattgcatat	ttagacagca	atatgagcgc	1500
	agttttgact	acataacttg	taattttgaa	aagggatatc	caaaaggatt	agttgacaaa	1560
45	gtaaatgaaa	attggaaaag	gtacggatat	aaatctggtg	gctcaggagg	ctctggagcc	1620
	atggatcaaa	tcgaagcaat	gttgtgcggc	ggaggagaga	agacaaaagt	ggcggttact	1680
50	acgaagactt	tggcagatcc	attgaattgg	ggtttagcag	cggatcaaata	gaaaggaagt	1740
	catttagatg	aagtgaagaa	gatggctgaa	gagtatcgta	gaccagtcgt	gaatcttggc	1800
	ggagaaacac	tgacgatcgg	acaagttgct	gccatctcca	ccgtaggagg	cagcgttaag	1860
55	gttgagttag	cggagacttc	aagagccggt	gtgaaagcta	gcagtgattg	ggttatggag	1920
	agcatgaaca	aaggtactga	cagttacgga	gtcaccaccg	gctttggtgc	tacttctcac	1980
60	cggagaacca	aaaacggcac	cgcattacaa	acagaactca	ttagattttt	gaacgccgga	2040
	atattcggaa	acacgaagga	gacatgtcac	acactgccgc	aatccgccac	aagagccgcc	2100
	atgctcgtca	gagtcaacac	tcttctccaa	ggatactccg	ggatccgatt	cgagatcctc	2160
65	gaagcgatta	caagtctcct	caaccacaac	atctctccgt	cactacctct	ccgtggaacc	2220
	attaccgcct	ccggcgatct	cgttcctctc	tcttacatcg	ccggacttct	caccggccgt	2280

ES 2 709 217 T3

cctaattcca aagccaccgg tcccgcggt gaatcgctaa ccgcgaaaga agcttttgag 2340
 aaagccggaa tcagtactgg attcttcgat ttacaaccta aggaagggtt agctctcggt 2400
 5 aatggcacgg cggttggatc tggaatggcg tcgatgggtc tattcgaagc gaatgtccaa 2460
 gcggtgttag cggagggttt atcagcgatc ttcgcggagg ttatgagcgg gaaacctgag 2520
 10 tttaccgatc atctgactca tcgttttaaa catcatcccc gacaaatcga agcggcggcg 2580
 ataatggagc acatactcga cggaaagtca tacatgaaat tagctcaaaa gggtcacgag 2640
 atggatccat tgcagaaacc aaaacaagat cgttacgctc ttcgtacatc tcctcaatgg 2700
 15 ctaggtcctc aaattgaagt aatccgtcaa gctacgaaat cgatagagcg tgaaatcaac 2760
 tccgttaacg ataatccgtt gatcgatggt tcgaggaaca aggcgattca cggtggtaac 2820
 20 ttccaaggaa caccaatcgg agtttctatg gataacacga gattggcgat tgctgcgatt 2880
 gggaagctaa tgtttgctca attctctgag cttgttaatg atttctacaa caatggactt 2940
 ccttcgaatc taactgcttc gagtaatcca agtttggtt atggattcaa aggagcagag 3000
 25 attgctatgg cttcttattg ttctgagctt caatacttgg ctaatccagt cacaagccat 3060
 gttcaatcag ctgagcaaca taatcaagat gtgaactctc ttggtttgat ctcgtctcgt 3120
 aaaacatctg aagctgtgga tattcttaag ctaatgtcaa caacgttcct tgtggggata 3180
 30 tgtcaagctg ttgatttgag acatttgag gagaatctga gacaaactgt gaagaacaca 3240
 gtttctcaag ttgctaagaa agtgtaacc actggaatca acggtgagtt acatccgtca 3300
 35 aggttttgcg agaaggactt gcttaagggt gttgatcgtg agcaagtgtt cacgtatgtg 3360
 gatgatcctt gtagcgctac gtaccggtt atgcagagac taagacaagt tattgttgat 3420
 cacgctttgt ccaacggtga gactgagaag aatgcagtga cttcgatctt tcaaaagatt 3480
 40 ggagcttttg aagaggagct taaggctgtg cttccaaagg aagttgaagc ggctagagcg 3540
 gcttatggga atggaactgc gccgattcct aaccggatta aggaatgtag gtcgtatccg 3600
 45 ttgtataggt tcgtgagggg agagcttgga acgaagtgtg tgactggaga aaaggttggtg 3660
 tctccgggag aggagtttga taaggctctc actgctatgt gtgaaggtaa acttattgat 3720
 50 ccgttgatgg attgtctcaa ggaatggaac ggagctccga ttccgatttg ctaa 3774

<210> 12
 <211> 1257
 <212> PRT
 55 <213> Secuencia artificial

<220>
 <223> Sintética

60 <400> 12

Met Ser Tyr Tyr His His His His His His Leu Glu Ser Thr Ser Leu
 1 5 10 15

65 Tyr Lys Lys Ala Gly Ser Ala Ala Ala Pro Phe Thr Met Arg Lys Leu
 20 25 30

ES 2 709 217 T3

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

ES 2 709 217 T3

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

ES 2 709 217 T3

	Met	Lys	Gly	Ser	His	Leu	Asp	Glu	Val	Lys	Lys	Met	Val	Glu	Glu	Tyr
				580					585					590		
5	Arg	Arg	Pro	Val	Val	Asn	Leu	Gly	Gly	Glu	Thr	Leu	Thr	Ile	Gly	Gln
			595					600					605			
10	Val	Ala	Ala	Ile	Ser	Thr	Val	Gly	Gly	Ser	Val	Lys	Val	Glu	Leu	Ala
		610					615					620				
15	Glu	Thr	Ser	Arg	Ala	Gly	Val	Lys	Ala	Ser	Ser	Asp	Trp	Val	Met	Glu
	625					630					635					640
20	Ser	Met	Asn	Lys	Gly	Thr	Asp	Ser	Tyr	Gly	Val	Thr	Thr	Gly	Phe	Gly
					645					650					655	
	Ala	Thr	Ser	His	Arg	Arg	Thr	Lys	Asn	Gly	Thr	Ala	Leu	Gln	Thr	Glu
				660					665					670		
25	Leu	Ile	Arg	Phe	Leu	Asn	Ala	Gly	Ile	Phe	Gly	Asn	Thr	Lys	Glu	Thr
			675					680					685			
30	Cys	His	Thr	Leu	Pro	Gln	Ser	Ala	Thr	Arg	Ala	Ala	Met	Leu	Val	Arg
		690					695					700				
35	Val	Asn	Thr	Leu	Leu	Gln	Gly	Tyr	Ser	Gly	Ile	Arg	Phe	Glu	Ile	Leu
	705					710					715					720
40	Glu	Ala	Ile	Thr	Ser	Leu	Leu	Asn	His	Asn	Ile	Ser	Pro	Ser	Leu	Pro
					725					730					735	
	Leu	Arg	Gly	Thr	Ile	Thr	Ala	Ser	Gly	Asp	Leu	Val	Pro	Leu	Ser	Tyr
				740					745					750		
45	Ile	Ala	Gly	Leu	Leu	Thr	Gly	Arg	Pro	Asn	Ser	Lys	Ala	Thr	Gly	Pro
			755					760					765			
50	Asp	Gly	Glu	Ser	Leu	Thr	Ala	Lys	Glu	Ala	Phe	Glu	Lys	Ala	Gly	Ile
	770						775					780				
55	Ser	Thr	Gly	Phe	Phe	Asp	Leu	Gln	Pro	Lys	Glu	Gly	Leu	Ala	Leu	Val
	785					790					795					800
60	Asn	Gly	Thr	Ala	Val	Gly	Ser	Gly	Met	Ala	Ser	Met	Val	Leu	Phe	Glu
					805					810					815	
	Ala	Asn	Val	Gln	Ala	Val	Leu	Ala	Glu	Val	Leu	Ser	Ala	Ile	Phe	Ala
				820					825					830		
65	Glu	Val	Met	Ser	Gly	Lys	Pro	Glu	Phe	Thr	Asp	His	Leu	Thr	His	Arg
			835					840					845			

ES 2 709 217 T3

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

ES 2 709 217 T3

	Glu	Gln	Val	Phe	Thr	Tyr	Val	Asp	Asp	Pro	Cys	Ser	Ala	Thr	Tyr	
	1115						1120					1125				
5	Pro	Leu	Met	Gln	Arg	Leu	Arg	Gln	Val	Ile	Val	Asp	His	Ala	Leu	
	1130						1135					1140				
10	Ser	Asn	Gly	Glu	Thr	Glu	Lys	Asn	Ala	Val	Thr	Ser	Ile	Phe	Gln	
	1145						1150					1155				
15	Lys	Ile	Gly	Ala	Phe	Glu	Glu	Glu	Leu	Lys	Ala	Val	Leu	Pro	Lys	
	1160						1165					1170				
20	Glu	Val	Glu	Ala	Ala	Arg	Ala	Ala	Tyr	Gly	Asn	Gly	Thr	Ala	Pro	
	1175						1180					1185				
	Ile	Pro	Asn	Arg	Ile	Lys	Glu	Cys	Arg	Ser	Tyr	Pro	Leu	Tyr	Arg	
	1190						1195					1200				
25	Phe	Val	Arg	Glu	Glu	Leu	Gly	Thr	Lys	Leu	Leu	Thr	Gly	Glu	Lys	
	1205						1210					1215				
30	Val	Val	Ser	Pro	Gly	Glu	Glu	Phe	Asp	Lys	Val	Phe	Thr	Ala	Met	
	1220						1225					1230				
35	Cys	Glu	Gly	Lys	Leu	Ile	Asp	Pro	Leu	Met	Asp	Cys	Leu	Lys	Glu	
	1235						1240					1245				
40	Trp	Asn	Gly	Ala	Pro	Ile	Pro	Ile	Cys							
	1250						1255									
	<210>	13														
	<211>	3774														
	<212>	ADN														
45	<213>	Secuencia artificial														
	<220>															
	<223>	Sintética														
50	<400>	13														
	atgtcgtact	accatcacca	tcaccatcac	ctcgaatcaa	caagtttgta	caaaaaagca										60
	ggctccgcgg	ccgccccctt	caccatgagg	aagctaaatc	cagctttaga	atttagagac										120
55	tttatccagg	tcttaaaaga	tgaagatgac	ttaatcgaaa	ttaccgaaga	gattgatcca										180
	aatctcgaag	taggtgcaat	tatgaggaag	gcctatgaat	cccacttacc	agccccgtta										240
	tttaaaaatc	tcaaaggtgc	ttcgaaggat	cttttcagca	ttttaggttg	cccagccggt										300
60	ttgagaagta	aggagaaagg	agatcatggt	agaattgccc	atcatctggg	gctcgacca										360
	aaaacaacta	tcaaggaaat	catagattat	ttgctggagt	gtaaggagaa	ggaacctctc										420
65	cccccaatca	ctgttcctgt	gtcatctgca	ccttgtaaaa	cacatatact	ttctgaagaa										480
	aaaatacatc	tacaaagcct	gccaacacca	tatctacatg	tttcagacgg	tggcaagtac										540

	ttacaaacgt	acggaatgtg	gattcttcaa	actccagata	aaaaatggac	taattgggtca	600
	attgctagag	gtatgggtgt	agatgacaag	catatcactg	gtctggtaat	tgagccacaa	660
5	catattagac	aaattgctga	ctcttgggca	gcaattggaa	aagcaaatga	aattcctttc	720
	gcgttatgtt	ttggcgttcc	cccagcagct	atttttagtta	gttccatgcc	aattcctgaa	780
10	gggtgtttctg	aatcggatta	tggtggcgca	atcttgggtg	agtcggttcc	agtagtaaaa	840
	tgtgagacca	acgatttaat	ggttcctgca	acgagtgaga	tggtatttga	gggtactttg	900
	tccttaacag	atacacatct	ggaaggccca	tttggtgaga	tgcatggata	tgttttcaaa	960
15	agccaaggtc	atccttgtcc	attgtacact	gtcaaggcta	tgagttacag	agacaatgct	1020
	attctacctg	tttcgaaccc	cggctcttgt	acggatgaga	cacatacctt	gattggttca	1080
	ctagtggcta	ctgaggccaa	ggagctggct	attgaatctg	gcttgccaat	tctggatgcc	1140
20	tttatgcctt	atgaggctca	ggctcttttg	cttatcttaa	aggtggattt	gaaagggctg	1200
	caagcattga	agacaacgcc	tgaagaattt	tgtaagaagg	taggtgatat	ttactttagg	1260
25	acaaaagttg	gttttatagt	ccatgaaata	attttggtgg	cagatgatat	cgacatattt	1320
	aacttcaaag	aagtcactctg	ggcctacggt	acaagacata	cacctgttgc	agatcagatg	1380
	gcttttgatg	atgtcacttc	ttttcctttg	gctccctttg	tttcgcagtc	atccagaagt	1440
30	aagactatga	aaggtggaaa	gtgcgttact	aattgcata	ttagacagca	atatgagcgc	1500
	agttttgact	acataacttg	taattttgaa	aagggatatc	caaaaggatt	agttgacaaa	1560
35	gtaaatgaaa	attggaaaag	gtacggatat	aaatctgggtg	gctcaggagg	ctctggagcc	1620
	atggatcaaa	tcgaagcaat	gttgtgcggc	ggaggagaga	agacaaaagt	ggcggttact	1680
	acgaagactt	tggcagatcc	attgaattgg	ggtttagcag	cggatcaaat	gaaaggaagt	1740
40	catttagatg	aagtgaagaa	gatggtcgaa	gagtatcgta	gaccagtcgt	gaatcttggc	1800
	ggagaaacac	tgacgatcgg	acaagttgct	gccatctcca	ccgtaggagg	cagcgттаag	1860
45	gttgagttag	cggagacttc	aagagccgggt	gtgaaagcta	gcagtgattg	ggttatggag	1920
	agcatgaaca	aaggtactga	cagttacgga	gtcaccaccg	gctttggtgc	tacttctcac	1980
	cggagaacca	aaaacggcac	cgcattacaa	acagaactca	ttagattttt	gaacgccgga	2040
50	atattcggaa	acacgaagga	gacatgtcac	acactgccgc	aatccgccac	aagagccgcc	2100
	atgctcgtca	gagtcaacac	tcttctccaa	ggatactccg	ggatccgatt	cgagatcctc	2160
55	gaagcgatta	caagtctcct	caaccacaac	atctctccgt	cactacctct	ccgtggaacc	2220
	attaccgcct	ccggcgatct	cgttcctctc	tcttacatcg	ccggacttct	caccggccgt	2280
	cctaattcca	aagccaccgg	tcccgcgggt	gaatcgctaa	ccgcgaaaga	agcttttgag	2340
60	aaagccggaa	tcagtactgg	attcttcgat	ttacaacct	aggaagggtt	agctctcggt	2400
	aatggcacgg	cggttggatc	tggaatggcg	tcgatggttc	tattcgaagc	gaatgtccaa	2460
65	gcggtgttag	cggaggtttt	atcagcgatc	ttcgcgagg	ttatgagcgg	gaaacctgag	2520
	tttaccgatc	atctgactca	tcgttttaaa	catcatcccc	gacaaatcga	agcggcgggc	2580

ES 2 709 217 T3

ataatggagc acatactcga cggaagctca tacatgaaat tagctcaaaa gggttcacgag 2640
 atggatccat tgcagaaacc aaaacaagat cgttacgctc ttcgtacatc tcctcaatgg 2700
 5 ctaggtcctc aaattgaagt aatccgtcaa gctacgaaat cgatagagcg tgaaatcaac 2760
 tccgttaacg ataatccgtt gatcgaatgt tcgaggaaca aggcgattca cgggtggtaac 2820
 10 ttccaaggaa caccaatcgg agtttctatg gataacacga gattggcgat tgctgcgatt 2880
 gggaagctaa tgtttgctca atttcttgag cttgttaatg atttctacaa caatggactt 2940
 ccttcgaatc taactgcttc gagtaatcca agtttggatt atggattcaa aggagcagag 3000
 15 attgctatgg cttcttattg ttctgagctt caatacttgg ctaatccagt cacaagccat 3060
 gttcaatcag ctgagcaaca taatcaagat gtgaactctc ttggtttgat ctcgtctcgt 3120
 aaaacatctg aagctgtgga tattcttaag ctaatgtcaa caacgttcct tgtggggata 3180
 20 tgtcaagctg ttgatttgag acatttgag gagaatctga gacaaactgt gaagaacaca 3240
 gtttctcaag ttgctaagaa agtgtaacc actggaatca acggtgagtt acatccgtca 3300
 25 aggttttgcg agaaggactt gcttaagggtt gttgatcgtg agcaagtgtt cacgtatgtg 3360
 gatgatcctt gtagcgctac gtacccgttg atgcagagac taagacaagt tattgttgat 3420
 cacgctttgt ccaacggtga gactgagaag aatgcagtga cttcgatctt tcaaaagatt 3480
 30 ggagcttttg aagaggagct taaggctgtg cttccaaagg aagttgaagc ggctagagcg 3540
 gcttatggga atggaactgc gccgattcct aaccggatta aggaatgtag gtcgtatccg 3600
 35 ttgtataggt tcgtgaggga agagcttgga acgaagttgt tgactggaga aaaggttggtg 3660
 tctccgggag aggagtttga taaggctctc actgctatgt gtgaaggtaa acttattgat 3720
 40 ccgttgatgg attgtctcaa ggaatggaac ggagctccga ttccgatttg ctaa 3774

<210> 14
 <211> 1257
 <212> PRT
 45 <213> Secuencia artificial

<220>
 <223> Sintética

50 <400> 14

Met Ser Tyr Tyr His His His His His His Leu Glu Ser Thr Ser Leu
 1 5 10 15

55 Tyr Lys Lys Ala Gly Ser Ala Ala Ala Pro Phe Thr Met Arg Lys Leu
 20 25 30

60 Asn Pro Ala Leu Glu Phe Arg Asp Phe Ile Gln Val Leu Lys Asp Glu
 35 40 45

65 Asp Asp Leu Ile Glu Ile Thr Glu Glu Ile Asp Pro Asn Leu Glu Val
 50 55 60

Gly Ala Ile Met Arg Lys Ala Tyr Glu Ser His Leu Pro Ala Pro Leu
 65

ES 2 709 217 T3

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

ES 2 709 217 T3

	340							345							350						
5	Glu	Thr	His 355	Thr	Leu	Ile	Gly	Ser 360	Leu	Val	Ala	Thr	Glu 365	Ala	Lys	Glu					
10	Leu	Ala 370	Ile	Glu	Ser	Gly	Leu 375	Pro	Ile	Leu	Asp	Ala 380	Phe	Met	Pro	Tyr					
15	Glu 385	Ala	Gln	Ala	Leu	Trp 390	Leu	Ile	Leu	Lys	Val 395	Asp	Leu	Lys	Gly	Leu 400					
20	Gln	Ala	Leu	Lys	Thr 405	Thr	Pro	Glu	Glu	Phe 410	Cys	Lys	Lys	Val	Gly 415	Asp					
25	Ile	Tyr	Phe	Arg 420	Thr	Lys	Val	Gly	Phe 425	Ile	Val	His	Glu	Ile 430	Ile	Leu					
30	Val	Ala	Asp 435	Asp	Ile	Asp	Ile	Phe 440	Asn	Phe	Lys	Glu	Val 445	Ile	Trp	Ala					
35	Tyr	Val 450	Thr	Arg	His	Thr	Pro 455	Val	Ala	Asp	Gln	Met 460	Ala	Phe	Asp	Asp					
40	Val 465	Thr	Ser	Phe	Pro	Leu 470	Ala	Pro	Phe	Val	Ser 475	Gln	Ser	Ser	Arg	Ser 480					
45	Lys	Thr	Met	Lys	Gly 485	Gly	Lys	Cys	Val	Thr 490	Asn	Cys	Ile	Phe	Arg 495	Gln					
50	Gln	Tyr	Glu	Arg 500	Ser	Phe	Asp	Tyr	Ile 505	Thr	Cys	Asn	Phe	Glu 510	Lys	Gly					
55	Tyr	Pro	Lys 515	Gly	Leu	Val	Asp	Lys 520	Val	Asn	Glu	Asn	Trp 525	Lys	Arg	Tyr					
60	Gly	Tyr 530	Lys	Ser	Gly	Gly	Ser 535	Gly	Gly	Ser	Gly	Ala 540	Met	Asp	Gln	Ile					
65	Glu 545	Ala	Met	Leu	Cys	Gly 550	Gly	Gly	Glu	Lys	Thr 555	Lys	Val	Ala	Val	Thr 560					
70	Thr	Lys	Thr	Leu	Ala 565	Asp	Pro	Leu	Asn	Trp 570	Gly	Leu	Ala	Ala	Asp 575	Gln					
75	Met	Lys	Gly	Ser 580	His	Leu	Asp	Glu	Val 585	Lys	Lys	Met	Val	Glu 590	Glu	Tyr					
80	Arg	Arg	Pro 595	Val	Val	Asn	Leu	Gly 600	Gly	Glu	Thr	Leu	Thr 605	Ile	Gly	Gln					
85	Val	Ala	Ala	Ile	Ser	Thr	Val	Gly	Gly	Ser	Val	Lys	Val	Glu	Leu	Ala					
	67																				

ES 2 709 217 T3

	610					615					620				
5	Glu 625	Thr	Ser	Arg	Ala	Gly 630	Val	Lys	Ala	Ser	Ser 635	Asp	Trp	Val	Glu 640
10	Ser	Met	Asn	Lys	Gly 645	Thr	Asp	Ser	Tyr	Gly 650	Val	Thr	Thr	Gly	Phe 655
15	Ala	Thr	Ser	His 660	Arg	Arg	Thr	Lys	Asn 665	Gly	Thr	Ala	Leu	Gln 670	Glu
20	Leu	Ile	Arg 675	Phe	Leu	Asn	Ala	Gly 680	Ile	Phe	Gly	Asn	Thr 685	Lys	Thr
25	Cys	His 690	Thr	Leu	Pro	Gln	Ser 695	Ala	Thr	Arg	Ala	Ala 700	Met	Leu	Val
30	Val 705	Asn	Thr	Leu	Leu	Gln 710	Gly	Tyr	Ser	Gly	Ile 715	Arg	Phe	Glu	Ile
35	Glu	Ala	Ile	Thr	Ser 725	Leu	Leu	Asn	His	Asn 730	Ile	Ser	Pro	Ser	Leu 735
40	Leu	Arg	Gly	Thr 740	Ile	Thr	Ala	Ser	Gly 745	Asp	Leu	Val	Pro	Leu 750	Tyr
45	Ile	Ala	Gly 755	Leu	Leu	Thr	Gly	Arg 760	Pro	Asn	Ser	Lys	Ala 765	Thr	Pro
50	Asp	Gly 770	Glu	Ser	Leu	Thr	Ala 775	Lys	Glu	Ala	Phe	Glu 780	Lys	Ala	Ile
55	Ser 785	Thr	Gly	Phe	Phe	Asp 790	Leu	Gln	Pro	Lys	Glu 795	Gly	Leu	Ala	Val 800
60	Asn	Gly	Thr	Ala	Val 805	Gly	Ser	Gly	Met	Ala 810	Ser	Met	Val	Leu	Phe
65	Ala	Asn	Val	Gln 820	Ala	Val	Leu	Ala	Glu 825	Val	Leu	Ser	Ala	Ile 830	Ala
70	Glu	Val	Met 835	Ser	Gly	Lys	Pro	Glu 840	Phe	Thr	Asp	His	Leu 845	Thr	Arg
75	Leu	Lys 850	His	His	Pro	Gly	Gln 855	Ile	Glu	Ala	Ala	Ala 860	Ile	Met	Glu
80	Ile	Leu	Asp	Gly	Ser	Ser 870	Tyr	Met	Lys	Leu	Ala 875	Gln	Lys	Val	Glu 880
85	Met	Asp	Pro	Leu	Gln	Lys	Pro	Lys	Gln	Asp	Arg	Tyr	Ala	Leu	Thr

ES 2 709 217 T3

	885	890	895
5	Ser Pro Gln Trp 900 Leu Gly Pro Gln Ile 905 Glu Val Ile Arg Gln 910 Ala Thr		
10	Lys Ser Ile 915 Glu Arg Glu Ile Asn 920 Ser Val Asn Asp Asn 925 Pro Leu Ile		
15	Asp Val 930 Ser Arg Asn Lys Ala 935 Ile His Gly Gly Asn 940 Phe Gln Gly Thr		
20	Pro 945 Ile Gly Val Ser Met 950 Asp Asn Thr Arg Leu 955 Ala Ile Ala Ala Ile 960		
25	Gly Lys Leu Met Phe 965 Ala Gln Phe Ser Glu 970 Leu Val Asn Asp Phe 975 Tyr		
30	Asn Asn Gly 980 Leu Pro Ser Asn Leu Thr 985 Ala Ser Ser Asn 990 Pro Ser Leu		
35	Asp Tyr Gly 995 Phe Lys Gly Ala Glu 1000 Ile Ala Met Ala Ser 1005 Tyr Cys Ser		
40	Glu Leu 1010 Gln Tyr Leu Ala Asn 1015 Pro Val Thr Ser His 1020 Val Gln Ser		
45	Ala Glu 1025 Gln His Asn Gln Asp 1030 Val Asn Ser Leu Gly 1035 Leu Ile Ser		
50	Ser Arg 1040 Lys Thr Ser Glu Ala 1045 Val Asp Ile Leu Lys 1050 Leu Met Ser		
55	Thr Thr 1055 Phe Leu Val Gly Ile 1060 Cys Gln Ala Val Asp 1065 Leu Arg His		
60	Leu Glu 1070 Glu Asn Leu Arg Gln 1075 Thr Val Lys Asn Thr 1080 Val Ser Gln		
65	Val Ala 1085 Lys Lys Val Leu Thr 1090 Thr Gly Ile Asn Gly 1095 Glu Leu His		
70	Pro Ser 1100 Arg Phe Cys Glu Lys 1105 Asp Leu Leu Lys Val 1110 Val Asp Arg		
75	Glu Gln 1115 Val Phe Thr Tyr Val 1120 Asp Asp Pro Cys Ser 1125 Ala Thr Tyr		
80	Pro Leu 1130 Met Gln Arg Leu Arg 1135 Gln Val Ile Val Asp 1140 His Ala Leu		
85	Ser Asn Gly Glu Thr Glu Lys Asn Ala Val Thr Ser Ile Phe Gln		

ES 2 709 217 T3

	1145		1150		1155	
5	Lys Ile 1160	Gly Ala Phe Glu	Glu 1165	Glu Leu Lys Ala	Val 1170	Leu Pro Lys
10	Glu Val 1175	Glu Ala Ala Arg	Ala 1180	Ala Tyr Gly Asn	Gly 1185	Thr Ala Pro
15	Ile Pro 1190	Asn Arg Ile Lys	Glu 1195	Cys Arg Ser Tyr	Pro 1200	Leu Tyr Arg
20	Phe Val 1205	Arg Glu Glu Leu	Gly 1210	Thr Lys Leu Leu	Thr 1215	Gly Glu Lys
25	Val Val 1220	Ser Pro Gly Glu	Glu 1225	Phe Asp Lys Val	Phe 1230	Thr Ala Met
30	Cys Glu 1235	Gly Lys Leu Ile	Asp 1240	Pro Leu Met Asp	Cys 1245	Leu Lys Glu
35	Trp Asn 1250	Gly Ala Pro Ile	Pro 1255	Ile Cys		
	<210> 15					
	<211> 1512					
	<212> ADN					
	<213> Secuencia artificial					
	<220>					
	<223> sintética					
40	<400> 15					
	atgaggaagc taaatccagc tttagaattt agagacttta tccaggctctt aaaagatgaa					60
	gatgacttaa tcgaaattac cgaagagatt gatccaaatc tcgaagtagg tgcaattatg					120
45	aggaaggcct atgaatccca cttaccagcc ccggtatttta aaaatctcaa aggtgcttcg					180
	aaggatcttt tcagcatttt aggttgccca gccggtttga gaagtaagga gaaaggagat					240
50	catggtagaa ttgcccata tctggggctc gacccaaaaa caactatcaa ggaaatcata					300
	gattattttgc tggagtgtaa ggagaaggaa cctctccccc caatcactgt tcctgtgtca					360
	tctgcacctt gtaaaacaca tatactttct gaagaaaaaa tacatctaca aagcctgcca					420
55	acaccatata tacatgtttc agacggtggc aagtacttac aaacgtacgg aatgtggatt					480
	cttcaaactc cagataaaaa atggactaat tgggtcaattg ctagaggat ggtttagat					540
60	gacaagcata tcaactggtct ggtaattgag ccacaacata ttagacaaat tgctgactct					600
	tgggcagcaa ttggaaaagc aaatgaaatt cttttcgcg tttgttttg cgttcccca					660
	gcagctatct tagttagttc catgccaaat cctgaagggtg tttctgaatc ggattatggt					720
65	ggcgcaatct tgggtgagtc ggttccagta gtaaaatgtg agaccaacga tttaatggtt					780
	cctgcaacga gtgagatggt atttgagggt actttgtcct taacagatac acatctggaa					840

ES 2 709 217 T3

ggcccatttg gtgagatgca tggatatggt ttcaaaagcc aaggtcatcc ttgtccattg 900
 tacactgtca aggctatgag ttacagagac aatgctattc tacctgtttc gaaccccggt 960
 5 ctttgtacgg atgagacaca taccttgatt ggttcactag tggctactga ggccaaggag 1020
 ctggctattg aatctggctt gccaattctg gatgccttta tgccttatga ggctcaggct 1080
 10 ctttggctta tcttaaaggt ggatttgaaa gggctgcaag cattgaagac aacgcctgaa 1140
 gaattttgta agaaggtagg tgatatttac tttaggacaa aagttggttt tatagtccat 1200
 gaaataattt tgggtggcaga tgatatcgac atatttaact tcaaagaagt catctggggc 1260
 15 tacgttacaa gacatacacc tgttgcagat cagatggctt ttgatgatgt cacttctttt 1320
 cctttggctc cctttgtttc gcagtcatcc agaagtaaga ctatgaaagg tggaaagtgc 1380
 20 gttactaatt gcatatttag acagcaatat gagcgcagtt ttgactacat aacttgtaat 1440
 tttgaaaagg gatatccaaa aggattagtt gacaaagtaa atgaaaattg gaaaaggtag 1500
 ggatataaat aa 1512
 25
 <210> 16
 <211> 503
 <212> PRT
 <213> Secuencia artificial
 30
 <220>
 <223> Sintética
 <400> 16
 35 Met Arg Lys Leu Asn Pro Ala Leu Glu Phe Arg Asp Phe Ile Gln Val
 1 5 10 15
 40 Leu Lys Asp Glu Asp Asp Leu Ile Glu Ile Thr Glu Glu Ile Asp Pro
 20 25 30
 45 Asn Leu Glu Val Gly Ala Ile Met Arg Lys Ala Tyr Glu Ser His Leu
 35 40 45
 50 Pro Ala Pro Leu Phe Lys Asn Leu Lys Gly Ala Ser Lys Asp Leu Phe
 50 55 60
 Ser Ile Leu Gly Cys Pro Ala Gly Leu Arg Ser Lys Glu Lys Gly Asp
 65 70 75 80
 55 His Gly Arg Ile Ala His His Leu Gly Leu Asp Pro Lys Thr Thr Ile
 85 90 95
 60 Lys Glu Ile Ile Asp Tyr Leu Leu Glu Cys Lys Glu Lys Glu Pro Leu
 100 105 110
 65 Pro Pro Ile Thr Val Pro Val Ser Ser Ala Pro Cys Lys Thr His Ile
 115 120 125
 Leu Ser Glu Glu Lys Ile His Leu Gln Ser Leu Pro Thr Pro Tyr Leu
 71

ES 2 709 217 T3

	130					135					140					
5	His 145	Val	Ser	Asp	Gly	Gly 150	Lys	Tyr	Leu	Gln	Thr 155	Tyr	Gly	Met	Trp	Ile 160
10	Leu	Gln	Thr	Pro	Asp 165	Lys	Lys	Trp	Thr	Asn 170	Trp	Ser	Ile	Ala	Arg 175	Gly
15	Met	Val	Val	Asp 180	Asp	Lys	His	Ile	Thr 185	Gly	Leu	Val	Ile	Glu 190	Pro	Gln
20	His	Ile	Arg 195	Gln	Ile	Ala	Asp	Ser 200	Trp	Ala	Ala	Ile	Gly 205	Lys	Ala	Asn
25	Glu	Ile 210	Pro	Phe	Ala	Leu	Cys 215	Phe	Gly	Val	Pro	Pro 220	Ala	Ala	Ile	Leu
30	Val 225	Ser	Ser	Met	Pro	Ile 230	Pro	Glu	Gly	Val	Ser 235	Glu	Ser	Asp	Tyr	Val 240
35	Gly	Ala	Ile	Leu	Gly 245	Glu	Ser	Val	Pro	Val 250	Val	Lys	Cys	Glu	Thr 255	Asn
40	Asp	Leu	Met	Val 260	Pro	Ala	Thr	Ser	Glu 265	Met	Val	Phe	Glu	Gly 270	Thr	Leu
45	Ser	Leu	Thr 275	Asp	Thr	His	Leu	Glu 280	Gly	Pro	Phe	Gly	Glu 285	Met	His	Gly
50	Tyr	Val 290	Phe	Lys	Ser	Gln	Gly 295	His	Pro	Cys	Pro	Leu 300	Tyr	Thr	Val	Lys
55	Ala 305	Met	Ser	Tyr	Arg	Asp 310	Asn	Ala	Ile	Leu	Pro 315	Val	Ser	Asn	Pro	Gly 320
60	Leu	Cys	Thr	Asp	Glu 325	Thr	His	Thr	Leu	Ile 330	Gly	Ser	Leu	Val	Ala 335	Thr
65	Glu	Ala	Lys	Glu 340	Leu	Ala	Ile	Glu	Ser 345	Gly	Leu	Pro	Ile	Leu 350	Asp	Ala
70	Phe	Met	Pro 355	Tyr	Glu	Ala	Gln	Ala 360	Leu	Trp	Leu	Ile	Leu 365	Lys	Val	Asp
75	Leu	Lys 370	Gly	Leu	Gln	Ala	Leu 375	Lys	Thr	Thr	Pro	Glu 380	Glu	Phe	Cys	Lys
80	Lys 385	Val	Gly	Asp	Ile	Tyr 390	Phe	Arg	Thr	Lys	Val 395	Gly	Phe	Ile	Val	His 400
85	Glu	Ile	Ile	Leu	Val	Ala	Asp	Asp	Ile	Asp	Ile	Phe	Asn	Phe	Lys	Glu

ES 2 709 217 T3

	405	410	415	
5	Val Ile Trp Ala Tyr Val Thr Arg His Thr Pro Val Ala Asp Gln Met 420 425 430			
10	Ala Phe Asp Asp Val Thr Ser Phe Pro Leu Ala Pro Phe Val Ser Gln 435 440 445			
15	Ser Ser Arg Ser Lys Thr Met Lys Gly Gly Lys Cys Val Thr Asn Cys 450 455 460			
20	Ile Phe Arg Gln Gln Tyr Glu Arg Ser Phe Asp Tyr Ile Thr Cys Asn 465 470 475 480			
25	Phe Glu Lys Gly Tyr Pro Lys Gly Leu Val Asp Lys Val Asn Glu Asn 485 490 495			
30	Trp Lys Arg Tyr Gly Tyr Lys 500			
35	<210> 17 <211> 1512 <212> ADN <213> Secuencia artificial			
40	<220> <223> sintética			
45	<400> 17 atgaggaagc taaatccagc tttagaatTT agagacttTa tccaggtcTt aaaagatgaa			60
50	gatgactTaa tcgaaattac cgaagagatt gatccaaatc tcgaagtagg tgcaattatg			120
55	aggaaggcct atgaatccca cttaccagcc ccgttattTa aaaatctcaa aggtgcttcg			180
60	aaggatctTt tcagcattTt aggttgccca gccggtTtga gaagtaagga gaaaggagat			240
65	catggtagaa ttgcccatca tctggggctc gacccaaaaa caactatcaa ggaaatcata			300
70	gattattTgc tggagtgtaa ggagaaggaa cctctccccc caatcactgt tcctgtgtca			360
75	tctgcacctt gtaaaacaca tatactTtct gaagaaaaaa tacatctaca aagcctgcca			420
80	acaccatatc tacatgtTtc agacggtggc aagtactTac aaacgtacgg aatgtggatt			480
85	cttcaaactc cagataaaaa atggactaat tggTcaattg ctagaggTat ggttgtagat			540
90	gacaagcata tcaTggtcTt ggtaattTgt ccacaacata ttagacaaat tgctgactcT			600
95	tgggcagcaa ttggaaaagc aaatgaaatt cttTtcgcgt tatgtTttTg cgttccccc			660
100	gcagctattT tagttagTtc catgccaaTt cctgaaggTg tttctgaatc ggattatgtt			720
105	ggcgcaatcT tgggtgagTc ggtTccagTa gtaaaatgtg agaccaacga tTtaatggTt			780
110	cctgcaacga gtgagatggT atttgagggt actTtgtcct taacagatac acatctggaa			840
115	ggcccatTtg gtgagatgca tggatatgtt ttcaaaagcc aaggTcatcc ttgtccattg			900
120	tacactgtca aggctatgag ttacagagac aatgctattc tacctgtTtc gaaccccggt			960

ES 2 709 217 T3

ctttgtacgg atgagacaca taccttgatt gggtcactag tggctactga ggccaaggag 1020
 ctggctattg aatctggcctt gccaatctctg gatgccttta tgccttatga ggctcaggct 1080
 5 ctttggctta tcttaaaggt ggatttgaaa gggctgcaag cattgaagac aacgcctgaa 1140
 gaattttgta agaaggtagg tgatatttac tttaggacaa aagttgggtt tatagtccat 1200
 10 gaaataattt tgggtggcaga tgatatcgac atattttaact tcaaagaagt catctgggcc 1260
 tacgttaciaa gacatacacc tgttcagat cagatggcct ttgatgatgt cacttctttt 1320
 cctttggctc cctttgtttc gcagtcaccc agaagtaaga ctatgaaagg tggaaagtgc 1380
 15 gttactaatt gcatatttag acagcaatat gagcgcagtt ttgactacat aacttgtaat 1440
 tttgaaaagg gatatccaaa aggattagtt gacaaagtaa atgaaaattg gaaaaggtag 1500
 ggatataaat aa 1512

<210> 18
 <211> 503
 <212> PRT
 25 <213> Secuencia artificial

<220>
 <223> sintética

30 <400> 18

Met Arg Lys Leu Asn Pro Ala Leu Glu Phe Arg Asp Phe Ile Gln Val
1 5 10 15

35 Leu Lys Asp Glu Asp Asp Leu Ile Glu Ile Thr Glu Glu Ile Asp Pro
20 25 30

40 Asn Leu Glu Val Gly Ala Ile Met Arg Lys Ala Tyr Glu Ser His Leu
35 40 45

45 Pro Ala Pro Leu Phe Lys Asn Leu Lys Gly Ala Ser Lys Asp Leu Phe
50 55 60

50 Ser Ile Leu Gly Cys Pro Ala Gly Leu Arg Ser Lys Glu Lys Gly Asp
65 70 75 80

His Gly Arg Ile Ala His His Leu Gly Leu Asp Pro Lys Thr Thr Ile
85 90 95

55 Lys Glu Ile Ile Asp Tyr Leu Leu Glu Cys Lys Glu Lys Glu Pro Leu
100 105 110

60 Pro Pro Ile Thr Val Pro Val Ser Ser Ala Pro Cys Lys Thr His Ile
115 120 125

65 Leu Ser Glu Glu Lys Ile His Leu Gln Ser Leu Pro Thr Pro Tyr Leu
130 135 140

His Val Ser Asp Gly Gly Lys Tyr Leu Gln Thr Tyr Gly Met Trp Ile
74

ES 2 709 217 T3

	145		150		155		160									
5	Leu	Gln	Thr	Pro	Asp 165	Lys	Lys	Trp	Thr	Asn 170	Trp	Ser	Ile	Ala	Arg 175	Gly
10	Met	Val	Val	Asp 180	Asp	Lys	His	Ile	Thr 185	Gly	Leu	Val	Ile	Cys 190	Pro	Gln
15	His	Ile	Arg 195	Gln	Ile	Ala	Asp	Ser 200	Trp	Ala	Ala	Ile	Gly 205	Lys	Ala	Asn
20	Glu	Ile 210	Pro	Phe	Ala	Leu	Cys 215	Phe	Gly	Val	Pro	Pro 220	Ala	Ala	Ile	Leu
25	Val 225	Ser	Ser	Met	Pro	Ile 230	Pro	Glu	Gly	Val	Ser 235	Glu	Ser	Asp	Tyr	Val 240
30	Gly	Ala	Ile	Leu	Gly 245	Glu	Ser	Val	Pro	Val 250	Val	Lys	Cys	Glu	Thr 255	Asn
35	Asp	Leu	Met	Val 260	Pro	Ala	Thr	Ser	Glu 265	Met	Val	Phe	Glu	Gly 270	Thr	Leu
40	Ser	Leu	Thr 275	Asp	Thr	His	Leu	Glu 280	Gly	Pro	Phe	Gly	Glu 285	Met	His	Gly
45	Tyr	Val 290	Phe	Lys	Ser	Gln	Gly 295	His	Pro	Cys	Pro	Leu 300	Tyr	Thr	Val	Lys
50	Ala 305	Met	Ser	Tyr	Arg	Asp 310	Asn	Ala	Ile	Leu	Pro 315	Val	Ser	Asn	Pro	Gly 320
55	Leu	Cys	Thr	Asp	Glu 325	Thr	His	Thr	Leu	Ile 330	Gly	Ser	Leu	Val	Ala 335	Thr
60	Glu	Ala	Lys	Glu 340	Leu	Ala	Ile	Glu	Ser 345	Gly	Leu	Pro	Ile	Leu 350	Asp	Ala
65	Phe	Met	Pro 355	Tyr	Glu	Ala	Gln	Ala 360	Leu	Trp	Leu	Ile	Leu 365	Lys	Val	Asp
70	Leu	Lys 370	Gly	Leu	Gln	Ala	Leu 375	Lys	Thr	Thr	Pro	Glu 380	Glu	Phe	Cys	Lys
75	Lys 385	Val	Gly	Asp	Ile	Tyr 390	Phe	Arg	Thr	Lys	Val 395	Gly	Phe	Ile	Val	His 400
80	Glu	Ile	Ile	Leu	Val 405	Ala	Asp	Asp	Ile	Asp 410	Ile	Phe	Asn	Phe	Lys 415	Glu
85	Val	Ile	Trp	Ala	Tyr	Val	Thr	Arg	His	Thr	Pro	Val	Ala	Asp	Gln	Met

ES 2 709 217 T3

	420	425	430	
5	Ala Phe Asp 435 Asp Val Thr Ser Phe 440 Pro Leu Ala Pro Phe 445 Val Ser Gln			
10	Ser Ser 450 Arg Ser Lys Thr Met 455 Lys Gly Gly Lys Cys 460 Val Thr Asn Cys			
15	Ile Phe Arg Gln Gln Tyr 470 Glu Arg Ser Phe Asp 475 Tyr Ile Thr Cys Asn 480			
20	Phe Glu Lys Gly Tyr 485 Pro Lys Gly Leu Val 490 Asp Lys Val Asn Glu 495 Asn			
25	Trp Lys Arg Tyr 500 Gly Tyr Lys			
30	<210> 19 <211> 1512 <212> ADN <213> Secuencia artificial			
35	<220> <223> Sintética			
40	<400> 19 atgaggaagc taaatccagc tttagaatTT agagacttta tccaggTctt aaaagatgaa			60
45	gatgacttaa tcgaaattac cgaagagatt gatccaaatc tcgaagtagg tgcaattatg			120
50	aggaaggcct atgaatccca cttaccagcc ccgTtatTTa aaaaTctcaa aggtgcttcg			180
55	aaggatcttt tcagcatttt aggttgccca gccggtttga gaagtaagga gaaaggagat			240
60	catggtagaa ttgcccatca tctggggctc gacccaaaaa caactatcaa ggaaatcata			300
65	gattatttgc tggagtgtaa ggagaaggaa cctctccccc caatcactgt tcctgtgtca			360
70	tctgcacctt gtaaaacaca tatactttct gaagaaaaaa tacatctaca aagcctgcca			420
75	acaccatatc tacatgtttc agacggtggc aagtacttac aaacgtacgg aatgtggatt			480
80	cttcaaactc cagataaaaa atggactaat tggTcaattg ctagaggTat ggttgtagat			540
85	gacaagcata tcaTggtct ggtaattcat ccacaacata ttagacaaat tgctgactct			600
90	tgggcagcaa ttggaaaagc aaatgaaatt cttttcgcgT tatgttttgg cgttcccca			660
95	gcagctatTT tagttagttc catgccaaTT cctgaaggTg tttctgaatc ggattatgtt			720
100	ggcgcaatct tgggtgagtc ggttccagta gtaaaatgtg agaccaacga tttaatggtt			780
105	cctgcaacga gtgagatggt atttgagggt actttgtcct taacagatac acatctggaa			840
110	ggcccatTTg gtgagatgca tggatatgtt ttcaaaagcc aaggTcatcc ttgtccattg			900
115	tacactgtca aggctatgag ttacagagac aatgctattc tacctgtttc gaaccccggt			960
120	ctttgtacgg atgagacaca taccttgatt ggttcactag tggctactga ggccaaggag			1020
125	ctggctattg aatctggctt gccaatTctg gatgccttta tgccttatga ggctcaggct			1080

ES 2 709 217 T3

ctttggctta tcttaaaggt ggatttgaaa gggctgcaag cattgaagac aacgcctgaa 1140
 gaattttgta agaaggtagg tgatattttac tttaggacaa aagttggttt tatagtccat 1200
 5 gaaataattt tgggtggcaga tgatatcgac atattttaact tcaaagaagt catctgggcc 1260
 tacgtttacaa gacatacacc tgttgcagat cagatggctt ttgatgatgt cacttctttt 1320
 10 cctttggctc cctttgtttc gcagtcattcc agaagtaaga ctatgaaagg tggaaagtgc 1380
 gttactaatt gcatatttag acagcaatat gagcgcagtt ttgactacat aacttgtaat 1440
 tttgaaaagg gatatccaaa aggattagtt gacaaaagtaa atgaaaattg gaaaaggtag 1500
 15 ggatataaat aa 1512

<210> 20
 <211> 503
 20 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Sintética
 25 <400> 20

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

ES 2 709 217 T3

	165					170					175				
5	Met	Val	Val	Asp 180	Asp	Lys	His	Ile	Thr 185	Gly	Leu	Val	Ile	His 190	Pro Gln
10	His	Ile	Arg 195	Gln	Ile	Ala	Asp	Ser 200	Trp	Ala	Ala	Ile	Gly 205	Lys	Ala Asn
15	Glu	Ile 210	Pro	Phe	Ala	Leu	Cys 215	Phe	Gly	Val	Pro	Pro 220	Ala	Ala	Ile Leu
20	Val 225	Ser	Ser	Met	Pro	Ile 230	Pro	Glu	Gly	Val	Ser 235	Glu	Ser	Asp	Tyr Val 240
25	Gly	Ala	Ile	Leu	Gly 245	Glu	Ser	Val	Pro	Val 250	Val	Lys	Cys	Glu	Thr Asn 255
30	Asp	Leu	Met	Val 260	Pro	Ala	Thr	Ser	Glu 265	Met	Val	Phe	Glu	Gly 270	Thr Leu
35	Ser	Leu	Thr 275	Asp	Thr	His	Leu	Glu 280	Gly	Pro	Phe	Gly	Glu 285	Met	His Gly
40	Tyr	Val 290	Phe	Lys	Ser	Gln	Gly 295	His	Pro	Cys	Pro	Leu 300	Tyr	Thr	Val Lys
45	Ala 305	Met	Ser	Tyr	Arg	Asp 310	Asn	Ala	Ile	Leu	Pro 315	Val	Ser	Asn	Pro Gly 320
50	Leu	Cys	Thr	Asp	Glu 325	Thr	His	Thr	Leu	Ile 330	Gly	Ser	Leu	Val	Ala Thr 335
55	Glu	Ala	Lys	Glu 340	Leu	Ala	Ile	Glu	Ser 345	Gly	Leu	Pro	Ile	Leu 350	Asp Ala
60	Phe	Met	Pro 355	Tyr	Glu	Ala	Gln	Ala 360	Leu	Trp	Leu	Ile	Leu 365	Lys	Val Asp
65	Leu	Lys 370	Gly	Leu	Gln	Ala	Leu 375	Lys	Thr	Thr	Pro	Glu 380	Glu	Phe	Cys Lys
70	Lys 385	Val	Gly	Asp	Ile	Tyr 390	Phe	Arg	Thr	Lys	Val 395	Gly	Phe	Ile	Val His 400
75	Glu	Ile	Ile	Leu	Val 405	Ala	Asp	Asp	Ile	Asp 410	Ile	Phe	Asn	Phe	Lys Glu 415
80	Val	Ile	Trp	Ala 420	Tyr	Val	Thr	Arg	His 425	Thr	Pro	Val	Ala	Asp 430	Gln Met
85	Ala	Phe	Asp	Asp	Val	Thr	Ser	Phe	Pro	Leu	Ala	Pro	Phe	Val	Ser Gln

ES 2 709 217 T3

	435	440	445	
5	Ser Ser Arg Ser Lys Thr Met Lys Gly Gly Lys Cys Val Thr Asn Cys 450 455 460			
10	Ile Phe Arg Gln Gln Tyr Glu Arg Ser Phe Asp Tyr Ile Thr Cys Asn 465 470 475 480			
15	Phe Glu Lys Gly Tyr Pro Lys Gly Leu Val Asp Lys Val Asn Glu Asn 485 490 495			
20	Trp Lys Arg Tyr Gly Tyr Lys 500			
25	<210> 21 <211> 1512 <212> ADN <213> Secuencia artificial			
30	<220> <223> Sintética			
35	<400> 21 atgaggaagc taaatccagc tttagaattt agagacttta tccaggctctt aaaagatgaa			60
40	gatgacttaa tcgaaattac cgaagagatt gatccaaatc tcgaagtagg tgcaattatg			120
45	aggaaggcct atgaatccca cttaccagcc ccgttatttta aaaatctcaa aggtgcttcg			180
50	aaggatcttt tcagcatttt aggttgccca gccggtttga gaagtaagga gaaaggagat			240
55	catggtagaa ttgcccata tctggggctc gacccaaaaa caactatcaa ggaaatcata			300
60	gattattttgc tggagtgtaa ggagaaggaa cctctccccc caatcactgt tcctgtgtca			360
65	tctgcacctt gtaaaacaca tatactttct gaagaaaaaa tacatctaca aagcctgcca			420
70	acaccatata tacatgtttc agacggtggc aagtacttac aaacgtacgg aatgtggatt			480
75	cttcaaactc cagataaaaa atggactaat tgggtcaattg ctagaggatg ggtttagatg			540
80	gacaagcata tcaactggtc ggtaattcct ccacaacata ttagacaaat tgctgactct			600
85	tgggcagcaa ttggaaaagc aaatgaaatt cctttcgcgt tatgtttttg cgttcccca			660
90	gcagctattt tagttagttc catgccaatt cctgaagggtg tttctgaatc ggattatggt			720
95	ggcgcaatct tgggtgagtc ggttccagta gtaaaatgtg agaccaacga tttaatggtt			780
100	cctgcaacga gtgagatggt atttgagggt actttgtcct taacagatac acatctggaa			840
105	ggcccatttg gtgagatgca tggatatggt ttcaaaagcc aaggtcatcc ttgtccattg			900
110	tacactgtca aggctatgag ttacagagac aatgctattc tacctgtttc gaaccccggt			960
115	ctttgtacgg atgagacaca taccttgatt gggtcactag tggctactga ggccaaggag			1020
120	ctggctattg aatctggctt gccaattctg gatgccttta tgccttatga ggctcaggct			1080
125	ctttggctta tcttaaagggt ggatttgaaa gggctgcaag cattgaagac aacgcctgaa			1140
130	gaattttgta agaaggtagg tgatatttac tttaggacaa aagttgggtt tatagtccat			1200

ES 2 709 217 T3

gaaataat tgggtggcaga tgatatcgac atattttaact tcaaagaagt catctgggcc 1260
 tacgttacaa gacatacacc tgttgcatgat cagatggcctt ttgatgatgt cacttctttt 1320
 5 cctttggctc cctttgtttc gcagtcaccc agaagtaaga ctatgaaagg tggaaagtgc 1380
 gttactaatt gcatatttag acagcaatat gagcgcagtt ttgactacat aacttgtaat 1440
 10 tttgaaaagg gatatccaaa aggattagtt gacaaagtaa atgaaaattg gaaaagggtac 1500
 ggatataaat aa 1512

15 <210> 22
 <211> 503
 <212> PRT
 <213> Secuencia artificial

20 <220>
 <223> Sintética

<400> 22

25 Met Arg Lys Leu Asn Pro Ala Leu Glu Phe Arg Asp Phe Ile Gln Val
 1 5 10 15

Leu Lys Asp Glu Asp Asp Leu Ile Glu Ile Thr Glu Glu Ile Asp Pro
 20 25 30

Asn Leu Glu Val Gly Ala Ile Met Arg Lys Ala Tyr Glu Ser His Leu
 35 40 45

35 Pro Ala Pro Leu Phe Lys Asn Leu Lys Gly Ala Ser Lys Asp Leu Phe
 50 55 60

40 Ser Ile Leu Gly Cys Pro Ala Gly Leu Arg Ser Lys Glu Lys Gly Asp
 65 70 75 80

45 His Gly Arg Ile Ala His His Leu Gly Leu Asp Pro Lys Thr Thr Ile
 85 90 95

Lys Glu Ile Ile Asp Tyr Leu Leu Glu Cys Lys Glu Lys Glu Pro Leu
 100 105 110

50 Pro Pro Ile Thr Val Pro Val Ser Ser Ala Pro Cys Lys Thr His Ile
 115 120 125

55 Leu Ser Glu Glu Lys Ile His Leu Gln Ser Leu Pro Thr Pro Tyr Leu
 130 135 140

60 His Val Ser Asp Gly Gly Lys Tyr Leu Gln Thr Tyr Gly Met Trp Ile
 145 150 155 160

65 Leu Gln Thr Pro Asp Lys Lys Trp Thr Asn Trp Ser Ile Ala Arg Gly
 165 170 175

Met Val Val Asp Asp Lys His Ile Thr Gly Leu Val Ile Pro Pro Gln
 80

ES 2 709 217 T3

	180	185	190
5	His Ile Arg 195 Gln Ile Ala Asp Ser 200 Trp Ala Ala Ile Gly 205 Lys Ala Asn		
10	Glu Ile 210 Pro Phe Ala Leu Cys 215 Phe Gly Val Pro Pro 220 Ala Ala Ile Leu		
15	Val 225 Ser Ser Met Pro Ile 230 Pro Glu Gly Val Ser 235 Glu Ser Asp Tyr Val 240		
20	Gly Ala Ile Leu Gly 245 Glu Ser Val Pro Val 250 Val Lys Cys Glu Thr 255 Asn		
25	Asp Leu Met Val 260 Pro Ala Thr Ser Glu 265 Met Val Phe Glu Gly 270 Thr Leu		
30	Ser Leu Thr 275 Asp Thr His Leu Glu 280 Gly Pro Phe Gly Glu 285 Met His Gly		
35	Tyr Val 290 Phe Lys Ser Gln Gly 295 His Pro Cys Pro Leu 300 Tyr Thr Val Lys		
40	Ala Met Ser Tyr Arg Asp 310 Asn Ala Ile Leu Pro 315 Val Ser Asn Pro Gly 320		
45	Leu Cys Thr Asp Glu 325 Thr His Thr Leu Ile 330 Gly Ser Leu Val Ala 335 Thr		
50	Glu Ala Lys Glu 340 Leu Ala Ile Glu Ser 345 Gly Leu Pro Ile Leu 350 Asp Ala		
55	Phe Met Pro 355 Tyr Glu Ala Gln Ala 360 Leu Trp Leu Ile Leu 365 Lys Val Asp		
60	Leu Lys 370 Gly Leu Gln Ala Leu 375 Lys Thr Thr Pro Glu 380 Glu Phe Cys Lys		
65	Lys 385 Val Gly Asp Ile Tyr 390 Phe Arg Thr Lys Val 395 Gly Phe Ile Val His 400		
	Glu Ile Ile Leu Val 405 Ala Asp Asp Ile Asp 410 Ile Phe Asn Phe Lys 415 Glu		
	Val Ile Trp Ala Tyr Val Thr Arg His 425 Thr Pro Val Ala Asp 430 Gln Met		
	Ala Phe Asp 435 Asp Val Thr Ser Phe 440 Pro Leu Ala Pro Phe 445 Val Ser Gln		
	Ser Ser Arg Ser Lys Thr Met Lys Gly Gly Lys Cys Val Thr Asn Cys		

ES 2 709 217 T3

	450	455	460	
5	Ile Phe Arg Gln Gln Tyr Glu Arg Ser Phe Asp Tyr Ile Thr Cys Asn 465 470 475 480			
10	Phe Glu Lys Gly Tyr Pro Lys Gly Leu Val Asp Lys Val Asn Glu Asn 485 490 495			
	Trp Lys Arg Tyr Gly Tyr Lys 500			
15	<210> 23 <211> 1512 <212> ADN <213> Secuencia artificial			
20	<220> <223> sintética			
25	<400> 23 atgaggaagc taaatccagc tttagaatTT agagactTTa tccaggTcTT aaaagatgaa			60
	gatgactTaa tcgaaattac cgaagagatt gatccaaatc tcgaagtagg tgcaattatg			120
30	aggaaggcct atgaatccca cttaccagcc ccgTtattTa aaaatctcaa aggtgcttcg			180
	aaggatctTT tcagcattTT aggttgccca gccggtTtga gaagtaagga gaaaggagat			240
	catggtagaa ttgcccatca tctggggctc gacccaaaaa caactatcaa ggaaatcata			300
35	gattattTgc tggagtgtaa ggagaaggaa cctctcccc caatcactgt tcctgtgtca			360
	tctgcacctt gtaaaacaca tatactTTct gaagaaaaaa tacatctaca aagcctgcca			420
40	acaccatatc tacatgtTtc agacggtggc aagtactTac aaacgtacgg aatgtggatt			480
	cttcaaactc cagataaaaa atggactaat tggTcaattg ctagaggTat ggttgtagat			540
	gacaagcata tcaTggtct ggtaattctg ccacaacata ttagacaaat tgctgactct			600
45	tgggcagcaa ttggaaaagc aaatgaaatt cttTtcgcgt tatgtTttTg cgtTccccca			660
	gcagctattt tagTtagTtc catgccaat cctgaaggTg tTtctgaatc ggattatgtt			720
50	ggcgcaatct tgggtgagTc ggtTccagTa gTaaaatgtg agaccaacga tTtaatggTt			780
	cctgcaacga gtgagatggT atttgagggT actTtgtcct taacagatac acatctggaa			840
	ggcccatTtg gtgagatgca tggatatgtt ttcaaaagcc aaggTcatcc ttgtccattg			900
55	tacactgtca aggctatgag ttacagagac aatgctattc tacctgtTtc gaaccccggt			960
	ctTtgtacgg atgagacaca taccttgatt ggtTcactag tggctactga ggccaaggag			1020
60	ctggctattg aatctggctt gccaatTctg gatgcctTta tgccttatga ggctcaggct			1080
	ctTtggctTa tctTaaaggT ggatttgaaa gggctgcaag cattgaagac aacgcctgaa			1140
	gaattTtgTa agaaggtagg tgatattTtac tTtaggacaa aagTtggtTt tatagtccat			1200
65	gaaataattt tggTggcaga tgatatcgac atattTtaact tcaaagaagT catctgggcc			1260
	tacgtTtaca gacatacacc tgtTgcagat cagatggctt ttgatgatgt cactTctTtt			1320

ES 2 709 217 T3

	cctttggctc cctttgtttc gcagtcattcc agaagtaaga ctatgaaagg tggaaagtgc	1380
	gttactaatt gcatatttag acagcaatat gagcgcagtt ttgactacat aacttgtaat	1440
5	tttgaaaagg gatatccaaa aggattagtt gacaaagtaa atgaaaattg gaaaaggtag	1500
	ggatataaat aa	1512
10	<210> 24 <211> 503 <212> PRT <213> Secuencia artificial	
15	<220> <223> sintética <400> 24	
20	Met Arg Lys Leu Asn Pro Ala Leu Glu Phe Arg Asp Phe Ile Gln Val 1 5 10 15	
25	Leu Lys Asp Glu Asp Asp Leu Ile Glu Ile Thr Glu Glu Ile Asp Pro 20 25 30	
30	Asn Leu Glu Val Gly Ala Ile Met Arg Lys Ala Tyr Glu Ser His Leu 35 40 45	
35	Pro Ala Pro Leu Phe Lys Asn Leu Lys Gly Ala Ser Lys Asp Leu Phe 50 55 60	
40	Ser Ile Leu Gly Cys Pro Ala Gly Leu Arg Ser Lys Glu Lys Gly Asp 65 70 75 80	
45	His Gly Arg Ile Ala His His Leu Gly Leu Asp Pro Lys Thr Thr Ile 85 90 95	
50	Lys Glu Ile Ile Asp Tyr Leu Leu Glu Cys Lys Glu Lys Glu Pro Leu 100 105 110	
55	Pro Pro Ile Thr Val Pro Val Ser Ser Ala Pro Cys Lys Thr His Ile 115 120 125	
60	Leu Ser Glu Glu Lys Ile His Leu Gln Ser Leu Pro Thr Pro Tyr Leu 130 135 140	
65	His Val Ser Asp Gly Gly Lys Tyr Leu Gln Thr Tyr Gly Met Trp Ile 145 150 155 160	
	Leu Gln Thr Pro Asp Lys Lys Trp Thr Asn Trp Ser Ile Ala Arg Gly 165 170 175	
	Met Val Val Asp Asp Lys His Ile Thr Gly Leu Val Ile Leu Pro Gln 180 185 190	
	His Ile Arg Gln Ile Ala Asp Ser Trp Ala Ala Ile Gly Lys Ala Asn	

ES 2 709 217 T3

	195						200						205					
5	Glu	Ile	Pro	Phe	Ala	Leu	Cys	Phe	Gly	Val	Pro	Pro	Ala	Ala	Ile	Leu		
	210						215					220						
10	Val	Ser	Ser	Met	Pro	Ile	Pro	Glu	Gly	Val	Ser	Glu	Ser	Asp	Tyr	Val		
	225					230					235					240		
15	Gly	Ala	Ile	Leu	Gly	Glu	Ser	Val	Pro	Val	Val	Lys	Cys	Glu	Thr	Asn		
					245					250					255			
20	Asp	Leu	Met	Val	Pro	Ala	Thr	Ser	Glu	Met	Val	Phe	Glu	Gly	Thr	Leu		
				260					265					270				
25	Ser	Leu	Thr	Asp	Thr	His	Leu	Glu	Gly	Pro	Phe	Gly	Glu	Met	His	Gly		
			275					280					285					
30	Tyr	Val	Phe	Lys	Ser	Gln	Gly	His	Pro	Cys	Pro	Leu	Tyr	Thr	Val	Lys		
	290						295					300						
35	Ala	Met	Ser	Tyr	Arg	Asp	Asn	Ala	Ile	Leu	Pro	Val	Ser	Asn	Pro	Gly		
	305					310					315					320		
40	Leu	Cys	Thr	Asp	Glu	Thr	His	Thr	Leu	Ile	Gly	Ser	Leu	Val	Ala	Thr		
					325					330					335			
45	Glu	Ala	Lys	Glu	Leu	Ala	Ile	Glu	Ser	Gly	Leu	Pro	Ile	Leu	Asp	Ala		
				340					345					350				
50	Phe	Met	Pro	Tyr	Glu	Ala	Gln	Ala	Leu	Trp	Leu	Ile	Leu	Lys	Val	Asp		
			355					360					365					
55	Leu	Lys	Gly	Leu	Gln	Ala	Leu	Lys	Thr	Thr	Pro	Glu	Glu	Phe	Cys	Lys		
	370						375					380						
60	Lys	Val	Gly	Asp	Ile	Tyr	Phe	Arg	Thr	Lys	Val	Gly	Phe	Ile	Val	His		
	385					390					395					400		
65	Glu	Ile	Ile	Leu	Val	Ala	Asp	Asp	Ile	Asp	Ile	Phe	Asn	Phe	Lys	Glu		
					405					410					415			
70	Val	Ile	Trp	Ala	Tyr	Val	Thr	Arg	His	Thr	Pro	Val	Ala	Asp	Gln	Met		
				420					425					430				
75	Ala	Phe	Asp	Asp	Val	Thr	Ser	Phe	Pro	Leu	Ala	Pro	Phe	Val	Ser	Gln		
			435					440					445					
80	Ser	Ser	Arg	Ser	Lys	Thr	Met	Lys	Gly	Gly	Lys	Cys	Val	Thr	Asn	Cys		
	450						455					460						
85	Ile	Phe	Arg	Gln	Gln	Tyr	Glu	Arg	Ser	Phe	Asp	Tyr	Ile	Thr	Cys	Asn		

ES 2 709 217 T3

	465		470		475		480									
5	Phe	Glu	Lys	Gly	Tyr	Pro	Lys	Gly	Leu	Val	Asp	Lys	Val	Asn	Glu	Asn
				485					490						495	
10	Trp	Lys	Arg	Tyr	Gly	Tyr	Lys									
				500												
15	<210>	25														
	<211>	1512														
	<212>	ADN														
	<213>	Secuencia artificial														
	<220>															
	<223>	Sintética														
20	<400>	25														
	atgaggaagc	taaatccagc	tttagaattt	agagacttta	tccagggtctt	aaaagatgaa									60	
	gatgacttaa	tcgaaattac	cgaagagatt	gatccaaatc	tcgaagtagg	tgcaattatg									120	
25	aggaaggcct	atgaatccca	cttaccagcc	ccgttattta	aaaatctcaa	aggtgcttcg									180	
	aaggatcttt	tcagcatttt	aggttgccca	gccggtttga	gaagtaagga	gaaaggagat									240	
30	catggtagaa	ttgcccatca	tctggggctc	gacccaaaaa	caactatcaa	ggaaatcata									300	
	gattatttgc	tggagtgtaa	ggagaaggaa	cctctccccc	caatcactgt	tcctgtgtca									360	
	tctgcacctt	gtaaaacaca	tatactttct	gaagaaaaaa	tacatctaca	aagcctgcca									420	
35	acaccatata	tacatgtttc	agacggtggc	aagtacttac	aaacgtacgg	aatgtggatt									480	
	cttcaaactc	cagataaaaa	atggactaat	tggtcaattg	ctagagggtat	ggttgtagat									540	
40	gacaagcata	tcactgggtct	ggtaattggg	ccacaacata	ttagacaaat	tgctgactct									600	
	tgggcagcaa	ttggaaaagc	aaatgaaatt	cctttcgcgt	tatgttttgg	cgttccccca									660	
	gcagctattt	tagttagttc	catgccaatt	cctgaagggtg	tttctgaatc	ggattatgtt									720	
45	ggcgcaatct	tgggtgagtc	ggttccagta	gtaaaatgtg	agaccaacga	tttaatgggt									780	
	cctgcaacga	gtgagatggg	atgtgagggt	actttgtcct	taacagatac	acatctggaa									840	
50	ggcccatattg	gtgagatgca	tggatatgtt	ttcaaaagcc	aagggtcatcc	ttgtccattg									900	
	tacactgtca	aggctatgag	ttacagagac	aatgctattc	tacctgtttc	gaaccccggt									960	
	ctttgtacgg	atgagacaca	taccttgatt	ggttcactag	tggctactga	ggccaaggag									1020	
55	ctggctattg	aatctggcct	gccaattctg	gatgccttta	tgcccttatga	ggctcaggct									1080	
	ctttggctta	tcttaaagggt	ggatttgaaa	gggctgcaag	cattgaagac	aacgcctgaa									1140	
60	gaattttgta	agaaggtagg	tgatattttac	tttaggacaa	aagttgggtt	tatagtccat									1200	
	gaaataat	tgggtggcaga	tgatatcgac	atattttaact	tcaaagaagt	catctgggcc									1260	
	tacgtttaca	gacatacacc	tgttgcagat	cagatggcctt	ttgatgatgt	cacttctttt									1320	
65	cctttggctc	cctttgtttc	gcagtcaccc	agaagtaaga	ctatgaaagg	tggaaagtgc									1380	
	gttactaatt	gcatat	tttag	acagcaatat	gagcgcagtt	ttgactacat	aacttgtaat								1440	

ES 2 709 217 T3

tttgaaaagg gatattccaaa aggattagtt gacaaagtaa atgaaaattg gaaaaggtag 1500
ggatataaat aa 1512

5
<210> 26
<211> 503
<212> PRT
<213> Secuencia artificial
10
<220>
<223> sintética
<400> 26
15
Met Arg Lys Leu Asn Pro Ala Leu Glu Phe Arg Asp Phe Ile Gln Val
1 5 10 15
20 Leu Lys Asp Glu Asp Asp Leu Ile Glu Ile Thr Glu Glu Ile Asp Pro
20 25 30
25 Asn Leu Glu Val Gly Ala Ile Met Arg Lys Ala Tyr Glu Ser His Leu
35 40 45
30 Pro Ala Pro Leu Phe Lys Asn Leu Lys Gly Ala Ser Lys Asp Leu Phe
50 55 60
35 Ser Ile Leu Gly Cys Pro Ala Gly Leu Arg Ser Lys Glu Lys Gly Asp
65 70 75 80
40 His Gly Arg Ile Ala His His Leu Gly Leu Asp Pro Lys Thr Thr Ile
85 90 95
45 Lys Glu Ile Ile Asp Tyr Leu Leu Glu Cys Lys Glu Lys Glu Pro Leu
100 105 110
50 Pro Pro Ile Thr Val Pro Val Ser Ser Ala Pro Cys Lys Thr His Ile
115 120 125
55 Leu Ser Glu Glu Lys Ile His Leu Gln Ser Leu Pro Thr Pro Tyr Leu
130 135 140
60 His Val Ser Asp Gly Gly Lys Tyr Leu Gln Thr Tyr Gly Met Trp Ile
145 150 155 160
65 Leu Gln Thr Pro Asp Lys Lys Trp Thr Asn Trp Ser Ile Ala Arg Gly
165 170 175
70 Met Val Val Asp Asp Lys His Ile Thr Gly Leu Val Ile Arg Pro Gln
180 185 190
75 His Ile Arg Gln Ile Ala Asp Ser Trp Ala Ala Ile Gly Lys Ala Asn
195 200 205
80 Glu Ile Pro Phe Ala Leu Cys Phe Gly Val Pro Pro Ala Ala Ile Leu
86

ES 2 709 217 T3

	210					215					220					
5	Val 225	Ser	Ser	Met	Pro	Ile 230	Pro	Glu	Gly	Val	Ser 235	Glu	Ser	Asp	Tyr	Val 240
10	Gly	Ala	Ile	Leu	Gly 245	Glu	Ser	Val	Pro	Val 250	Val	Lys	Cys	Glu	Thr 255	Asn
15	Asp	Leu	Met	Val 260	Pro	Ala	Thr	Ser	Glu 265	Met	Val	Phe	Glu	Gly 270	Thr	Leu
20	Ser	Leu	Thr 275	Asp	Thr	His	Leu	Glu 280	Gly	Pro	Phe	Gly	Glu 285	Met	His	Gly
25	Tyr	Val 290	Phe	Lys	Ser	Gln	Gly 295	His	Pro	Cys	Pro	Leu 300	Tyr	Thr	Val	Lys
30	Ala 305	Met	Ser	Tyr	Arg	Asp 310	Asn	Ala	Ile	Leu	Pro 315	Val	Ser	Asn	Pro	Gly 320
35	Leu	Cys	Thr	Asp	Glu 325	Thr	His	Thr	Leu	Ile 330	Gly	Ser	Leu	Val	Ala 335	Thr
40	Glu	Ala	Lys	Glu 340	Leu	Ala	Ile	Glu	Ser 345	Gly	Leu	Pro	Ile	Leu 350	Asp	Ala
45	Phe	Met	Pro 355	Tyr	Glu	Ala	Gln	Ala 360	Leu	Trp	Leu	Ile	Leu 365	Lys	Val	Asp
50	Leu	Lys 370	Gly	Leu	Gln	Ala	Leu 375	Lys	Thr	Thr	Pro	Glu 380	Glu	Phe	Cys	Lys
55	Lys 385	Val	Gly	Asp	Ile	Tyr 390	Phe	Arg	Thr	Lys	Val 395	Gly	Phe	Ile	Val	His 400
60	Glu	Ile	Ile	Leu	Val 405	Ala	Asp	Asp	Ile	Asp 410	Ile	Phe	Asn	Phe	Lys 415	Glu
65	Val	Ile	Trp	Ala 420	Tyr	Val	Thr	Arg	His 425	Thr	Pro	Val	Ala	Asp 430	Gln	Met
70	Ala	Phe	Asp 435	Asp	Val	Thr	Ser	Phe 440	Pro	Leu	Ala	Pro	Phe 445	Val	Ser	Gln
75	Ser	Ser 450	Arg	Ser	Lys	Thr	Met 455	Lys	Gly	Gly	Lys	Cys 460	Val	Thr	Asn	Cys
80	Ile 465	Phe	Arg	Gln	Gln	Tyr 470	Glu	Arg	Ser	Phe	Asp 475	Tyr	Ile	Thr	Cys	Asn 480
85	Phe	Glu	Lys	Gly	Tyr	Pro	Lys	Gly	Leu	Val	Asp	Lys	Val	Asn	Glu	Asn

ES 2 709 217 T3

	485	490	495	
5	Trp Lys Arg Tyr Gly Tyr Lys			
	500			
10	<210> 27			
	<211> 1512			
	<212> ADN			
	<213> Secuencia artificial			
15	<220>			
	<223> Sintética			
	<400> 27			
	atgaggaagc taaatccagc tttagaattt agagacttta tccaggctctt aaaagatgaa		60	
20	gatgacttaa tcgaaattac cgaagagatt gatccaaatc tcgaagtagg tgcaattatg		120	
	aggaaggcct atgaatccca cttaccagcc ccgttattta aaaatctcaa aggtgcttcg		180	
	aaggatcttt tcagcatttt aggttgccca gccggtttga gaagtaagga gaaaggagat		240	
25	catggtagaa ttgcccata tctggggctc gacccaaaaa caactatcaa ggaaatcata		300	
	gattattttgc tggagtgtaa ggagaaggaa cctctccccc caatcactgt tcctgtgtca		360	
30	tctgcacctt gtaaaacaca tatactttct gaagaaaaaa tacatctaca aagcctgcca		420	
	acaccatata tacatgtttc agacggtggc aagtacttac aaacgtacgg aatgtggatt		480	
	cttcaaactc cagataaaaa atggactaat tgggtcaattg ctagagggtat ggttgtagat		540	
35	gacaagcata tcaactggtct ggtaattgat ccacaacata ttagacaaat tgctgactct		600	
	tgggcagcaa ttggaaaagc aaatgaaatt cctttcgcgt tatgttttgg cgttcccca		660	
40	gcagctatct tagttagttc catgccaat cctgaagggtg tttctgaatc ggattatggt		720	
	ggcgcaatct tgggtgagtc ggttccagta gtaaaatgtg agaccaacga ttaaatgggt		780	
	cctgcaacga gtgagatggt atttgaggggt actttgtcct taacagatac acatctggaa		840	
45	ggcccatctt gtgagatgca tggatatggt ttcaaaaagcc aagggtcatcc ttgtccattg		900	
	tacactgtca aggctatgag ttacagagac aatgctattc tacctgtttc gaaccccggt		960	
50	ctttgtacgg atgagacaca taccttgatt gggtcactag tggctactga ggccaaggag		1020	
	ctggctattg aatctggctt gccaatctct gatgccttta tgccttatga ggctcaggct		1080	
	ctttggctta tcttaaagggt ggatttgaaa gggctgcaag cattgaagac aacgcctgaa		1140	
55	gaattttgta agaaggtagg tgatatctac tttaggacaa aagttgggtt tatagtccat		1200	
	gaaataattt tgggtggcaga tgatatcgac atattttaact tcaaagaagt catctgggcc		1260	
60	tacgttacaa gacatacacc tgttgcagat cagatggctt ttgatgatgt cacttctttt		1320	
	cctttggctc cctttgtttc gcagtcaccc agaagtaaga ctatgaaagg tggaaagtgc		1380	
	gttactaatt gcatatttag acagcaatat gagcgcagtt ttgactacat aacttgtaat		1440	
65	tttgaaaagg gatatccaaa aggattagtt gacaaagtaa atgaaaattg gaaaaggtag		1500	
	ggatataaat aa		1512	

ES 2 709 217 T3

<210> 28
 <211> 503
 <212> PRT
 5 <213> Secuencia artificial

 <220>
 <223> Sintética
 10 <400> 28
 Met Arg Lys Leu Asn Pro Ala Leu Glu Phe Arg Asp Phe Ile Gln Val
 1 5 10 15
 15 Leu Lys Asp Glu Asp Asp Leu Ile Glu Ile Thr Glu Glu Ile Asp Pro
 20 25 30
 20 Asn Leu Glu Val Gly Ala Ile Met Arg Lys Ala Tyr Glu Ser His Leu
 35 40 45
 25 Pro Ala Pro Leu Phe Lys Asn Leu Lys Gly Ala Ser Lys Asp Leu Phe
 50 55 60
 30 Ser Ile Leu Gly Cys Pro Ala Gly Leu Arg Ser Lys Glu Lys Gly Asp
 65 70 75 80
 35 His Gly Arg Ile Ala His His Leu Gly Leu Asp Pro Lys Thr Thr Ile
 85 90 95
 40 Lys Glu Ile Ile Asp Tyr Leu Leu Glu Cys Lys Glu Lys Glu Pro Leu
 100 105 110
 45 Pro Pro Ile Thr Val Pro Val Ser Ser Ala Pro Cys Lys Thr His Ile
 115 120 125
 50 Leu Ser Glu Glu Lys Ile His Leu Gln Ser Leu Pro Thr Pro Tyr Leu
 130 135 140
 55 His Val Ser Asp Gly Gly Lys Tyr Leu Gln Thr Tyr Gly Met Trp Ile
 145 150 155 160
 60 Leu Gln Thr Pro Asp Lys Lys Trp Thr Asn Trp Ser Ile Ala Arg Gly
 165 170 175
 65 Met Val Val Asp Asp Lys His Ile Thr Gly Leu Val Ile Asp Pro Gln
 180 185 190
 70 His Ile Arg Gln Ile Ala Asp Ser Trp Ala Ala Ile Gly Lys Ala Asn
 195 200 205
 75 Glu Ile Pro Phe Ala Leu Cys Phe Gly Val Pro Pro Ala Ala Ile Leu
 210 215 220
 Val Ser Ser Met Pro Ile Pro Glu Gly Val Ser Glu Ser Asp Tyr Val
 89

ES 2 709 217 T3

	225		230		235		240									
5	Gly	Ala	Ile	Leu	Gly 245	Glu	Ser	Val	Pro	Val 250	Val	Lys	Cys	Glu	Thr 255	Asn
10	Asp	Leu	Met	Val 260	Pro	Ala	Thr	Ser	Glu 265	Met	Val	Phe	Glu	Gly 270	Thr	Leu
15	Ser	Leu	Thr 275	Asp	Thr	His	Leu	Glu 280	Gly	Pro	Phe	Gly	Glu 285	Met	His	Gly
20	Tyr	Val 290	Phe	Lys	Ser	Gln	Gly 295	His	Pro	Cys	Pro	Leu 300	Tyr	Thr	Val	Lys
25	Ala	Met	Ser	Tyr	Arg	Asp 310	Asn	Ala	Ile	Leu	Pro 315	Val	Ser	Asn	Pro	Gly 320
30	Leu	Cys	Thr	Asp	Glu 325	Thr	His	Thr	Leu	Ile 330	Gly	Ser	Leu	Val	Ala 335	Thr
35	Glu	Ala	Lys	Glu 340	Leu	Ala	Ile	Glu	Ser 345	Gly	Leu	Pro	Ile	Leu 350	Asp	Ala
40	Phe	Met	Pro 355	Tyr	Glu	Ala	Gln	Ala 360	Leu	Trp	Leu	Ile	Leu 365	Lys	Val	Asp
45	Leu	Lys 370	Gly	Leu	Gln	Ala	Leu 375	Lys	Thr	Thr	Pro	Glu 380	Glu	Phe	Cys	Lys
50	Lys	Val	Gly	Asp	Ile	Tyr 390	Phe	Arg	Thr	Lys	Val 395	Gly	Phe	Ile	Val	His 400
55	Glu	Ile	Ile	Leu	Val 405	Ala	Asp	Asp	Ile	Asp 410	Ile	Phe	Asn	Phe	Lys 415	Glu
60	Val	Ile	Trp	Ala 420	Tyr	Val	Thr	Arg	His 425	Thr	Pro	Val	Ala	Asp 430	Gln	Met
65	Ala	Phe	Asp 435	Asp	Val	Thr	Ser	Phe 440	Pro	Leu	Ala	Pro	Phe 445	Val	Ser	Gln
70	Ser	Ser 450	Arg	Ser	Lys	Thr	Met 455	Lys	Gly	Gly	Lys	Cys 460	Val	Thr	Asn	Cys
75	Ile	Phe	Arg	Gln	Gln	Tyr 470	Glu	Arg	Ser	Phe	Asp 475	Tyr	Ile	Thr	Cys	Asn 480
80	Phe	Glu	Lys	Gly	Tyr 485	Pro	Lys	Gly	Leu	Val 490	Asp	Lys	Val	Asn	Glu 495	Asn
85	Trp	Lys	Arg	Tyr	Gly	Tyr	Lys									

500

5	<210> 29 <211> 1512 <212> ADN <213> Secuencia artificial	
10	<220> <223> Sintética	
	<400> 29	
	atgaggaagc taaatccagc tttagaatth agagacttta tccaggtctt aaaagatgaa	60
15	gatgacttaa tcgaaattac cgaagagatt gatccaaatc tcgaagtagg tgcaattatg	120
	aggaaggcct atgaatccca cttaccagcc ccgttathta aaaatctcaa aggtgcttcg	180
20	aaggatcttt tcagcathtt aggttgccca gccggtttga gaagtaagga gaaaggagat	240
	catggtagaa ttgcccatca tctggggctc gacccaaaaa caactatcaa ggaaatcata	300
	gattatthtc tggagtgtaa ggagaaggaa cctctcccc caatcactgt tcctgtgtca	360
25	tctgcacctt gtaaaacaca tatactthct gaagaaaaaa tacatctaca aagcctgcca	420
	acaccatata tacatgtthc agacggtggc aagtacttac aaacgtacgg aatgtggatt	480
30	cttcaaactc cagataaaaa atggactaat tggatcaattg ctagaggat ggtttagat	540
	gacaagcata tcactggtct ggtaattgtt ccacaacata ttagacaaat tgctgactct	600
	tgggcagcaa ttggaaaagc aaatgaaatt cctttcgcgt tatgtthtgg cgttccccca	660
35	gcagctatth tagttagttc catgccaat cctgaagggtg thtctgaatc ggattatgtt	720
	ggcgcaatct tgggtgagtc ggttccagta gtaaaatgtg agaccaacga tthaatggtt	780
40	cctgcaacga gtgagatggt atthtaggggt actthgtcct taacagatac acatctggaa	840
	ggcccattht gtgagatgca tggatatgtt thcaaaagcc aaggtcatcc thgtccattg	900
	tacactgtca aggctatgag ttacagagac aatgttatth tacctgtthc gaaccccggt	960
45	ctthgtacgg atgagacaca taccttgatt ggttcactag tggctactga ggccaaggag	1020
	ctggctatth aatctggctt gccaatthctg gatgccthta tgccttatga ggctcaggct	1080
50	ctthggctta thttaaagggt ggatthgaaa gggctgcaag cattgaagac aacgcctgaa	1140
	gaatthtgta agaaggtagg tgataththac thtaggacaa aagttggtht tatagtccat	1200
	gaaataatth tgggtggcaga tgatatcgac ataththact tcaaagaagt catctgggcc	1260
55	tacgttacaa gacatacacc tgttgcatg cagatggctt thgatgatgt cactththt	1320
	cctthggctc cctthgtthc gcagtcattc agaagtaaga ctatgaaagg tggaaagtgc	1380
60	gttactaatt gcatatthtag acagcaatat gagcgcagtt thgactacat aacttgtaat	1440
	thtgaaaagg gatathcaaa aggattagtt gacaaagtaa atgaaaattg gaaaaggtag	1500
	ggatataaat aa	1512
65	<210> 30 <211> 503 <212> PRT	

ES 2 709 217 T3

<213> Secuencia artificial

<220>

<223> sintética

<400> 30

5

10

15

20

25

30

35

40

45

50

55

60

65

Met Arg Lys Leu Asn Pro Ala Leu Glu Phe Arg Asp Phe Ile Gln Val
1 5 10 15

Leu Lys Asp Glu Asp Asp Leu Ile Glu Ile Thr Glu Glu Ile Asp Pro
20 25 30

Asn Leu Glu Val Gly Ala Ile Met Arg Lys Ala Tyr Glu Ser His Leu
35 40 45

Pro Ala Pro Leu Phe Lys Asn Leu Lys Gly Ala Ser Lys Asp Leu Phe
50 55 60

Ser Ile Leu Gly Cys Pro Ala Gly Leu Arg Ser Lys Glu Lys Gly Asp
65 70 75 80

His Gly Arg Ile Ala His His Leu Gly Leu Asp Pro Lys Thr Thr Ile
85 90 95

Lys Glu Ile Ile Asp Tyr Leu Leu Glu Cys Lys Glu Lys Glu Pro Leu
100 105 110

Pro Pro Ile Thr Val Pro Val Ser Ser Ala Pro Cys Lys Thr His Ile
115 120 125

Leu Ser Glu Glu Lys Ile His Leu Gln Ser Leu Pro Thr Pro Tyr Leu
130 135 140

His Val Ser Asp Gly Gly Lys Tyr Leu Gln Thr Tyr Gly Met Trp Ile
145 150 155 160

Leu Gln Thr Pro Asp Lys Lys Trp Thr Asn Trp Ser Ile Ala Arg Gly
165 170 175

Met Val Val Asp Asp Lys His Ile Thr Gly Leu Val Ile Val Pro Gln
180 185 190

His Ile Arg Gln Ile Ala Asp Ser Trp Ala Ala Ile Gly Lys Ala Asn
195 200 205

Glu Ile Pro Phe Ala Leu Cys Phe Gly Val Pro Pro Ala Ala Ile Leu
210 215 220

Val Ser Ser Met Pro Ile Pro Glu Gly Val Ser Glu Ser Asp Tyr Val
225 230 235 240

Gly Ala Ile Leu Gly Glu Ser Val Pro Val Val Lys Cys Glu Thr Asn
92

ES 2 709 217 T3

	245	250	255
5	Asp Leu Met Val 260	Pro Ala Thr Ser Glu 265	Met Val Phe Glu Gly 270 Thr Leu
10	Ser Leu Thr 275	Asp Thr His Leu Glu 280	Gly Pro Phe Gly Glu 285 Met His Gly
15	Tyr Val 290	Phe Lys Ser Gln Gly 295	His Pro Cys Pro Leu 300 Tyr Thr Val Lys
20	Ala Met Ser Tyr Arg 305	Asp 310 Asn Ala Ile Leu Pro 315	Val Ser Asn Pro Gly 320
25	Leu Cys Thr Asp Glu 325	Thr His Thr Leu Ile 330	Gly Ser Leu Val Ala Thr 335
30	Glu Ala Lys Glu 340	Leu Ala Ile Glu Ser 345	Gly Leu Pro Ile Leu 350 Asp Ala
35	Phe Met Pro 355	Tyr Glu Ala Gln Ala 360	Leu Trp Leu Ile Leu 365 Lys Val Asp
40	Leu Lys Gly Leu Gln Ala Leu 375	Lys Thr Thr Pro Glu 380	Glu Phe Cys Lys
45	Lys Val Gly Asp Ile Tyr 390	Phe Arg Thr Lys Val 395	Gly Phe Ile Val His 400
50	Glu Ile Ile Leu Val 405	Ala Asp Asp Ile Asp 410	Ile Phe Asn Phe Lys 415 Glu
55	Val Ile Trp Ala Tyr Val Thr Arg His 425	Thr Pro Val Ala Asp 430	Gln Met
60	Ala Phe Asp 435	Asp Val Thr Ser Phe 440	Pro Leu Ala Pro Phe 445 Val Ser Gln
65	Ser Ser Arg Ser Lys Thr Met 455	Lys Gly Gly Lys Cys 460	Val Thr Asn Cys
70	Ile Phe Arg Gln Gln Tyr 470	Glu Arg Ser Phe Asp 475	Tyr Ile Thr Cys Asn 480
75	Phe Glu Lys Gly Tyr 485	Pro Lys Gly Leu Val 490	Asp Lys Val Asn Glu 495 Asn
80	Trp Lys Arg Tyr 500	Gly Tyr Lys	

	<211>	1512	
	<212>	ADN	
	<213>	Secuencia artificial	
5	<220>		
	<223>	sintética	
	<400>	31	
10	atgaggaagc	taaatccagc	tttagaattt agagacttta tccaggctctt aaaagatgaa 60
	gatgacttaa	tcgaaattac	cgaagagatt gatccaaatc tcgaagtagg tgcaattatg 120
	aggaaggcct	atgaatccca	cttaccagcc ccgttatTTta aaaatctcaa aggtgcttcg 180
15	aaggatcttt	tcagcatttt	aggttgccca gccggtttga gaagtaagga gaaaggagat 240
	catggtagaa	ttgcccatca	tctggggctc gacccaaaaa caactatcaa ggaaatcata 300
20	gattatttgc	tgagtgtaa	ggagaaggaa cctctccccc caatcactgt tcctgtgtca 360
	tctgcacctt	gtaaaacaca	tatactttct gaagaaaaaa tacatctaca aagcctgcca 420
	acaccatatc	tacatgtttc	agacggtggc aagtacttac aaacgtacgg aatgtggatt 480
25	cttcaaactc	cagataaaaa	atggactaat tggTcaattg ctagaggTat ggttgtagat 540
	gacaagcata	tcactggTct	ggtaatttcg ccacaacata ttagacaaat tgctgactct 600
30	tgggcagcaa	ttggaaaagc	aaatgaaatt cctttcgcgT tatgttttgg cgttcccca 660
	gcagctatTT	tagttagttc	catgccaaat cctgaaggTg tttctgaatc ggattatgtt 720
	ggcgcaatct	tgggtgagtc	ggttccagta gtaaaatgtg agaccaacga tttaatggtt 780
35	cctgcaacga	gtgagatggT	atttgagggt actttgtcct taacagatac acatctggaa 840
	ggcccatTTg	gtgagatgca	tggatatgtt ttcaaaagcc aaggTcatcc ttgtccattg 900
40	tacactgtca	aggctatgag	ttacagagac aatgctattc tacctgtttc gaaccccggt 960
	cTTtgTacgg	atgagacaca	taccttgatt ggttcactag tggctactga ggccaaggag 1020
	ctggctattg	aatctggctt	gccaattctg gatgccttta tgccttatga ggctcaggct 1080
45	cTTtggttta	tcttaaaggT	ggatttgaaa gggctgcaag cattgaagac aacgcctgaa 1140
	gaattttgta	agaaggtagg	Tgatatttac tttaggacaa aagttggttt tatagtccat 1200
50	gaaataatTT	tggTggcaga	Tgatatcgac atattTaaT tcaaagaagT catctgggcc 1260
	tacgtttaca	gacatacacc	TgtTgcagat cagatggctt ttgatgatgt cacttctttt 1320
	cctttggctc	cctttgtttc	gcagTcatcc agaagtaaga ctatgaaagg tggaaagtgc 1380
55	gttactaatt	gcataTTtag	acagcaatat gagcgagtt ttgactacat aacttgtaat 1440
	tttgaaaagg	gatatccaaa	aggattagtt gacaaaagtaa atgaaaattg gaaaaggTac 1500
60	ggatataaat	aa	1512
	<210>	32	
	<211>	503	
	<212>	PRT	
65	<213>	Secuencia artificial	
	<220>		
	<223>	sintética	

ES 2 709 217 T3

<400> 32

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

ES 2 709 217 T3

	260	265	270
5	Ser Leu Thr 275 Asp Thr His Leu	Glu 280 Gly Pro Phe Gly	Glu 285 Met His Gly
10	Tyr Val 290 Phe Lys Ser Gln	Gly 295 His Pro Cys Pro	Leu 300 Tyr Thr Val Lys
15	Ala 305 Met Ser Tyr Arg	Asp 310 Asn Ala Ile Leu	Pro 315 Val Ser Asn Pro Gly 320
20	Leu Cys Thr Asp	Glu 325 Thr His Thr Leu	Ile 330 Gly Ser Leu Val Ala 335 Thr
25	Glu Ala Lys	Glu 340 Leu Ala Ile Glu	Ser 345 Gly Leu Pro Ile Leu 350 Asp Ala
30	Phe Met 355 Pro Tyr Glu Ala Gln	Ala 360 Leu Trp Leu Ile	Leu 365 Lys Val Asp
35	Leu Lys 370 Gly Leu Gln Ala	Leu 375 Lys Thr Thr Pro	Glu 380 Glu Phe Cys Lys
40	Lys Val 385 Gly Asp Ile Tyr	Phe 390 Arg Thr Lys Val	Gly 395 Phe Ile Val His 400
45	Glu Ile Ile Leu	Val 405 Ala Asp Asp Ile	Asp 410 Ile Phe Asn Phe Lys 415 Glu
50	Val Ile Trp	Ala 420 Tyr Val Thr Arg	His 425 Thr Pro Val Ala 430 Asp Gln Met
55	Ala Phe 435 Asp Asp Val Thr Ser	Phe 440 Pro Leu Ala Pro	Phe 445 Val Ser Gln
60	Ser Ser 450 Arg Ser Lys Thr	Met 455 Lys Gly Gly Lys	Cys 460 Val Thr Asn Cys
65	Ile Phe Arg Gln Gln	Tyr 470 Glu Arg Ser Phe	Asp 475 Tyr Ile Thr Cys Asn 480
70	Phe Glu Lys Gly	Tyr 485 Pro Lys Gly Leu	Val 490 Asp Lys Val Asn Glu 495 Asn
75	Trp Lys Arg	Tyr 500 Gly Tyr Lys	
80	<210> 33	<211> 1512	<212> ADN
85	<213> Secuencia artificial		

ES 2 709 217 T3

<220>

<223> Sintética

<400> 33

5	atgaggaagc taaatccagc tttagaat	ttt agagacttta tccaggctctt aaaagatgaa	60
	gatgacttaa tcgaaattac cgaagagatt	gatccaaatc tcgaagtagg tgcaattatg	120
10	aggaaggcct atgaatccca cttaccagcc	ccgttatttta aaaatctcaa aggtgcttcg	180
	aaggatcttt tcagcatttt aggttgccca	gccggtttga gaagtaagga gaaaggagat	240
	catggtagaa ttgcccatca tctggggctc	gacccaaaaa caactatcaa ggaaatcata	300
15	gattatttgc tggagtgtaa ggagaaggaa	cctctccccc caatcactgt tcctgtgtca	360
	tctgcacctt gtaaaacaca tatactttct	gaagaaaaaa tacatctaca aagcctgcca	420
20	acaccatata tacatgtttc agacggtggc	aagtacttac aaacgtacgg aatgtggatt	480
	cttcaaactc cagataaaaa atggactaat	tgggtcaattg ctagaggtagt ggtttagat	540
	gacaagcata tcaactggtct ggtaattaat	ccacaacata ttagacaaat tgctgactct	600
25	tgggcagcaa ttggaaaagc aaatgaaatt	cctttcgcgt tatgttttgg cgttcccca	660
	gcagctat	ttt tagttagttc catgccaat cctgaagg	720
30	ggcgcaatct tgggtgagtc ggttccagta	gtaaaatgtg agaccaacga ttaaatggtt	780
	cctgcaacga gtgagatggt atttgagggt	actttgtcct taacagatac acatctggaa	840
	ggcccat	ttt gtgagatgca tggatatgtt ttcaaaagcc	900
35	tacactgtca aggctatgag ttacagagac	aatgctattc tacctgtttc gaaccccggt	960
	ccttgtacgg atgagacaca taccttgatt	ggttcactag tggctactga ggccaaggag	1020
40	ctggctattg aatctggctt gccaat	ttctg gatgccttta tgccttatga ggctcaggct	1080
	ccttggctta tcttaaagggt ggatttgaaa	gggctgcaag cattgaagac aacgcctgaa	1140
	gaattttgta agaaggtagg tgatat	tttac tttaggacaa aagttggtt tatagtccat	1200
45	gaaataat	ttt tggtggcaga tgatatcgac atatttaact	1260
	tacgttacaa gacatacacc tgttgcagat	cagatggctt ttgatgatgt cacttctttt	1320
50	ccttggctc ccttgtttc gcagtc	atcc agaagtaaga ctatgaaagg tggaaagtgc	1380
	gttactaatt gcatat	tttag acagcaatat gagcgcagtt ttgactacat	1440
	tttgaaaagg gatatccaaa aggattagtt	gacaaagtaa atgaaaattg gaaaaggtag	1500
55	ggatataaat aa		1512

<210> 34

<211> 503

60 <212> PRT

<213> Secuencia artificial

<220>

<223> Sintética

65 <400> 34

Met Arg Lys Leu Asn Pro Ala Leu Glu Phe Arg Asp Phe Ile Gln Val

ES 2 709 217 T3

	1					5					10					15				
5	Leu	Lys	Asp	Glu 20	Asp	Asp	Leu	Ile	Glu 25	Ile	Thr	Glu	Glu	Ile 30	Asp	Pro				
10	Asn	Leu	Glu 35	Val	Gly	Ala	Ile	Met 40	Arg	Lys	Ala	Tyr	Glu 45	Ser	His	Leu				
	Pro	Ala 50	Pro	Leu	Phe	Lys	Asn 55	Leu	Lys	Gly	Ala	Ser 60	Lys	Asp	Leu	Phe				
15	Ser 65	Ile	Leu	Gly	Cys	Pro 70	Ala	Gly	Leu	Arg	Ser 75	Lys	Glu	Lys	Gly	Asp 80				
20	His	Gly	Arg	Ile	Ala 85	His	His	Leu	Gly	Leu 90	Asp	Pro	Lys	Thr	Thr 95	Ile				
25	Lys	Glu	Ile	Ile 100	Asp	Tyr	Leu	Leu	Glu 105	Cys	Lys	Glu	Lys	Glu 110	Pro	Leu				
30	Pro	Pro	Ile 115	Thr	Val	Pro	Val	Ser 120	Ser	Ala	Pro	Cys	Lys 125	Thr	His	Ile				
	Leu	Ser 130	Glu	Glu	Lys	Ile	His 135	Leu	Gln	Ser	Leu	Pro 140	Thr	Pro	Tyr	Leu				
35	His 145	Val	Ser	Asp	Gly	Gly 150	Lys	Tyr	Leu	Gln	Thr 155	Tyr	Gly	Met	Trp	Ile 160				
40	Leu	Gln	Thr	Pro	Asp 165	Lys	Lys	Trp	Thr	Asn 170	Trp	Ser	Ile	Ala	Arg 175	Gly				
45	Met	Val	Val	Asp 180	Asp	Lys	His	Ile	Thr 185	Gly	Leu	Val	Ile	Asn 190	Pro	Gln				
50	His	Ile	Arg 195	Gln	Ile	Ala	Asp	Ser 200	Trp	Ala	Ala	Ile	Gly 205	Lys	Ala	Asn				
	Glu	Ile 210	Pro	Phe	Ala	Leu	Cys 215	Phe	Gly	Val	Pro	Pro 220	Ala	Ala	Ile	Leu				
55	Val 225	Ser	Ser	Met	Pro	Ile 230	Pro	Glu	Gly	Val	Ser 235	Glu	Ser	Asp	Tyr	Val 240				
60	Gly	Ala	Ile	Leu	Gly 245	Glu	Ser	Val	Pro	Val 250	Val	Lys	Cys	Glu	Thr 255	Asn				
65	Asp	Leu	Met	Val 260	Pro	Ala	Thr	Ser	Glu 265	Met	Val	Phe	Glu	Gly 270	Thr	Leu				
	Ser	Leu	Thr	Asp	Thr	His	Leu	Glu	Gly	Pro 98	Phe	Gly	Glu	Met	His	Gly				

ES 2 709 217 T3

		275					280					285				
5	Tyr	Val 290	Phe	Lys	Ser	Gln	Gly 295	His	Pro	Cys	Pro	Leu 300	Tyr	Thr	Val	Lys
10	Ala	Met	Ser	Tyr	Arg	Asp 310	Asn	Ala	Ile	Leu	Pro 315	Val	Ser	Asn	Pro	Gly 320
15	Leu	Cys	Thr	Asp	Glu 325	Thr	His	Thr	Leu	Ile 330	Gly	Ser	Leu	Val	Ala 335	Thr
20	Glu	Ala	Lys	Glu 340	Leu	Ala	Ile	Glu	Ser 345	Gly	Leu	Pro	Ile	Leu 350	Asp	Ala
25	Phe	Met	Pro 355	Tyr	Glu	Ala	Gln	Ala 360	Leu	Trp	Leu	Ile	Leu 365	Lys	Val	Asp
30	Leu	Lys 370	Gly	Leu	Gln	Ala	Leu 375	Lys	Thr	Thr	Pro	Glu 380	Glu	Phe	Cys	Lys
35	Lys 385	Val	Gly	Asp	Ile	Tyr 390	Phe	Arg	Thr	Lys	Val 395	Gly	Phe	Ile	Val	His 400
40	Glu	Ile	Ile	Leu	Val 405	Ala	Asp	Asp	Ile	Asp 410	Ile	Phe	Asn	Phe	Lys 415	Glu
45	Val	Ile	Trp	Ala 420	Tyr	Val	Thr	Arg	His 425	Thr	Pro	Val	Ala	Asp 430	Gln	Met
50	Ala	Phe	Asp 435	Asp	Val	Thr	Ser	Phe 440	Pro	Leu	Ala	Pro	Phe 445	Val	Ser	Gln
55	Ser	Ser 450	Arg	Ser	Lys	Thr	Met 455	Lys	Gly	Gly	Lys	Cys 460	Val	Thr	Asn	Cys
60	Ile	Phe	Arg	Gln	Gln	Tyr 470	Glu	Arg	Ser	Phe	Asp 475	Tyr	Ile	Thr	Cys	Asn 480
65	Phe	Glu	Lys	Gly	Tyr 485	Pro	Lys	Gly	Leu	Val 490	Asp	Lys	Val	Asn	Glu 495	Asn
70	Trp	Lys	Arg	Tyr 500	Gly	Tyr	Lys									
75	<210>	35														
80	<211>	1512														
85	<212>	ADN														
90	<213>	Secuencia artificial														
95	<220>															
100	<223>	sintética														
105	<400>	35														

ES 2 709 217 T3

	atgaggaagc taaatccagc tttagaat	ttt agagacttta tccaggtc	ttt aaaagatgaa	60
	gatgacttaa tcgaaattac cgaagagatt	gatccaaatc tcgaagtagg tgcaattatg	120	
5	aggaaggcct atgaatccca cttaccagcc	ccgttat	180	
	aaggatcttt tcagcatttt aggttgccca	gccggtttga gaagtaagga gaaaggagat	240	
10	catggtagaa ttgcccatca tctggggctc	gacccaaaaa caactatcaa ggaaatcata	300	
	gattatttgc tggagtgtaa ggagaaggaa	cctctcccc caatcactgt tcctgtgtca	360	
	tctgcacctt gtaaaacaca tatactttct	gaagaaaaaa tacatctaca aagcctgcca	420	
15	acaccatata tacatgtttc agacggtggc	aagtacttac aaacgtacgg aatgtggatt	480	
	cttcaaactc cagataaaaa atggactaat	tggatcaattg ctattggtat ggtttagat	540	
20	gacaagcata tcaactggtc ggtaattaaa	ccacaacata ttagacaaat tgctgactct	600	
	tgggcagcaa ttggaaaagc aaatgaaatt	cctttcgcgt tatgttttgg cgttcccca	660	
	gcagctat	tagttagttc catgccaatt cctgaagg	720	
25	ggcgcaatct tgggtgagtc gggtccagta	gtaaaatgtg agaccaacga ttaaatggtt	780	
	cctgcaacga gtgagatggt atttgaggg	actttgtcct taacagatac acatctggaa	840	
30	ggcccat	gtgagatgca tggatatgtt ttcaaaagcc	900	
	tacactgtca aggtatgag ttacagagac	aatgctattc tacctgtttc gaacccggt	960	
	ctttgtacgg atgagacaca taccttgatt	ggttcactag tggctactga ggccaaggag	1020	
35	ctggctattg aatctggctt gccaattctg	gatgccttta tgccttatga ggctcaggct	1080	
	ctttggctta tcttaaagggt ggatttgaaa	gggctgcaag cattgaagac aacgcctgaa	1140	
40	gaattttgta agaaggtagg tgatatttac	tttaggacaa aagttggtt tatagtccat	1200	
	gaaataattt tgggtggcaga tgatatcgac	atatttaact tcaaagaagt catctgggcc	1260	
	tacgttacaa gacatacacc tgttgcagat	cagatggctt ttgatgatgt cacttctttt	1320	
45	cctttggctc cctttgtttc gcagtcaccc	agaagtaaga ctatgaaagg tggaaagtgc	1380	
	gttactaatt gcatatttag acagcaatat	gagcgcagtt ttgactacat aacttgtaat	1440	
50	tttggaaagg gatatccaaa aggattagtt	gacaaagtaa atgaaaattg gaaaaggtag	1500	
	ggatataaat aa		1512	

55 <210> 36
 <211> 503
 <212> PRT
 <213> Secuencia artificial

60 <220>
 <223> Sintética
 <400> 36

65 Met Arg Lys Leu Asn Pro Ala Leu Glu Phe Arg Asp Phe Ile Gln Val
 1 5 10 15

Leu Lys Asp Glu Asp Asp Leu Ile Glu Ile Thr Glu Glu Ile Asp Pro
 100

ES 2 709 217 T3

	20	25	30
5	Asn Leu Glu Val Gly Ala Ile Met Arg Lys Ala Tyr Glu Ser His Leu		
	35	40	45
10	Pro Ala Pro Leu Phe Lys Asn Leu Lys Gly Ala Ser Lys Asp Leu Phe		
	50	55	60
15	Ser Ile Leu Gly Cys Pro Ala Gly Leu Arg Ser Lys Glu Lys Gly Asp		
	65	70	75
20	His Gly Arg Ile Ala His His Leu Gly Leu Asp Pro Lys Thr Thr Ile		
	85	90	95
25	Lys Glu Ile Ile Asp Tyr Leu Leu Glu Cys Lys Glu Lys Glu Pro Leu		
	100	105	110
30	Pro Pro Ile Thr Val Pro Val Ser Ser Ala Pro Cys Lys Thr His Ile		
	115	120	125
35	Leu Ser Glu Glu Lys Ile His Leu Gln Ser Leu Pro Thr Pro Tyr Leu		
	130	135	140
40	His Val Ser Asp Gly Gly Lys Tyr Leu Gln Thr Tyr Gly Met Trp Ile		
	145	150	155
45	Leu Gln Thr Pro Asp Lys Lys Trp Thr Asn Trp Ser Ile Ala Ile Gly		
	165	170	175
50	Met Val Val Asp Asp Lys His Ile Thr Gly Leu Val Ile Lys Pro Gln		
	180	185	190
55	His Ile Arg Gln Ile Ala Asp Ser Trp Ala Ala Ile Gly Lys Ala Asn		
	195	200	205
60	Glu Ile Pro Phe Ala Leu Cys Phe Gly Val Pro Pro Ala Ala Ile Leu		
	210	215	220
65	Val Ser Ser Met Pro Ile Pro Glu Gly Val Ser Glu Ser Asp Tyr Val		
	225	230	235
70	Gly Ala Ile Leu Gly Glu Ser Val Pro Val Val Lys Cys Glu Thr Asn		
	245	250	255
75	Asp Leu Met Val Pro Ala Thr Ser Glu Met Val Phe Glu Gly Thr Leu		
	260	265	270
80	Ser Leu Thr Asp Thr His Leu Glu Gly Pro Phe Gly Glu Met His Gly		
	275	280	285
85	Tyr Val Phe Lys Ser Gln Gly His Pro Cys Pro Leu Tyr Thr Val Lys		
		101	

ES 2 709 217 T3

	290	295	300	
5	Ala 305 Met Ser Tyr Arg Asp 310 Asn Ala Ile Leu Pro 315 Val Ser Asn Pro Gly 320			
10	Leu Cys Thr Asp Glu 325 Thr His Thr Leu Ile 330 Gly Ser Leu Val Ala Thr 335			
15	Glu Ala Lys Glu 340 Leu Ala Ile Glu Ser 345 Gly Leu Pro Ile Leu 350 Asp Ala			
20	Phe Met Pro 355 Tyr Glu Ala Gln Ala 360 Leu Trp Leu Ile Leu 365 Lys Val Asp			
25	Leu Lys 370 Gly Leu Gln Ala Leu 375 Lys Thr Thr Pro Glu 380 Glu Phe Cys Lys			
30	Lys Val 385 Gly Asp Ile Tyr 390 Phe Arg Thr Lys Val 395 Gly Phe Ile Val His 400			
35	Glu Ile Ile Leu Val 405 Ala Asp Asp Ile Asp 410 Ile Phe Asn Phe Lys 415 Glu			
40	Val Ile Trp Ala 420 Tyr Val Thr Arg His 425 Thr Pro Val Ala Asp 430 Gln Met			
45	Ala Phe Asp 435 Asp Val Thr Ser Phe 440 Pro Leu Ala Pro Phe 445 Val Ser Gln			
50	Ser Ser 450 Arg Ser Lys Thr Met 455 Lys Gly Gly Lys Cys 460 Val Thr Asn Cys			
55	Ile Phe Arg Gln Gln Tyr 470 Glu Arg Ser Phe Asp 475 Tyr Ile Thr Cys Asn 480			
60	Phe Glu Lys Gly Tyr 485 Pro Lys Gly Leu Val 490 Asp Lys Val Asn Glu 495 Asn			
65	Trp Lys Arg Tyr 500 Gly Tyr Lys			
	<210> 37			
	<211> 1512			
	<212> ADN			
	<213> Secuencia artificial			
	<220>			
	<223> sintética			
	<400> 37			
	atgaggaagc taaatccagc tttagaattt agagacttta tccaggcttt aaaagatgaa			60
	gatgacttaa tcgaaattac cgaagagatt gatccaaatc tcgaagtagg tgcaattatg			120

ES 2 709 217 T3

```

    aggaaggcct atgaatccca cttaccagcc ccgttatTTa aaaatctcaa aggtgcttcg      180
    aaggatcttt tcagcatttt aggttgccca gccggtttga gaagtaagga gaaaggagat      240
5   catggtagaa ttgcccatca tctggggctc gacccaaaaa caactatcaa ggaaatcata      300
    gattattttgc tggagtgtaa ggagaaggaa cctctccccc caatcactgt tcctgtgtca      360
10  tctgcacctt gtaaaacaca tatactttct gaagaaaaaa tacatctaca aagcctgcca      420
    acaccatata tacatgtttc agacggtggc aagtacttac aaacgtacgg aatgtggatt      480
    cttcaaactc cagataaaaa atggactaat tggTcaattg ctagaggTat ggTtgtagat      540
15  gacaagcata tcaTtggtct ggtaattaaa ccacaaccta ttagacaaat tgctgactct      600
    tgggcagcaa ttggaaaagc aaatgaaatt cttttcgcgT tatgtttttg cgttcccca      660
20  gcagctatTT tagttagttc catgccaaTt cctgaaggTg tttctgaatc ggattatgtt      720
    ggcgcaatct tgggtgagtc ggttccagta gtaaaatgtg agaccaacga tttaatggTt      780
    cctgcaacga gtgagatggT atttgagggt actttgtcct taacagatac acatctggaa      840
25  ggcccatTTg gtgagatgca tggatatgtt ttcaaaagcc aaggTcatcc ttgtccattg      900
    tacactgtca aggctatgag ttacagagac aatgctattc tacctgtttc gaaccccggt      960
30  ctttgtacgg atgagacaca taccttgatt ggttcactag tggctactga ggccaaggag     1020
    ctggctattg aatctggctt gccaattctg gatgcctTTa tgccttatga ggctcaggct     1080
    ctttggctTa tcttaaaggT ggatttgaaa gggctgcaag cattgaagac aacgcctgaa     1140
35  gaattttgTa agaaggtagg tgatattTat tttaggacaa aagTtggttT tatagtccat     1200
    gaaataattt tggTggcaga tgatatcgac atattTaaTt tcaaagaagT catctgggcc     1260
40  tacgttaciaa gacatacacc tgTtgCagat cagatggctt ttgatgatgt cacttctttt     1320
    cctttggctc cctttgtttc gcagTcatcc agaagtaaga ctatgaaagg tggaaagtgc     1380
    gttactaatt gcatattTtag acagcaatat gagcgCagtt ttgactacat aacttgtaat     1440
45  tttgaaaagg gatatccaaa aggattagtt gacaaagTaa atgaaaattg gaaaaggTat     1500
    ggatataaat aa                                                                1512

```

```

50  <210> 38
    <211> 503
    <212> PRT
    <213> Secuencia artificial

```

```

55  <220>
    <223> Sintética
    <400> 38

```

```

60  Met Arg Lys Leu Asn Pro Ala Leu Glu Phe Arg Asp Phe Ile Gln Val
    1           5           10           15

```

```

65  Leu Lys Asp Glu Asp Asp Leu Ile Glu Ile Thr Glu Glu Ile Asp Pro
    20           25           30

```

```

    Asn Leu Glu Val Gly Ala Ile Met Arg Lys Ala Tyr Glu Ser His Leu
                                103

```

ES 2 709 217 T3

		35					40					45					
5	Pro	Ala 50	Pro	Leu	Phe	Lys	Asn 55	Leu	Lys	Gly	Ala	Ser 60	Lys	Asp	Leu	Phe	
10	Ser 65	Ile	Leu	Gly	Cys	Pro 70	Ala	Gly	Leu	Arg	Ser 75	Lys	Glu	Lys	Gly	Asp 80	
	His	Gly	Arg	Ile	Ala 85	His	His	Leu	Gly	Leu 90	Asp	Pro	Lys	Thr	Thr 95	Ile	
15	Lys	Glu	Ile	Ile 100	Asp	Tyr	Leu	Leu	Glu 105	Cys	Lys	Glu	Lys	Glu 110	Pro	Leu	
20	Pro	Pro	Ile 115	Thr	Val	Pro	Val	Ser 120	Ser	Ala	Pro	Cys	Lys 125	Thr	His	Ile	
25	Leu	Ser 130	Glu	Glu	Lys	Ile	His 135	Leu	Gln	Ser	Leu	Pro 140	Thr	Pro	Tyr	Leu	
30	His 145	Val	Ser	Asp	Gly	Gly 150	Lys	Tyr	Leu	Gln	Thr 155	Tyr	Gly	Met	Trp	Ile 160	
	Leu	Gln	Thr	Pro	Asp 165	Lys	Lys	Trp	Thr	Asn 170	Trp	Ser	Ile	Ala	Arg 175	Gly	
35	Met	Val	Val	Asp 180	Asp	Lys	His	Ile	Thr 185	Gly	Leu	Val	Ile	Lys 190	Pro	Gln	
40	Pro	Ile	Arg 195	Gln	Ile	Ala	Asp	Ser 200	Trp	Ala	Ala	Ile	Gly 205	Lys	Ala	Asn	
45	Glu	Ile 210	Pro	Phe	Ala	Leu	Cys 215	Phe	Gly	Val	Pro	Pro 220	Ala	Ala	Ile	Leu	
50	Val 225	Ser	Ser	Met	Pro	Ile 230	Pro	Glu	Gly	Val	Ser 235	Glu	Ser	Asp	Tyr	Val 240	
	Gly	Ala	Ile	Leu	Gly 245	Glu	Ser	Val	Pro	Val 250	Val	Lys	Cys	Glu	Thr 255	Asn	
55	Asp	Leu	Met	Val 260	Pro	Ala	Thr	Ser	Glu 265	Met	Val	Phe	Glu	Gly 270	Thr	Leu	
60	Ser	Leu	Thr 275	Asp	Thr	His	Leu	Glu 280	Gly	Pro	Phe	Gly	Glu 285	Met	His	Gly	
65	Tyr	Val 290	Phe	Lys	Ser	Gln	Gly 295	His	Pro	Cys	Pro	Leu 300	Tyr	Thr	Val	Lys	
	Ala	Met	Ser	Tyr	Arg	Asp	Asn	Ala	Ile	Leu	Pro	Val	Ser	Asn	Pro	Gly	
										104							

ES 2 709 217 T3

305 310 315 320
 5 Leu Cys Thr Asp Glu Thr His Thr Leu Ile Gly Ser Leu Val Ala Thr
 325 330 335
 10 Glu Ala Lys Glu Leu Ala Ile Glu Ser Gly Leu Pro Ile Leu Asp Ala
 340 345 350
 15 Phe Met Pro Tyr Glu Ala Gln Ala Leu Trp Leu Ile Leu Lys Val Asp
 355 360 365
 20 Leu Lys Gly Leu Gln Ala Leu Lys Thr Thr Pro Glu Glu Phe Cys Lys
 370 375 380
 25 Lys Val Gly Asp Ile Tyr Phe Arg Thr Lys Val Gly Phe Ile Val His
 385 390 395 400
 30 Glu Ile Ile Leu Val Ala Asp Asp Ile Asp Ile Phe Asn Phe Lys Glu
 405 410 415
 35 Val Ile Trp Ala Tyr Val Thr Arg His Thr Pro Val Ala Asp Gln Met
 420 425 430
 40 Ala Phe Asp Asp Val Thr Ser Phe Pro Leu Ala Pro Phe Val Ser Gln
 435 440 445
 45 Ser Ser Arg Ser Lys Thr Met Lys Gly Gly Lys Cys Val Thr Asn Cys
 450 455 460
 50 Ile Phe Arg Gln Gln Tyr Glu Arg Ser Phe Asp Tyr Ile Thr Cys Asn
 465 470 475 480
 55 Phe Glu Lys Gly Tyr Pro Lys Gly Leu Val Asp Lys Val Asn Glu Asn
 485 490 495
 60 Trp Lys Arg Tyr Gly Tyr Lys
 500
 65 <210> 39
 <211> 4
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Sintética
 <400> 39
 Gly Gly Ser Gly
 1
 <210> 40
 <211> 6
 <212> PRT

<213> Secuencia artificial
 <220>
 <223> sintética
 5 <400> 40
 Gly Ser Gly Ser Gly Ser
 1 5
 10
 <210> 41
 <211> 8
 <212> PRT
 15 <213> Secuencia artificial
 <220>
 <223> sintética
 20 <400> 41
 Gly Gly Ser Gly Gly Gly Ser Gly
 1 5
 25
 <210> 42
 <211> 9
 <212> PRT
 30 <213> Secuencia artificial
 <220>
 <223> sintética
 <400> 42
 35 Ser Gly Gly Ser Gly Gly Ser Gly Gly
 1 5
 40 <210> 43
 <211> 12
 <212> PRT
 <213> Secuencia artificial
 45 <220>
 <223> sintética
 <400> 43
 50 Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly
 1 5 10
 55 <210> 44
 <211> 15
 <212> PRT
 <213> Secuencia artificial
 <220>
 60 <223> sintética
 <400> 44
 65 Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
 1 5 10 15

REIVINDICACIONES

- 5 1. Una célula hospedadora que comprende: (a) una proteína de fusión expresada recombinantemente que comprende un primer dominio que comprende una fenilalanina amonio liasa, y un segundo dominio que comprende una cinamato descarboxilasa; y (b) un transportador unido a la membrana expresado recombinantemente.
- 10 2. La célula hospedadora de la reivindicación 1, donde la cinamato descarboxilasa producida por dicha célula es una cinamato descarboxilasa mutada que comprende una mutación en una posición de residuo aminoacídico que corresponde a una posición seleccionada del grupo que consiste en: 155, 156, 159, 162, 163, 164, 172, 173, 174, 175, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 226, 227, 280, 285, 286, 287, 291, 326, 331, 360, 361, 395, 396, 398, 440, 441 de la SEC. N.º ID: 8, y combinaciones de las mismas.
- 15 3. La célula hospedadora de la reivindicación 1, donde la fenilalanina amonio liasa producida por dicha célula es una fenilalanina amonio liasa que comprende una secuencia aminoacídica seleccionada del grupo que consiste en: SEC. N.º ID: 2, SEC. N.º ID: 4, SEC. N.º ID: 6, y un fragmento funcional o variante de las mismas.
- 20 4. Un método para cristalizar una proteína de fusión que comprende un primer dominio que comprende una fenilalanina amonio liasa, y un segundo dominio que comprende una cinamato descarboxilasa expresada por la célula hospedadora de cualesquiera de las reivindicaciones anteriores, el método comprende:
 - a) obtener la proteína de fusión expresada por la célula hospedadora de cualesquiera de las reivindicaciones anteriores;
 - 25 b) proporcionar una solución de la proteína de fusión a una concentración de aproximadamente 1 mg/ml a aproximadamente 50 mg/ml;
 - c) mezclar la solución de la proteína de fusión con una solución de depósito en una relación de volúmenes de aproximadamente 1:10 a aproximadamente 10:1;
 - d) añadir ácido 3-hidroxicinámico a la solución del paso (a) o del paso (b) a una concentración final de ácido 3-hidroxicinámico del 0,01% en p/v al 1% en p/v; y
 - 30 e) mantener la mezcla del paso (c) a una temperatura de 1 °C a 20 °C adecuada para la formación del cristal de la proteína de fusión, de manera que la proteína de fusión cristalizada se encuentra en un complejo con el ácido 3-hidroxicinámico.

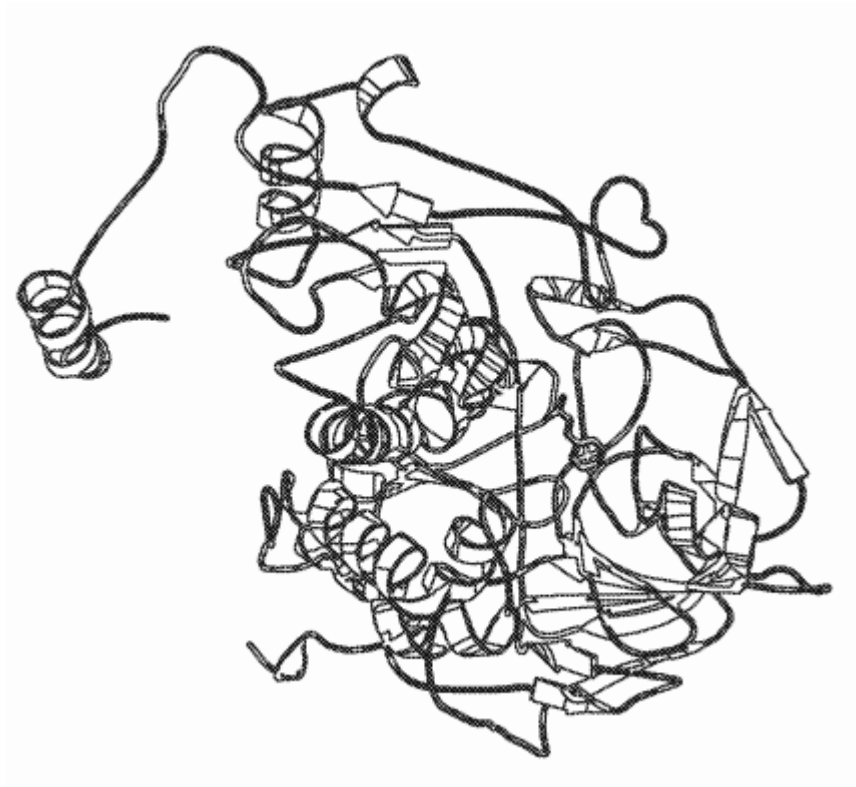


FIG. 1A

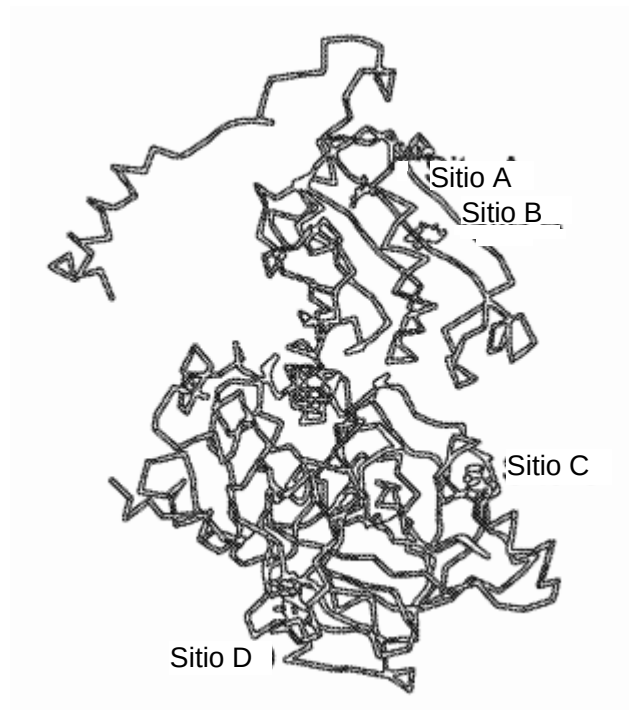


FIG. 1B

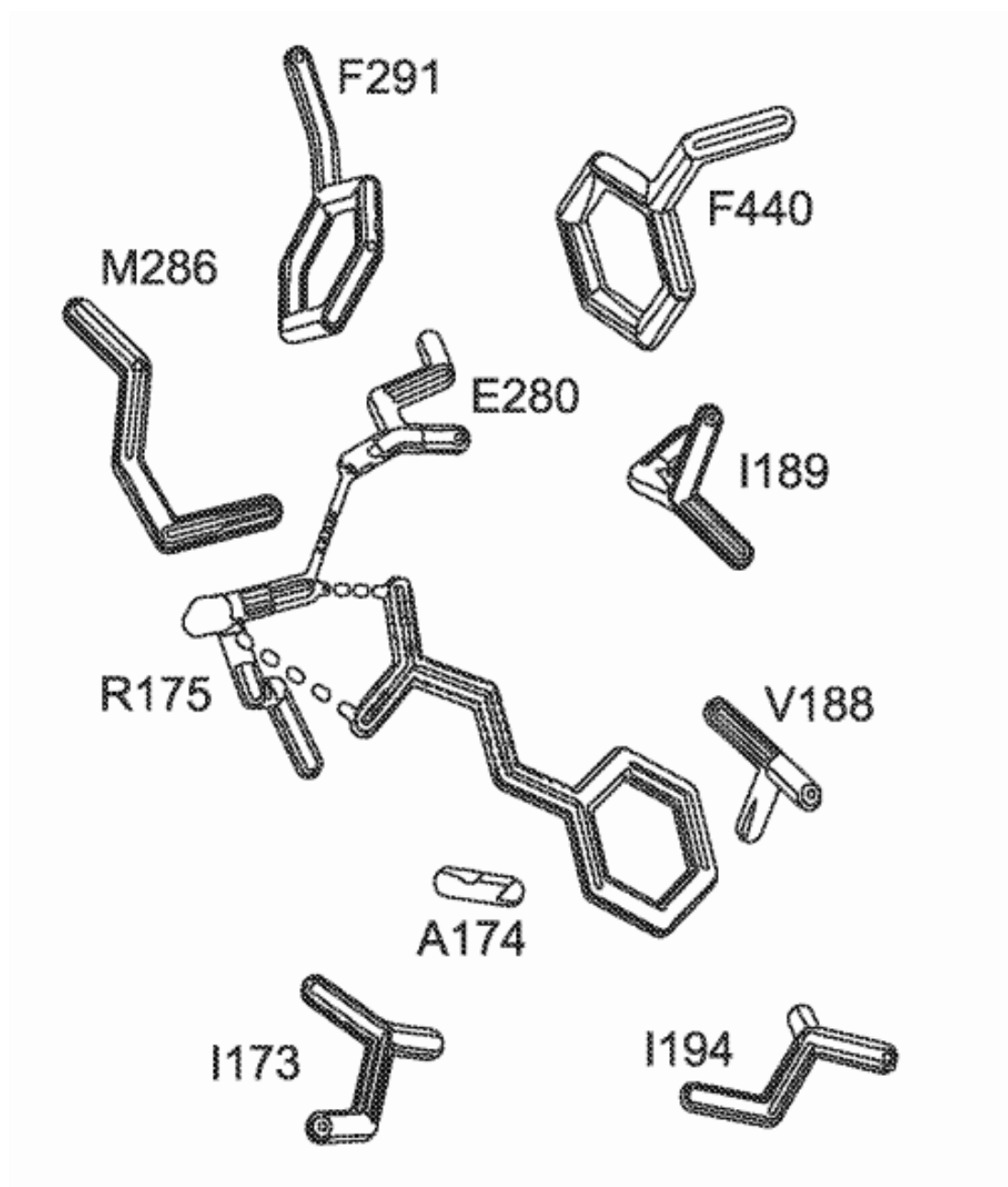


FIG. 1C

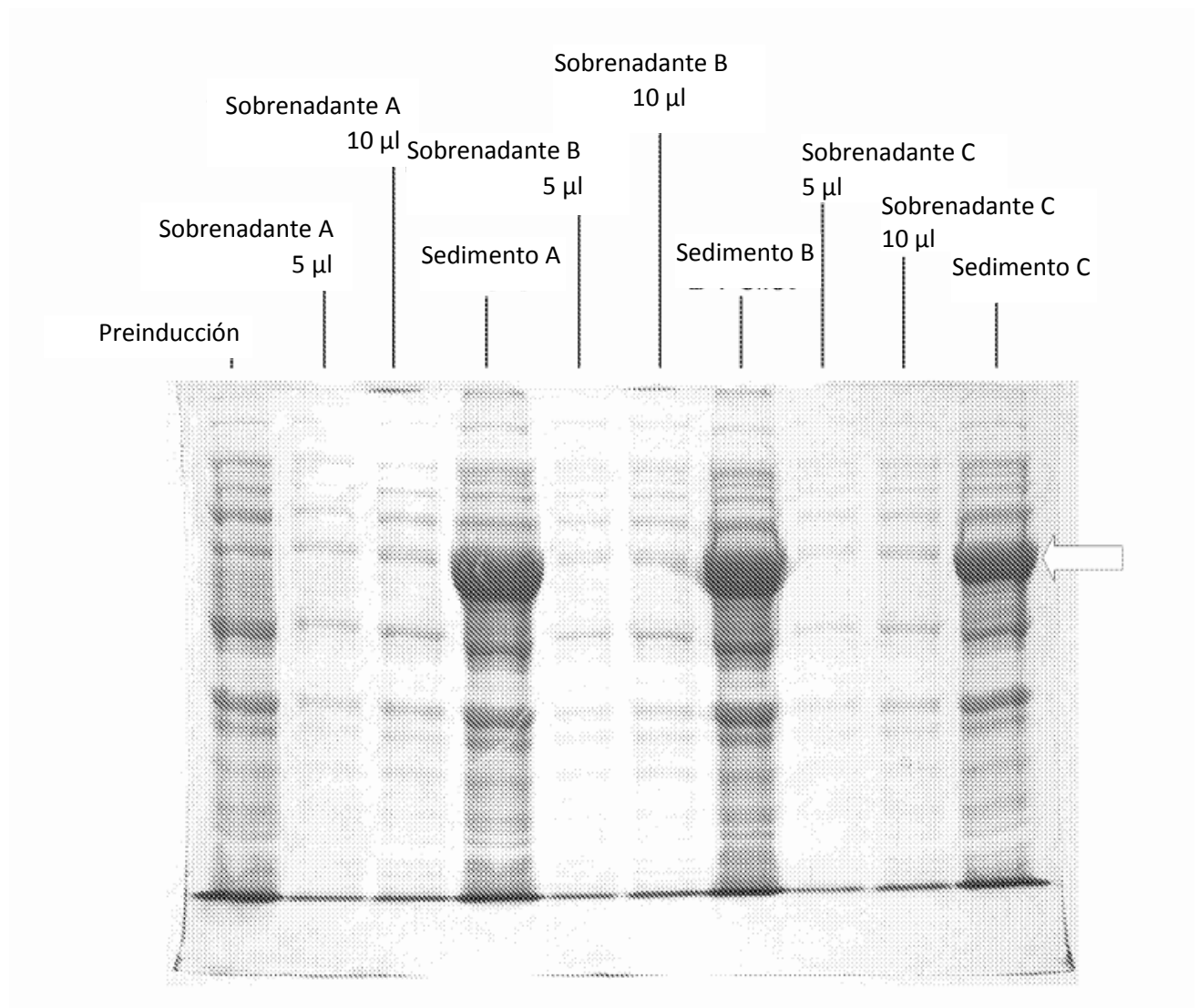


FIG. 2A

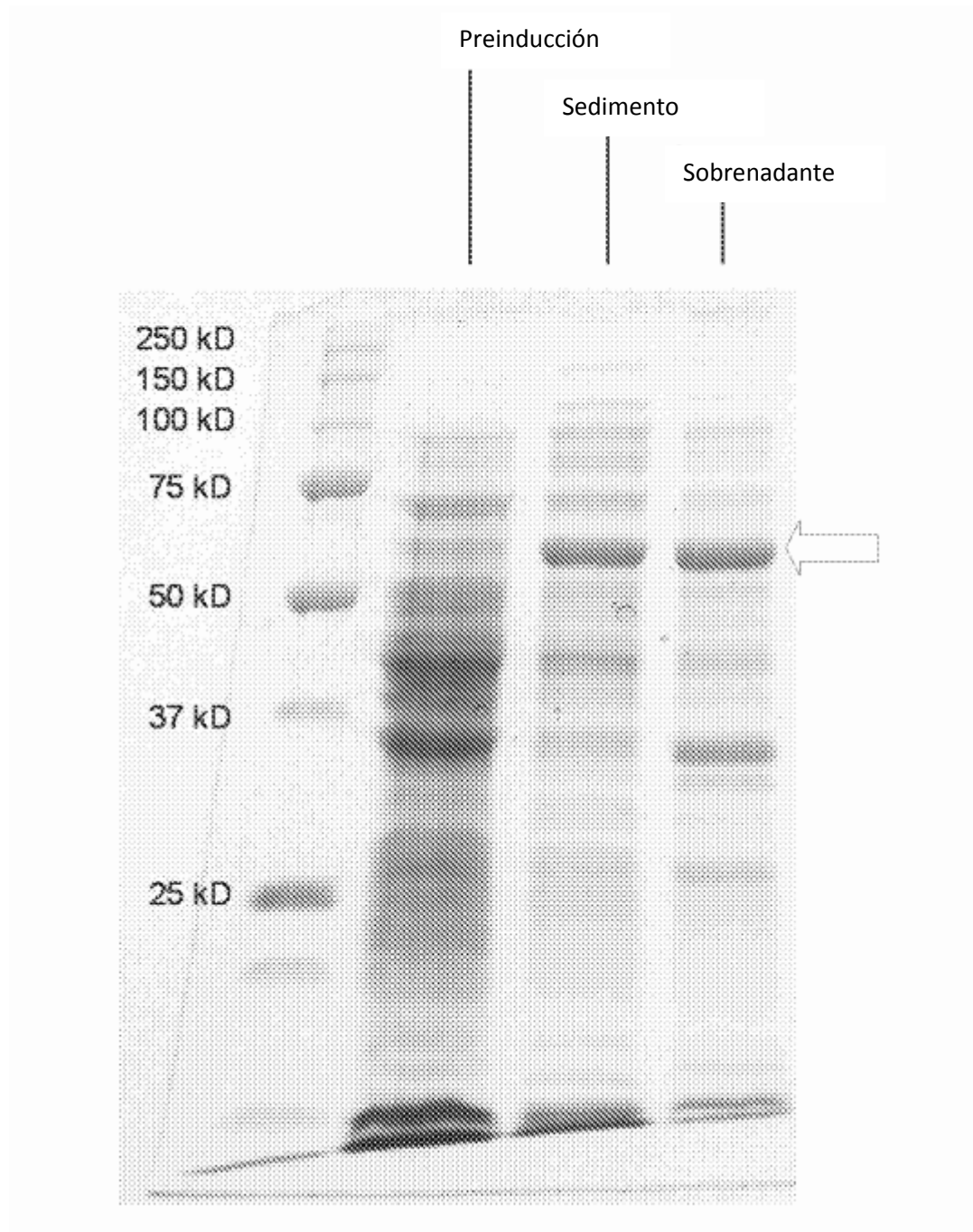


FIG. 2B

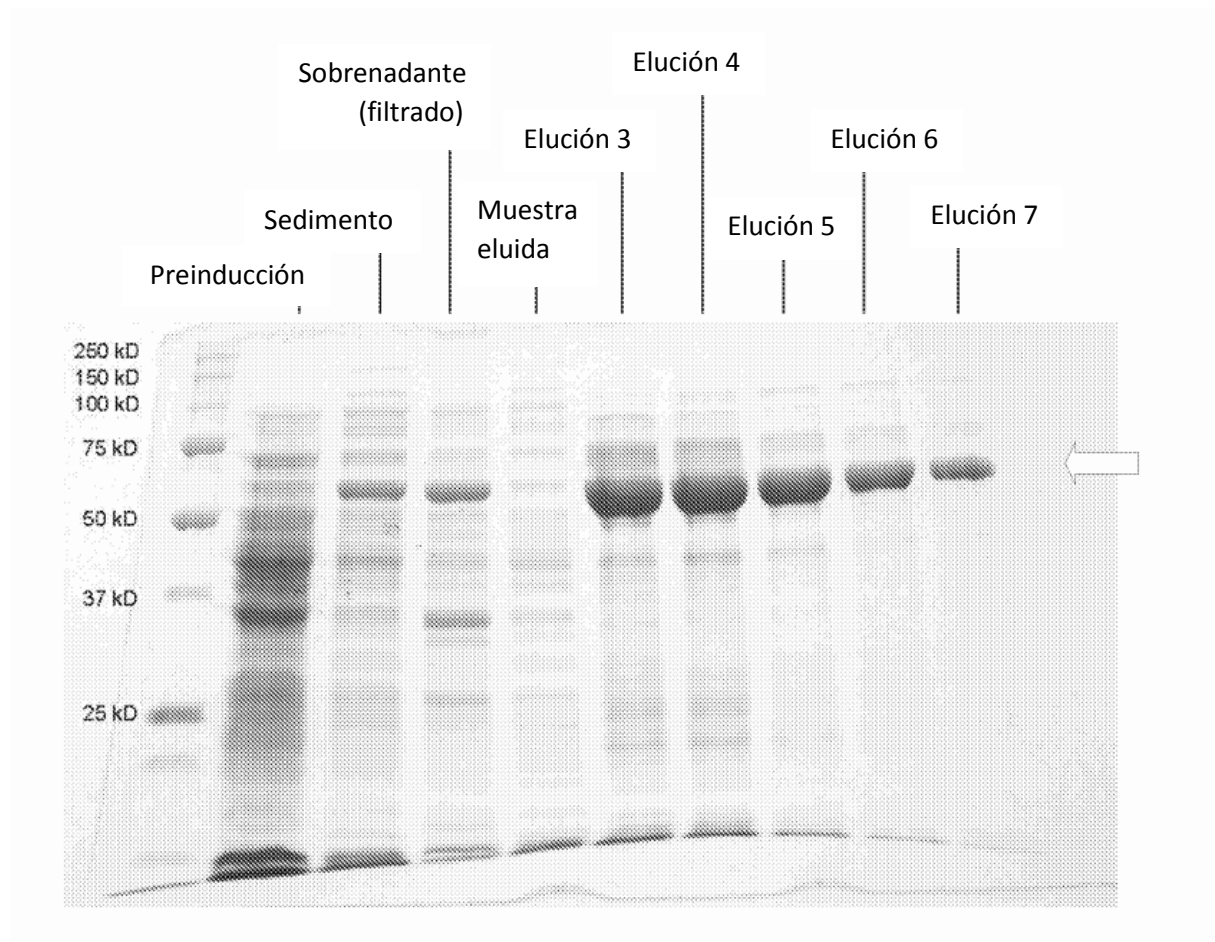


FIG. 3

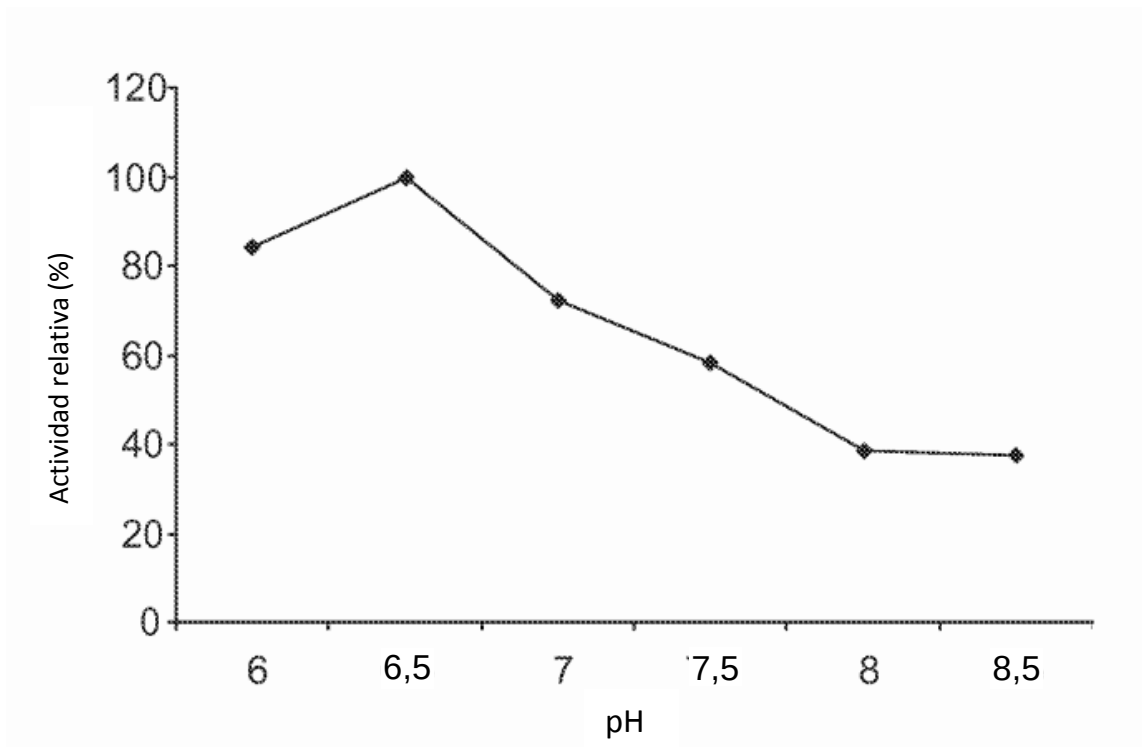


FIG. 4

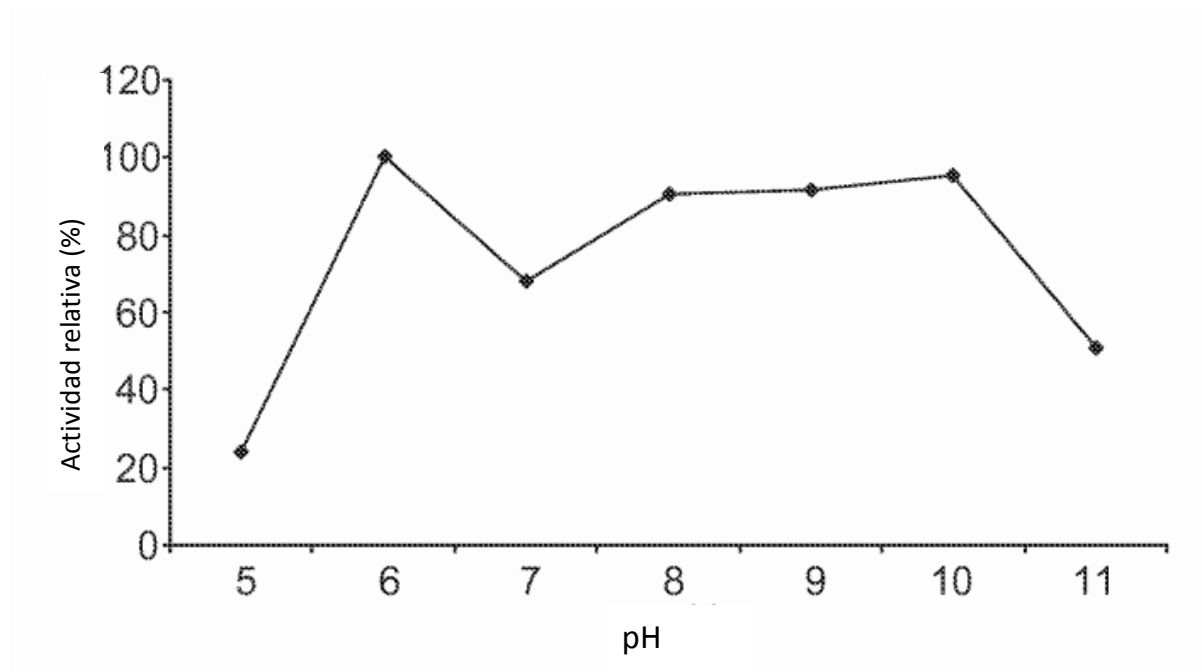


FIG. 5

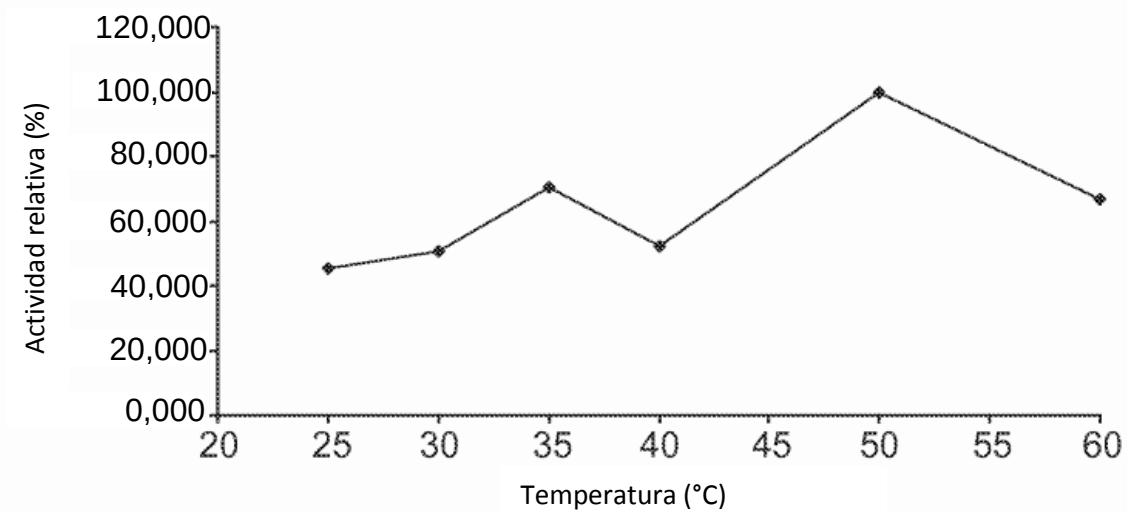


FIG. 6

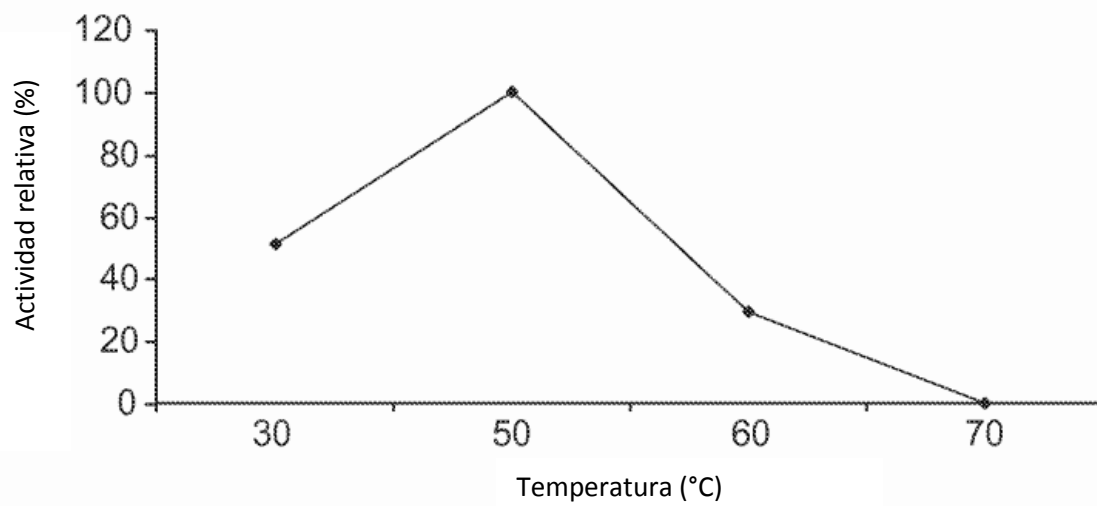


FIG. 7

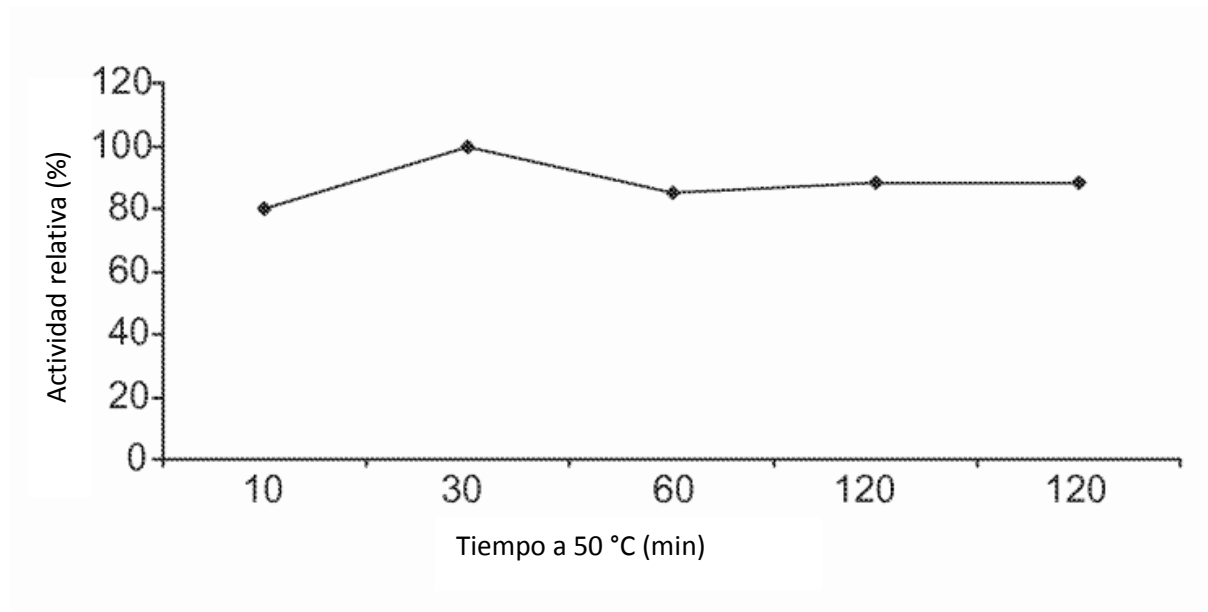
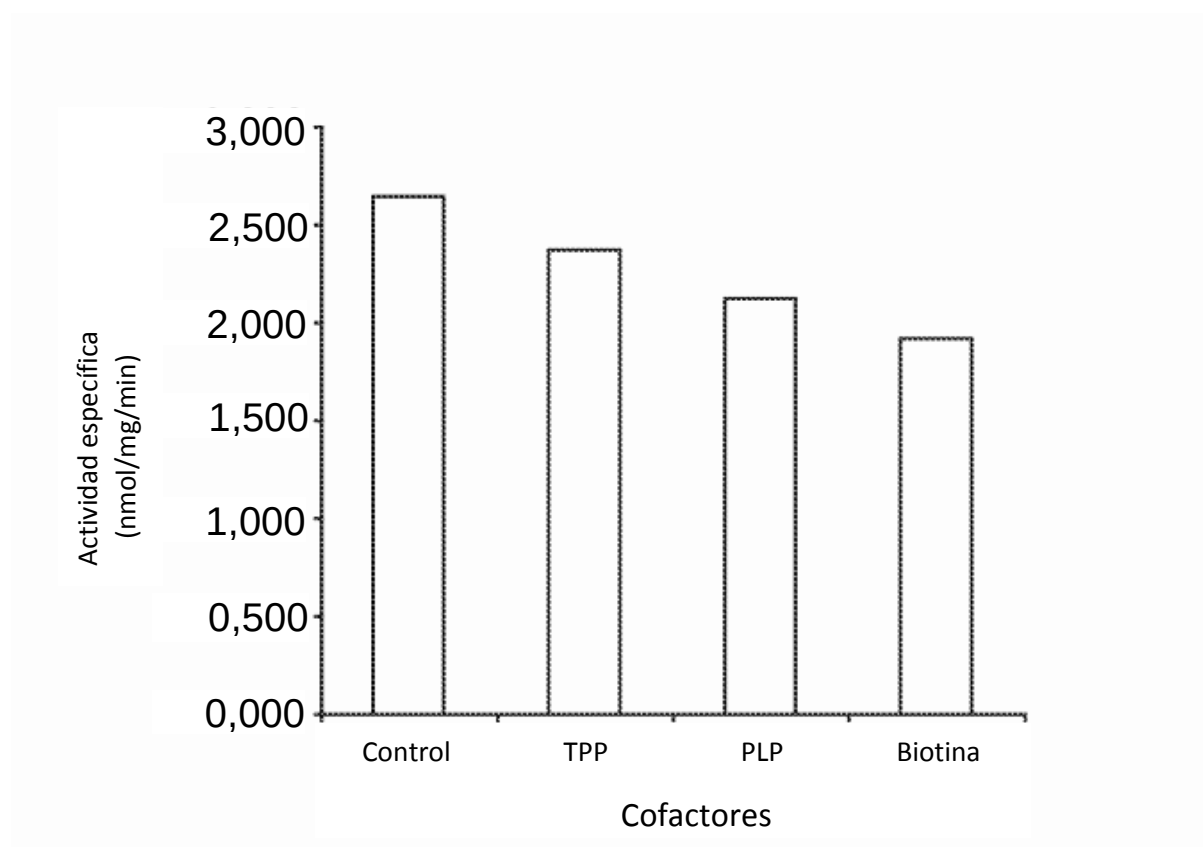


FIG. 8

**FIG. 9**

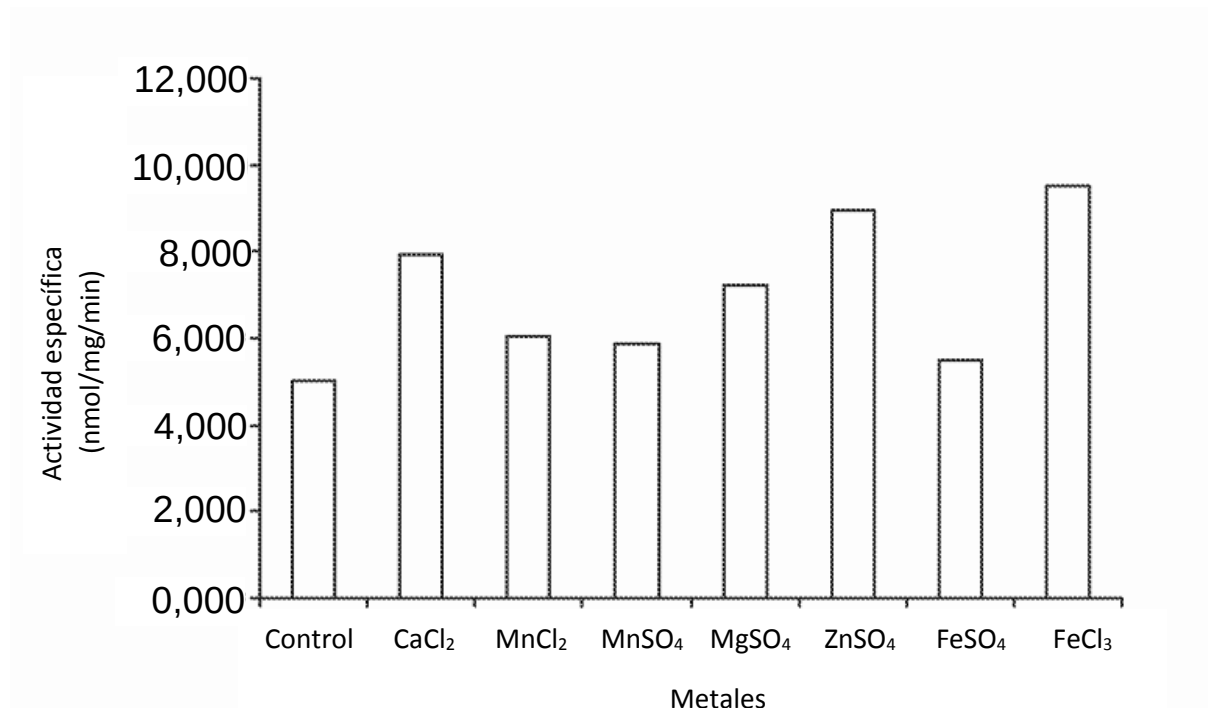


FIG. 10A

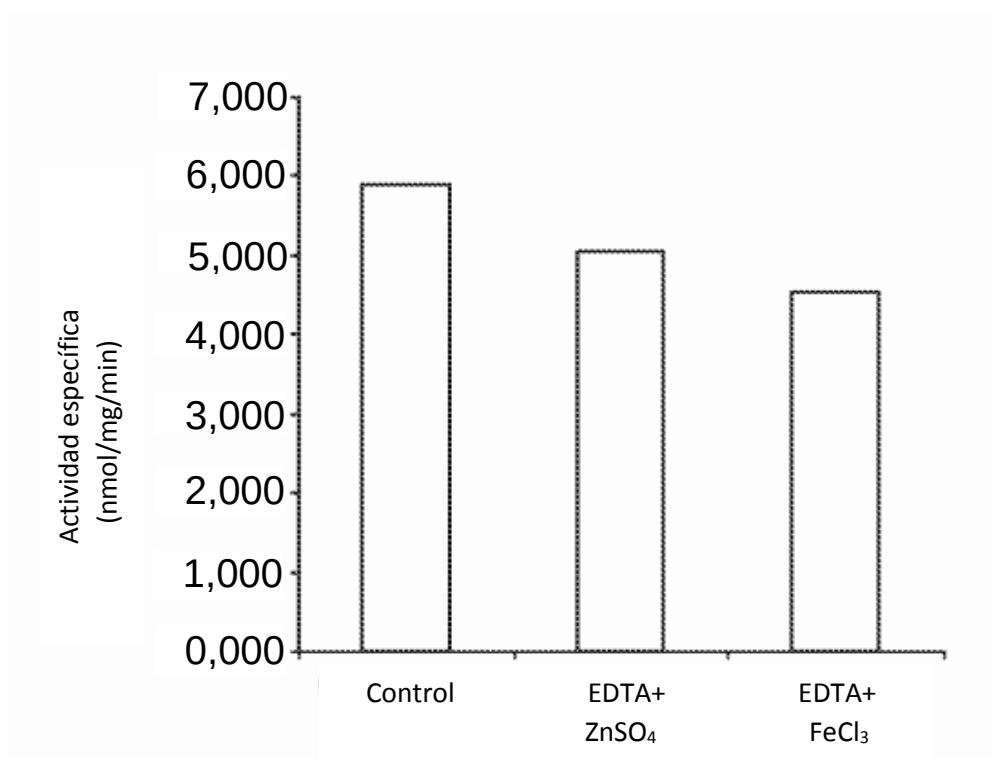


FIG. 10B

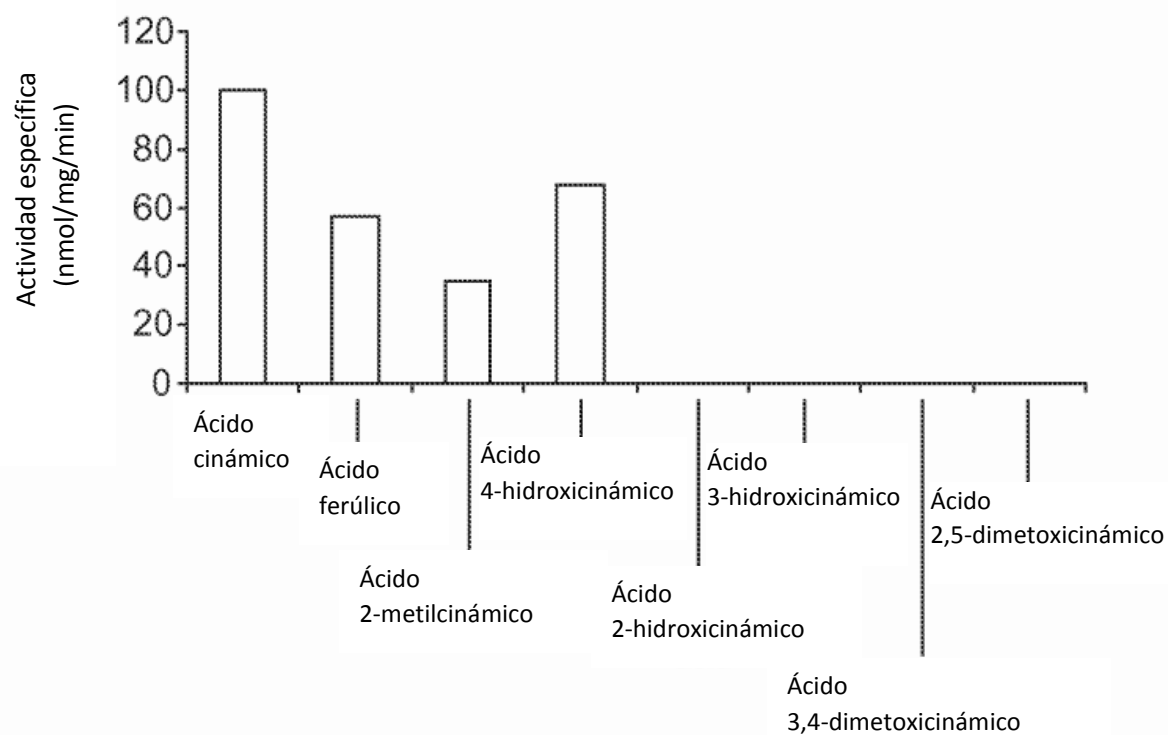


FIG. 11

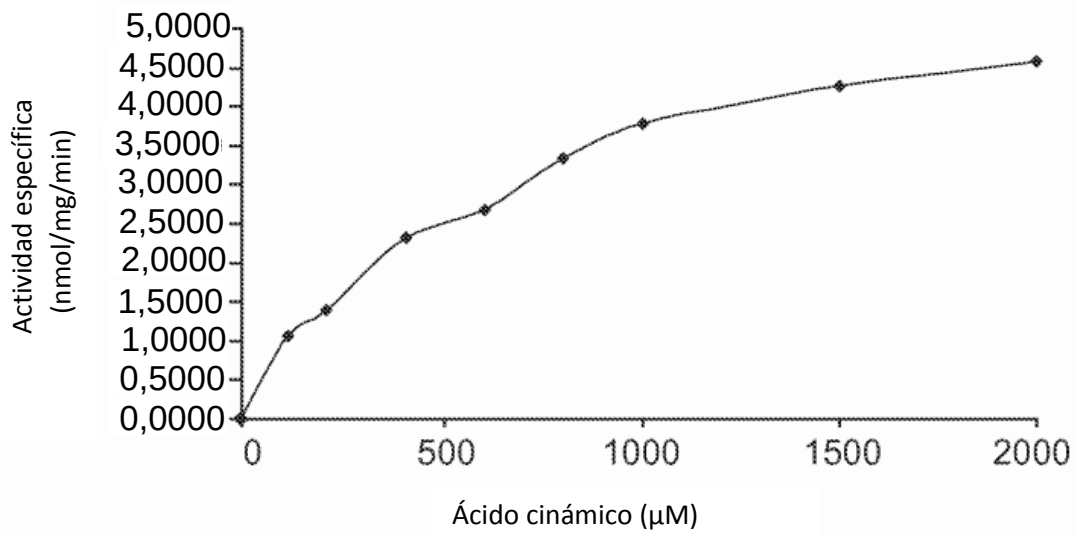


FIG. 12A

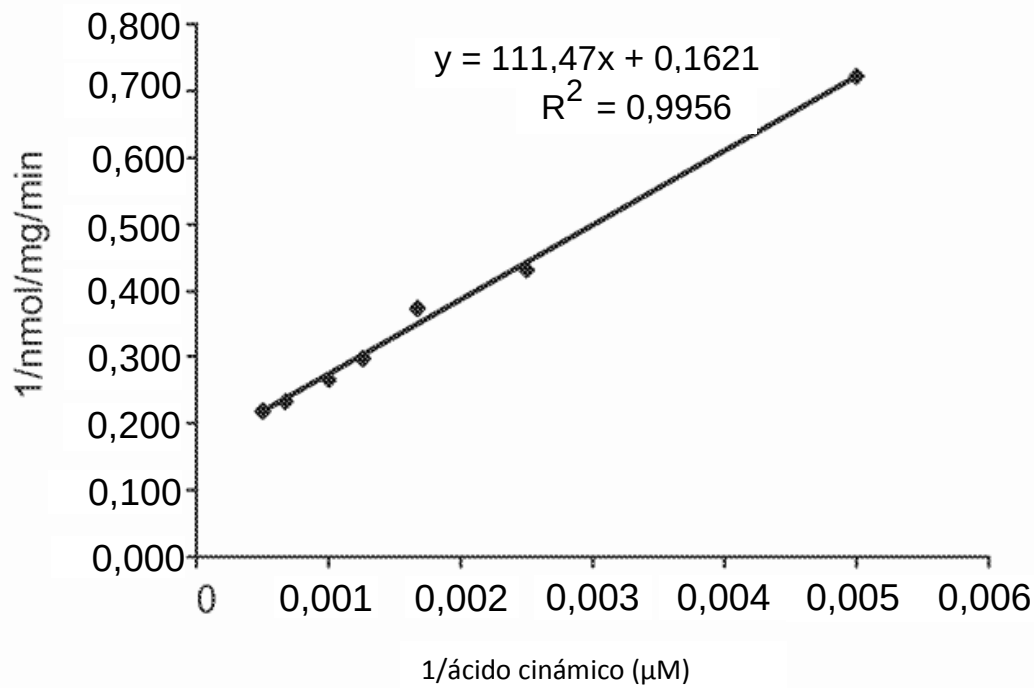


FIG. 12B

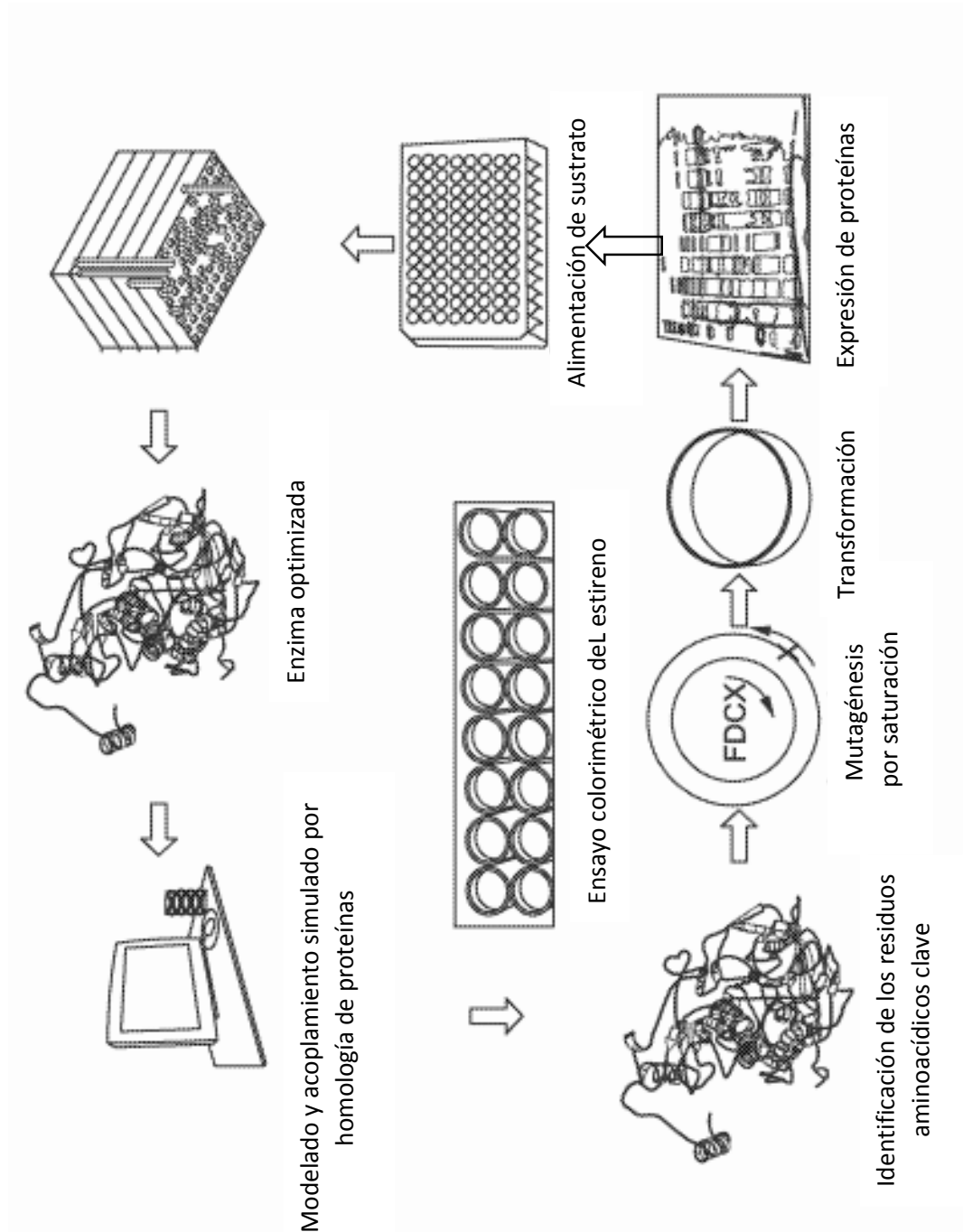


FIG. 13

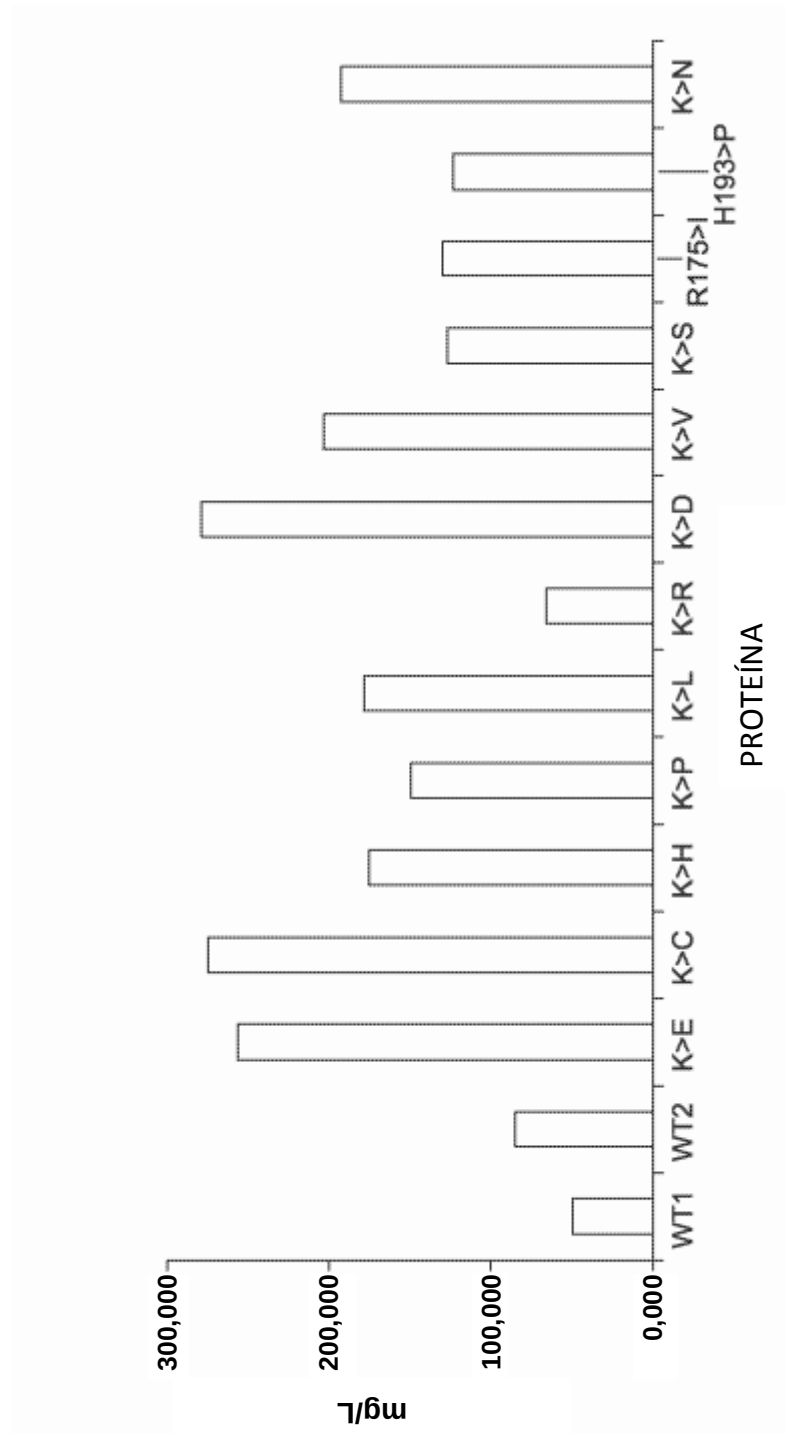


FIG. 14

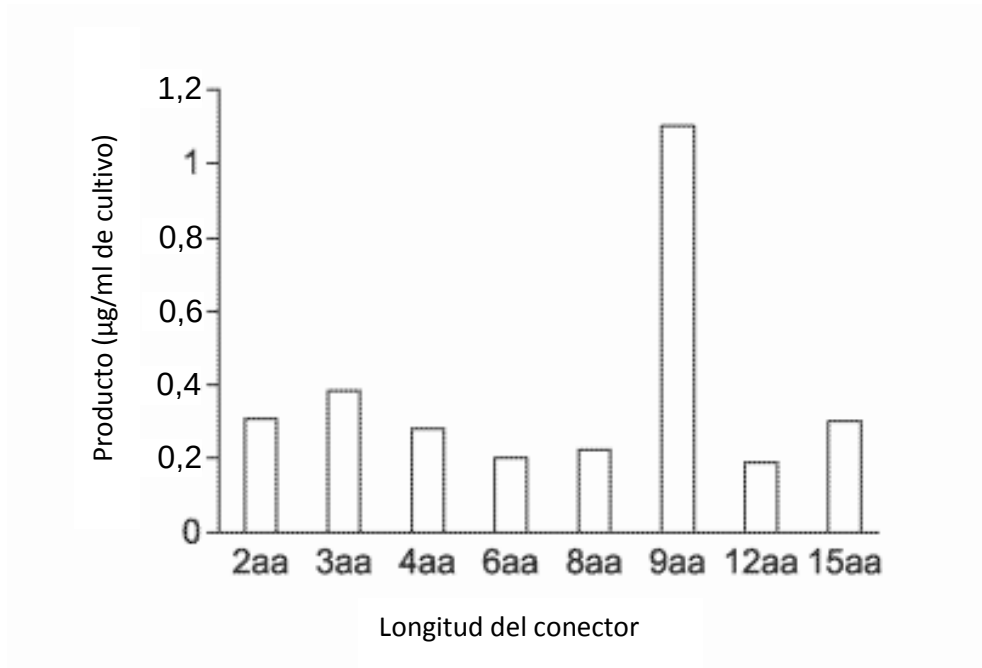


FIG. 15

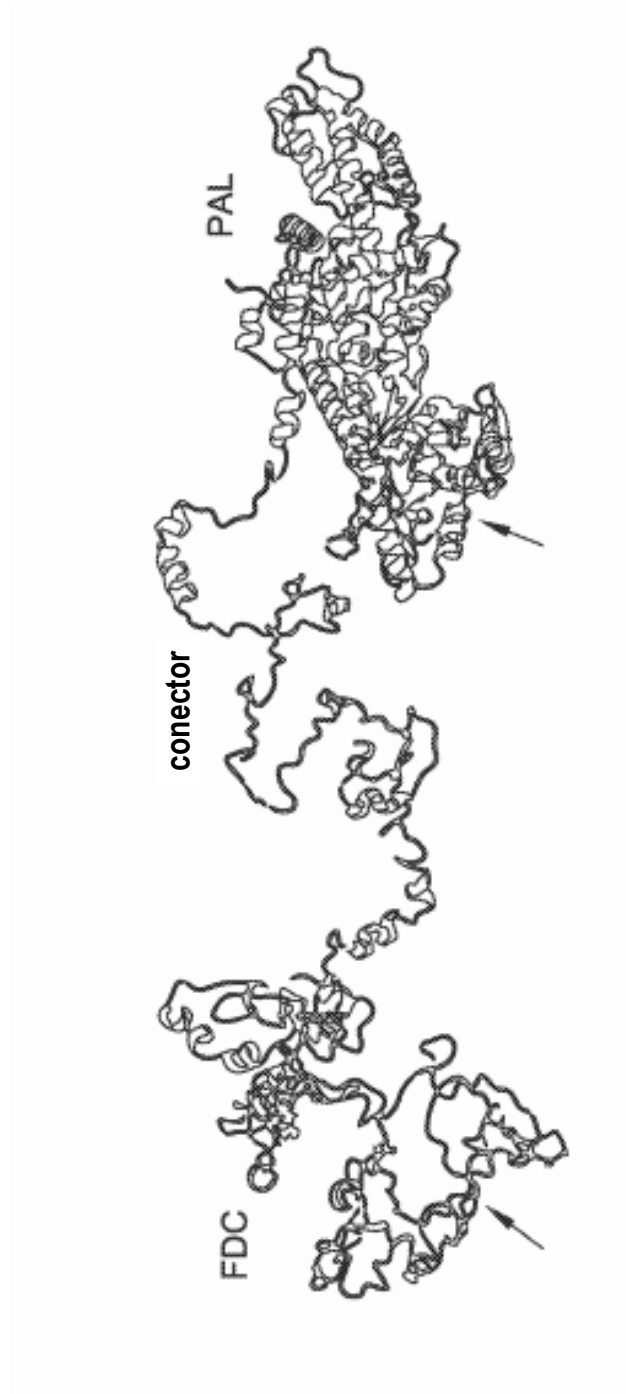


FIG. 16

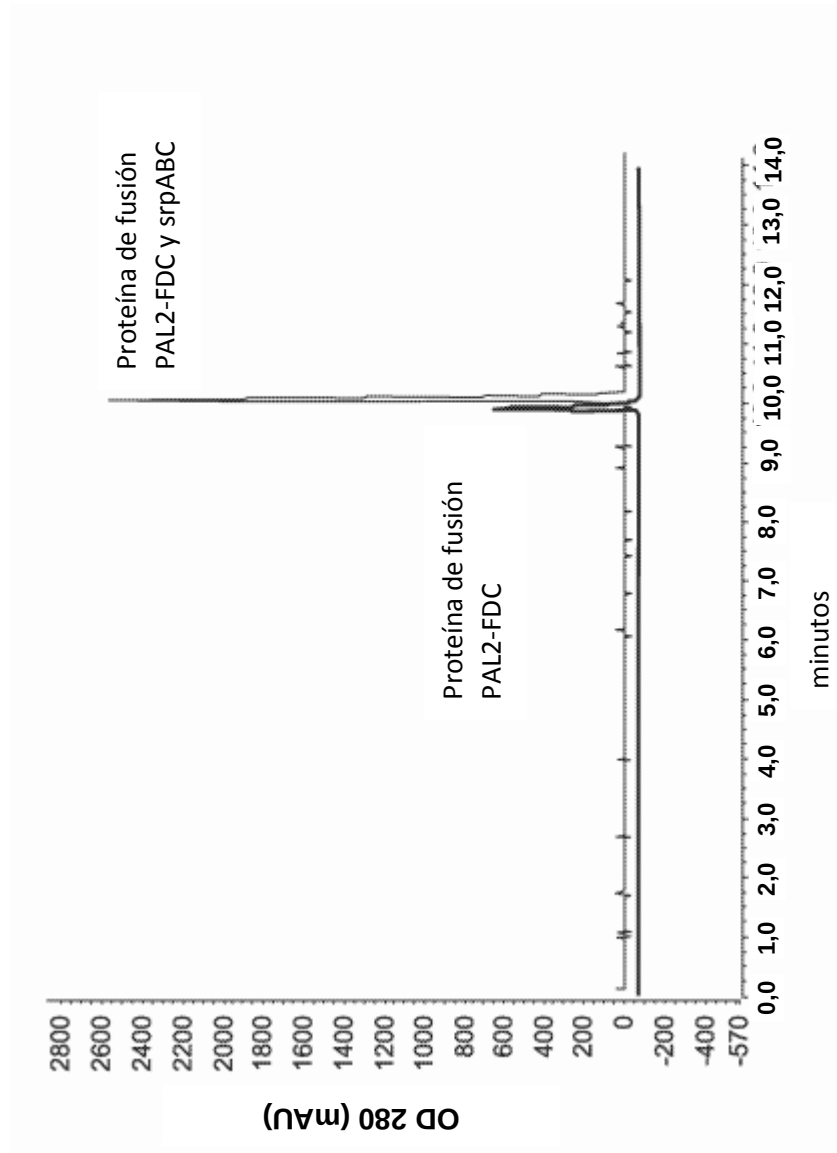


FIG. 17

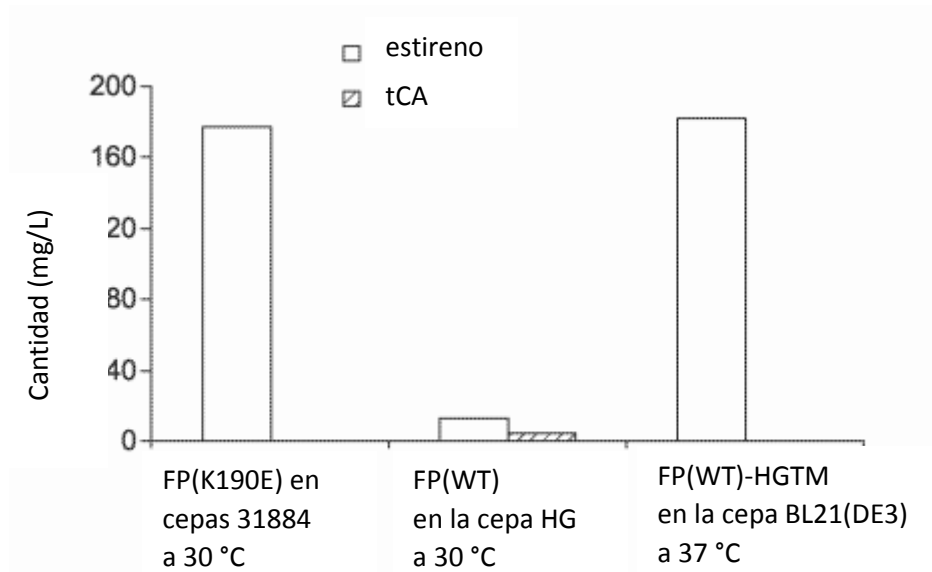


FIG. 18A

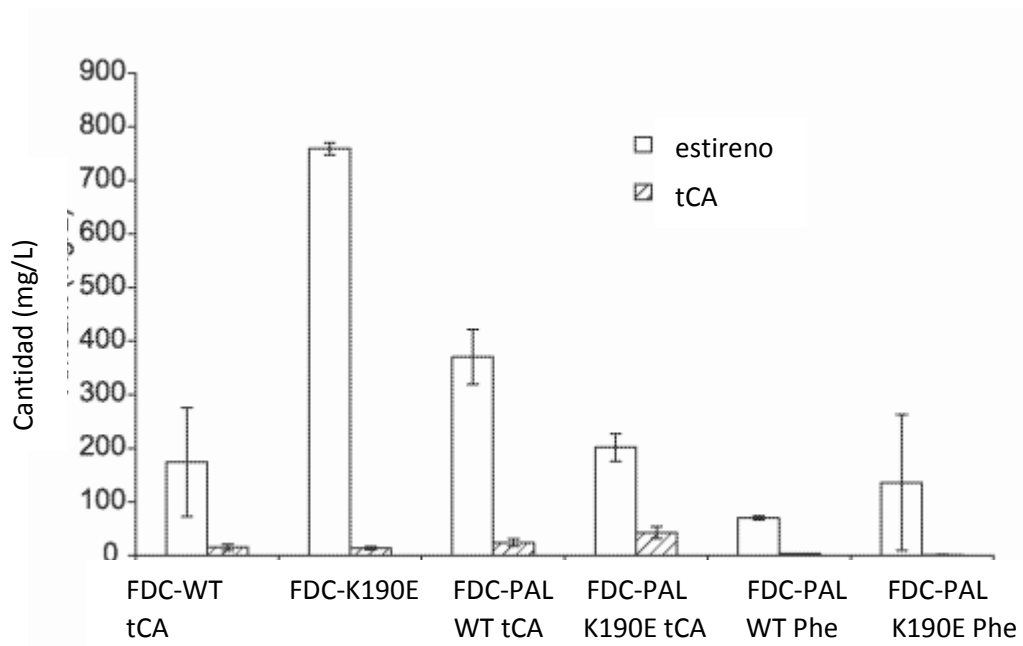


FIG. 18B



FIG. 19

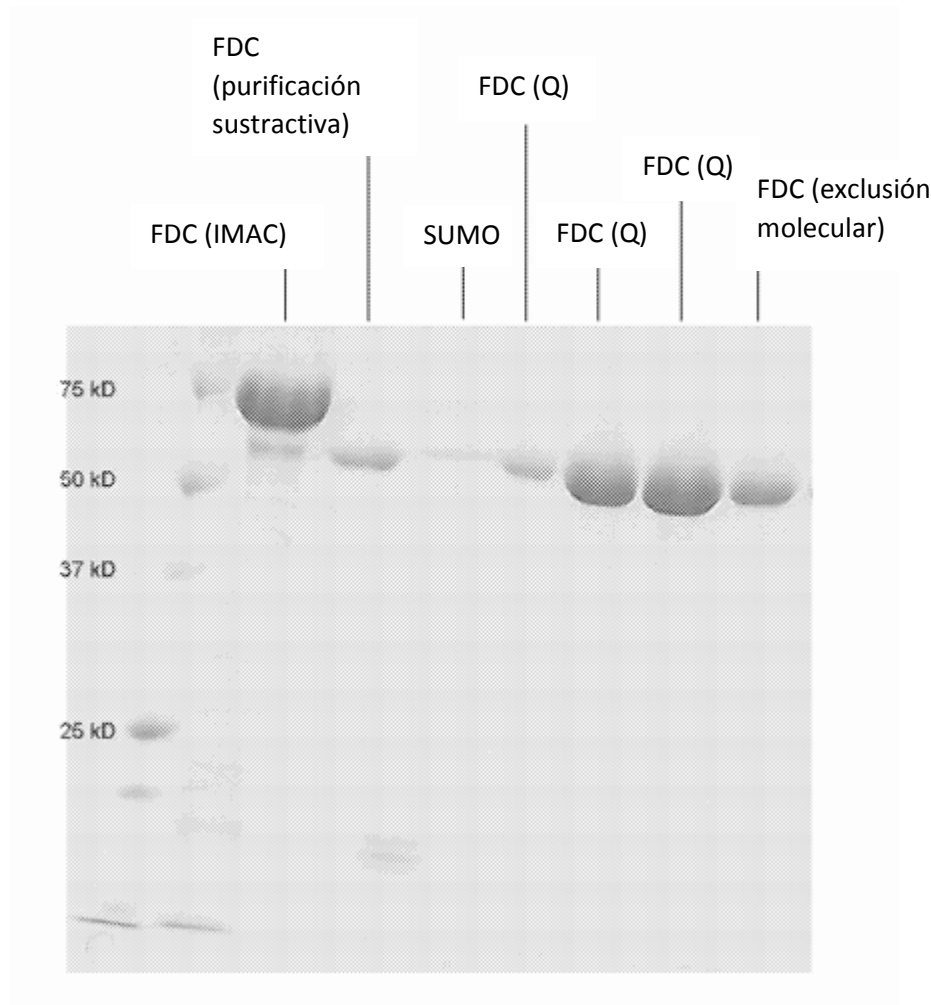


FIG. 20

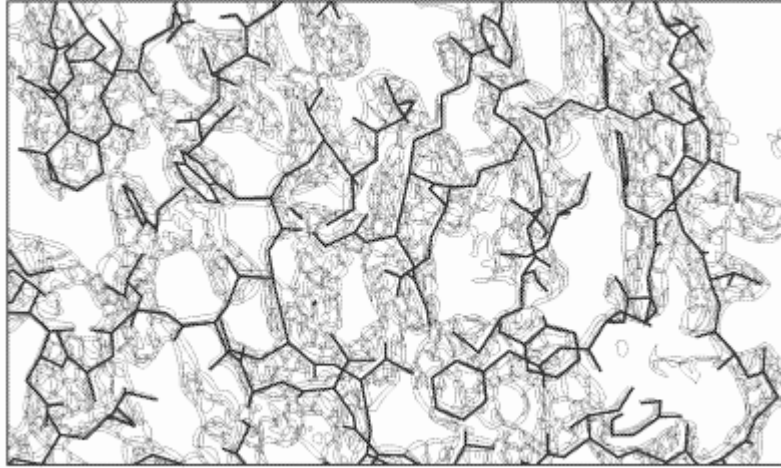


FIG. 21A



FIG. 21B