

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 712 682**

51 Int. Cl.:

C12N 9/04 (2006.01)
C12N 5/00 (2006.01)
C12P 17/18 (2006.01)
C12N 9/02 (2006.01)
C12P 13/04 (2006.01)
C12P 13/24 (2006.01)
C12P 17/12 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **07.05.2013 PCT/US2013/039874**
87 Fecha y número de publicación internacional: **14.11.2013 WO13169725**
96 Fecha de presentación y número de la solicitud europea: **07.05.2013 E 13788385 (6)**
97 Fecha y número de publicación de la concesión europea: **26.12.2018 EP 2847327**

54 Título: **Biocatalizadores y métodos para la hidroxilación de compuestos químicos**

30 Prioridad:

08.05.2012 US 201261644135 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
14.05.2019

73 Titular/es:

**CODEXIS, INC. (100.0%)
200 Penobscot Drive
Redwood City, CA 94063, US**

72 Inventor/es:

**CHEN, HAIBIN;
BONG, YONG KOY;
CABIROL, FABIEN;
GOHEL, ANUPAM;
LI, TAO;
MOORE, JEFFREY C.;
QUINTANAR-AUDELO, MARTINA;
YANG, HONG;
COLLIER, STEVEN, J. y
SMITH, DEREK**

74 Agente/Representante:

IZQUIERDO BLANCO, María Alicia

ES 2 712 682 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

Biocatalizadores y métodos para la hidroxilación de compuestos químicos

1. CAMPO TÉCNICO

[0001] La descripción se refiere a biocatalizadores para la hidroxilación de compuestos químicos.

3. ANTECEDENTES

[0002] Los derivados de prolina con grupos funcionales en los carbonos del anillo son bloques de construcción útiles para la síntesis de compuestos farmacéuticos debido a la conformación restringida de la prolina. Uno de estos derivados, la prolina hidroxilada, es un material de partida para la síntesis de varios compuestos terapéuticos, incluidos los antibióticos carbapenem (véase, por ejemplo, Altamura et al., 1995, J Med Chem. 38 (21): 4244-56), inhibidores de la enzima convertidores de angiotensina, inhibidores de la proteasa (ver, por ejemplo, Chen et al., 2002, J Org Chem. 67 (8):2730-3; Chen et al., 2006, J Med Chem. 49 (3):995-1005), análogos de ácido nucleico (efimov et al., 2006, Nucleic Acids Res. 34 (8):2247-2257), inhibidores de la isopreniltransferasa (O'Connell et al., 2000, Chem Pharm Bull. 48 (5):740-742), y la construcción de la biblioteca de fármacos (Vergnon et al., 2004, J Comb Chem. 6 (1):91-8; Remuzon P., 1996, Tetrahedron 52:13803-13835). De manera similar, los derivados hidroxilados de un homólogo de prolina, el ácido L-pipecólico, también conocido como homoprolina, también sirven como bloques de construcción para compuestos farmacéuticos. Por ejemplo, el ácido hidroxipipecólico es un intermediario en la síntesis de los inhibidores de la β -lactamasa (ver, por ejemplo, los documentos WO2009091856, WO2010126820 y US20110046102) y los inhibidores de la enzima de conversión de TNF-alfa (TACE) (Levatic et al., 2002, Bioorg Medicinal Chem Lett. 12 (10):1387-1390).

[0003] La hidroxiprolina puede obtenerse a partir de fuentes naturales, tales como materiales vegetales y hidrolizados de colágeno. La hidroxiprolina también se puede sintetizar químicamente, como a partir de los materiales de partida bromuro de alilo y ácido dietilacetamida-omalónico (Kyun Lee et al., 1973, Bull. Chem. Soc. Japón, 46:2924), ácido D-glutámico (eguchi et al., 1974, Bull. Chem. Soc. Japón, 47:1704-08), glioxal y ácido oxaloacético (Ramaswamy et al., 1977, J. Org. Chem. 42 (21):3440-3443) y β -alanina (Sinha et al., 2000, Proc. ECSOC-4, The Fourth International Electronic Conference on Synthetic Organic Chemistry, ISBN 3-906980-05-7).

[0004] El ácido hidroxipipecólico también se puede obtener de plantas y otras fuentes naturales (ver, por ejemplo, Romeo et al., 1983, Phytochemistry 22 (7):1615-1617; Fowden, L., 1958, Biochem J. 70 (4):629-33; Clark-Lewis y Mortimer, 1959, Nature 184 (Suppl 16):1234-5). La síntesis química del ácido hidroxipipecólico se describe en Callens et al., 2010, Bulletin des Sociétés Chimiques Belges 91 (8):713-723; Adams et al., 1996, Chem. Commun. 3:349-350; Botman et al., 2004, Organic Letters 6 (26):4941-4944; Cohen et al., 1956, Science 123 (3202):842-843; Beyerman et al., 1959, Recueil des Travaux Chimiques des Pays-Bas, 78 (9):648-658; Marin et al., 2004, J Org Chem. 69 (1):130-41; Kumar et al., 2005, J Org Chem. 70 (1):360-3; Liang et al., 2005, J Org Chem. 70 (24):10182-5; Kalamkar et al., 2008, J Org Chem. 73 (9):3619-22; Chiou et al., 2010, J Org Chem. 75 (5):1748-51; Lemire et al., 2010, J Org Chem. 75 (6):2077-80; y Angelique et al., 2000, Tetrahedron Lett. 41 (36):7033-7036.

[0005] El aislamiento de fuentes naturales está limitado por la disponibilidad de materias primas, requiere la purificación de una cantidad significativa de contaminantes de fondo y carece de ciertos diastereómeros deseados. Los métodos químicos sintéticos pueden requerir pasos complejos, ser difíciles de escalar a niveles de escala industrial y requerir pasos de purificación adicionales debido a la formación de múltiples productos hidroxilados.

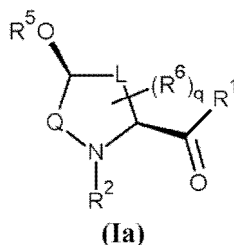
[0006] Otro enfoque para preparar prolina hidroxilada utiliza hidroxilasas de prolina, que son dioxigenasas dependientes de 2-oxoglutarato, que utilizan arato de 2-oxoglut (α -cetoglutarato) y O_2 como co-sustratos e ion ferroso como cofactor (véase, por ejemplo, Klein et al., 2011, Adv Synth. Catal. 353:1375-1383; Patente de EE.UU. N° 5364775; y Shibasaki et al., 1999, Appl Environ Microbiol. 65 (9):4028-4031). A diferencia de las hidroxilasa de prolilos que reconocen específicamente la peptidilprolina en el pro-colágeno y los péptidos relacionados, las hidroxilasas de prolina son capaces de convertir la prolina libre en hidroxiprolina. Se conocen varias enzimas microbianas que producen cis-3-, cis-4-o trans-4-hidroxiprolina (ver, por ejemplo, Patente de EE.UU. N° 5962292; Patente de EE.UU. N° 5963254; Patente de EE.UU. N° 5854040; WO2009139365; y EP2290065) y una enzima que produce trans-3-hidroxiprolina se ha identificado en extractos del hongo *Glarea lozoyensis*. Muchas de las hidroxilasas de la prolina se encuentran en bacterias, donde están asociadas con la biosíntesis de antibióticos peptídicos. La enzima hidroxilasa cis-4-prolina también muestra actividad al convertir ácido L-pipecólico (es decir, ácido (2S)-piperidina-2-carboxílico) en ácido cis-5-hidroxipipecólico (es decir, ácido (2S, 5S)-5-hidroxipiperidina-2-carboxílico; Klein et al., Supra). Se han demostrado conversiones *in vitro* para preparar ácido 5-hidroxipipecólico usando estas enzimas, pero se ha descubierto que las hidroxilasas de prolina aisladas se desnaturalizan en condiciones de reacción y tienen una actividad específica relativamente baja, lo que hace que los usos *in vitro* sean impracticables para aplicaciones comerciales (Klein et al., supra). Mientras que las células completas recombinantes que expresan hidroxilasas de prolina clonadas son más adecuadas para los procesos industriales a gran escala, el uso de células completas limita las variaciones en las condiciones de reacción, tales como altas concentraciones de sustrato; restringe los tipos de sustratos que se pueden usar con células enteras a aquellos que son permeables a

las células; y da como resultado subproductos indeseables que deben separarse del producto final. Además, los sistemas *in vivo* pueden requerir medios de crecimiento definidos que no sean óptimos ni rentables debido a que el uso de medios de crecimiento ricos preparados a partir de hidrolizados de proteínas contienen prolina libre, que puede ser un inhibidor competitivo cuando sustratos distintos de la prolina son el objetivo. Son deseables los métodos alternativos para sintetizar formas hidroxiladas de prolina y análogos de prolina, así como otros compuestos químicos, que se pueden ampliar fácilmente y dar como resultado un producto estereomérico sustancialmente puro.

4. RESUMEN

[0007] La presente invención proporciona un polipéptido diseñado que tiene actividad de prolina hidroxilasa, que comprende una secuencia de aminoácidos que tiene al menos un 80% de identidad de secuencia con la SEQ ID NO:2 en la que la hidroxilasa de prolina tiene al menos 1,2 veces más actividad en comparación con el polipéptido de SEQ ID NO:2 y en donde la secuencia de aminoácidos comprende una diferencia de residuos en comparación con la secuencia de la SEQ ID NO:2 en la posición de residuo X166.

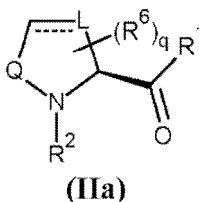
[0008] La presente invención también proporciona un procedimiento para preparar un compuesto de producto de fórmula (Ia),



en donde

Q se selecciona del grupo que consiste en un alquileo (C₁-C₅) y alquenileo (C₂-C₅);
 L se selecciona del grupo que consiste en un enlace, alquileo (C₁-C₄) y alquenileo (C₂-C₄);
 R¹ se selecciona del grupo que consiste en hidroxilo, amino, alquilo (C₁-C₆), arilo, alquilo (C₁-C₆) y arilo;
 R² se selecciona del grupo que consiste en hidrógeno y alquilo (C₁-C₆) opcionalmente sustituido, alqueno (C₂-C₆) y alquilo (C₂-C₆);
 R⁵ es hidrógeno, o un enlace directo a un átomo de carbono de L para formar un epóxido;
 cada aparición de R₆ se selecciona del grupo que consiste en halo, alquilo (C₁-C₆) y alquilo (C₁-C₆); y
 q es un número entero de 0 a 4;
 en donde la suma de los átomos de carbono del anillo para Q + L es un número entero de 2 a 5;
 con la condición de que

(i) cuando la suma de los átomos de carbono del anillo para Q + L es 2, entonces L es un metileno; y
 (ii) cuando la suma de los átomos de carbono del anillo para Q + L es 3, entonces L es un enlace o etileno; el proceso que comprende poner en contacto el compuesto de sustrato de fórmula (IIa),



en donde L, Q, R¹, R², R⁶ y q son como se definieron anteriormente para el compuesto de fórmula (Ia); y -
 --- representa un enlace opcional a un átomo de carbono de L para formar un doble enlace;

con un polipéptido diseñado de la invención en presencia de un co-sustrato en condiciones de reacción adecuadas.

[0009] La presente descripción proporciona biocatalizadores de hidroxilasa de prolina diseñados, polinucleótidos que codifican los biocatalizadores, métodos de preparación y procedimientos para preparar compuestos hidroxilados utilizando estos biocatalizadores diseñados. Las hidroxilasas de prolina de la presente descripción se han diseñado para tener propiedades mejoradas con respecto a la hidroxilasa de cis-4-prolina natural (SEQ ID NO:2) de

Sinorhizobium meliloti, una bacteria Gram negativa fijadora de nitrógeno. Las propiedades mejoradas del biocatalizador de las hidroxilasas de prolina diseñadas incluyen, entre otras, actividad, regioselectividad, tolerancia al sustrato y estabilidad. También se encuentra que las hidroxilasas de prolina modificadas hidroxilan una variedad de compuestos de sustrato, incluida la conversión selectiva del ácido (2S)-piperidina-2-carboxílico (es decir, ácido L-pipecólico) a ácido (2S, 5S)-5-hidroxi-piperidina-2-carboxílico (es decir, ácido cis-5-hidroxi-pipecólico). Las enzimas ingenierizadas con propiedades mejoradas tienen una o más diferencias de residuos en comparación con la hidroxilasa de prolina natural, donde las diferencias de residuos ocurren en las posiciones de residuos que afectan las propiedades enzimáticas anteriores.

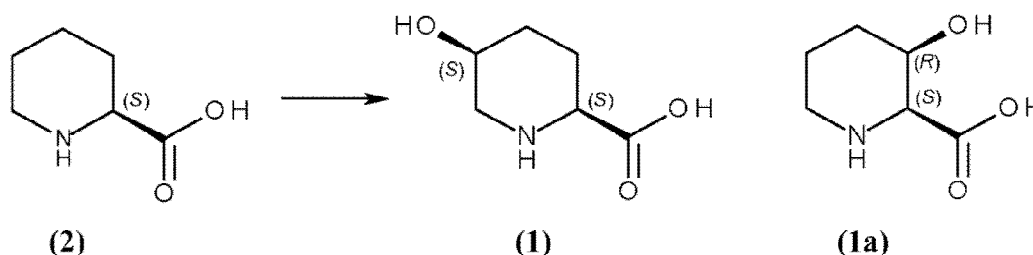
[0010] Por consiguiente, en un aspecto, la presente descripción proporciona polipéptidos modificados que tienen actividad de prolina hidroxilasa, donde los polipéptidos comprenden una secuencia de aminoácidos que tiene al menos un 80% de identidad con la SEQ ID NO:2 y una o más diferencias de residuos en comparación con la SEQ. ID NO:2 en las posiciones de residuos seleccionadas de: X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X186; X188; X205; X225; X230; X270; y X271.

[0011] En algunas realizaciones, las diferencias de residuos en las posiciones de residuos X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X186; X188; X205; X225; X230; X270; y X271 se pueden seleccionar de X2K; X2T; X3S; X4Q; X4L; X4E; X4S; X5I; X5L; X5M; X9I; X13T; X17V; X24R; X24S; X25R; X26R; X26T; X26W; X29A; X30V; X30P; X36T; X42E; X52P; X57T; X57A; X58A; X59G; X62Q; X66Q; X86S; X88R; X92V; X95M; X98F; X98T; X103L; X103Q; X112T; X112V; X113E; X114N; X115E; X115H; X115D; X115G; X115S; X115A; X116L; X121F; X131Y; X131F; X140L; X150S; X151A; X151H; X151S; X166T; X166L; X166Q; X186G; X188G; X205V; X225L; X225Y; X225W; X230V; X270E; X271K; y X271R. La siguiente descripción detallada proporciona una guía sobre las opciones de diferencias de residuos que se pueden usar para preparar hidroxilasas de prolina diseñadas con las propiedades biocatalíticas mejoradas deseadas.

[0012] En algunas realizaciones, el polipéptido de hidroxilasa de prolina diseñado tiene una secuencia de aminoácidos que comprende al menos una combinación de características seleccionadas de: (a) X103L y X166Q; (b) X52P y X255Y; (c) X4E/L/S y X115A; (d) X25R y X58A; (e) X29A y X166T/Q/L; (f) X115H/D/G y X121F; (g) X3S, X103L y X166Q; (h) X103L, X131Y/F, y X166T/Q/L; (i) X26T, X103L y X166T/Q/L; (j) X25R, X66Q, X92V y X115E; (k) X25R, X66Q, X92V, X103L, X115E y X166Q; y (l) X3S, X25R, X66Q, X92V, X103L, X115E y X166Q.

[0013] Como se indicó anteriormente, los polipéptidos diseñados que tienen actividad hidroxilasa de prolina también son capaces de convertir el compuesto de sustrato (2), ácido (2S)-piperidina-2-carboxílico en el compuesto del producto (1), (2S, 5S)-5-hidroxi-piperidina, como se muestra en el **Esquema 1**,

Esquema 1



con propiedades mejoradas en comparación con la enzima natural. En algunas realizaciones, los polipéptidos diseñados pueden convertir el compuesto de sustrato (2) en el compuesto del producto (1) con al menos 1,2 veces, 1,5 veces, 2 veces, 3 veces, 4 veces, 5 veces o 10 veces o más. La actividad de la enzima natural, y con más del 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más de exceso diastereomérico de ácido (2S, 5R)-5-hidroxi-piperidina-2-carboxílico.

[0014] En algunas realizaciones, los polipéptidos diseñados que tienen actividad hidroxilasa de prolina también muestran una regioselectividad incrementada para formar ácido (2S, 5S)-5-hidroxi-piperidina-2-carboxílico en exceso de otros regioisómeros, por ejemplo (2S,3R)-3-ácido hidroxi-piperidina-2-carboxílico, mostrado en el **Esquema 1** como compuesto (1a). Así, en algunas realizaciones, las hidroxilasas de prolina son capaces de convertir el compuesto de sustrato (2) en un compuesto de producto (1) en exceso del compuesto de producto (1a), donde la relación de compuesto de producto (1) se formó sobre el compuesto de producto (1a) es al menos 1,5, 2, 3, 4, 5 o 6 o más. En algunas realizaciones, las hidroxilasas de prolina diseñadas con una mayor selectividad para formar el compuesto de producto (1) en exceso del compuesto de producto (1a) comprenden una secuencia de aminoácidos que tiene una o más de las siguientes características: X103L; X115E; X131Y y X166Q, particularmente la

combinación de las características X103L y X166Q.

[0015] En algunas realizaciones, los polipéptidos diseñados que tienen propiedades mejoradas tienen una secuencia de aminoácidos que comprende una secuencia seleccionada del grupo que consiste en la SEQ ID NO:8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226, y 228.

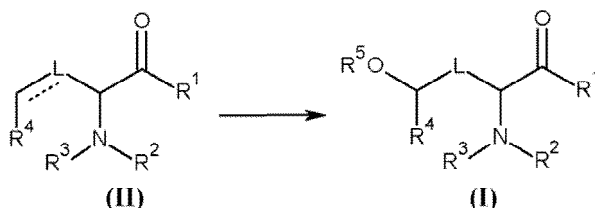
[0016] En otro aspecto, la presente descripción proporciona polinucleótidos que codifican los polipéptidos de hidroxilasa de prolina modificados con propiedades mejoradas. Las secuencias de polinucleótidos ejemplares se proporcionan en el Listado de Secuencias e incluyen las SEQ ID NO:7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177, 179, 181, 183, 185, 187, 189, 191, 193, 195, 197, 199, 201, 203, 205, 207, 209, 211, 213, 215, 217, 219, 221, 223, 225 y 227.

[0017] En un aspecto adicional, la presente descripción también proporciona polinucleótidos con codones optimizados que codifican un polipéptido que comprende una secuencia de aminoácidos de la hidroxilasa de prolina de tipo salvaje (SEQ ID NO:2). En algunas realizaciones, los polinucleótidos con codón optimizado tienen una expresión aumentada en una célula huésped bacteriana en comparación con el polinucleótido de origen natural que codifica la enzima hidroxilasa de prolina. En algunas realizaciones, los polinucleótidos con codones optimizados pueden tener al menos 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más de identidad a la secuencia de ácido nucleico optimizada de codón de la SEQ ID NO:1, 3 o 5, que codifica las secuencias polipeptídicas idénticas de la SEQ ID NO:2, 4 o 6, respectivamente. La secuencia optimizada de codón de la SEQ ID NO:1, 3 o 5 puede mejorar la expresión de la hidroxilasa de prolina de tipo salvaje codificada en al menos 1,2 veces, 1,5 veces o 2 veces o más en comparación con la secuencia de polinucleótidos de origen natural.

[0018] En otro aspecto, los polinucleótidos de la descripción pueden incorporarse en vectores de expresión y células huésped para la expresión de los polinucleótidos y los polipéptidos de hidroxilasa de prolina codificados correspondientes. Como tal, en algunas realizaciones, la presente descripción proporciona métodos para preparar los polipéptidos de hidroxilasa de prolina mediante el cultivo de una célula huésped que comprende el polinucleótido o vector de expresión capaz de expresar una hidroxilasa de prolina de la descripción en condiciones adecuadas para la expresión del polipéptido modificado. En algunas realizaciones, el método de preparación de la hidroxilasa de prolina puede comprender la etapa adicional de aislar el polipéptido expresado.

[0019] En algunas realizaciones, la presente descripción también proporciona métodos para fabricar un polipéptido de hidroxilasa de prolina modificado, en el que el método puede comprender: (a) sintetizar un polinucleótido que codifica un polipéptido que comprende una secuencia de aminoácidos seleccionada de la SEQ ID NO:8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228, y que tienen una o más diferencias de residuos en comparación con la SEQ ID NO:2 en las posiciones de residuos seleccionadas de: X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X186; X188; X205; X225; X230; X270; y X271, y (b) que expresan el polipéptido de hidroxilasa de prolina codificado por el polinucleótido. Como se señaló anteriormente, las diferencias de residuos en las posiciones X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X186; X188; X205; X225; X230; X270; y X271 se seleccionan de X2K; X2T; X3S; X4Q; X4L; X4E; X4S; X5I; X5L; X5M; X9I; X13T; X17V; X24R; X24S; X25R; X26R; X26T; X26W; X29A; X30V; X30P; X36T; X42E; X52P; X57T; X57A; X58A; X59G; X62Q; X66Q; X86S; X88R; X92V; X95M; X98F; X98T; X103L; X103Q; X112T; X112V; X113E; X114N; X115E; X115H; X115D; X115G; X115S; X115A; X116L; X121F; X131Y; X131F; X140L; X150S; X151A; X151H; X151S; X166T; X166L; X166Q; X186G; X188G; X205V; X225L; X225Y; X225W; X230V; X270E; X271K; y X271R. Como se proporciona adicionalmente en la descripción detallada, pueden incorporarse variaciones adicionales durante la síntesis del polinucleótido para preparar hidroxilasas de prolina modificadas con las correspondientes diferencias en las secuencias de aminoácidos expresadas.

[0020] En otro aspecto, los polipéptidos de hidroxilasa de prolina modificados pueden usarse en un proceso para preparar diversos compuestos hidroxilados, tales como prolina hidroxilada o ácidos piperidina-2-carboxílicos hidroxilados. Por consiguiente, en algunas realizaciones, los polipéptidos de hidroxilasa de prolina modificados se pueden usar en un proceso para la conversión de un compuesto de sustrato de fórmula (I) en un compuesto de producto de fórmula (II), como se muestra a continuación:

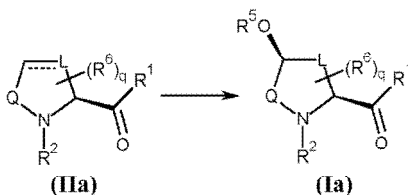


en donde

L se selecciona del grupo que consiste en un enlace, alquileo (C₁-C₄) y alquilenilo (C₂-C₄);
 R¹ se selecciona del grupo que consiste en hidroxilo, amino, alquilo (C₁-C₆), arilo, alquilo (C₁-C₆) y arilo;
 R² y R³ se seleccionan cada uno independientemente del grupo que consiste en hidrógeno y alquilo (C₁-C₆) opcionalmente sustituido, alquileo (C₂-C₆) y alquilenilo (C₂-C₆);
 R⁴ se selecciona del grupo que consiste en alquilo (C₁-C₆), alquileo (C₂-C₆) opcionalmente sustituido, alquilenilo (C₂-C₆), arilo, heteroarilo, cicloalquilo, heterocicloalquilo; o R⁴ junto con uno de R¹ o R² es un alquileo (C₁-C₆) o alquilenilo (C₂-C₅) y forma un anillo heterocíclico de 5 a 8 miembros que contiene el átomo de nitrógeno, en donde el anillo está opcionalmente sustituido con 1 a 4 grupos R⁶ seleccionados independientemente;
 R⁵ es hidrógeno o un enlace que forma un epóxido con un átomo de carbono de L;
 cada aparición de R⁶ se selecciona independientemente del grupo que consiste en halo, alquilo (C₁-C₆) y alquilo (C₁-C₆); y
 ----- representa un enlace opcional a un átomo de carbono de L para formar un doble enlace;
 con la condición de que

- (i) cuando R⁴ no forma un anillo con un R² o R³, o cuando R⁴ forma un anillo heterocíclico de 5 miembros que contiene el átomo de nitrógeno con uno de R² o R³, entonces L es un metileno;
- (ii) cuando R⁴ forma un anillo heterocíclico de 6 miembros que contiene el átomo de nitrógeno con uno de R² o R³, entonces L es un enlace o etileno; y
- (iii) cuando R⁵ es un enlace a un átomo de carbono de L para formar un epóxido, entonces R⁴ forma el anillo heterocíclico que contiene el átomo de nitrógeno con uno de R² o R³ y L es un alquileo (C₁-C₄) o alquilenilo (C₂-C₄).

[0021] En algunas realizaciones, la hidroxilasa de prolina modificada se puede usar para la conversión de un compuesto de anillo de fórmula estructural (IIa) al compuesto producto hidroxilado de fórmula estructural (Ia);



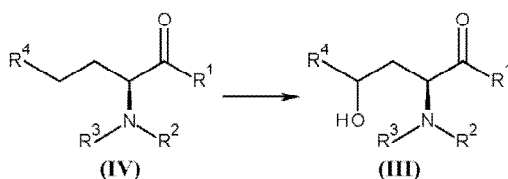
en donde

Q se selecciona del grupo que consiste en un alquileo (C₁-C₅) y alquilenilo (C₂-C₅);
 L se selecciona del grupo que consiste en un enlace, alquileo (C₁-C₄) y alquilenilo (C₂-C₄);
 R¹ se selecciona del grupo que consiste en hidroxilo, amino, alquilo (C₁-C₆), arilo, alquilo (C₁-C₆) y arilo;
 R² se selecciona del grupo que consiste en hidrógeno y alquilo (C₁-C₆) opcionalmente sustituido, alquileo (C₂-C₆) y alquilenilo (C₂-C₆);
 R⁵ es hidrógeno, o un enlace directo a un átomo de carbono de L para formar un epóxido;
 cada aparición de R⁶ se selecciona del grupo que consiste en halo, alquilo (C₁-C₆) y alquilo (C₁-C₆);
 q es un número entero de 0 a 4; y
 ----- representa un enlace opcional a un átomo de carbono de L para formar un doble enlace;
 en donde la suma de los átomos de carbono del anillo para Q + L es un número entero de 2 a 5;
 con la condición de que

- (i) cuando la suma de los átomos de carbono del anillo para Q + L es 2, entonces L es un metileno; y
- (ii) cuando la suma de los átomos de carbono del anillo para Q + L es 3, entonces L es un enlace o etileno.

[0022] En algunas realizaciones, la hidroxilasa de prolina modificada se puede usar para la conversión de un compuesto de sustrato de fórmula estructural (IV) en el compuesto de producto hidroxilado de fórmula estructural

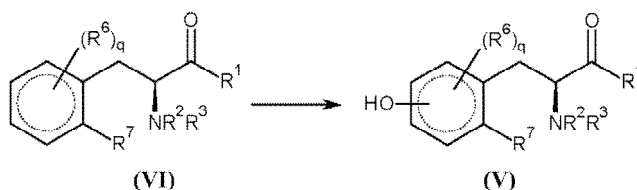
(III);



en donde,

R^1 se selecciona del grupo que consiste en hidroxilo, amino, alquilo (C₁-C₆), arilo, alquilo (C₁-C₆) y arilo; R^2 y R^3 se seleccionan cada uno independientemente del grupo que consiste en hidrógeno y alquilo (C₁-C₆) opcionalmente sustituido, alqueno (C₂-C₆) y alquino (C₂-C₆); y R^4 se selecciona del grupo que consiste en alquilo, alqueno, alquino, arilo, heteroarilo, cicloalquilo y heterocicloalquilo opcionalmente sustituidos.

[0023] En algunas realizaciones, la hidroxilasa de prolina modificada se puede usar para la conversión de un compuesto de sustrato de fórmula estructural (VI) en el compuesto de producto hidroxilado de fórmula estructural (V);



en donde,

R^1 se selecciona del grupo que consiste en hidroxilo, amino, alquilo (C₁-C₆), arilo, alquilo (C₁-C₆) y arilo; R^2 y R^3 se seleccionan independientemente del grupo que consiste en hidrógeno y alquilo (C₁-C₆) opcionalmente sustituido, alqueno (C₂-C₆) y alquino (C₂-C₆); cada aparición de R^6 se selecciona independientemente del grupo que consiste en halo, alquilo (C₁-C₆) y alquilo (C₁-C₆); R^7 se selecciona del grupo que consiste en hidrógeno, halo y alquilo (C₁-C₆) opcionalmente sustituido y alquilo (C₁-C₆); o R^7 junto con uno de R^2 o R^3 forma un anillo heterocíclico de 5 a 7 miembros que contiene el átomo de nitrógeno; q es un número entero de 0 a 4; y --- representan dobles enlaces opcionales que forman un anillo arilo.

[0024] Las reacciones de hidroxilación anteriores que utilizan las hidroxilasas de prolina modificadas se llevan a cabo bajo condiciones de reacción adecuadas en presencia de un co-sustrato (por ejemplo, α -cetoglutarato), un metal de transición divalente (por ejemplo, Fe²⁺) y oxígeno molecular (es decir, O₂). Las condiciones de reacción adecuadas pueden comprender rangos de co-sustrato, metal de transición divalente, oxígeno molecular, así como rangos de otros parámetros, como la concentración de reductor (por ejemplo, ácido ascórbico), concentración de detergente, pH, temperatura, tampón, sistema de solvente, carga de sustrato, carga de polipéptidos, presión y tiempo de reacción. En algunas realizaciones, la reacción de hidroxilación se puede llevar a cabo, en la que la hidroxilasa de prolina modificada se inmoviliza sobre un soporte sólido.

[0025] En algunas realizaciones, las condiciones de reacción adecuadas pueden comprender lo siguiente: (a) carga de sustrato a aproximadamente 5 g/L a 30 g/L; (b) aproximadamente 0,1 g/L a 10 g/L del polipéptido diseñado; (c) aproximadamente 19 g/L (0,13 M) a 57 g/L (0,39 M) de α -cetoglutarato; (d) aproximadamente 14 g/L (0,08 M) a 63 g/L (0,36 M) de ácido ascórbico; (e) aproximadamente 1,5 g/L (3,8 mM) a 4,5 g/L (11,5 mM) de FeSO₄; (f) un pH de aproximadamente 6 a 7; (g) temperatura de aproximadamente 20° a 40°C; y (h) tiempo de reacción de 2-24 h. En algunas realizaciones, las condiciones de reacción adecuadas comprenden la aireación forzada de la solución de reacción con O₂ a una velocidad de aproximadamente 3 l/h.

[0026] En algunas realizaciones, las condiciones de reacción adecuadas pueden comprender lo siguiente: (a) carga de sustrato de aproximadamente 10 g/L a 100 g/L; (b) de aproximadamente 1 g/L a aproximadamente 50 g/L de polipéptido diseñado; (c) α -cetoglutarato en aproximadamente 1 a 2 equivalentes molares de compuesto de sustrato;

(d) ácido ascórbico en aproximadamente 0,25 a 0,75 equivalentes molares de compuesto de sustrato; (e) de aproximadamente 0,5 mM a aproximadamente 12 mM de FeSO_4 ; (f) pH de aproximadamente 6 a 8; (g) temperatura de aproximadamente 20° a 40°C; y (h) tiempo de reacción de 6 a 120 h. En algunos experimentos, las condiciones de reacción adecuadas comprenden la aireación forzada de la solución de reacción con O_2 a una velocidad de aproximadamente 2 L/ha aproximadamente 5 L/h.

[0027] La orientación sobre la elección de hidroxilasas de prolina modificadas, la preparación de los biocatalizadores, la elección de sustratos enzimáticos y los parámetros para llevar a cabo los procesos se describen con más detalle en la descripción detallada que sigue.

5. DESCRIPCION DETALLADA

[0028] Como se usa en esta memoria descriptiva y en las reivindicaciones adjuntas, las formas singulares "un", "una", "el" y "la" incluyen referentes plurales a menos que el contexto indique claramente lo contrario. Así, por ejemplo, la referencia a "un polipéptido" incluye más de un polipéptido.

[0029] De manera similar, "comprende", "comprenden", "que comprende" "incluye", "incluyen" y "que incluye" son intercambiables y no pretenden ser limitativos.

[0030] Debe entenderse adicionalmente que cuando las descripciones de diversas realizaciones usan el término "que comprende", los expertos en la técnica entenderían que en algunos casos específicos, una realización puede describirse alternativamente usando el lenguaje "que consiste esencialmente en" o "que consiste en."

[0031] Se debe entender que tanto la descripción general anterior como la siguiente descripción detallada son solo ilustrativas y explicativas y no son restrictivas de esta descripción.

[0032] Los encabezados de sección utilizados en el presente documento tienen únicamente fines organizativos y no deben interpretarse como limitantes de la materia descrita.

5.1 Abreviaturas

[0033] Las abreviaturas utilizadas para los aminoácidos codificados genéticamente son convencionales y son las siguientes:

Aminoácido	Abreviatura de tres letras	Abreviatura de una letra
Alanina	Ala	A
Arginina	Arg	R
Asparagina	Asn	N
Aspartato	Asp	D
Cisteína	Cys	C
Glutamato	Glu	E
Glutamina	Gln	Q
Glicina	Gly	G
Histidina	HIS	H
Isoleucina	Ile	I
Leucina	Leu	L
Lisina	Lys	K
Metionina	Met	M
Fenilalanina	Phe	F
Prolina	Pro	P
Serina	Ser	S
Treonina	Thr	T
Triptofano	Trp	W
Tirosina	Tyr	Y
Valina	Val	V

[0034] Cuando se usan las abreviaturas de tres letras, a menos que estén precedidas específicamente por una "L" o una "D" o que estén claras en el contexto en el que se usa la abreviatura, el aminoácido puede estar en la configuración L o D sobre el carbono α (C_α). Por ejemplo, mientras que "Ala" designa alanina sin especificar la

configuración sobre el carbono α , "D-Ala" y "L-Ala" designan D-alanina y L-alanina, respectivamente. Cuando se usan las abreviaturas de una letra, las letras mayúsculas designan aminoácidos en la configuración L sobre el carbono α y las letras minúsculas designan aminoácidos en la configuración D sobre el carbono α . Por ejemplo, "A" designa L-alanina y "a" designa D-alanina. Cuando las secuencias de polipéptidos se presentan como una cadena de abreviaturas de una o tres letras (o mezclas de las mismas), las secuencias se presentan en la dirección de amino (N) a carboxi (C) de acuerdo con la convención común.

[0035] Las abreviaturas utilizadas para los nucleósidos genéticamente encolados son convencionales y son las siguientes: adenosina (A); guanosina (G); citidina (C); timidina (T); y uridina (U). A menos que esté específicamente delineado, los nucleósidos abreviados pueden ser ribonucleósidos o 2'-desoxirribonucleósidos. Los nucleósidos pueden especificarse como ribonucleósidos o 2'-desoxirribonucleósidos en una base individual o en una base agregada. Cuando las secuencias de ácido nucleico se presentan como una cadena de abreviaturas de una letra, las secuencias se presentan en la dirección 5' a 3' de acuerdo con la convención común, y los fosfatos no están indicados.

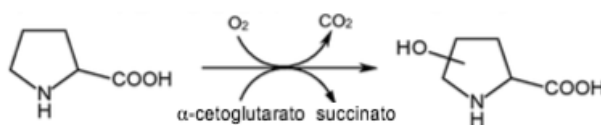
5.2 Definiciones

[0036] En referencia a la presente descripción, los términos técnicos y científicos usados en las descripciones en este documento tendrán los significados comúnmente entendidos por un experto en la técnica, a menos que se defina específicamente lo contrario. En consecuencia, se pretende que los siguientes términos tengan los siguientes significados:

"Proteína", "polipéptido" y "péptido" se usan indistintamente en el presente documento para indicar un polímero de al menos dos aminoácidos unidos covalentemente por un enlace amida, independientemente de la longitud o la modificación postraduccional (por ejemplo, glicosilación, fosforilación, lipidación, miristilación, ubiquitinación, etc.). Dentro de esta definición se incluyen los aminoácidos D y L, y las mezclas de los aminoácidos D y L.

[0037] "Polinucleótido" o "ácido nucleico" se refiere a dos o más nucleósidos que están unidos por enlace covalente. El polinucleótido puede estar compuesto por ribonucleótidos (es decir, un ARN), totalmente compuesto por 2' desoxirribonucleótidos (es decir, un ADN) o mezclas de desoxirribonucleótidos ribo y 2'. Mientras que los nucleósidos normalmente se unirán entre sí mediante enlaces fosfodiéster estándar, los polinucleótidos pueden incluir uno o más enlaces no estándar. El polinucleótido puede ser monocatenario o bicatenario, o incluye tanto regiones monocatenarias como regiones bicatenarias. Además, mientras que un polinucleótido estará compuesto típicamente de las nucleobases codificantes naturales (es decir, adenina, guanina, uracilo, timina y citosina), puede incluir una o más modificaciones o nucleobases sintéticas, tales como, por ejemplo, inosina, xantina, hipoxantina, etc. Preferiblemente, tales nucleobases modificadas o sintéticas estarán codificando nucleobases.

[0038] "Hidroxilasa de prolina" se refiere a un polipéptido que tiene una capacidad enzimática para convertir la prolina libre en hidroxiprolina en presencia de co-sustrato α -cetoglutarato y dioxígeno, como se ilustra a continuación:



Debe entenderse que las hidroxilasas de prolina no se limitan a la reacción anterior con prolina, sino que pueden hidroxilar otros sustratos, por ejemplo ácido pipercolico. Las hidroxilasas de prolina como se usan en el presente documento incluyen hidroxilasa de prolina natural (de tipo salvaje) así como polipéptidos modificados genéticamente no naturales generados por la manipulación humana.

[0039] "Co-sustrato" de una hidroxilasa de prolina se refiere a α -cetoglutarato y análogos de co-sustrato que pueden reemplazar a α -cetoglutarato en la hidroxilación de prolina y análogos de sustrato de prolina. Los análogos de co-sustrato incluyen, a modo de ejemplo y no de limitación, 2-oxoadipato (véase, por ejemplo, Majamaa et al., 1985, Biochem. J. 229:127-133).

[0040] La "secuencia codificante" se refiere a la porción de un ácido nucleico (por ejemplo, un gen) que codifica una secuencia de aminoácidos de una proteína.

[0041] "Que ocurre naturalmente" o "tipo salvaje" se refiere a la forma que se encuentra en la naturaleza. Por ejemplo, una secuencia de polipéptido o polinucleótido de tipo natural o natural es una secuencia presente en un organismo que puede aislarse de una fuente en la naturaleza y que no ha sido modificada intencionalmente por la manipulación humana.

[0042] "Recombinante" o "diseñado" o "de origen no natural" cuando se usa con referencia a, por ejemplo, una

célula, ácido nucleico o polipéptido, se refiere a un material, o un material correspondiente a la forma natural o nativa del material, que se ha modificado de una manera que de otra manera no existiría en la naturaleza, o es idéntico al mismo pero producido o derivado de materiales sintéticos y/o mediante manipulación utilizando técnicas recombinantes. Los ejemplos no limitantes incluyen, entre otros, células recombinantes que expresan genes que no se encuentran dentro de la forma nativa (no recombinante) de la célula o expresan genes nativos que de otro modo se expresan en un nivel diferente.

[0043] "Porcentaje de identidad de secuencia" y "homología de porcentaje" se usan de manera intercambiable en el presente documento para referirse a comparaciones entre polinucleótidos y polipéptidos, y se determinan comparando dos secuencias alineadas óptimamente en una ventana de comparación, en donde la porción del polinucleótido o la secuencia de polipéptidos en la ventana de comparación puede comprender adiciones o deleciones (es decir, huecos) en comparación con la secuencia de referencia para el alineamiento óptimo de las dos secuencias. El porcentaje puede calcularse determinando el número de posiciones en las que se produce la base de ácido nucleico o aminoácido idéntico en ambas secuencias para obtener el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes por el número total de posiciones en la ventana de comparar y multiplicar el resultado por 100 para obtener el porcentaje de identidad de secuencia. Alternativamente, el porcentaje puede calcularse determinando el número de posiciones en las cuales ocurre la misma base de ácido nucleico o el residuo de aminoácido en ambas secuencias o una base de ácido nucleico o un residuo de aminoácido se alinea con un hueco para obtener el número de posiciones emparejadas, dividiendo el número de posiciones coincidentes por el número total de posiciones en la ventana de comparación y multiplicando el resultado por 100 para obtener el porcentaje de identidad de secuencia. Los expertos en la técnica aprecian que hay muchos algoritmos establecidos disponibles para alinear dos secuencias. La alineación óptima de las secuencias para la comparación se puede realizar, por ejemplo, mediante el algoritmo de homología local de Smith y Waterman, 1981, Adv. Appl. Math. 2:482, por el algoritmo de alineación de homología de Needleman y Wunsch, 1970, J. Mol. Biol. 48:443, por el método de búsqueda de similitud de Pearson y Lipman, 1988, Proc. Natl Acad Sci. EE.UU. 85:2444, por implementaciones computarizadas de estos algoritmos (GAP, BE-STFIT, FASTA y TFASTA en el paquete de software GCG Wisconsin), o por inspección visual (ver en general, Current Protocols in Molecular Biology, FM Ausubel et al., eds., Current Protocols, Greene Publishing Associates, Inc. y John Wiley & Sons, Inc., (Suplemento de 1995) (Ausubel)). Ejemplos de algoritmos que son adecuados para determinar el porcentaje de identidad de secuencia y la similitud de secuencia son los algoritmos BLAST y BLAST 2.0, que se describen en Altschul et al., 1990, J. Mol. Biol. 215:403-410 y Altschul et al., 1977, Nucleic Acids Res. 3389-3402, respectivamente. El software para realizar los análisis de BLAST está disponible públicamente a través del sitio web del Centro Nacional de Información Biotecnológica. Este algoritmo involucra primero la identificación de pares de secuencias de alta puntuación (HSP) identificando palabras cortas de longitud W en la secuencia de consulta, que coinciden o satisfacen alguna puntuación de umbral de valor positivo T cuando se alinean con una palabra de la misma longitud en una secuencia de base de datos. T se conoce como el umbral de puntuación de palabras vecinas (Altschul et al, supra). Estos aciertos de palabras vecinas iniciales actúan como semillas para iniciar búsquedas para encontrar HSP más largos que los contengan. Los aciertos de palabras se extienden en ambas direcciones a lo largo de cada secuencia hasta donde se puede aumentar la puntuación de alineación acumulativa. Las puntuaciones acumulativas se calculan utilizando, para secuencias de nucleótidos, los parámetros M (puntuación de recompensa para un par de residuos coincidentes; siempre >0) y N (puntuación de penalización para residuos no coincidentes; siempre <0). Para las secuencias de aminoácidos, se utiliza una matriz de puntuación para calcular la puntuación acumulativa. La extensión de los aciertos de palabra en cada dirección se detiene cuando: la puntuación de alineación acumulada cae en la cantidad X de su valor máximo alcanzado; el puntaje acumulativo llega a cero o por debajo, debido a la acumulación de una o más alineaciones de residuos de puntaje negativo; o se llega al final de cualquiera de las secuencias. Los parámetros del algoritmo BLAST W, T y X determinan la sensibilidad y la velocidad de la alineación. El programa BLASTN (para secuencias de nucleótidos) usa como predeterminado una longitud de palabra (W) de 11, una expectativa (E) de 10, M = 5, N = -4 y una comparación de ambas cadenas. Para las secuencias de aminoácidos, el programa BLASTP utiliza como valor predeterminado una longitud de palabra (W) de 3, una expectativa (E) de 10, y la matriz de puntuación BLOSUM62 (véase Henikoff y Henikoff, 1989, Proc Natl Acad Sci EE.UU. 89:10915). La determinación ejemplar de alineación de secuencia y % de identidad de secuencia puede emplear los programas BESTFIT o GAP en el paquete de software GCG Wisconsin (Accelrys, Madison WI), usando los parámetros predeterminados provistos.

[0044] La "secuencia de referencia" se refiere a una secuencia definida utilizada como base para una comparación de secuencias. Una secuencia de referencia puede ser un subconjunto de una secuencia más grande, por ejemplo, un segmento de un gen o secuencia polipeptídica de longitud completa. En general, una secuencia de referencia tiene una longitud de al menos 20 residuos de nucleótidos o aminoácidos, una longitud de al menos 25 residuos, una longitud de al menos 50 residuos, o la longitud completa del ácido nucleico o polipéptido. Dado que dos polinucleótidos o polipéptidos pueden cada uno (1) comprender una secuencia (es decir, una porción de la secuencia completa) que es similar entre las dos secuencias, y (2) puede comprender además una secuencia que es divergente entre las dos secuencias, comparaciones de secuencias entre dos (o más) polinucleótidos o polipéptidos se realizan típicamente comparando secuencias de los dos polinucleótidos o polipéptidos en una "ventana de comparación" para identificar y comparar regiones locales de similitud de secuencias. En algunas realizaciones, una "secuencia de referencia" puede basarse en una secuencia de aminoácidos primaria, donde la secuencia de referencia es una secuencia que puede tener uno o más cambios en la secuencia primaria.

[0045] La "ventana de comparación" se refiere a un segmento conceptual de al menos aproximadamente 20 posiciones de nucleótidos contiguos o residuos de aminoácidos en los que se puede comparar una secuencia con una secuencia de referencia de al menos 20 nucleótidos contiguos o aminoácidos y en donde la porción de la secuencia en la ventana de comparación puede comprender adiciones o eliminaciones (es decir, huecos) de 20 por ciento o menos en comparación con la secuencia de referencia (que no comprende adiciones o eliminaciones) para el alineamiento óptimo de las dos secuencias. La ventana de comparación puede tener más de 20 residuos contiguos e incluye, opcionalmente, 30, 40, 50, 100 o más ventanas.

[0046] "Identidad sustancial" se refiere a una secuencia de polinucleótidos o polipéptidos que tiene al menos 80 por ciento de identidad de secuencia, al menos 85 por ciento de identidad y 89 a 95 por ciento de identidad de secuencia, más generalmente al menos 99 por ciento de identidad de secuencia en comparación con una secuencia de referencia sobre una ventana de comparación de al menos 20 posiciones de residuos, con frecuencia sobre una ventana de al menos 30-50 residuos, en donde el porcentaje de identidad de secuencia se calcula comparando la secuencia de referencia con una secuencia que incluye delecciones o adiciones que totalizan 20 por ciento o menos de la secuencia de referencia en la ventana de comparación. En realizaciones específicas aplicadas a polipéptidos, el término "identidad sustancial" significa que dos secuencias de polipéptidos, cuando están alineadas de manera óptima, como los programas GAP o BESTFIT que usan pesos de brecha predeterminados, comparten al menos 80 por ciento de identidad de secuencia, preferiblemente al menos 89 por ciento de identidad de secuencia, al menos 95 por ciento de identidad de secuencia o más (por ejemplo, 99 por ciento de identidad de secuencia). Preferiblemente, las posiciones de residuos que no son idénticas difieren por sustituciones conservativas de aminoácidos.

[0047] "Correspondiente a", "referencia a" o "relativo a" cuando se usan en el contexto de la numeración de un aminoácido dado o secuencia polinucleotídica se refiere a la numeración de los residuos de una secuencia de referencia específica cuando el aminoácido dado o secuencia de polinucleótidos se compara con la secuencia de referencia. En otras palabras, el número de residuo o la posición de residuo de un polímero dado se designa con respecto a la secuencia de referencia en lugar de por la posición numérica real del residuo dentro de la secuencia de aminoácidos o polinucleótido dada. Por ejemplo, una secuencia de aminoácidos dada, como la de una hidroxilasa de prolina modificada, puede alinearse con una secuencia de referencia introduciendo espacios para optimizar las coincidencias de residuos entre las dos secuencias. En estos casos, aunque los huecos están presentes, la numeración del residuo en la secuencia de aminoácidos o polinucleótidos dada se realiza con respecto a la secuencia de referencia con la que se ha alineado.

[0048] "Diferencia de aminoácidos" o "diferencia de residuos" se refiere a un cambio en el residuo de aminoácido en una posición de una secuencia polipeptídica con respecto al residuo de aminoácido en una posición correspondiente en una secuencia de referencia. En algunos casos, las diferencias de residuos también se conocen como "características" de la secuencia del polipéptido. Las posiciones de las diferencias de aminoácidos generalmente se denominan en este documento "Xn", donde n se refiere a la posición correspondiente en la secuencia de referencia en la que se basa la diferencia de residuos. Por ejemplo, una "diferencia de residuos en la posición X103 en comparación con la SEQ ID NO:2" se refiere a un cambio del residuo de aminoácido en la posición del polipéptido correspondiente a la posición 103 de la SEQ ID NO:2. Por lo tanto, si el polipéptido de referencia de la SEQ ID NO:2 tiene una isoleucina en la posición 103, entonces una "diferencia de residuos en la posición X103 en comparación con la SEQ ID NO:2" se refiere a una sustitución de aminoácidos de cualquier resto que no sea isoleucina en la posición del polipéptido correspondiente a la posición 103 de la SEQ ID NO:2. En la mayoría de los casos en este documento, la diferencia específica de residuos de aminoácidos en una posición se indica como "XnY", donde "Xn" especifica la posición correspondiente como se describe anteriormente, e "Y" es el identificador de una sola letra del aminoácido encontrado en el polipéptido diseñado (es decir, el residuo diferente que en el polipéptido de referencia). En algunas realizaciones, cuando más de un aminoácido puede aparecer en una posición de residuo específica, los aminoácidos alternativos pueden enumerarse en la forma XnY/Z, donde Y y Z representan residuos de aminoácidos alternativos, o presentarse como "Xn" seguido por una lista de los residuos especificados. En algunos casos (por ejemplo, en la Tabla **2A**, **2B**, **2C**, **2D** y **2E**), la presente descripción también proporciona diferencias específicas de aminoácidos indicadas por la notación convencional "AnB", donde A es el identificador de una sola letra del residuo en la referencia secuencia, "n" es el número de la posición del residuo en la secuencia de referencia, y B es el identificador de una sola letra de la sustitución del residuo en la secuencia del polipéptido diseñado. La presente descripción incluye secuencias polipeptídicas diseñadas que comprenden una o más diferencias de aminoácidos que incluyen sustituciones de aminoácidos tanto conservativas como no conservativas.

[0049] "Sustitución de aminoácidos conservativa" se refiere a una sustitución de un residuo con un residuo diferente que tiene una cadena lateral similar, y por lo tanto típicamente implica la sustitución del aminoácido en el polipéptido con aminoácidos dentro de la misma clase de aminoácidos o similar definida. A modo de ejemplo y no de limitación, un aminoácido con una cadena lateral alifática puede estar sustituido con otro aminoácido alifático, por ejemplo, alanina, valina, leucina e isoleucina; un aminoácido con cadena lateral hidroxilo está sustituido con otro aminoácido con una cadena lateral hidroxilo, por ejemplo, serina y treonina; un aminoácido que tiene cadenas laterales aromáticas está sustituido con otro aminoácido que tiene una cadena lateral aromática, por ejemplo, fenilalanina, tirosina, triptófano e histidina; un aminoácido con una cadena lateral básica está sustituido con otro aminoácido con una cadena lateral básica, por ejemplo, lisina y arginina; un aminoácido con una cadena lateral ácida está sustituido

con otro aminoácido con una cadena lateral ácida, por ejemplo, ácido aspártico o ácido glutámico; y un aminoácido hidrófobo o hidrófilo se reemplaza con otro aminoácido hidrófobo o hidrófilo, respectivamente. Las sustituciones conservativas ejemplares se proporcionan en la **Tabla 1** a continuación.

Tabla 1

Residuo	Sustituciones conservativas posibles
A, L, V, I	Otros (A, L, V, I) alifáticos Otros (A, L, V, I, G, M) no polares
G, M	Otros (A, L, V, I, G, M) no polares
D, E	Otros (D, E) ácidos
K, R	Otros (K, R) básicos
N, Q, S, T	Otros polares
H, Y, W, F	Otros (H, Y, W, F) aromáticos
C, P	Ninguno

[0050] "Sustitución no conservativa" se refiere a la sustitución de un aminoácido en el polipéptido con un aminoácido con propiedades de cadena lateral significativamente diferentes. Las sustituciones no conservativas pueden usar aminoácidos entre los grupos definidos, y no dentro de ellos, y afectan (a) la estructura del esqueleto peptídico en el área de la sustitución (p. ej., prolina para glicina), (b) la carga o hidrofobicidad, o (c) la mayor parte de la cadena lateral. A modo de ejemplo y no de limitación, una sustitución no conservativa ejemplar puede ser un aminoácido sustituido con un aminoácido básico o alifático; un aminoácido aromático sustituido con un pequeño aminoácido; y un aminoácido hidrófilo sustituido con un aminoácido hidrófobo.

[0051] "Eliminación" se refiere a la modificación del polipéptido mediante la eliminación de uno o más aminoácidos del polipéptido de referencia. Las eliminaciones pueden comprender la eliminación de 1 o más aminoácidos, 2 o más aminoácidos, 5 o más aminoácidos, 10 o más aminoácidos, 15 o más aminoácidos, o 20 o más aminoácidos, hasta el 10% del número total de aminoácidos, o hasta el 20% del número total de aminoácidos que forman la enzima de referencia, al tiempo que conserva la actividad enzimática y/o retiene las propiedades mejoradas de una enzima hidroxilasa de prolina modificada. Las eliminaciones pueden dirigirse a las porciones internas y/o porciones terminales del polipéptido. En diversas realizaciones, la eliminación puede comprender un segmento continuo o puede ser discontinua.

[0052] "Inserción" se refiere a la modificación del polipéptido mediante la adición de uno o más aminoácidos del polipéptido de referencia. En algunas realizaciones, las enzimas de hidroxilasa de prolina modificadas comprenden inserciones de uno o más aminoácidos en el polipéptido de hidroxilasa de prolina que se produce naturalmente, así como inserciones de uno o más aminoácidos en otros polipéptidos de hidroxilasa de prolina mejorados. Las inserciones pueden estar en las partes internas del polipéptido, o en el extremo carboxi o amino. Las inserciones como se usan en este documento incluyen proteínas de fusión como se conoce en la técnica. La inserción puede ser un segmento contiguo de aminoácidos o separado por uno o más de los aminoácidos en la secuencia de referencia.

[0053] "Fragmento" como se usa en el presente documento se refiere a un polipéptido que tiene una eliminación amino-terminal y/o carboxi-terminal, pero donde la secuencia de aminoácidos restante es idéntica a las posiciones correspondientes en la secuencia. Los fragmentos pueden tener una longitud de al menos 14 aminoácidos, una longitud de al menos 20 aminoácidos, una longitud de al menos 50 aminoácidos o más, y hasta un 70%, 80%, 90%, 95%, 98% y 99% del total polipéptido de hidroxilasa de prolina de longitud, por ejemplo el polipéptido de la SEQ ID NO:2 o polipéptido diseñado de la SEQ ID NO:34.

[0054] "Polipéptido aislado" se refiere a un polipéptido que está sustancialmente separado de otros contaminantes que naturalmente lo acompañan, por ejemplo, proteínas, lípidos y polinucleótidos. El término abarca polipéptidos que se han eliminado o purificado de su entorno natural o sistema de expresión (por ejemplo, célula huésped o síntesis *in vitro*). Las enzimas hidroxilasas de prolina mejoradas pueden estar presentes dentro de una célula, presentes en el medio celular o preparadas en diversas formas, tales como lisados o preparaciones aisladas. Como tal, en algunas formas de realización, la enzima hidroxilasa de prolina mejorada puede ser un polipéptido aislado.

[0055] "Polipéptido sustancialmente puro" se refiere a una composición en la que la especie polipeptídica es la especie predominante presente (es decir, en base molar o en peso, es más abundante que cualquier otra especie macromolecular individual en la composición), y es generalmente una composición sustancialmente purificada cuando la especie diana comprende al menos aproximadamente el 50 por ciento de las especies macromoleculares presentes en moles o % en peso. En general, una composición de hidroxilasa de prolina sustancialmente pura comprenderá aproximadamente el 60% o más, aproximadamente el 70% o más, aproximadamente el 80% o más, aproximadamente el 90% o más, aproximadamente el 95% o más y aproximadamente el 98% o más de todas las especies macromoleculares por mol o % en peso presente en la composición. En algunas realizaciones, la especie diana se purifica hasta homogeneidad esencial (es decir, las especies contaminantes no pueden detectarse en la

composición mediante métodos de detección convencionales), en donde la composición consiste esencialmente en una sola especie macromolecular. Las especies solventes, las moléculas pequeñas (<500 Daltons) y las especies de iones elementales no se consideran especies macromoleculares. En algunas realizaciones, el polipéptido de hidroxilasas de prolina mejorado aislado es una composición de polipéptido sustancialmente pura.

[0056] "Estereoselectividad" se refiere a la formación preferencial en una reacción química o enzimática de un estereoisómero sobre otro. La estereoselectividad puede ser parcial, donde la formación de un estereoisómero se favorece sobre el otro, o puede ser completa donde se forma un solo estereoisómero. Cuando los estereoisómeros son enantiómeros, la estereoselectividad se conoce como enantioselectividad, la fracción (típicamente reportada como un porcentaje) de un enantiómero en la suma de ambos. Comúnmente se reporta alternativamente en la técnica (típicamente como un porcentaje) como el exceso enantiomérico (e.e.) calculado de acuerdo con la fórmula $[\text{enantiómero mayor} - \text{enantiómero menor}] / [\text{enantiómero mayor} + \text{enantiómero menor}]$. Cuando los estereoisómeros son diastereoisómeros, la estereoselectividad se denomina diastereoselectividad, la fracción (típicamente informada como un porcentaje) de un diastereómero en una mezcla de dos diastereoisómeros, comúnmente reportada alternativamente como el exceso diastereomérico (d.e.).

[0057] El exceso enantiomérico y el exceso diastereomérico son tipos de exceso estereomérico. "Altamente estereoselectivo" se refiere a una reacción química o enzimática que es capaz de convertir un sustrato, por ejemplo, compuesto (2), en su producto hidroxilado correspondiente, por ejemplo, compuesto (1), con al menos aproximadamente el 85% de exceso estereomérico.

[0058] "Regioselectividad" o "reacción regioselectiva" se refiere a una reacción en la que una dirección de unión o ruptura se produce preferentemente sobre todas las demás direcciones posibles. Las reacciones pueden ser completamente (100%) regioselectivas si la discriminación es completa, sustancialmente regioselectivas (al menos 75%) o parcialmente regioselectivas (x%), si el producto de la reacción en un sitio predomina sobre el producto de la reacción en otros sitios, por ejemplo, formación preferencial del producto compuesto (1) sobre el producto compuesto (1a)).

[0059] La "propiedad enzimática mejorada" se refiere a un polipéptido de hidroxilasa de prolina que muestra una mejora en cualquier propiedad enzimática en comparación con una hidroxilasa de prolina de referencia. Para los polipéptidos de hidroxilasa de prolina descritos en el presente documento, la comparación generalmente se realiza con la enzima hidroxilasa de prolina de tipo salvaje, aunque en algunas realizaciones, la hidroxilasa de prolina de referencia puede ser otra hidroxilasa de prolina. Las propiedades enzimáticas para las cuales es deseable una mejora incluyen, entre otras, la actividad enzimática (que puede expresarse en términos de conversión porcentual del sustrato), estabilidad térmica, estabilidad del disolvente, perfil de actividad del pH, requisitos de cofactor, refractariedad a los inhibidores (por ejemplo, inhibición de sustrato o producto), y estereoselectividad.

[0060] "Actividad enzimática aumentada" se refiere a una propiedad mejorada de los polipéptidos de hidroxilasa de prolina modificados, que puede representarse por un aumento en la actividad específica (por ejemplo, producto producido/tiempo/peso de proteína) o un aumento en el porcentaje de conversión del sustrato al producto (p. ej., porcentaje de conversión de la cantidad inicial de sustrato a producto en un período de tiempo específico utilizando una cantidad específica de hidroxilasa de prolina) en comparación con la enzima de hidroxilasa de prolina de referencia. En los ejemplos se proporcionan métodos ejemplares para determinar la actividad de la enzima. Cualquier propiedad relacionada con la actividad enzimática puede verse afectada, incluidas las propiedades enzimáticas clásicas de K_m , V_{max} o k_{cat} , cuyos cambios pueden conducir a un aumento de la actividad enzimática. Las mejoras en la actividad de la enzima pueden ser desde aproximadamente 1,2 veces la actividad enzimática de la enzima de hidroxilasa de prolina de tipo salvaje correspondiente, hasta 2 veces, 5 veces, 10 veces, 20 veces, 25 veces, 50 veces o más actividad enzimática que la la hidroxilasa de prolina de origen natural u otra hidroxilasa de prolina modificada a partir de la cual se derivaron los polipéptidos de la hidroxilasa de prolina. La actividad de la hidroxilasa de prolina se puede medir mediante uno de los ensayos estándar, como por ejemplo, al monitorear los cambios en las propiedades espectrofotométricas de los reactivos o productos. En algunas realizaciones, la cantidad de productos producidos se puede medir mediante separación por cromatografía líquida de alto rendimiento (HPLC) combinada con absorbancia UV o detección fluorescente después de la derivación, tal como con o-ftalaldehído (OPA) o cloruro de dansilo. Las comparaciones de las actividades enzimáticas se realizan utilizando una preparación definida de enzima, un ensayo definido bajo una condición establecida y uno o más sustratos definidos, como se describe con más detalle en este documento. En general, cuando se comparan los lisados, se determina el número de células y la cantidad de proteína analizada, así como el uso de sistemas de expresión idénticos y células huésped idénticas para minimizar las variaciones en la cantidad de enzima producida por las células huésped y presente en los lisados.

[0061] "Conversión" se refiere a la conversión enzimática del (de los) sustrato(s) en el (los) producto(s) correspondiente(s). "Porcentaje de conversión" se refiere al porcentaje del sustrato que se convierte al producto dentro de un período de tiempo en condiciones específicas. Por lo tanto, la "actividad enzimática" o "actividad" de un polipéptido de hidroxilasa de prolina se puede expresar como "porcentaje de conversión" del sustrato al producto.

[0062] "Termoestable" se refiere a un polipéptido de hidroxilasa de prolina que mantiene una actividad similar (más

del 60% al 80%, por ejemplo) después de la exposición a temperaturas elevadas (por ejemplo, 40-80°C) durante un período de tiempo (por ejemplo, 0,5-24 h) en comparación con la enzima de tipo salvaje.

[0063] "Solvente estable" se refiere a un polipéptido de hidroxilasa de prolina que mantiene una actividad similar (más de, por ejemplo, 60% a 80%) después de la exposición a concentraciones variables (por ejemplo, 5-99%) de disolvente (etanol, alcohol isopropílico, dimetilsulfóxido (DMSO), tetrahidrofurano, 2-metiltetrahidrofurano, acetona, tolueno, acetato de butilo, metilo *terc*-butilo éter, etc.) durante un período de tiempo (p. ej., 0,5-24 h) en comparación con la enzima de tipo salvaje.

[0064] "Estable termo y solvente" se refiere a un polipéptido de hidroxilasa de prolina que es tanto termoestable como estable a los solventes.

[0065] La "hibridación rigurosa" se usa en el presente documento para referirse a las condiciones en las que los híbridos de ácido nucleico son estables. Como saben los expertos en la técnica, la estabilidad de los híbridos se refleja en la temperatura de fusión (T_m) de los híbridos. En general, la estabilidad de un híbrido es una función de la fuerza iónica, la temperatura, el contenido de G/C y la presencia de agentes caotrópicos. Los valores de T_m para polinucleótidos pueden calcularse utilizando métodos conocidos para predecir las temperaturas de fusión (véase, por ejemplo, Baldino et al., *Methods Enzymology* 168:761-777; Bolton et al., 1962, *Proc. Natl. Acad. Sci. EE.UU.* 48:1390; Bresslauer et al., 1986, *Proc. Natl. Acad. Sci. EE.UU.* 83:8893-8897; Freier et al., 1986, *Proc. Natl. Acad. Sci. EE.UU.* 83:9373-9377; Kierzek et al., *Biochemistry* 25:7840-7846; Rychlik et al., 1990, *Nucleic Acids Res* 18:6409-6412 (erratum, 1991, *Nucleic Acids Res.* 19:698); Sambrook et al., *Supra*); Suggs et al., 1981, en *Developmental Biology Using Purified Genes* (Brown et al., Eds.), Pp. 683-693, Academic Press; y Wetmur, 1991, *Crit Rev Biochem Mol Biol* 26:227-259.). En algunas realizaciones, el polinucleótido codifica el polipéptido descrito en el presente documento e hibrida en condiciones definidas, tales como condiciones moderadamente rigurosas o altamente rigurosas, al complemento de una secuencia que codifica una enzima hidroxilasa de prolina modificada de la presente descripción.

[0066] La "rigurosidad de hibridación" se refiere a condiciones de hibridación, tales como condiciones de lavado, en la hibridación de ácidos nucleicos. En general, las reacciones de hibridación se realizan en condiciones de menor rigor, seguidas de lavados de mayor rigor variable pero más alto. El término "hibridación moderadamente rigurosa" se refiere a las condiciones que permiten que el ADN diana se una a un ácido nucleico complementario que tiene una identidad de aproximadamente el 60%, preferiblemente una identidad de aproximadamente el 75%, una identidad de aproximadamente el 85% con el ADN diana, con una identidad superior a aproximadamente el 90% al objetivo-polinucleótido. Las condiciones ejemplares moderadamente rigurosas son condiciones equivalentes a la hibridación en formamida al 50%, solución de 5X Denhart, 5XSSPE, SDS al 0,2% a 42°C, seguido de lavado en 0,2XSSPE, SDS al 0,2%, a 42°C. "Hibridación de alta rigurosidad" se refiere en general a condiciones que son aproximadamente 10°C o menos a partir de la temperatura de fusión térmica T_m según se determina bajo la condición de solución para una secuencia de polinucleótidos definida. En algunas realizaciones, una condición de rigurosidad alta se refiere a las condiciones que permiten la hibridación de solo aquellas secuencias de ácido nucleico que forman híbridos estables en NaCl 0,018M a 65°C (es decir, si un híbrido no es estable en NaCl 0,018M a 65°C, no será estable en condiciones de alta rigurosidad, como se contempla en este documento). Se pueden proporcionar condiciones de alta rigurosidad, por ejemplo, mediante hibridación en condiciones equivalentes a 50% de formamida, solución de 5X Denhart, 5XSSPE, SDS al 0,2% a 42°C, seguido de lavado en 0,1XSSPE y SDS al 0,1% a 65°C. otra condición de alta rigurosidad es la hibridación en condiciones equivalentes a la hibridación en 5X SSC que contiene 0,1% (p:v) SDS a 65°C y lavado en 0,1x SSC que contiene 0,1% SDS a 65°C. Otras condiciones de hibridación de alta rigurosidad, así como condiciones moderadamente rigurosas, se describen en las referencias citadas anteriormente.

[0067] El polinucleótido "heterólogo" se refiere a cualquier polinucleótido que se introduce en una célula huésped mediante técnicas de laboratorio, e incluye polinucleótidos que se eliminan de una célula huésped, se someten a manipulación de laboratorio y luego se reintroducen en una célula huésped.

[0068] "Codón optimizado" se refiere a cambios en los codones del polinucleótido que codifican una proteína a los utilizados preferentemente en un organismo particular, de modo que la proteína codificada se expresa de manera eficiente en el organismo de interés. Aunque el código genético está degenerado porque la mayoría de los aminoácidos están representados por varios codones, llamados "sinónimos" o codones "sinónimos", es bien sabido que el uso de codones por parte de organismos particulares no es aleatorio y está sesgado hacia tripletes de codones particulares. Este sesgo de uso del codón puede ser mayor en referencia a un gen dado, genes de función común u origen ancestral, proteínas altamente expresadas frente a proteínas de bajo número de copias y regiones codificantes de proteínas agregadas del genoma de un organismo. En algunas realizaciones, los polinucleótidos que codifican las enzimas hidroxinasas de la prolina pueden ser codones optimizados para la producción óptima en el organismo huésped seleccionado para la expresión.

[0069] "Codones de sesgo de uso de codones preferidos óptimos" se refiere de manera intercambiable a los codones que se usan a mayor frecuencia en las regiones codificantes de proteínas que otros codones que codifican para el mismo aminoácido. Los codones preferidos se pueden determinar en relación con el uso de

codones en un solo gen, un conjunto de genes de función u origen común, genes altamente expresados, la frecuencia del codón en las regiones codificantes de proteínas agregadas de todo el organismo, la frecuencia de codones en la proteína agregada codificando regiones de organismos relacionados, o combinaciones de los mismos. Los codones cuya frecuencia aumenta con el nivel de expresión génica son típicamente codones óptimos para la expresión. Se conoce una variedad de métodos para determinar la frecuencia del codón (por ejemplo, el uso del codón, el uso relativo del codón sinónimo) y la preferencia del codón en organismos específicos, incluido el análisis multivariado, por ejemplo, el análisis de conglomerados o el análisis de correspondencia, y el número efectivo de codones utilizados. en un gen (consulte GCG CodonPreference, Genetics Computer Group Wisconsin Package; CodonW, John Peden, Universidad de Nottingham; McInerney, J. O, 1998, Bioinformatics 14:372-73; Stenico et al., 1994, Nucleic Acids Res. 22:437-46; Wright, F., 1990, Gene 87:23-29). Las tablas de uso de codones están disponibles para muchos organismos diferentes (ver, por ejemplo, Wada et al., 1992, Nucleic Acids Res. 20:2111-2118; Nakamura et al., 2000, Nucl. Acids Res. 28:292; Duret, et al., supra; Henaut y Danchin, "Escherichia coli y Salmonella", 1996, Neidhardt, et al. Eds., ASM Press, Washington DC, p. 2047-2066). La fuente de datos para obtener el uso de codones puede depender de cualquier secuencia de nucleótidos disponible capaz de codificar una proteína. Estos conjuntos de datos incluyen secuencias de ácido nucleico que se sabe que codifican proteínas expresadas (p. ej., Secuencias codificantes de proteínas completas - CDS), etiquetas de secuencias expresadas (eSTS) o regiones codificadoras predichas de secuencias genómicas (consulte, por ejemplo, Mount, D. Bioinformatics: Secuencia y análisis del genoma, Capítulo 8, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001; Uberbacher, EC, 1996, Methods Enzymol. 266:259-281; Tiwari et al., 1997. Comput. Appl. Biosci. 13:263-270).

[0070] La "secuencia de control" se define aquí para incluir todos los componentes, que son necesarios o ventajosos para la expresión de un polinucleótido y/o polipéptido de la presente descripción. Cada secuencia de control puede ser nativa o ajena a la secuencia de ácido nucleico que codifica el polipéptido. Dichas secuencias de control incluyen, pero no se limitan a, promotor, terminador de transcripción, secuencia líder (es decir, iniciación de la traducción), secuencia de poliadenilación, secuencia de propéptidos y una secuencia peptídica señal. Las secuencias de control pueden proporcionarse con enlazadores con el fin de introducir sitios de restricción específicos que faciliten la ligadura de las secuencias de control con la región codificante de la secuencia de ácido nucleico que codifica un polipéptido.

[0071] "Conectado operativamente" se define en el presente documento como una configuración en la que una secuencia de control se coloca adecuadamente (es decir, en una relación funcional) en una posición relativa a un polinucleótido de interés tal que la secuencia de control dirija o regule la expresión del polinucleótido y/o polipéptido de interés.

[0072] "Secuencia promotora" se refiere a una secuencia de ácido nucleico que es reconocida por una célula huésped para la expresión de un polinucleótido de interés, tal como una secuencia de codificación. La secuencia promotora contiene secuencias de control transcripcional, que median la expresión de un polinucleótido de interés. El promotor puede ser cualquier secuencia de ácido nucleico que muestre actividad transcripcional en la célula huésped de elección, incluidos los promotores mutantes, truncados e híbridos, y puede obtenerse a partir de genes que codifican polipéptidos extracelulares o intracelulares, ya sea homólogos o heterólogos a la célula huésped.

[0073] Las "condiciones de reacción adecuadas" se refieren a esas condiciones en la solución de reacción biocatalítica (por ejemplo, rangos de carga de enzimas, carga de sustrato, carga de co-sustrato, carga de cofactor, temperatura, pH, tampones, codisolventes, etc.) bajo el cual un polipéptido de hidroxilasa de prolina de la presente descripción es capaz de convertir un compuesto de sustrato en un compuesto de producto (por ejemplo, conversión del compuesto (2) en compuesto (1)). Las "condiciones de reacción adecuadas" ejemplares se proporcionan en la presente descripción y se ilustran mediante los ejemplares.

[0074] "Carga", en el sentido de "carga de compuestos" o "carga de enzimas" o "carga de cofactor" se refiere a la concentración o cantidad de un componente en una mezcla de reacción al comienzo de la reacción.

[0075] "Sustrato" en el contexto de un proceso mediado por biocatalizadores se refiere al compuesto o molécula sobre la que actúa el biocatalizador. Por ejemplo, un compuesto ejemplar para el biocatalizador de hidroxilasa de prolina en el proceso descrito aquí es el compuesto (2).

[0076] "Producto" en el contexto de un proceso mediado por un biocatalizador se refiere al compuesto o molécula resultante de la acción del biocatalizador. Por ejemplo, un producto ejemplar para el biocatalizador de hidroxilasa de prolina en el proceso descrito aquí es el compuesto (1).

[0077] "Reductante" se refiere a un compuesto o agente capaz de convertir Fe^{+3} en Fe^{+2} . Un reductor ejemplar es el ácido ascórbico, que generalmente está en forma de ácido L-ascórbico.

[0078] "Alquilo" se refiere a grupos hidrocarbonados saturados de 1 a 18 átomos de carbono inclusive, ya sea de cadena lineal o ramificada, más preferiblemente de 1 a 8 átomos de carbono inclusive, y lo más preferiblemente de 1 a 6 átomos de carbono inclusive. Un alquilo con un número específico de átomos de carbono se indica entre

paréntesis, por ejemplo, alquilo (C_1-C_6) se refiere a un alquilo de 1 a 6 átomos de carbono.

[0079] "Alquenilo" se refiere a grupos hidrocarburo de 2 a 12 átomos de carbono inclusive, ya sea lineal o ramificado que contiene al menos un doble enlace, pero que opcionalmente contiene más de un doble enlace.

[0080] "Alquinilo" se refiere a grupos hidrocarbonados de 2 a 12 átomos de carbono inclusive, ya sea lineal o ramificado que contiene al menos un triple enlace pero que opcionalmente contiene más de un triple enlace, y adicionalmente opcionalmente que contiene uno o más restos dobles.

[0081] "Alquilenilo" se refiere a un radical hidrocarbonado divalente de cadena lineal o ramificada que tiene de 1 a 18 átomos de carbono inclusive, más preferiblemente de 1 a 8 átomos de carbono inclusive, y lo más preferiblemente de 1 a 6 átomos de carbono inclusivamente, opcionalmente sustituido con uno o sustituyentes más adecuados. Los "alquilenos" ejemplares incluyen, pero no se limitan a, metileno, etileno, propileno, butileno y similares.

[0082] "Alquenilenilo" se refiere a un radical hidrocarbonado divalente de cadena lineal o ramificada que tiene 2 a 12 átomos de carbono inclusive y uno o más dobles enlaces carbono-carbono, más preferiblemente de 2 a 8 átomos de carbono inclusive, y lo más preferiblemente 2 a 6 carbono átomos inclusive, opcionalmente sustituidos con uno o más sustituyentes adecuados.

[0083] "Heteroalquilo," heteroalquenilo "y heteroalquinilo" se refieren a alquilo, alquenilo y alquinilo como se definen en el presente documento en el que uno o más de los átomos de carbono están cada uno independientemente reemplazados con los mismos o diferentes heteroátomos o grupos heteroatómicos. Los heteroátomos y/o grupos heteroatómicos que pueden reemplazar a los átomos de carbono en la lista, pero no se limitan a, -O-, -S-, -SO-, -NR_y-, -PH-, -S(O)-, -S(O)₂-, -S(O)NR_y-, -S(O)₂NR_y-, y similares, incluyendo combinaciones de los mismos, donde cada R_y se selecciona independientemente de hidrógeno, alquilo, heteroalquilo, cicloalquilo, heterocicloalquilo, arilo y heteroarilo.

[0084] "Arilo" se refiere a un grupo carbocíclico aromático insaturado de 6 a 12 átomos de carbono, incluyendo inclusive un solo anillo (por ejemplo, fenilo) o múltiples anillos condensados (por ejemplo, naftilo o antrilo). Los arilos ejemplares incluyen fenilo, piridilo, naftilo y similares.

[0085] "Arlalquilo" se refiere a un alquilo sustituido con un arilo, es decir, grupos arilo-alquilo, que tienen preferiblemente de 1 a 6 átomos de carbono inclusive en el resto alquilo y de 6 a 12 átomos de carbono inclusive en el resto arilo. Tales grupos arilalquilo están ejemplificados por bencilo, fenetilo y similares.

[0086] "Arloxi" se refiere a grupos -OR_λ, donde R_λ es un grupo arilo, que puede estar opcionalmente sustituido.

[0087] "Cicloalquilo" se refiere a grupos alquilo cíclicos de 3 a 12 átomos de carbono que tienen inclusive un solo anillo cíclico o anillos condensados múltiples que pueden estar opcionalmente sustituidos con 1 a 3 grupos alquilo. Los grupos cicloalquilo ejemplares incluyen, pero no se limitan a, estructuras de un solo anillo tales como ciclopropilo, ciclobutilo, ciclopentilo, ciclooctilo, 1-metilo-ciclopropilo, 2-metilciclopentilo, 2-metilciclooctilo y similares, o estructuras de anillos múltiples, incluyendo anillos múltiples sistemas, tales como adamantilo, y similares.

[0088] "Cicloalquilalquilo" se refiere a un alquilo sustituido con un cicloalquilo, es decir, grupos cicloalquilo-alquilo-, que tienen preferiblemente de 1 a 6 átomos de carbono inclusive en el resto alquilo y de 3 a 12 átomos de carbono inclusive en el resto cicloalquilo. Dichos grupos cicloalquilalquilo están ejemplificados por ciclopropilmetilo, ciclohexiletilo y similares.

[0089] "Amino" se refiere al grupo -NH₂. Amino sustituido se refiere al grupo -NHR_η, NR_ηR_η y NR_ηR_ηR_η, donde cada R se selecciona independientemente entre alquilo, cicloalquilo, cicloheteroalquilo, alcoxi, arilo, heteroarilo, heteroarilo, heteroarilalquilo, acilo, alcocarbonilo, sulfanilo, sulfanilo, sulfato, y similares. Los grupos amino típicos incluyen, pero están limitados a, dimetilamino, dietilamino, trimetilamonio, trietilamonio, metilsulfonilamino, furanilo-oxi-sulfamino, y similares.

[0090] "Aminoalquilo" se refiere a un grupo alquilo en el que uno o más de los átomos de hidrógeno están reemplazados con uno o más grupos amino, incluyendo grupos amino sustituidos.

[0091] "Aminocarbonilo" se refiere a -C(O)NH₂. Aminocarbonilo sustituido se refiere a -C(O)NR_ηR_η, donde el grupo amino NR_ηR_η es como se define aquí.

[0092] "Oxi" se refiere a un grupo divalente -O-, que puede tener varios sustituyentes para formar diferentes grupos oxi, incluyendo éteres y ésteres.

[0093] En el presente documento, "alcoxi" o "alquilxi" se usan de manera intercambiable para referirse al grupo -OR_ζ, en donde R_ζ es un grupo alquilo, que incluye grupos alquilo opcionalmente sustituidos.

[0094] "Carboxi" se refiere a -COOH.

[0095] "Carbonilo" se refiere a -C(O)-, que puede tener una variedad de sustituyentes para formar diferentes grupos carbonilo incluyendo ácidos, halogenuros de ácido, aldehídos, amidas, ésteres y cetonas.

[0096] "Carboxialquilo" se refiere a un alquilo en el que uno o más de los átomos de hidrógeno están reemplazados con uno o más grupos carboxi.

[0097] "Aminocarbonilalquilo" se refiere a un alquilo sustituido con un grupo aminocarbonilo, como se define en el presente documento.

[0098] "Halógeno" o "halo" se refiere a flúor, cloro, bromo y yodo.

[0099] "Haloalquilo" se refiere a un grupo alquilo en el que uno o más de los átomos de hidrógeno se reemplazan con un halógeno. Por lo tanto, el término "haloalquilo" pretende incluir monohaloalquilos, dihaloalquilos, trihaloalquilos, etc. hasta perhaloalquilos. Por ejemplo, la expresión "haloalquilo (C₁ C₂)" incluye 1-fluorometilo, difluorometilo, trifluorometilo, 1-fluoroetilo, 1,1-difluoroetilo, 1,2-difluoroetilo, 1,1,1 trifluoroetilo, perfluoroetilo, etc.

[0100] "Hidroxi" se refiere a -OH.

[0101] "Hidroxi-alquilo" se refiere a un grupo alquilo en el que uno o más de los átomos de hidrógeno se reemplazan con uno o más grupos hidroxilo.

[0102] "Tiol" o "sulfanilo" se refiere a -SH. Tiol o sulfanilo sustituido se refiere a -S-R_n, donde R_n es un alquilo, arilo u otro sustituyente adecuado.

[0103] "Alquiltio" se refiere a -SR_z, donde R_z es un alquilo, que puede estar opcionalmente sustituido. El grupo alquiltio a modo de ejemplo incluye, pero no se limita a, metiltio, etiltio, n-propiltio y similares.

[0104] "Arlitio" se refiere a -SR_λ, donde R_λ es un arilo, que puede estar opcionalmente sustituido. Los grupos ariltio ejemplares incluyen, pero no se limitan a, feniltio, (4-metilfenilo)tio, piridiniltio y similares.

[0105] Alquiltioalquilo se refiere a un alquilo sustituido con un grupo alquiltio, -SR_z, donde R_z es un alquilo, que puede estar opcionalmente sustituido.

[0106] "Tiolalquilo" se refiere a un grupo alquilo en el que uno o más de los átomos de hidrógeno están reemplazados con uno o más grupos SH.

[0107] "Sulfonilo" se refiere a -SO₂-. Sulfonilo sustituido se refiere a -SO₂-R_n, donde R_n es un alquilo, arilo u otro sustituyente adecuado.

[0108] "Alquilosulfonilo" se refiere a -SO₂-R_z, donde R_z es un alquilo, que puede estar opcionalmente sustituido. Los grupos Alquilosulfonilo típicos incluyen, pero no se limitan a, metilsulfonilo, etilsulfonilo, n-propilsulfonilo y similares.

[0109] "Alquilosulfonilalquilo" se refiere a un alquilo sustituido con un grupo alquilosulfonilo, -SO₂-R_z, donde R_z es un alquilo, que puede estar opcionalmente sustituido.

[0110] "Heteroarilo" se refiere a un grupo heterocíclico aromático de 1 a 10 átomos de carbono inclusive y de 1 a 4 heteroátomos seleccionados inclusivamente entre oxígeno, nitrógeno y azufre cuando se adelgaza el anillo. Dichos grupos heteroarilo pueden tener un solo anillo (por ejemplo, piridilo o furilo) o múltiples anillos condensados (por ejemplo, indolizínilo o benzotienilo).

[0111] "Heteroarilalquilo" se refiere a un alquilo sustituido con un heteroarilo, es decir, grupos heteroarilo-alquilo, que tienen preferiblemente de 1 a 6 átomos de carbono inclusive en el resto alquilo y de 5 a 12 átomos en el anillo inclusive en el resto heteroarilo. Dichos grupos heteroarilalquilo están ejemplificados por piridilmetilo y similares.

[0112] "Heterociclo", "heterocíclico" e indistintamente "heterocicloalquilo" se refiere a un grupo saturado o insaturado que tiene un solo anillo o múltiples anillos condensados, de 2 a 10 átomos de anillo de carbono inclusive y de 1 a 4 átomos de anillo hetero seleccionados de forma inclusiva de nitrógeno, azufre u oxígeno dentro del anillo. Dichos grupos heterocíclicos pueden tener un solo anillo (por ejemplo, piperidinilo o tetrahidrofurilo) o múltiples anillos condensados (por ejemplo, indolinilo, dihidrobenzofurano o quinclidinilo). Los ejemplos de heterociclos incluyen, pero no se limitan a, furano, tiofeno, tiazol, oxazol, pirrol, imidazol, pirazol, piridina, pirazina, pirimidina, piridazina, indolizina, isoindol, indolol, purina, quinolizina, isoquinina, ftalazina, naftilpiridina, quinoxalina, quinazolina, cinolina, pteridina, carbazol, carbolina, fenantridina, acridina, fenantrolina, isotiazol, fenazina, isoxazol, fenoxazina, fenotiazina, imidazolidina, imidazolidina, piperidina, piperazina, pirrolidina, indolina y similares.

[0113] "Heterocicloalquilalquilo" se refiere a un alquilo sustituido con un heterocicloalquilo, es decir, grupos heterocicloalquilo-alquilo, que tienen preferiblemente de 1 a 6 átomos de carbono inclusive en el resto alquilo y de 3 a 12 átomos de anillo, inclusive en el resto heterocicloalquilo.

[0114] "Anillo miembro" pretende abarcar cualquier estructura cíclica. El número que precede al término "miembro" denota el número de átomos esqueléticos que constituyen el anillo. Así, por ejemplo, ciclohexilo, piridina, pirano y tiopirano son anillos de 6 miembros y ciclopentilo, pirrol, furano y tiofeno son anillos de 5 miembros.

[0115] "Anillo bicíclico fusionado", como se usa en el presente documento, se refiere a restos de anillos heterocíclicos y/o carbocíclicos no sustituidos y sustituidos que tienen 5 a 8 átomos en cada anillo, los anillos que tienen 2 átomos comunes.

[0116] A menos que se especifique lo contrario, las posiciones ocupadas por hidrógeno en los grupos anteriores se pueden sustituir adicionalmente con sustituyentes ejemplificados por, pero no limitados a, hidroxilo, oxo, nitro, metoxi, etoxi, alcoxi, alcoxi sustituido, trifluorometoxi, haloalcoxi, fluoro, cloro, bromo, yodo, halo, metilo, etilo, propilo, butilo, alquilo, alqueno, alquino, alquilo sustituido, trifluorometilo, haloalquilo, hidroxialquilo, alcoxialquilo, tio, alquiltio, acilo, carboxi, alcoxycarbonilo, carboxiladino, sustituido carboxamido, alquilosulfonilo, alquilosulfonilo, alquilosulfonilamino, sulfonamido, sulfonamido sustituido, ciano, amino, amino sustituido, alquilamino, dialquilamino, aminoalquilo, acilamino, amidino, amidoximo, hidroxamilo, fenilo, arilo, arilo sustituido, ariloxi, arilalquilo, arilalqueno, arilalquino, piridilo, imidazolilo, heteroarilo, heteroarilo sustituido, heteroariloxi, heteroarilalquilo, heteroarilalqueno, heteroarilalquino, ciclobutilo, ciclopentilo, ciclohexilo, cicloalquilo, cicloalqueno, cicloalquilalquilo, cicloalquilo sustituido, cicloalquilxilo, pirrolidinilo, piperidinilo, morfolino, heterociclo, (heterociclo)oxi, y (heterociclo)alquilo; y los heteroátomos preferidos son oxígeno, nitrógeno y azufre. Se entiende que cuando existen valencias abiertas en estos sustituyentes, se pueden sustituir adicionalmente con grupos alquilo, cicloalquilo, arilo, heteroarilo y/o heterociclo, y que cuando existen estas valencias abiertas en el carbono, se pueden sustituir adicionalmente con halógeno y oxígeno, sustituyentes unidos a nitrógeno, o azufre, y donde existen múltiples valencias abiertas, estos grupos pueden unirse para formar un anillo, ya sea por formación directa de un enlace o por formación de enlaces a un nuevo heteroátomo, preferiblemente oxígeno, nitrógeno, o azufre. Se entiende además que las sustituciones anteriores se pueden realizar siempre que la sustitución del hidrógeno con el sustituyente no introduzca una inestabilidad inaceptable en las moléculas de la presente descripción y, por lo demás, sea químicamente razonable.

[0117] "Opcional" u "opcionalmente" significa que el evento o circunstancia que se describe posteriormente puede o no ocurrir, y que la descripción incluye casos en los que ocurre el evento o circunstancia y casos en los que no. Un experto en la técnica entenderá que con respecto a cualquier molécula descrita que contenga uno o más sustituyentes opcionales, solo se pretende incluir compuestos estéricamente prácticos y/o sintéticamente viables. "Opcionalmente sustituido" se refiere a todos los modificadores posteriores en un término o serie de grupos químicos. Por ejemplo, en el término "arilalquilo opcionalmente sustituido, la porción" alquilo "y la porción" arilo "de la molécula pueden o no estar sustituidas, y para la serie" alquilo, cicloalquilo, arilo y heteroarilo opcionalmente sustituidos," los grupos alquilo, cicloalquilo, arilo y heteroarilo, independientemente de los otros, pueden estar o no sustituidos.

5.3 Polipéptidos de hidroxilasa de prolina diseñados

[0118] La presente descripción proporciona polipéptidos que tienen actividad de hidroxilasa de prolina, polinucleótidos que codifican los polipéptidos, métodos de preparación de los polipéptidos y métodos para usar los polipéptidos. Cuando la descripción se refiere a polipéptidos, debe entenderse que también describe los polinucleótidos que codifican los polipéptidos.

[0119] Las hidroxilasas de prolina pertenecen a una clase de enzimas diooxigenasa que catalizan la hidroxilación de prolina en presencia de α -cetoglutarato y oxígeno (O_2). El α -cetoglutarato se descarboxila estequiométricamente durante la hidroxilación, con un átomo de la molécula de O_2 incorporado en el succinato y el otro en el grupo hidroxilo formado en el residuo de prolina. Como se señaló anteriormente, las hidroxilasas de prolina se distinguen de la hidroxilasa de prolilo por su capacidad para hidroxilar la prolina libre.

[0120] Se han identificado varios tipos de hidroxilasas de prolina basadas en los principales productos diastereoméricos formados en la reacción enzimática: cis-3-hidroxilasa de prolina (cis-P3H), cis-4-hidroxilasa de prolina (cis-P4H), trans-3-prolina-hidroxilasa (trans-P3H), y trans-4-hidroxilasa de prolina (trans-P4H). Las enzimas cis-P3H se han identificado en *Streptomyces* sp. TH1, *Streptomyces canus* y *Bacillus* sp. TH2 y TH3 (Mori H. et al., 1996, Appl. Environ. Microbiol. 62 (6):1903-1907). trans-P3H se ha identificado en *Glarea lozoyensis* (Petersen, L. et al., 2003, Appl Microbiol Biotechnol. 62 (2-3):263-7). cis-P4H se han identificado en *Lotus corniculatus rhizobia*, *Mesorhizobium loti*, *Sinorhizobium meliloti* y *Medicago sativa rhizobia*, (Hara y Kino, 2009, Biochem Biophys Res Commun. 379 (4):882-6; patente de EE.UU. n° de publicación. 20110091942). trans-P4H se han identificado en *Dactylosporangium* sp., *Amycolatopsis* sp., *Streptomyces griseoviridis*, *Streptomyces* sp. y *Glarea lozoyensis* (Shibasaki T. et al., 1999, Appl. Environ. Microbiol. 65 (9):4028-31; 2003, Petersen, L. et al., 2003, Appl Microbiol

Biotechnol. 62 (2-3):263-7; Mori, H. et al., 1996, Appl. Environ. Microbiol. 62:1903-1907; Lawrence, C-C, et al., 1996, Biochem. J. 313:185-191; y EP0641862).

[0121] La cis-4-hidroxilasa de prolina de *Sinorhizobium meliloti* convierte la prolina libre en el producto primario cis-4-hidroxiprolina. Según Klein et al., Supra, la enzima también reconoce el ácido L-pipecólico, convirtiéndolo en una mezcla de ácido cis-5-y cis-3-hidroxipipecólico. Sin embargo, la actividad en el ácido pipecólico es menor que en la prolina, y se informa que la enzima tiene una actividad específica baja y desnaturalización en condiciones de reacción (Klein et al., Supra). En consecuencia, las reacciones de conversión *in vitro* para preparar hidroxiprolina y ácido hidroxipipecólico con una enzima de tipo salvaje recombinante expresada en *E. coli* no fueron adecuadas como estrategia sintética para preparaciones a escala comercial. Se encontró que las células completas que expresaban la enzima eran más efectivas, pero requerían el uso de un medio de crecimiento definido que carece de prolina para minimizar la competencia de la prolina libre y también simplificar la purificación del producto de ácido hidroxipipecólico (Klein et al., supra).

[0122] En la presente descripción, se describen hidroxilasas de prolina modificadas que superan las deficiencias de la cis-4-hidroxilasa de prolina de tipo salvaje de *Sinorhizobium meliloti*. Los polipéptidos de hidroxilasa de prolina derivados de la enzima de tipo salvaje de *Sinorhizobium meliloti* son capaces de convertir eficazmente *in vitro* prolina libre en cis-4-hidroxiprolina, pero también pueden convertir eficientemente una gama de sustratos, incluida la conversión de ácido L-pipecólico (es decir, ácido (2S)-piperidina-2-carboxílico) al ácido cis-5-hidroxipipecólico (es decir, ácido (2S, 5S)-5-hidroxipiperidina-2-carboxílico). Significativamente, la presente descripción identifica las posiciones de los residuos de aminoácidos y las mutaciones correspondientes en la secuencia del polipéptido de la hidroxilasa de prolina que mejoran las propiedades enzimáticas en comparación con la enzima natural, incluyendo, entre otras, actividad, estabilidad, expresión, regioselectividad, estereoselectividad, tolerancia del sustrato, y especificidad del sustrato. En particular, la presente descripción proporciona polipéptidos diseñados para convertir de manera eficiente el compuesto de sustrato (2), ácido (2S)-piperidina-2-carboxílico, al compuesto de producto (1), ácido (2S, 5S)-5-hidroxipiperidona-2-carboxílico (como se ilustra en el **Esquema 1** anterior) en presencia de un co-sustrato en condiciones de reacción adecuadas.

[0123] En algunas realizaciones, los polipéptidos de hidroxilasa de prolina modificados muestran una actividad incrementada en la conversión de prolina y ácido (2S)-piperidina-2-carboxílico al producto cis-4-hidroxi prolina y ácido (2S, 5S)-5-hidroxipiperidina-2-carboxílico, respectivamente, en un tiempo definido con la misma cantidad de enzima en comparación con la enzima de tipo salvaje. En algunas realizaciones, el polipéptido de hidroxilasa de prolina diseñado tiene al menos aproximadamente 1,2 veces, 1,5 veces, 2 veces, 3 veces, 4 veces, 5 veces o 10 veces o más en condiciones de reacción adecuadas en comparación con el polipéptido representado por la SEQ ID. NO:2.

[0124] En algunas realizaciones, los polipéptidos de hidroxilasa de prolina modificados han aumentado la regioselectividad en comparación con la hidroxilasa de prolina de tipo salvaje. Específicamente, la enzima natural convierte la prolina en, principalmente, si no exclusivamente, cis-4-hidroxiprolina y convierte el ácido (2S)-piperidina-2-carboxílico en una mezcla de productos diastereoméricos que comprenden el compuesto (1), ácido (2S, 5S)-5-hidroxipiperidina-2-carboxílico y compuesto (**1a**), ácido (2S,3R)-3-hidroxi-piperidina-2-carboxílico. En algunas realizaciones, los polipéptidos de hidroxilasa de prolina diseñados en la presente invención son capaces de formar selectivamente el compuesto (1) en exceso del compuesto del producto (**1a**). En algunas realizaciones, los polipéptidos modificados pueden formar selectivamente el compuesto (1) en exceso del producto compuesto (**1a**), donde la relación del compuesto (1) formado sobre el compuesto (**1a**) en condiciones de reacción adecuadas es al menos 1,5, 2, 3, 4, 5 o 6 o más.

[0125] En algunas realizaciones, los polipéptidos de hidroxilasa de prolina diseñados son capaces de convertir el compuesto de sustrato (2) en el compuesto de producto (1) sin formar cantidades significativas de ácido trans-5-hidroxipipecólico (es decir, ácido (2S, 5R)-5-hidroxipiperidina-2-carboxílico). En algunas realizaciones, los polipéptidos de hidroxilasa de prolina modificados son capaces de convertir el compuesto de sustrato (2) en el compuesto del producto (1) en condiciones de reacción adecuadas en más del 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5 o más exceso diastereomérico de ácido (2S, 5R)-5-hidroxipiperidina-2-carboxílico.

[0126] En algunas realizaciones, los polipéptidos de hidroxilasa de prolina modificados son capaces de convertir el compuesto de sustrato (2) en el compuesto de producto (1) en condiciones de reacción adecuadas con mayor tolerancia para la presencia de sustrato en relación con el polipéptido de referencia de la SEQ ID NO:2. De este modo, en algunas realizaciones, los polipéptidos de hidroxilasa de prolina modificados son capaces de convertir el compuesto de sustrato (2) en el compuesto del producto (1) a una concentración de carga de sustrato de al menos aproximadamente 10 g/L, aproximadamente 20 g/L, aproximadamente 30 g/L, aproximadamente 40 g/L, aproximadamente 50 g/L, aproximadamente 70 g/L, aproximadamente 100 g/L, aproximadamente 125 g/L, aproximadamente 150 g/L, aproximadamente 175 g/L o aproximadamente 200 g/L o más con un porcentaje de conversión de al menos aproximadamente el 40%, al menos aproximadamente el 50%, al menos aproximadamente el 60%, al menos aproximadamente el 70%, al menos aproximadamente el 80%, al menos aproximadamente el 90%, al menos aproximadamente el 95%, al menos aproximadamente el 98%, o al menos aproximadamente el 99%, en un tiempo de reacción de aproximadamente 120 horas o menos, 72 horas o menos, aproximadamente 48 horas o

menos, aproximadamente 36 horas o menos, o aproximadamente 24 h menos, en condiciones de reacción adecuadas.

[0127] Las condiciones de reacción adecuadas bajo las cuales las propiedades mejoradas descritas anteriormente de los polipéptidos diseñados llevan a cabo la reacción de hidroxilación se pueden determinar con respecto a las concentraciones o cantidades de polipéptido, sustrato, co-sustrato, cofactor de metal de transición, reductor, tampón, co-disolvente, pH, condiciones que incluyen temperatura y tiempo de reacción, y/o condiciones con el polipéptido inmovilizado sobre un soporte sólido, como se describe más adelante y en los ejemplos.

[0128] Los polipéptidos diseñados a modo de ejemplo que tienen actividad hidroxilasa de prolina con propiedades mejoradas, particularmente en la conversión del compuesto (2) en compuesto (1), comprenden una secuencia de aminoácidos que tiene una o más diferencias de residuos en comparación con la SEQ ID NO:2 en las siguientes posiciones de residuos: X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X186; X188; X205; X225; X230; X270; y X271. Las diferencias de aminoácidos específicas en cada una de estas posiciones que están asociadas con las propiedades mejoradas de los polipéptidos ejemplares de las **Tablas 2A, 2B, 2C, 2D, 2E, 2F, 2G y 2H** incluyen: X2K; X2T; X3S; X4Q; X4L; X4E; X4S; X5I; X5L; X5M; X9I; X13T; X17V; X24R; X24S; X25R; X26R; X26T; X26W; X29A; X30V; X30P; X36T; X42E; X52P; X57T; X57A; X58A; X59G; X62Q; X66Q; X86S; X88R; X92V; X95M; X98F; X98T; X103L; X103Q; X112T; X112V; X113E; X114N; X115E; X115H; X115D; X115G; X115S; X115A; X116L; X121F; X131Y; X131F; X140L; X150S; X151A; X151H; X151S; X166T; X166L; X166Q; X186G; X188G; X205V; X225L; X225Y; X225W; X230V; X270E; X271K; y X271R.

[0129] La información de la estructura y función de los polipéptidos de hidroxilasa de prolina de origen no natural (o diseñados) de la presente descripción se basa en la conversión del compuesto (2) en el compuesto (1), cuyos resultados se muestran a continuación en las **Tablas 2A, 2B, 2C, 2D, 2E, 2F, 2G y 2H**. Los identificadores de secuencia numerada impar (es decir, las SEQ ID NO) se refieren a la secuencia de nucleótidos que codifica la secuencia de aminoácidos proporcionada por las SEQ ID NO de número par. Las secuencias ejemplares se proporcionan en el archivo de listado de secuencia electrónica que acompaña a esta descripción. Las diferencias de residuos de aminoácidos se basan en la comparación con la secuencia de referencia de SEQ ID NO:2 (o SEQ ID NO:4, o 6), que representan la secuencia de aminoácidos de origen natural de la cis-4-hidroxilasa de prolina de *Sinorhizobium meliloti*. La actividad de cada polipéptido diseñado con respecto al polipéptido de referencia de la SEQ ID NO:2 se determinó como la conversión del sustrato, ácido (2S)-piperidina-2-carboxílico, al producto, ácido (2S, 5S)-5-hidroxipiperidina-2-carboxílico durante un período de tiempo y temperatura establecidos en un ensayo de alto rendimiento (HTP), que se usó como pantalla principal. Los valores del ensayo HTP en las **Tablas 2A, 2B y 2F** se determinaron usando lisados de células claras de *E. coli* en formato de placa de 96 pocillos de ~ 200 µl de volumen por pocillo siguientes a condiciones de reacción de ensayo como se indica en la tabla y los Ejemplos. En algunos casos, se utilizó un ensayo en polvo en matraz de agitación (SFP) o en polvo procesado posteriormente (DSP) como cribado secundario para evaluar las propiedades de las hidroxilasas de prolina, cuyos resultados se proporcionan en las **Tablas 2C, 2D, 2E, 2G, y 2H**. Las formas SFP proporcionan una preparación en polvo más purificada de los polipéptidos diseñados y pueden contener los polipéptidos diseñados que representan hasta aproximadamente el 30% de la proteína total. Las preparaciones de DSP pueden proporcionar una forma aún más purificada del polipéptido diseñado ya que las preparaciones pueden contener las hidroxilasas de prolina diseñadas que son hasta aproximadamente el 80% de la proteína total. Las hidroxilasas de prolina modificadas también se examinaron para determinar sus regioselectividades midiendo la relación (expresada como la relación de selectividad) del compuesto del producto (1), ácido (2S, 5S)-5-hidroxipiperidina-2-carboxílico, al compuesto del producto (1a), ácido (2S,3R)-3-hidroxipiperidina-2-carboxílico, formado en las reacciones.

Tabla 2A: Polipéptidos diseñados y mejoras de enzimas relativas utilizando preparaciones de HTP

	SEQ ID NO: (nt/aa)	Diferencias de aminoácido (en relación con SEQ ID NO: 2)	Actividad^{1,2} (Condición A) (en relación con SEQ ID NO: 2)
5	1/2	N/A	1
	3/4	N/A	1 ³
	5/6	N/A	1 ³
10	7/8	A42E	1,53
	9/10	I103L	2,37
	11/12	I103Q	2,26
15	13/14	F116L	1,15
	15/16	N131Y	1,7
	17/18	A150S	2,07
	19/20	S29A; H166T	1,47
20	21/22	H166T	1,90
	23/24	H166Q	4,96
	25/26	S30V	1,33
	27/28	S30P	2,75
25	29/30	A36T	1,8
	31/32	S59G	1,54
	33/34	V57T	1,54
	35/36	Q52P; S225Y	1,3
30	37/38	V57A	1,43
	39/40	T115H	1,68
	41/42	L112T	1,87
	43/44	A66Q	1,21
35	45/46	H271K	1,24
	47/48	T115D	1,25
	49/50	T115G; L121F	1,4
	51/52	H271R	1,31
40	53/54	L112V	1,32
	55/56	T115S	1,31
	57/58	V25R; V58A	1,49
	59/60	S2K	1,55
45	61/62	S2T	1,52
	63/64	E13T	1,6
	65/66	H4Q	6,27
	67/68	H4L	2,44
50	69/70	F5I	2,32
	71/72	F5L	3,9
	73/74	F5M	1,94
	75/76	V9I	2,81
55	77/78	H4E	1,28
	79/80	H4L; T115A	1,22
	81/82	H4S	1,75
	83/84	S225W	1,47
60	85/86	S225L	1,6
	87/88	L230V	1,28
	89/90	N131F	1,26
	91/92	T92V	1,37
65	93/94	V95M	1,73
	95/96	T115G	1,37
	97/98	T3S	6,43

(Continuado)

SEQ ID NO: (nt/aa)	Diferencias de aminoácido (en relación con SEQ ID NO: 2)	Actividad ^{1,2} (Condición A) (en relación con SEQ ID NO: 2)
5	¹ Condición A del ensayo HTP A (ver también Ejemplo 4): Las células cultivadas en placas de 96 pocillos se lisaron con 100 µl de tampón de lisis (1 mg/ml de lisozima, 0,5 mg/ml de sulfato de polimixina B (PMBS) y 50 mM de tampón fosfato, pH 6,3). Las condiciones de reacción para una reacción de 200 µl comprendieron: 10 g/L de compuesto de sustrato (2); 19 g/L de ácido α-cetoglutarico; 21 g/L ácido L-ascórbico; Sal de Mohr 1,5 mM; Tampón de fosfato de potasio 50 mM, pH 6,3 (pH ajustado con KOH); 100 µl de lisado crudo; y la temperatura de reacción a aproximadamente 25°C (temperatura ambiente) durante aproximadamente 24 h. Las placas se cubrieron con un sello permeable a O ₂ y se agitaron en un agitador de placa de títulos a la velocidad N° 2,5.	
10	² La actividad relativa a SEQ ID NO:2 se calcula como el % de conversión del producto formado por % de conversión de la SEQ ID NO:2 correspondiente en las condiciones de reacción especificadas. El % de conversión se cuantificó dividiendo primero las áreas del pico del producto por la suma de las áreas del sustrato, producto e impurezas/producto lateral pico determinado por análisis de HPLC y luego restándolo con el % de conversión del control negativo (pico de superposición).	
15	³ Las secuencias de aminoácidos de los polipéptidos representados por la SEQ ID NO:4 y 6 son idénticas a la secuencia de aminoácidos de la hidroxilasa de prolina que se produce naturalmente en la SEQ ID NO:2, pero que se prepararon en E. coli utilizando los diferentes genes de codones optimizados de SEQ ID NO:3 y 5, respectivamente. Los genes optimizados de codón de SEQ ID NO:3 y 5 mostraron una expresión incrementada de 1,2 veces del polipéptido de hidroxilasa de prolina en E. coli en comparación con la expresión de SEQ ID NO:1,	
20		
25		

Tabla 2B: Polipéptidos diseñados y mejoras de enzimas relativas utilizando preparaciones de HTP

SEQ ID NO: (nt/aa)	Diferencias de aminoácidos (en relación con SEQ ID NO: 2)	Actividad (Condición B) ¹ (en relación con SEQ ID NO: 112) ²
113/114	C86S; T92V; I103L; M151S; H166Q; G270E	2,3
115/116	I103Q; F116L; H166L; S225L	1,53
117/118	A66Q; I103L; H166Q	1,48
119/120	A66Q; I103L; H166Q	1,81
121/122	A66Q; T92V; I103L; H166Q	1,67
123/124	T92V; I103L; H166Q	1,51
125/126	A66Q; T92V; I103L; D113E; T115S; H166Q	1,86
127/128	V25R; A66Q; T92V; I103L; T115E; H166Q	1,54
45	¹ Condición B del Ensayo HTP (ver también Ejemplo 4): Las células cultivadas en placas de 96 pocillos se lisaron con 100 µl de tampón de lisis (1 mg/ml de lisozima, 0,5 mg/ml de sulfato de polimixina B (PMBS) y 50 mM de tampón fosfato, pH 6,3). Las condiciones de reacción para una reacción de 200 µl comprendieron: 10 g/L de compuesto de sustrato (2); 19 g/L de ácido α-cetoglutarico; 21 g/L de ácido ascórbico; Sal de Mohr 1,5 mM; Tampón fosfato 50 mM, pH 6,3 (pH ajustado con KOH); 100 ml de lisado crudo; y una temperatura de reacción de aprox. 25°C (temperatura ambiente) durante aproximadamente 24 h. Las placas se cubrieron con un sello permeable a O ₂ , y las placas se agitaron en un agitador de placas de titulación a la velocidad N° 2.5.	
50	² La actividad relativa a SEQ ID NO:112 se calcula como el % de conversión del producto formado en comparación con el % de conversión de la SEQ ID NO:112 correspondiente en las condiciones de reacción especificadas. el % de conversión se cuantificó dividiendo primero las áreas del pico del producto por la suma de las áreas del sustrato, producto e impurezas/producto lateral pico determinado por análisis de HPLC y luego restándolo con el % de conversión del control negativo (pico de superposición).	
55		
60		
65		

Tabla 2C: Polipéptidos diseñados y mejoras de enzimas relativas que utilizan preparaciones enzimáticas "Mini-DSP"

SEQ ID NO: (nt/aa)	Diferencias de Aminoácidos (en relación con SEQ ID NO: 2)	%Conv ¹	Relación de selectividad: (1)/(1a) ²
1/2	N/A	16,6 ³	1,6
3/4	N/A	49,6 ³	2,1
9/10	I103L	21,2 ³	5
23/24	H166Q	53,7 ³	2,5
99/100	H166Q	51,8 ³	2,4
101/102	I103L	24,5 ³	3,1
103/104	I103L; H166Q	67,3 ³	5,5
105/106	I103L; N131Y; H166Q	40 ³	5,7
107/108	T3S; I103L; H166Q	66,9 ⁴	5,2
129/130	T3S; A66Q; T92V; I103L; H166Q	67,8 ⁴	5,1
131/132	T3S; V25R; A66Q; T92V; I103L; T115E; H166Q	79,7 ⁴	5,6
133/134	T3S; V25R; A66Q; T92V; I103L; H166Q	69,9 ⁴	5,1

¹ % de conversión se cuantificó dividiendo el área del pico del producto por la suma de las áreas del compuesto de sustrato (2), el compuesto del producto (1) y el pico del compuesto del producto (1a) según lo determinado por análisis de HPLC.

² La relación de selectividad se refiere a la relación de las áreas de HPLC para el compuesto del producto (1) sobre el compuesto del producto (1a).

³ Se utilizó la condición de reacción D: la reacción se llevó a cabo en un recipiente de reacción de 2 ml. La mezcla de reacción comprendía: 10 g/l de compuesto (2); 19 g/L de ácido α-cetoglutarico; 21 g/L de ácido L-ascorbico; Sal de Mohr 1,5 mM; Tampón de fosfato de potasio 50 mM, pH 6,3; 20 g/l de proteína de preparación de polvo Mini-DSP; y una temperatura de reacción de aprox. 25°C (temperatura ambiente) durante aproximadamente 24 h. Las reacciones se agitaron a 1200 rpm y la reacción se mantuvo abierta a la atmósfera.

⁴ Se utilizó la condición de reacción E: La reacción se llevó a cabo en un matraz de fondo redondo de dos bocas y 250 ml. La mezcla inicial de 60 ml de tampón de fosfato de potasio 50 mM se cargó con 0,146 g (1,5 g/L) de sal de Mohr (hierro sulfato de amonio [II] hexahidrato), 3,46 g (35 g/L) de ácido α-cetoglutarico y 1,36 g (14 g/L) de ácido ascórbico con pH ajustado a 6,3 con 50% de KOH. La enzima se añadió a la mezcla en forma de 0,2 g (2 g/l) de mini DSP preparado previamente disuelto en 16 ml de tampón fosfato de potasio 50 mM, pH 6,3, con 1 ml de antiespumante, con agitación durante 15 minutos. El sustrato se añadió a la mezcla en forma de 2 g (20 g/L) de compuesto (2) (disuelto previamente en 24 ml de tampón de fosfato de potasio 50 mM, pH 6,3) con agitación a 25°C bajo inyección de oxígeno (caudal = 2 L/h).

Tabla 2D: Polipéptidos diseñados y mejoras de enzimas relativas utilizando preparaciones completas de DSP

SEQ ID NO: (nt/aa)	Diferencias de Aminoácidos (en relación con SEQ ID NO: 2)	Condición D ¹		Condición F ²		Condición E ⁴	
		%Conv ⁵	Relación de selectividad (1)/(1a) ⁶	%Conv ⁵	Relación de selectividad (1)/(1a) ⁶	%Conv ⁵	Relación de selectividad (1)/(1a) ⁶
1/2	N/A	83,5	5,7	46,4	1,9	-	-
103/104	I103L; H166Q	86,7	8,1	71,3	4,7	-	-
107/108	T3S; I103L; H166Q	-	-	74,7	6	-	-
109/110	A26T; I103L; H166Q	-	-	73,5	4,6	-	-
131/132	T3S; V25R; A66Q; T92V; I103L; T115E; H166Q	-	-	-	-	66,9	6,5:1

¹ Se utilizó una condición de reacción D: la reacción se llevó a cabo en un recipiente de reacción de 2 ml. La mezcla de reacción comprendía: 10 g/l de compuesto de sustrato (2); 19 g/L de ácido α-cetoglutarico; 21 g/L de ácido L-ascórbico; Sal de Mohr 1,5 mM; Tampón de fosfato de potasio 50 mM, pH 6,3; 20 g/L polvo de enzima DSP completa; y una temperatura de reacción de aproximadamente 25°C durante aproximadamente 24 h. Las reacciones se agitaron a 1200 rpm, y el vial se mantuvo abierto a la atmósfera.

(Continuado)

SEQ ID NO: (nt/aa)	Diferencias de Aminoácidos (en relación con SEQ ID NO: 2)	Condición D ¹		Condición F ²		Condición E ⁴	
		%Conv ⁵	Relación de selectividad (1)/ (1a) ⁶	%Conv ⁵	Relación de selectividad (1)/ (1a) ⁶	%Conv ⁵	Relación de selectividad (1)/ (1a) ⁶
² Se utilizó la condición de reacción F: la reacción se llevó a cabo en un recipiente de reacción de 1 ml. La mezcla de reacción comprende 20 g/l de compuesto de sustrato (2); 38 g/L de ácido α-cetoglutarico; 21 g/L de ácido L-ascórbico; sal de Mohr 1,5 mM; Tampón de fosfato de potasio 50 mM, pH 6,3; 10 g/L polvo de enzima DSP completa; y una temperatura de reacción de aproximadamente 25°C durante aproximadamente 24 h. Las reacciones se agitaron a 1200 rpm y el vial se mantuvo abierto a la atmósfera. ⁴ Se usó la condición de reacción E de DSP: la reacción se llevó a cabo en un matraz de fondo redondo de dos cuellos de 250 ml. La mezcla inicial de 60 ml de tampón de fosfato de potasio 50 mM se cargó con 0,146 g (1,5 g/L) de sal de Mohr (hierro sulfato de amonio [II] hexahidrato), 3,46 g (35 g/L) de ácido α-cetoglutarico y 1,36 g (14 g/L) de ácido L-ascórbico con pH ajustado a 6,3 con KOH al 50%. La enzima se añadió a la mezcla en forma de 0,2 g (2 g/l) de polvo de DSP completo previamente disuelto en 16 ml de tampón de fosfato de potasio 50 mM, pH 6,3, con 1 ml de antiespumante, con agitación durante 15 minutos. El sustrato se añadió a la mezcla en forma de 2 g (20 g/L) de compuesto (2) (disuelto previamente en 24 ml de tampón fosfato 50 mM, pH 6,3) con agitación a 25°C bajo inyección de oxígeno (velocidad de flujo = 2 L/h). ⁵ El % de conversión se cuantificó dividiendo el área del pico del producto por la suma de las áreas del sustrato, el compuesto del producto (1) y el pico del compuesto del producto (1a) según lo determinado por el análisis de HPLC. ⁶ La relación de selectividad se refiere a la proporción de áreas de HPLC para el compuesto del producto (1) sobre el compuesto del producto (1a).							

Tabla 2E: Polipéptidos diseñados y mejoras de enzimas relativas utilizando preparaciones de SFP

SEQ ID NO: (nt/aa)	Diferencias de aminoácidos (en relación con SEQ ID NO: 2)	Actividad ¹	%Conv ⁴	Relación de selectividad (1)/ (1a) ⁴
111/112	I103L; H166Q	0,19 ²	3,9	3,1
125/126	A66Q; T92V; I103L; D113E; T115S; H166Q	1,8 ³	11,1	3,1
127/128	V25R; A66Q; T92V; I103L; T115E; H166Q	1,8 ¹	5,9	3,5
¹ Se utilizó la condición C de reacción SFP: la reacción se llevó a cabo en un recipiente de reacción de 200 µl. La mezcla de reacción comprendía: 30 g/l de compuesto de sustrato (2); 52,5 g/L de ácido α-cetoglutarico; 21 g/L de ácido L-ascórbico; Sal de Mohr 2,25 mM; tampón de fosfato de potasio 50 mM, pH 6,3 (pH ajustado con KOH); 5 g/L de proteína de preparación de polvo de enzima SFP; y una temperatura de reacción de aprox. 25°C durante 24 h. Las placas se cubrieron con un sello permeable a O₂ y se agitaron en un agitador de placa de títulos a la velocidad N° 2.5. ² La actividad relativa a SEQ ID NO:108 se calcula como el % de conversión del producto formado en comparación con el % de conversión de la SEQ ID NO:108 correspondiente en las condiciones de reacción especificadas. El % de conversión se cuantificó dividiendo primero las áreas del pico del producto por la suma de las áreas del sustrato, producto e impurezas/producto lateral pico determinado por análisis de HPLC y luego restándolo con el % de conversión del control negativo (pico de superposición)). ³ La actividad relativa a SEQ ID NO:124 se calcula como el % de conversión del producto formado por % de conversión de la SEQ ID NO:124 correspondiente en las condiciones de reacción especificadas. El % de conversión se cuantificó dividiendo primero las áreas del pico del producto por la suma de las áreas del sustrato, el producto y el pico de impurezas/producto secundario según lo determinado por HPLC. ⁴ Las condiciones de reacción utilizadas para el porcentaje de conv y porcentaje de selectividad diferían de las condiciones de reacción de SFP anteriores como se indica a continuación. Reacción en un recipiente de 100 ml. La mezcla de reacción comprendía: 20 g/l de compuesto de sustrato (2); 2 g/L de preparación de enzima en polvo SFP 14 g/L de ácido ascórbico; 35 g/L de ácido cetoglutarico; 1,5 g/L de sal de Mohr; antiespumante al 1% (v/v); y oxígeno a 2 l/h.				

Tabla 2F: Polipéptidos diseñados y mejoras en la enzima relativa usando preparaciones de HTP

SEQ ID NO: (nt/aa)	Diferencias de Aminoácidos (en relación con SEQ ID NO: 2)	Actividad de pliegue mejorado ¹	Actividad de pliegue ² (+ácido ascórbico preinc.) ²	Selectividad de pliegue mejorado ³
137/138	V25R; A66Q; T92V; I103L; E114N; T115E; H166Q;	1,5	n,d,	1
139/140	V25R; A66Q; T92V; I103L; T115E; M140L; H166Q;	1,7	n,d,	1,1
141/142	V25R; A66Q; T92V; S98T; I103L; T115E; H166Q;	4,1	n,d,	1,1
143/144	V25R; A42L; A66Q; T92V; I103L; T115E; H166Q;	0,8	n,d,	1,3
145/146	A17V; V25R; A26W; A66Q; T92V; I103L; T115E; H166Q;	3,8	n,d,	0,8
147/148	V25R; A62Q; A66Q; T92V; I103L; T115E; H166Q;	1,5	n,d,	0,9
149/150	V25R; A26R; A66Q; T92V; I103L; T115E; H166Q;	1,7	n,d,	0,9
151/152	V25R; A66Q; T92V; I103L; T115E; M151A; H166Q;	1,4	n,d,	1
153/154	V25R; A66Q; A88R; T92V; I103L; T115E; H166Q;	1,4	n,d,	0,9
155/156	V25R; A66Q; T92V; I103L; T115E; H166Q; V188G;	4,4	n,d,	0,9
157/158	E24R; V25R; A66Q; T92V; I103L; T115E; H166Q;	1,5	n,d,	0,9
159/160	E24S; V25R; A66Q; T92V; I103L; T115E; H166Q;	1,6	n,d,	0,9
161/162	V25R; A66Q; T92V; S98F; I103L; T115E; H166Q;	0,9	n,d,	1,3
163/164	T3S; V25R; A42L; A66Q; T92V; I103L; T115E; H166Q;	+	n,d,	+
165/166	T3S; V25R; A42L; A66Q; T92V; S98T; I103L; T115E; H166Q;	+	n,d,	+
167/168	T3S; V25R; A42L; A66Q; T92V; S98T; I103L; T115E; H166Q; V188G;	+	n,d,	+
169/170	E24S; V25R; A42L; A66Q; T92V; I103L; T115E; H166Q; V188G;	1,43	1,66	4,46
171/172	E24S; V25R; A42L; A62Q; A66Q; T92V; S98T; I103L; E114N; T115E; M140L; H166Q;	1,60	1,54	4,30
173/174	E24S; V25R; A42L; A66Q; T92V; S98T; I103L; E114N; T115E; H166Q;	1,77	2,41	5,92
175/176	E24S; V25R; A26R; A42L; A66Q; T92V; I103L; T115E; H166Q;	1,73	1,70	4,50
177/178	V25R; A42L; A62Q; A66Q; T92V; I103L; T115E; M140L; H166Q; V188G;	1,41	1,62	4,98
179/180	E24S; V25R; A26R; A42L; A66Q; A88R; T92V; I103L; T115E; M140L; H166Q;	2,59	1,81	3,94
181/182	E24S; V25R; A26R; A42L; A66Q; A88R; T92V; S98T; I103L; E114N; T115E; M140L; H166Q; V188G;	1,68	1,41	4,31
183/184	E24S; V25R; A42L; A62Q; A66Q; T92V; I103L; E114N; T115E; M140L; H166Q;	1,46	3,13	4,84

(Continuado)

SEQ ID NO: (nt/aa)	Diferencias de Aminoácidos (en relación con SEQ ID NO: 2)	Actividad de pliegue mejorado ¹	Actividad de pliegue ² (+ácido ascórbico preinc.) ²	Selectividad de pliegue mejorado ³
185/186	E24S; V25R; A26R; A42L; A66Q; A88R; T92V; I103L; E114N; T115E; H166Q;	1,71	2,08	4,17
187/188	E24S; V25R; A42L; A66Q; T92V; I103L; T115E; M140L; H166Q;	1,68	2,11	4,47
189/190	E24S; V25R; A42L; A66Q; T92V; S98T; I103L; E114N; T115E; H166Q;	+	1,58	4,21
191/192	V25R; A26R; A42L; A66Q; T92V; I103L; E114N; T115E; M151H; H166Q;	1,87	1,53	3,77
193/194	E24S; V25R; A42L; A62Q; A66Q; T92V; S98T; I103L; T115E; M140L; M151H; H166Q;	2,19	1,53	4,08
195/196	V25R; A42L; A66Q; T92V; I103L; E114N; T115E; H166Q;	1,62	1,87	5,45
197/198	V25R; A42L; A62Q; A66Q; T92V; I103L; T115E; M140L; M151H; H166Q; I205V;	+	1,73	3,74
199/200	E24S; V25R; A42L; A66Q; A88R; T92V; I103L; E114N; T115E; H166Q;	1,51	2,29	4,32
201/202	E24S; V25R; A26R; A42L; A62Q; A66Q; T92V; I103L; T115E; M151H; H166Q;	1,16	1,85	5,34
203/204	E24S; V25R; A26R; A42L; A66Q; A88R; T92V; I103L; E114N; T115E; H166Q;	2,02	2,01	4,11
205/206	E24S; V25R; A26R; A42L; A62Q; A66Q; T92V; I103L; T115E; M140L; M151H; H166Q;	+	1,58	3,76
207/208	E24S; V25R; A26R; A42L; A66Q; A88R; T92V; I103L; E114N; T115E; H166Q; Q186G;	+	1,50	3,59
209/210	E24S; V25R; A26W; A42L; A66Q; A88R; T92V; I103L; T115E; M140L; H166Q;	1,23	1,50	3,64
211/212	E24S; V25R; A26R; A42L; A62Q; A66Q; A88R; T92V; I103L; E114N; T115E; M140L; M151H; H166Q; Q186G;	+	2,28	3,95
213/214	E24S; V25R; A42L; A66Q; T92V; S98F; I103L; T115E; M140L; H166Q; V188G;	+	1,51	6,2
215/216	E24S; V25R; A26R; A42L; A62Q; A66Q; A88R; T92V; S98F; I103L; E114S; T115E; M140L; M151H; H166Q;	+	1,32	6,3
217/218	E24S; V25R; A26R; A42L; A66Q; T92V; S98F; I103L; T115E; M140L; M151H; H166Q;	+	1,33	6,5
219/220	E24S; V25R; A42L; A66Q; T92V; S98F; I103L; T115E; H166Q;	+	1,53	8,8

(Continuado)

SEQ ID NO: (nt/aa)	Diferencias de Aminoácidos (en relación con SEQ ID NO: 2)	Actividad de pliegue mejorado ¹	Actividad de pliegue ² (+ácido ascórbico preinc.) ²	Selectividad de pliegue mejorado ³
221/222	T3S; E24S; V25R; A42L; A66Q; T92V; S98T; I103L; E114N; T115E; H166Q;	+	n.d.	+
223/224	T3S; E24S; V25R; A26R; A42L; A66Q; A88R; T92V; I103L; T115E; M140L; H166Q;	+	n.d.	+
225/226	T3S; E24S; V25R; A42L; A62Q; A66Q; T92V; I103L; E114N; T115E; M140L; H166Q;	+	n.d.	+
227/228	T3S; E24S; V25R; A26R; A42L; A66Q; A88R; T92V; I103L; T115E; M140L; H166Q;	+	n.d.	+
¹Ensayo HTP - Condiciones F Para polipéptidos de la SEQ ID NO:138-168: Las células cultivadas en placas de 96 pocillos se lisaron con 100 µl de tampón de lisis (1 mg/ml de lisozima, 0,5 mg/ml de sulfato de polimixina B (PMBS) y tampón fosfato 50 mM, pH 6,3). Las condiciones de reacción para una reacción de 200 µl comprendieron: 30 g/L de compuesto de sustrato (2); 52,5 g/L de ácido α-cetoglutarico; 21 g/L de ácido L-ascórbico; 2,25 g/L de sal de Mohr; Tampón de fosfato de potasio 50 mM, pH 6,3 (pH ajustado con KOH); 100 µl de lisado crudo; y la temperatura de reacción a aproximadamente 25°C (temperatura ambiente) durante aproximadamente 24 h. Las placas se cubrieron con un sello permeable a O ₂ y se agitaron en un agitador de placa de títulos a la velocidad N° 2,5. La actividad se calculó como el % de conversión del producto formado en comparación con el % de conversión del polipéptido de referencia correspondiente de la SEQ ID NO:128 en las condiciones de reacción especificadas. % La conversión se cuantificó dividiendo en primer lugar las áreas del pico del producto por la suma de las áreas del sustrato, producto e impurezas/pico del producto secundario según lo determinó el análisis de HPLC y luego restándolo con el % de conversión del control negativo (pico de superposición). Para polipéptidos de la SEQ ID NO:170-228: Al igual que para la SEQ ID NO:138-168, excepto la temperatura de reacción a 30°C. La actividad se calculó como el % de conversión del producto formado en comparación con el % de conversión del polipéptido de referencia correspondiente de la SEQ ID NO:144 en las condiciones de reacción especificadas. el % de conversión se cuantificó dividiendo primero las áreas del pico del producto por la suma de las áreas del sustrato, producto e impurezas/producto lateral pico determinado por análisis de HPLC y luego restándolo con el % de conversión del control negativo (pico de superposición). ²Condiciones G del ensayo HTP - Preincubación con ácido ascórbico Los ensayos de HTP de los polipéptidos de la SEQ ID NO:170 - 228 también se llevaron a cabo después de la preincubación de 2 horas y media del lisado en ácido ascórbico. Procedimiento de preincubación: se transfirieron 100 µl de lisado a una nueva placa de pocillos profundos, seguido de la adición de 20 µl/pocillo de solución madre de ácido ascórbico. La placa se agitó a temperatura ambiente, velocidad 2,5 durante 2 h. El ensayo de actividad de HTP se llevó a cabo utilizando las Condiciones de ensayo de HTP a 30°C para la SEQ ID NO:170-228 como se describe anteriormente. ³Selectividad Mejorada: Selectividad mejorada para el producto 2,5-regioisómero del compuesto (1) frente al producto 2,3-regioisómero del compuesto (1a) producido en los Ensayos de HTP realizados bajo la Condición F se determinó utilizando LC-MS. Preparación de la muestra de LC-MS: se pipetearon 10 µl de la mezcla de reacción del ensayo HTP y se diluyeron veinte mil veces (en tres etapas) con una mezcla 1:1 de MeCN:agua. La muestra se agitó en vórtex y se centrifugó a 12.000 rpm durante 5 minutos. El sobrenadante se transfirió a un vial de 2 µl para el análisis de LC-MS utilizando los parámetros de análisis descritos en el Ejemplo 7. "+" indica que la actividad o la selectividad mejoraron de 1 a 5 veces en relación con la SEQ ID NO:2. "n.d." indica "no determinado"				

Tabla 2G: Polipéptidos diseñados y mejoras en la enzima relativa utilizando preparaciones de SFP

SEQ ID NO: (nt/aa)	Diferencias de Aminoácidos (en relación con SEQ ID NO: 2)	Actividad de pliegue mejorado ¹	% conversión ²	Relación de selectividad (1)/(1a) ³
127/128	V25R; A66Q; T92V; I103L; T115E; H166Q	1	8,9	3,9 : 1
137/138	V25R; A66Q; T92V; I103L; E114N; T115E; H166Q;	0,9	8,2	3,2 : 1
139/140	V25R; A66Q; T92V; I103L; T115E; M140L; H166Q;	1,2	10,4	3,4 : 1
141/142	V25R; A66Q; T92V; S98T; I103L; T115E; H166Q;	1,9	16,7	3,5 : 1
143/144	V25R; A42L; A66Q; T92V; I103L; T115E; H166Q;	1,3	11,8	4,5 : 1
173/174	E24S; V25R; A42L; A66Q; T92V; S98T; I103L; E114N; T115E; H166Q;	1,8	8,9	3,9 : 1
179/180	E24S; V25R; A26R; A42L; A66Q; A88R; T92V; I103L; T115E; M140L; H166Q;	2,9	8,2	3,2 : 1
183/184	E24S; V25R; A42L; A62Q; A66Q; T92V; I103L; E114N; T115E; M140L; H166Q;	1,8	10,4	3,4 : 1
185/186	E24S; V25R; A26R; A42L; A66Q; A88R; T92V; I103L; E114N; T115E; H166Q;	1,9	16,7	3,5 : 1
187/188	E24S; V25R; A42L; A66Q; T92V; I103L; T115E; M140L; H166Q;	1,2	11,8	4,5 : 1
199/200	E24S; V25R; A42L; A66Q; A88R; T92V; I103L; E114N; T115E; H166Q;	1,7	8,9	3,9 : 1
203/204	E24S; V25R; A26R; A42L; A66Q; A88R; T92V; I103L; E114N; T115E; H166Q;	1,8	8,2	3,2 : 1
¹Ensayo de actividad SFP - Condiciones G				
Este ensayo se llevó a cabo en una escala de 5 ml con las condiciones finales de la mezcla de reacción: 10 g/L de SFP del polipéptido, 30 g/L de sustrato del compuesto (2), 52,5 g/L de ácido α -cetoglutárico, 21 g/L de ácido ascórbico ácido, 2,25 g/L de sal de Mohr en tampón fosfato 50 mM pH 6,3. Se preparó una solución madre de premezcla que consta de lo siguiente: 37,5 g/L (187,5 mg) de sustrato, 65,63 g/L (328 mg) de ácido α -cetoglutárico y 26,25 g/L (131,3 mg) de ácido ascórbico. Después de ajustar el pH, se agregaron 2,81 g/L (14 mg) de sal de Mohr y se disolvieron en 5 ml de tampón fosfato 50 mM, pH 6,3 (para cada reacción de vial) y el pH se ajustó a pH 6,3 con KOH al 50%. Se preparó una solución madre de 50 g/l de enzima SFP disolviendo 100 mg de la preparación de SFP de la variante polipeptídica deseada en 2 μ l de tampón fosfato 50 mM, pH 6,3. El ensayo se inició agregando 4 ml de la solución madre de premezcla a un vial de vidrio seguido de 1 ml de solución madre de enzima. El vial se agitó a 250 rpm, temperatura ambiente usando un agitador de placa caliente magnética durante 24 h. Se añadieron 10 ml de acetonitrilo al 75% (MeCN) en agua para apagar la reacción con agitación. Luego se llevó 1 ml de la reacción inactivada a un bloque de ensayo de 2 ml de Costar 96 pozo profundo y se centrifugó durante 10 minutos a 4.000 rpm. Se llevó a cabo un análisis de dansilación y HPLC como se describe para los ensayos de SFP en el Ejemplo 3.				
La actividad se calculó como el % de conversión del producto formado en comparación con el % de conversión del polipéptido de referencia correspondiente de la SEQ ID NO:128 en las condiciones de reacción especificadas. El % de conversión se cuantificó dividiendo primero las áreas del pico del producto por la suma de las áreas del sustrato, producto e impurezas/producto lateral pico determinado por análisis de HPLC y luego restándolo con el % de conversión del control negativo (pico de superposición).				
² El % de conversión se cuantificó dividiendo el área del pico del producto por la suma de las áreas del compuesto de sustrato (2), el compuesto del producto (1) y el pico del compuesto del producto (1a) según lo determinado por análisis de HPLC.				
³ La relación de selectividad se refiere a la proporción de áreas de HPLC para el compuesto del producto (1) sobre el compuesto del producto (1a).				

Tabla 2H: Polipéptidos diseñados y mejoras de enzimas relativas utilizando preparaciones DSP

SEQ ID NO: (nt/aa)	Diferencias de aminoácidos (en relación con SEQ ID NO: 2)	%conversión ³	Relación de selectividad (1)/(1a) ⁴
Condición H¹			
131/132	T3S; V25R; A66Q; T92V; I103L; T115E; H166Q	30,0	5,1 : 1
163/164	T3S; V25R; A42L; A66Q; T92V; I103L; T115E; H166Q;	21,9	8,6 : 1
165/166	T3S; V25R; A42L; A66Q; T92V; S98T; I103L; T115E; H166Q;	21,0	9,4 : 1
167/168	T3S; V25R; A42L; A66Q; T92V; S98T; I103L; T115E; H166Q; V188G;	18,3	9,7 : 1
221/222	T3S; E24S; V25R; A42L; A66Q; T92V; S98T; I103L; E114N; T115E; H166Q;	21,0	9:1
223/224	T3S; E24S; V25R; A26R; A42L; A66Q; A88R; T92V; I103L; T115E; M140L; H166Q;	33,0	6,7 : 1
225/226	T3S; E24S; V25R; A42L; A62Q; A66Q; T92V; I103L; E114N; T115E; M140L; H166Q;	24,9	6,5 : 1
Condición I²			
223/224	T3S; E24S; V25R; A26R; A42L; A66Q; A88R; T92V; I103L; T115E; M140L; H166Q;	51,7	6,7 : 1
227/228	T3S; E24S; V25R; A26R; A42L; A66Q; A88R; T92V; I103L; T115E; M140L; H166Q;	77,02	6,7 : 1
Ensayo ¹DSP - Condición H:			
Este ensayo de DSP se llevó a cabo en una escala de 10 ml. Se preparó una solución de sustrato del compuesto (2) (0,30 g) en 2 µl de tampón fosfato de potasio 50 mM (pH 6,3) y se añadió a una mezcla previamente agitada de la enzima polipeptídica DSP (0,05 g), antiespumante Y-30. Emulsión (100 a 300 ml), sal de Mohr (0,021 g), ácido α-cetoglutarico (0,51 g) y ácido ascórbico (0,204 g) en 7 ml de tampón de fosfato de potasio 50 mM (pH 6,3). La mezcla resultante se roció continuamente con oxígeno (2 l/h) a 25°C durante 24 h. El progreso de la reacción se monitorizó mediante HPLC como se describe en el Ejemplo 6.			
²Ensayo DSP - Condición I:			
Este ensayo de DSP se llevó a cabo en una escala de 100 ml. Se preparó una solución de sustrato del compuesto (2) (3,0 g) en 20 ml de tampón de fosfato de potasio 50 mM (pH 6,3) y se añadió a una mezcla previamente agitada de la enzima polipeptídica DSP (0,5 g), antiespumante emulsión Y-30. (1 ml), sal de Mohr (0,21 g), ácido α-cetoglutarico (5,1 g) y ácido ascórbico (2,04 g) en 70 ml de tampón de fosfato de potasio 50 mM (pH 6,3). La mezcla resultante se roció continuamente con oxígeno (2 l/h) y se agitó a 25°C durante 24 h. El progreso de la reacción se controló por HPLC como se describe en el Ejemplo 6.			
³ El % de conversión se cuantificó dividiendo el área del pico del producto por la suma de las áreas del compuesto de sustrato (2), el compuesto del producto (1) y el pico del compuesto del producto (1a) según lo determinado por análisis de HPLC.			
⁴ La relación de selectividad se refiere a la proporción de áreas de HPLC para el compuesto del producto (1) sobre el compuesto del producto (1a).			

[0130] A partir de un análisis de los polipéptidos ejemplares, las mejoras en las propiedades enzimáticas se asocian con diferencias de residuos en comparación con la SEQ ID NO:2 en las posiciones de residuos X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X186; X188; X205; X225; X230; X270; y X271. Las diferencias específicas de residuos en cada una de estas posiciones que están asociadas con las propiedades mejoradas incluyen: X2K; X2T; X3S; X4Q; X4L; X4E; X4S; X5I; X5L; X5M; X9I; X13T; X17V; X24R; X24S; X25R; X26R; X26T; X26W; X29A; X30V; X30P; X36T; X42E; X52P; X57T; X57A; X58A; X59G; X62Q; X66Q; X86S; X88R; X92V; X95M; X98F; X98T; X103L; X103Q; X112T; X112V; X113E; X114N; X115E; X115H; X115D; X115G; X115S; X115A; X116L; X121F; X131Y; X131F; X140L; X150S; X151A; X151H; X151S; X166T; X166L; X166Q; X186G; X188G; X205V; X225L; X225Y; X225W; X230V; X270E; X271K; y X271R.

[0131] Las propiedades enzimáticas específicas asociadas con las diferencias de residuos en comparación con la SEQ ID NO:2 en las posiciones de residuos anteriores incluyen, entre otras, la actividad enzimática, la regioselectividad, la expresión de polipéptidos y la tolerancia del sustrato. Las mejoras en la actividad enzimática y la

tolerancia del sustrato están asociadas con diferencias de residuos en las posiciones de los residuos X3; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X188; X225; X230; X270; y X271. Las mejoras en la regioselectividad están asociadas con diferencias de residuos en las posiciones de residuos: X3; X25; X42; X66; X92; X98; X103; X115; X131; y X166. Las mejoras en la expresión de polipéptidos están asociadas con diferencias de residuos en las posiciones de residuos: X2; X4; X5; X9; y X13. Por consiguiente, las diferencias de residuos en las posiciones de residuos anteriores se pueden usar individualmente o en varias combinaciones para producir polipéptidos de hidroxilasa de prolina diseñados que tengan las propiedades mejoradas deseadas, que incluyen, entre otras, la actividad enzimática, la regioselectividad, la estereoselectividad y la tolerancia al sustrato. Se pueden usar otras diferencias de residuos que afectan a la expresión del polipéptido para aumentar la expresión de la hidroxilasa de prolina diseñada.

[0132] A la luz de la guía proporcionada en este documento, se contempla además que cualquiera de los polipéptidos diseñados a modo de ejemplo de SEQ ID NO:8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228 pueden ser utilizados como la secuencia de aminoácidos de partida para sintetizar otros polipéptidos de hidroxilasa de prolina modificados, por ejemplo mediante rondas de evolución posteriores que incorporan nuevas combinaciones de diversas diferencias de aminoácidos de otros polipéptidos en las Tablas **2A, 2B, 2C, 2D, 2E, 2F, 2G y 2H**, y otras posiciones de residuos descritas en este documento. Se pueden generar mejoras adicionales al incluir las diferencias de aminoácidos en las posiciones de los residuos que se han mantenido sin cambios durante las rondas anteriores de evolución.

[0133] Por consiguiente, en algunas realizaciones, el polipéptido diseñado que tiene actividad hidroxilasa de prolina comprende una secuencia de aminoácidos que tiene al menos 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más de identidad de secuencia a la secuencia de referencia SEQ ID NO:2 y una o más diferencias de residuos en comparación con SEQ ID NO:2 en las posiciones de residuos seleccionadas de: X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X186; X188; X205; X225; X230; X270; y X271.

[0134] En algunas realizaciones, el polipéptido diseñado que tiene actividad hidroxilasa de prolina con propiedades mejoradas en comparación con la SEQ ID NO:2, comprende una secuencia de aminoácidos que tiene al menos 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más de identidad con una secuencia de referencia seleccionada de la SEQ ID NO:8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228, y una o más diferencias de residuos en comparación con la SEQ ID NO:2 en las posiciones de residuos seleccionadas de: X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X186; X188; X205; X225; X230; X270; y X271. En algunas realizaciones, la secuencia de aminoácidos de referencia se selecciona de la SEQ ID NO:10, 24, 104, 106, 108, 110, 132, 164, 222, 224, 226 y 228. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:10. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:24. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:104. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:110. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:132. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:164. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:222. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:224. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:226. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:228.

[0135] En algunas realizaciones, el polipéptido diseñado que tiene actividad hidroxilasa de prolina con propiedades mejoradas en comparación con la SEQ ID NO:2, comprende una secuencia de aminoácidos seleccionada de la SEQ ID NO:8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228, y que tienen una o más diferencias de residuos en comparación con la SEQ ID NO:2 en las posiciones de residuos seleccionadas de: X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X186; X188; X205; X225; X230; X270; y X271. En algunas realizaciones, la secuencia de aminoácidos se selecciona de la SEQ ID NO:10, 24, 104, 106, 108, 110, 132, 164, 222, 224, 226 y

228. En algunas realizaciones, la secuencia de aminoácidos es la SEQ. ID NO:10. En algunas realizaciones, la secuencia de aminoácidos es la SEQ ID NO:24. En algunas realizaciones, la secuencia de aminoácidos es la SEQ ID NO:104. En algunas realizaciones, la secuencia de aminoácidos es la SEQ ID NO:108. En algunas realizaciones, la secuencia de aminoácidos es la SEQ ID NO:110. En algunas realizaciones, la secuencia de aminoácidos es la SEQ ID NO:132. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:164. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:222. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:224. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:226. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:228.

[0136] En algunas realizaciones, las diferencias de residuos en las posiciones de residuos X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X186; X188; X205; X225; X230; X270; y X271 se seleccionan de X2K; X2T; X3S; X4Q; X4L; X4E; X4S; X5I; X5L; X5M; X9I; X13T; X17V; X24R; X24S; X25R; X26R; X26T; X26W; X29A; X30V; X30P; X36T; X42E; X52P; X57T; X57A; X58A; X59G; X62Q; X66Q; X86S; X88R; X92V; X95M; X98F; X98T; X103L; X103Q; X112T; X112V; X113E; X114N; X115E; X115H; X115D; X115G; X115S; X115A; X116L; X121F; X131Y; X131F; X140L; X150S; X151A; X151H; X151S; X166T; X166L; X166Q; X186G; X188G; X205V; X225L; X225Y; X225W; X230V; X270E; X271K; y X271R.

[0137] Por consiguiente, en algunas realizaciones, los polipéptidos de hidroxilasa de prolina modificados que muestran una o más de las propiedades mejoradas descritas en este documento pueden comprender una secuencia de aminoácidos que tiene la identidad de la secuencia de aminoácidos a una secuencia de referencia como se describe anteriormente, y una o más diferencias de residuos en comparación con la SEQ ID NO:2 seleccionada de: X2K; X2T; X3S; X4Q; X4L; X4E; X4S; X5I; X5L; X5M; X9I; X13T; X17V; X24R; X24S; X25R; X26R; X26T; X26W; X29A; X30V; X30P; X36T; X42E; X52P; X57T; X57A; X58A; X59G; X62Q; X66Q; X86S; X88R; X92V; X95M; X98F; X98T; X103L; X103Q; X112T; X112V; X113E; X114N; X115E; X115H; X115D; X115G; X115S; X115A; X116L; X121F; X131Y; X131F; X140L; X150S; X151A; X151H; X151S; X166T; X166L; X166Q; X186G; X188G; X205V; X225L; X225Y; X225W; X230V; X270E; X271K; y X271R.

[0138] En algunas realizaciones, la hidroxilasa de prolina modificada tiene una secuencia de aminoácidos que comprende al menos una o más diferencias de residuos en comparación con la SEQ ID NO:2 seleccionada de: X25R; X26T; X103L; X115E; X131Y/F; y X166Q.

[0139] En algunas realizaciones, el polipéptido de hidroxilasa de prolina diseñado comprende una secuencia de aminoácidos que tiene al menos una combinación de diferencias de residuos en comparación con la SEQ ID NO:2 seleccionada de: (a) X103L y X166Q; (b) X52P y X255Y; (c) X4E/L/S y X115A; (d) X25R y X58A; (e) X29A y X166T/Q/L; (f) X115H/D/G y X121F; (g) X3S, X103L y X166Q; (h) X103L, X131Y/F, y X166T/Q/L; (i) X26T, X103L y X166T/Q/L; (j) X25R, X66Q, X92V y X115E; (k) X25R, X66Q, X92V, X103L, X115E y X166Q; y (l) X3S, X25R, X66Q, X92V, X103L, X115E y X166Q.

[0140] Como apreciará el experto en la materia, en algunas realizaciones, una o una combinación de las diferencias de residuos anteriores que se seleccionan pueden mantenerse constantes (es decir, mantenerse) en las hidroxilasas de prolina como una característica central, y diferencias adicionales de residuos. En otras posiciones de residuos incorporadas en la secuencia para generar polipéptidos de hidroxilasa de prolina modificados adicionales con propiedades mejoradas. Por consiguiente, debe entenderse para cualquier hidroxilasa de prolina incorporada que contenga una o un subconjunto de las diferencias de residuos anteriores, la presente descripción se refiere a otras hidroxilasas de prolina modificadas que comprenden una o un subconjunto de las diferencias de residuos, y adicionalmente una o más diferencias de residuos en las otras posiciones de residuos descritas en este documento. A modo de ejemplo y no de limitación, una hidroxilasa de prolina diseñada que comprende una diferencia de residuos en la posición del residuo X103, puede incorporar además una o más diferencias de residuos en las otras posiciones del residuo, por ejemplo, X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X186; X188; X205; X225; X230; X270; y X271. Otro ejemplo es una hidroxilasa de prolina diseñada que comprende una diferencia de residuos en la posición del residuo X166, que puede comprender además una o más diferencias de residuos en las otras posiciones del residuo, por ejemplo, X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X186; X188; X205; X225; X230; X270; y X271.

[0141] Como se indicó anteriormente, los polipéptidos diseñados que tienen actividad de hidroxilasa de prolina también son capaces de convertir el compuesto de sustrato (2) en el compuesto de producto (1). En algunas realizaciones, el polipéptido de hidroxilasa de prolina diseñado es capaz de convertir el compuesto de sustrato (2) en el compuesto del producto (1) con al menos 1,2 veces, 1,5 veces, 2 veces, 3 veces, 4 veces, 5 veces, 10 veces o más actividad en relación con la actividad del polipéptido de referencia de la SEQ ID NO:2. En algunas realizaciones, el polipéptido de hidroxilasa de prolina diseñado para convertir el compuesto de sustrato (2) en el compuesto del producto (1) con al menos 1,2 veces, 1,5 veces La actividad 2 veces, 3 veces, 4 veces, 5 veces, 10 veces o más en relación con la actividad del polipéptido de referencia de la SEQ ID NO:2 comprende una secuencia

de aminoácidos que tiene una o más características seleccionadas de: X2K; X2T; X3S; X4Q; X4L; X4E; X4S; X5I; X5L; X5M; X9I; X13T; X17V; X24R; X24S; X25R; X26R; X26T; X26W; X29A; X30V; X30P; X36T; X42E; X52P; X57T; X57A; X58A; X59G; X62Q; X66Q; X86S; X88R; X92V; X95M; X98F; X98T; X103L; X103Q; X112T; X112V; X113E; X114N; X115E; X115H; X115D; X115G; X115S; X115A; X116L; X121F; X131Y; X131F; X140L; X150S; X151A; X151H; X151S; X166T; X166L; X166Q; X186G; X188G; X205V; X225L; X225Y; X225W; X230V; X270E; X271K; y X271R.

[0142] En algunas realizaciones, el polipéptido de hidroxilasa de prolina modificado es capaz de convertir el compuesto de sustrato (2) en el compuesto de producto (1) con al menos 1,2 veces la actividad con respecto a la SEQ ID NO:2 y comprende una secuencia de aminoácidos seleccionada de: SEQ ID NO:8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 210, 212, 214, 216, 218, 220, 222, 224, 226, y 228.

[0143] En algunas realizaciones, el polipéptido de hidroxilasa de prolina diseñado es capaz de convertir el compuesto de sustrato (2) en el compuesto de producto (1) con al menos 2 veces la actividad con respecto a la SEQ ID NO:2 y comprende una secuencia de aminoácidos que tiene una o más diferencias de residuos seleccionadas de: X3S; X30P; X86S; X103L; X103Q; X113E; X115E; X150S; X166Q; X151S; X225L; y 270E.

[0144] En algunas realizaciones, el polipéptido de hidroxilasa de prolina diseñado para convertir el compuesto de sustrato (2) en el compuesto de producto (1) con al menos 2 veces la actividad con respecto a la SEQ ID NO:2 comprende una secuencia de aminoácidos seleccionada de: SEQ ID NO:10, 12, 18, 24, 28, 66, 68, 70, 72, 76, 98, 100, 102, 104, 106, 108, 110; 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228.

[0145] En algunas realizaciones, el polipéptido de hidroxilasa de prolina modificado es capaz de convertir al menos el 50% o más, el 60% o más, el 70% o más, el 89% o más, el 90% o más, el 91% o más, el 92% o más, 93% o más, 94% o más, o 95% o más del compuesto (2) al compuesto (1) en 120 horas o menos, 72 horas o menos, 48 horas o menos, o 24 o menos, a una carga de sustrato de aproximadamente 100 g/L, aproximadamente 50 g/L, o aproximadamente 20 g/L en condiciones de ensayo HTP, en condiciones de ensayo SFP, o condiciones de ensayo DSP. En algunas realizaciones, el polipéptido de hidroxilasa de prolina diseñado es capaz de convertir al menos el 50% o más del compuesto (2) en el compuesto (1) en 24 horas o menos a una carga de sustrato de aproximadamente 20 g/L en condiciones de ensayo DSP a aproximadamente 25°C.

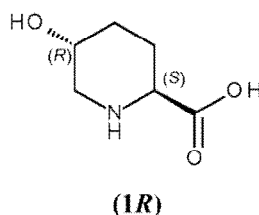
[0146] En algunas realizaciones, el polipéptido de hidroxilasa de prolina diseñado es capaz de convertir el compuesto de sustrato (2) en el compuesto de producto (1) en exceso del compuesto (1a). Las diferencias de residuos identificadas en las hidroxilasas de prolina modificadas a modo de ejemplo de las **Tablas 2A, 2B, 2C, 2D, 2E, 2F, 2G y 2H** se muestran para mantener o aumentar la regioselectividad para el compuesto (1) sobre el compuesto (1a) en la reacción de conversión. En algunas realizaciones, los polipéptidos de hidroxilasa de prolina modificados son capaces de convertir el compuesto de sustrato (2) en el compuesto del producto (1) en exceso del compuesto (1a), donde la proporción del compuesto (1) formado sobre el compuesto (1a) es al menos una relación de 1,5, 2, 3, 4, 5 o 6 o más, en particular bajo las condiciones del ensayo HTP, el ensayo SFP o el ensayo DSP.

[0147] En algunas realizaciones, el polipéptido de hidroxilasa de prolina diseñado para convertir el compuesto de sustrato (2) en el compuesto del producto (1) en exceso del compuesto (1a) en al menos una relación de 2 o más comprende una secuencia de aminoácidos con al menos una o más de las siguientes funciones: X103L; X115E; X166Q y X131Y. En algunas realizaciones, la hidroxilasa de prolina modificada es capaz de convertir el compuesto de sustrato (2) en un compuesto de producto (1) en exceso del compuesto (1a) en al menos una relación de 2 o más comprende una secuencia de aminoácidos seleccionada de la SEQ ID NO:10, 24, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228.

[0148] En algunas realizaciones, el polipéptido de hidroxilasa de prolina diseñado para convertir el compuesto de sustrato (2) en el compuesto del producto (1) en exceso del compuesto (1a) en al menos una relación de 4 o más comprende una secuencia de aminoácidos. Al menos las características X103L y X166Q. En algunas realizaciones, la hidroxilasa de prolina modificada es capaz de convertir el compuesto de sustrato (2) en el compuesto del producto (1) en exceso del compuesto (1a) en al menos una relación de 4 o mayor comprende una secuencia de aminoácidos seleccionada de la SEQ ID NO:104, 106, 108, 110, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226, y 228.

[0149] En algunas realizaciones, la hidroxilasa de prolina modificada es capaz de convertir el compuesto de sustrato (2) en el compuesto del producto (1) en exceso estereomérico del compuesto (1R), ácido (2S, 5R)-5-

hidroxipiperidina-2-carboxílico,



[0150] La enzima de tipo salvaje se caracteriza por su capacidad para convertir el ácido (2S)-piperidina-2-carboxílico en ácido (2S, 5S)-5-hidroxipiperidina-2-carboxílico, con poco o nada del producto trans hidroxí. (1R). Como se muestra en el presente documento, las diferencias de residuos en los polipéptidos de hidroxilasa de prolina modificados ejemplares de las **Tablas 2A, 2B, 2C, 2D, 2E, 2F, 2G y 2H** mantienen la alta diastereoselectividad, incluidos aquellos polipéptidos con cambios no conservadores en la secuencia de aminoácidos. En algunas realizaciones, la hidroxilasa de prolina modificada capaz de convertir el compuesto de sustrato (2) en el compuesto del producto (1) en exceso diastereomérico del compuesto (1R) comprende una secuencia de aminoácidos que tiene una o más características seleccionadas de: X2K; X2T; X3S; X4Q; X4L; X4E; X4S; X5I; X5L; X5M; X9I; X13T; X17V; X24R; X24S; X25R; X26R; X26T; X26W; X29A; X30V; X30P; X36T; X42E; X52P; X57T; X57A; X58A; X59G; X62Q; X66Q; X86S; X88R; X92V; X95M; X98F; X98T; X103L; X103Q; X112T; X112V; X113E; X114N; X115E; X115H; X115D; X115G; X115S; X115A; X116L; X121F; X131Y; X131F; X140L; X150S; X151A; X151H; X151S; X166T; X166L; X166Q; X186G; X188G; X205V; X225L; X225Y; X225W; X230V; X270E; X271K; y X271R. En algunas realizaciones, el producto compuesto (1) se forma en al menos un 90%, 95%, 96%, 97%, 98%, 99% o un exceso diastereomérico mayor del compuesto (1R). En algunas realizaciones, no se forma una cantidad detectable de producto trans hidroxí (1R) por los polipéptidos diseñados bajo condiciones de reacción adecuadas.

[0151] En algunas realizaciones, la hidroxilasa de prolina modificada tiene una secuencia de aminoácidos que comprende una o más diferencias de residuos en comparación con la SEQ ID NO:2 que aumentan la expresión de la actividad de hidroxilasa de prolina modificada en una célula huésped bacteriana, particularmente en E. coli. En algunas realizaciones, la secuencia de aminoácidos que muestra una expresión incrementada en una célula huésped bacteriana comprende una o más diferencias de residuos seleccionadas de: X2K; X2T; X4Q; X4L; X4E; X4S; X5I; X5L; X5M; X9I; y X13T.

[0152] En algunas realizaciones, el polipéptido de hidroxilasa de prolina modificado con propiedades mejoradas en la versión del compuesto (2) al compuesto (1) tiene una secuencia de aminoácidos que comprende una secuencia seleccionada de la SEQ ID NO:8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228.

[0153] En algunas realizaciones, el polipéptido diseñado que tiene actividad de hidroxilasa de prolina, comprende una secuencia de aminoácidos que tiene al menos 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% de identidad con una de las SEQ ID NO:8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228, y las diferencias de residuos de aminoácidos en comparación con la SEQ ID NO:2 presentes en cualquiera de la SEQ ID NO:8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228, según lo dispuesto en las **Tablas 2A, 2B, 2C, 2D, 2E, 2F, 2G y 2H**.

[0154] Además de las posiciones de residuos especificadas anteriormente, cualquiera de los polipéptidos de hidroxilasa de prolina descritos aquí puede comprender además otras diferencias de residuos en relación con la SEQ ID NO:2 en otras posiciones de residuos, es decir, posiciones de residuos distintas de X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X186; X188; X205; X225; X230; X270; y X271. Las diferencias de residuos en estas otras posiciones de residuos pueden proporcionar variaciones adicionales en la secuencia de aminoácidos sin afectar adversamente la capacidad del polipéptido para llevar a cabo la conversión de prolina en cis-4-hidroxiprolina, así como la conversión del compuesto (2) en compuesto (1). Por consiguiente, en algunas realizaciones, además de las diferencias de residuos de aminoácidos presentes en cualquiera de los

polipéptidos de hidroxilasa de prolina seleccionados seleccionados de la SEQ ID NO:8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228, la secuencia puede comprender además 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-26, 1-30, 1-35, 1-40, 1-45, o 1-50 diferencias de residuos en otras posiciones de residuos de aminoácidos en comparación con la SEQ ID NO:2. En algunas realizaciones, el número de diferencias de residuos de aminoácidos en comparación con la secuencia de referencia puede ser 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, 40, 45 o 50 posiciones de residuos. En algunas realizaciones, el número de diferencias de residuos de aminoácidos en comparación con la secuencia de referencia puede ser 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 20, 21, 22, 23, 24 o 25 posiciones de residuos. La diferencia de residuos en estas otras posiciones puede ser cambios conservativos o cambios no conservativos. En algunas realizaciones, las diferencias de residuos pueden comprender sustituciones conservativas y sustituciones no conservativas en comparación con el polipéptido de hidroxilasa de prolina natural de la SEQ ID NO:2.

[0155] En algunas realizaciones, la presente descripción también proporciona polipéptidos diseñados que comprenden un fragmento de cualquiera de los polipéptidos de hidroxilasa de prolina diseñados que retiene la actividad funcional y/o la propiedad mejorada de esa hidroxilasa de prolina. Por consiguiente, en algunas realizaciones, la presente descripción proporciona un fragmento de polipéptido capaz de convertir el compuesto (2) en el compuesto (1) en condiciones de reacción adecuadas, en donde el fragmento comprende al menos aproximadamente 80%, 90%, 95%, 96%, 97%, 98% o 99% de una secuencia de aminoácidos de longitud completa de un polipéptido de hidroxilasa de prolina de la presente descripción, tal como un polipéptido de hidroxilasa de prolina de ingeniería seleccionado de SEQ ID NO:8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228.

[0156] En algunas realizaciones, el polipéptido de hidroxilasa de prolina modificado puede tener una secuencia de aminoácidos que comprende una eliminación de uno cualquiera de los polipéptidos de hidroxilasa de prolina de descritos en este documento, tales como los polipéptidos ejemplares de la SEQ ID NO:8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228. Así, para cada una de las realizaciones de los polipéptidos de hidroxilasa de prolina modificados de la descripción, la secuencia de aminoácidos puede comprender delecciones de uno o más aminoácidos, 2 o más aminoácidos, 3 o más aminoácidos, 4 o más aminoácidos, 5 o más aminoácidos, 6 o más aminoácidos, 8 o más aminoácidos, 10 o más aminoácidos, 15 o más aminoácidos, o 20 o más aminoácidos, hasta el 10% del número total de aminoácidos, hasta el 10% del número total de aminoácidos, hasta el 20% del número total de aminoácidos, o hasta el 30% del número total de aminoácidos de los polipéptidos de hidroxilasa de prolina, donde se mantienen la actividad funcional asociada y/o las propiedades mejoradas de la hidroxilasa de prolina diseñada en el presente documento. En algunas realizaciones, las eliminaciones pueden comprender 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-15, 1-20, 1-21, 1-22, 1-23, 1-24, 1-25, 1-30, 1-35, 1-40, 1-45, o 1-50 aminoácidos residuos. En algunas realizaciones, el número de eliminaciones puede ser 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, 40, 45 o 50 residuos de aminoácidos. En algunas realizaciones, las supresiones pueden comprender delecciones 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 20, 21, 22, 23, 24, o 25 residuos de aminoácidos.

[0157] En algunas realizaciones, el polipéptido de hidroxilasa de prolina modificado genéticamente en el presente documento puede tener una secuencia de aminoácidos que comprende una inserción en comparación con uno cualquiera de los polipéptidos de hidroxilasa de prolina modificado genéticamente descritos en el presente documento, tales como los polipéptidos modificados ejemplares de la SEQ ID NO:8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228. De este modo, para cada una de las realizaciones de los polipéptidos de hidroxilasa de prolina de la descripción, las inserciones pueden comprender uno o más aminoácidos, 2 o más aminoácidos, 3 o más aminoácidos, 4 o más aminoácidos, 5 o más aminoácidos, 6 o más aminoácidos, 8 o más aminoácidos, 10 o más aminoácidos, 15 o más aminoácidos, 20 o más aminoácidos, 30 o más aminoácidos, 40 o más aminoácidos, o 50 o más aminoácidos, donde se mantiene la actividad funcional asociada y/o las propiedades mejoradas de la hidroxilasa de prolina diseñada aquí descrita. Las inserciones pueden ser en el extremo amino o carboxi, o en porciones internas del polipéptido de la hidroxilasa de prolina.

[0158] En algunas realizaciones, el polipéptido de hidroxilasa de prolina manipulado en el presente documento

puede tener una secuencia de aminoácidos que comprende una secuencia seleccionada de la SEQ ID NO:8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228, y opcionalmente una o varias (por ejemplo, hasta 3, 4, 5 o hasta 10) delecciones, inserciones y/o sustituciones de residuos de aminoácidos. En algunas realizaciones, la secuencia de aminoácidos tiene opcionalmente 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-15, 1-20, 1-21, 1-22, 1-23, 1-24, 1-25, 1-30, 1-35, 1-40, 1-45, o 1-50 eliminaciones, inserciones y/o sustituciones de residuos de aminoácidos. En algunas realizaciones, el número de secuencias de aminoácidos tiene opcionalmente 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, 40, 45 o 50 delecciones, inserciones y/o sustituciones de residuos de aminoácidos. En algunas realizaciones, la secuencia de aminoácidos tiene opcionalmente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 20, 21, 22, 23, 24 o 25 eliminaciones, inserciones y/o sustituciones de residuos de aminoácidos. En algunas realizaciones, las sustituciones pueden ser sustituciones conservativas o no conservativas.

[0159] En las realizaciones anteriores, las condiciones de reacción adecuadas para los polipéptidos diseñados pueden ser las descritas en las **Tablas 2A, 2B, 2C, 2D, 2E, 2F, 2G y 2H**. Por consiguiente, en algunas realizaciones, las condiciones de reacción adecuadas son las condiciones del ensayo HTP, que pueden comprender: 10 g/L o 20 g/L de carga de compuestos de sustrato; 19 g/L de ácido α -cetoglutarico; 21 g/L de ácido L-ascórbico; Sal de Mohr 1,5 mM; tampón de fosfato de potasio 50 mM, pH 6,3 (pH ajustado con KOH); 100 μ l de lisado crudo; y una temperatura de reacción a aproximadamente 25°C (temperatura ambiente) durante un tiempo de reacción de aproximadamente 24 h.

[0160] En algunas realizaciones, las condiciones de reacción adecuadas son las descritas para los ensayos en polvo de matraz de agitación (SFP), que pueden comprender: 30 g/L de carga de compuesto de sustrato; 52,5 g/L de ácido α -cetoglutarico; 21 g/L de ácido L-ascórbico; sal de Mohr 2,25 mM; tampón de fosfato de potasio 50 mM, pH 6,3 (pH ajustado con KOH); 5 g/L de proteína de preparación de polvo de enzima SFP; y una temperatura de reacción de aprox. 25°C (temperatura ambiente) durante un tiempo de reacción de aproximadamente 24 h.

[0161] En algunas realizaciones, las condiciones de reacción adecuadas son las descritas para los mini ensayos de polvo de proceso descendente (DSP), que comprenden: 10 g/L o 20 g/L de carga de sustrato; 19 g/L de ácido α -cetoglutarico; 21 g/L de ácido L-ascórbico; sal de Mohr 1,5 mM; tampón de fosfato de potasio 50 mM, pH 6,3; 20 g/L de proteína de preparación de polvo DSP; y una temperatura de reacción de aproximadamente 25°C (temperatura ambiente) durante aproximadamente el tiempo de reacción de aproximadamente 24 h.

[0162] La guía para el uso de estas condiciones de reacción anteriores y los polipéptidos de hidroxilasa de prolina se dan, entre otras, en las **Tablas 2A, 2B, 2C, 2D, 2E, 2F, 2G y 2H**, y los Ejemplos.

[0163] En algunas realizaciones, los polipéptidos de la descripción pueden estar en forma de polipéptidos de fusión en los que los polipéptidos diseñados se fusionan con otros polipéptidos, tales como, a modo de ejemplo y sin limitación, etiquetas de anticuerpos (por ejemplo, epítipo myc), secuencias de purificación (p. ej., etiquetas His para unirse a metales) y señales de localización celular (p. ej., señales de secreción). Por lo tanto, los polipéptidos diseñados aquí descritos se pueden usar con o sin fusiones a otros polipéptidos.

[0164] Debe entenderse que los polipéptidos descritos en el presente documento no están restringidos a los aminoácidos codificados genéticamente. Además de los aminoácidos codificados genéticamente, los polipéptidos descritos en el presente documento pueden estar compuestos, en su totalidad o en parte, de aminoácidos no codificados de origen natural y/o sintéticos. Ciertos aminoácidos no codificados comúnmente encontrados de los cuales pueden estar comprendidos los polipéptidos descritos en el presente documento incluyen, pero no se limitan a: los estereómeros D de los aminoácidos codificados genéticamente; ácido 2,3-diaminopropiónico (Dpr); ácido α -aminoisobutírico (Aib); ácido ϵ -aminohexanoico (Aha); ácido d-aminovalérico (Ava); N-metilglicina o sarcosina (MeGly o Sar); ornitina (Orn); citrulina (Cit); t-butilalanina (Bua); t-butilglicina (Bug); N-metilisoleucina (Melle); fenilglicina (Phg); ciclohexilalanina (Cha); norleucina (Nle); naftilalanina (Nal); 2-clorofenilalanina (Ocf); 3-clorofenilalanina (Mcf); 4-clorofenilalanina (Pcf); 2-fluorofenilalanina (Off); 3-fluorofenilalanina (Mff); 4-fluorofenilalanina (Pff); 2-bromofenilalanina (Obf); 3-bromofenilalanina (Mbf); 4-bromofenilalanina (Pbf); 2-metilfenilalanina (Omf); 3-metilfenilalanina (Mmf); 4-metilfenilalanina (Pmf); 2-nitrofenilalanina (Onf); 3-nitrofenilalanina (Mnf); 4-nitrofenilalanina (Pnf); 2-cianofenilalanina (Ocf); 3-cianofenilalanina (Mcf); 4-cianofenilalanina (Pcf); 2-trifluorometilfenilalanina (Otf); 3-trifluorometilfenilalanina (Mtf); 4-trifluorometilfenilalanina (Ptf); 4-aminofenilalanina (Paf); 4-yodofenilalanina (Pif); 4-aminometilfenilalanina (Pamf); 2,4-diclorofenilalanina (Opef); 3,4-diclorofenilalanina (Mpcf); 2,4-difluorofenilalanina (Opff); 3,4-difluorofenilalanina (Mpff); pirid-2-ilalanina (2pAla); pirid-3-ilalanina (3pAla); pirid-4-ilalanina (4pAla); naft-1-ilalanina (1nAla); naft-2-ilalanina (2nAla); tiazolilalanina (taAla); benzotienilalanina (bAla); tienilalanina (tAla); furilalanina (fAla); homofenilalanina (hPhe); homotirosina (hTyr); homotriptófano (hTrp); pentafluorofenilalanina (5ff); esterilcalanina (sAla); autilalanina (aAla); 3,3-difenilalanina (Dfa); ácido 3-amino-5-fenipentanoico (Afp); penicilamina (pluma); ácido 1,2,3,4-tetrahidroisoquinolina-3-carboxílico (Tic); β -2-tienilalanina (Thi); sulfóxido de metionina (Mso); N(w)-nitroarginina (Narg); homolisina (hLys); fosfinometilfenilalanina (pmPhe); Fosfoserina (pSer); fosfotreonina (pThr); ácido homoaspártico (hAsp); ácido

homoglutánico (hGlu); ácido 1-aminociclopent-(2 o 3)-en-4-carboxílico; ácido pipecólico (PA), ácido azetidina-3-carboxílico (ACA); ácido 1-aminociclopentano-3-carboxílico; alilglicina (aGly); propargilglicina (pgGly); homoalanina (hAla); norvalina (nVal); homoleucina (hLeu), homovalina (hVal); homoisoleucina (hle); homoarginina (hArg); N-acetilo lisina (AcLys); ácido 2,4-diaminobutírico (Dbu); ácido 2,3-diaminobutírico (Dab); N-metilvalina (MeVal); homocisteína (hCys); homoserina (hSer); Hidroxiprolina (Hyp) y homoprolina (hPro). Los aminoácidos no codificados adicionales de los cuales los polipéptidos descritos en el presente documento pueden estar comprendidos serán evidentes para los expertos en la técnica (véase, por ejemplo, los diversos aminoácidos proporcionados en Fasman, 1989, CRC Practical Handbook of Biochemistry and Molecular Biology, CRC). Press, Boca Raton, FL, en las páginas 3-70 y las referencias citadas en el mismo). Estos aminoácidos pueden estar en la configuración L o D.

[0165] Los expertos en la técnica reconocerán que los aminoácidos o residuos que llevan grupos protectores de la cadena lateral también pueden comprender los polipéptidos descritos en el presente documento. Los ejemplos no limitativos de dichos aminoácidos protegidos, que en este caso pertenecen a la categoría aromática, incluyen (grupos protectores enumerados entre paréntesis), pero no se limitan a: Arg (tos), Cys (metilbencilo), Cys (nitropiridinesulfenilo), Glu (δ-benciléster), Gln (xantilo), Asn (N-δ-xantilo), His (bom), His (bencilo), His (tos), Lys (fmoc), Lys (tos), Ser (O-bencilo), Thr (O-bencilo) y Tyr (O-bencilo).

[0166] Los aminoácidos no codificantes que están limitados conformacionalmente de los cuales pueden componerse los polipéptidos descritos en este documento incluyen, pero no se limitan a, N-metilaminoácidos (configuración L); ácido 1-aminociclopent-(2 o 3)-en-4-carboxílico; ácido pipecólico; ácido azetidina-3-carboxílico; homoprolina (hPro); y ácido 1-aminociclopentano-3-carboxílico.

[0167] En algunas realizaciones, los polipéptidos modificados pueden estar en diversas formas, por ejemplo, como una preparación aislada, como una enzima sustancialmente purificada, células completas transformadas con genes que codifican la enzima, y/o como extractos celulares y/o lisados de tales células. Las enzimas pueden liofilizarse, secarse por pulverización, precipitarse o estar en forma de una pasta cruda, como se explica más adelante.

[0168] En algunas realizaciones, los polipéptidos diseñados pueden proporcionarse sobre un soporte sólido, tal como una membrana, resina, portador sólido u otro material en fase sólida. Un soporte sólido puede estar compuesto de polímeros orgánicos tales como poliestireno, polietileno, polipropileno, polifluoroetileno, polietilenoxi y poliacrilamida, así como también copolímeros e injertos de los mismos. Un soporte sólido también puede ser inorgánico, como vidrio, sílice, vidrio de poro controlado (CPG), sílice de fase inversa o metal, como oro o platino. La configuración de un soporte sólido puede ser en forma de perlas, esferas, partículas, gránulos, un gel, una membrana o una superficie. Las superficies pueden ser planas, sustancialmente planas o no planas. Los soportes sólidos pueden ser porosos o no porosos, y pueden tener características de hinchamiento o no hinchamiento. Un soporte sólido puede configurarse en forma de pozo, depresión u otro contenedor, recipiente, característica o ubicación.

[0169] En algunas realizaciones, los polipéptidos diseñados que tienen actividad hidroxilasa de prolina de la presente descripción pueden inmovilizarse sobre un soporte sólido de modo que conserven su actividad mejorada, estereoselectividad y/u otras propiedades mejoradas con respecto al polipéptido de referencia de la SEQ ID NO.:2. En tales realizaciones, los polipéptidos inmovilizados pueden facilitar la conversión biocatalítica de los compuestos de sustrato de fórmula (II), (VI) u otros sustratos adecuados, al compuesto de producto de fórmula (I), (V), o productos correspondientes, respectivamente (por ejemplo, como se muestra en los **Esquemas 1, 2 y 3**), y una vez completada la reacción, se retienen fácilmente (por ejemplo, al retener las perlas en las que se inmoviliza el polipéptido) y luego se reutilizan o reciclan en reacciones posteriores. Tales procesos enzimáticos inmovilizados permiten una mayor eficiencia y reducción de costos. Por consiguiente, se contempla además que cualquiera de los métodos de uso de los polipéptidos de hidroxilasa de prolina de la presente descripción se puede llevar a cabo utilizando los mismos polipéptidos de hidroxilasa de prolina unidos o inmovilizados sobre un soporte sólido.

[0170] Los métodos de inmovilización de enzimas son bien conocidos en la técnica. Los polipéptidos diseñados pueden unirse de forma no covalente o covalente. Varios métodos para la conjugación e inmovilización de enzimas a soportes sólidos (por ejemplo, resinas, membranas, perlas, vidrio, etc.) son bien conocidos en la técnica y se describen en, por ejemplo: Yi et al., "Covalent immobilization of ω-transaminase from *Vibrio fluvialis* JS17 on chitosan beads," *Process Biochemistry* 42 (5):895-898 (mayo de 2007); Martín et al., "Characterization of free and immobilized (S)-aminotransferase for acetophenone production," *Applied Microbiology and Biotechnology* 76 (4):843-851 (septiembre de 2007); Koszelewski et al., "Immobilization of ω-transaminases by en- capsulation in a sol-gel/celite matrix", *Journal of Molecular Catalysis B: Enzymatic*, 63:39-44 (abril de 2010); Truppo et al., "Development of an Improved Immobilized CAL-B for the Enzymatic Resolution of a Key Intermediate to Odanacatib," *Organic Process Research & Development*, publicado en línea:dx.doi.org/10.1021/op200157c; Hermanson, GT, *Bioconjugate Techniques*, Segunda Edición, Academic Press (2008); Mateo et al., "Epoxy sephabeads: a novel epoxy support for stabilization of industrial enzymes via very intense multipoint covalent attachment," *Biotechnology Progress* 18 (3):629-34 (2002); y *Bioconjugation Protocols: Strategies and Methods*, In *Methods in Molecular Biology*, CM Niemeyer **E-d.**, Humana Press (2004). Los soportes sólidos útiles para inmovilizar las hidroxilasas de prolina modificadas de la presente descripción incluyen, pero no se limitan a, perlas o resinas que comprenden polimetacrilato con grupos funcionales epóxido, polimetacrilato con grupos funcionales amino epóxido, copolímero

de estireno/DVB o polimetacrilato con grupos funcionales octadecilo. Soportes sólidos ejemplares útiles para inmovilizar los polipéptidos de hidroxilasa de prolina modificados de la presente descripción incluyen, pero no se limitan a, perlas de quitosán, Eupergit C y SEPABEADS (Mitsubishi), incluidos los siguientes tipos diferentes de SEPABEAD:EC-EP, EC -HFA/S, EXA252, EXE119 y EXE120.

[0171] En algunas realizaciones, los polipéptidos descritos en el presente documento pueden proporcionarse en forma de kits. Las enzimas en los kits pueden estar presentes individualmente o como una pluralidad de enzimas. Los kits pueden incluir además reactivos para llevar a cabo las reacciones enzimáticas, sustratos para evaluar la actividad de las enzimas, así como reactivos para detectar los productos. Los kits también pueden incluir dispensadores de reactivos e instrucciones para el uso de los kits.

[0172] En algunas realizaciones, los kits de la presente descripción incluyen matrices que comprenden una pluralidad de polipéptidos de hidroxilasa de prolina diferentes en una posición direccionable diferente, en la que los diferentes polipéptidos son variantes diferentes de una secuencia de referencia, cada una con al menos una propiedad enzimática mejorada diferente. En algunas realizaciones, una pluralidad de polipéptidos inmovilizados sobre soportes sólidos se puede configurar en una matriz en varias ubicaciones, direccionables para el suministro robótico de reactivos, o mediante métodos de detección y/o instrumentos. La matriz se puede usar para ensayar una variedad de compuestos de sustrato para la conversión de los polipéptidos. Dichas matrices que comprenden una pluralidad de polipéptidos diseñados y métodos de su uso se describen, por ejemplo, en el documento WO2009/008908A2.

5.4 Polinucleótidos que codifican las hidroxilasas de prolina, los vectores de expresión y las células hospedadoras

[0173] En otro aspecto, la presente descripción proporciona polinucleótidos que codifican los polipéptidos de hidroxilasa de prolina modificados descritos en el presente documento. Los polinucleótidos pueden unirse operativamente a una o más secuencias reguladoras heterólogas que controlan la expresión génica para crear un polinucleótido recombinante capaz de expresar el polipéptido. Las construcciones de expresión que contienen un polinucleótido heterólogo que codifica la hidroxilasa de prolina modificada se pueden introducir en células huésped apropiadas para expresar el polipéptido de hidroxilasa de prolina correspondiente.

[0174] Como será evidente para el experto en la materia, la disponibilidad de una secuencia proteica y el conocimiento de los codones correspondientes a los diversos aminoácidos proporcionan una descripción de todos los polinucleótidos capaces de codificar los polipéptidos en cuestión. La degeneración del código genético, donde los mismos aminoácidos están codificados por codones alternativos o sinónimo, permite que se produzca una gran cantidad de ácidos nucleicos, todos los cuales codifican las enzimas hidroxilasas de prolina mejoradas. Por lo tanto, al tener conocimiento de una secuencia de aminoácidos particular, los expertos en la técnica podrían producir cualquier número de ácidos nucleicos diferentes simplemente modificando la secuencia de uno o más codones de una manera que no cambie la secuencia de aminoácidos de la proteína. En este sentido, la presente descripción contempla específicamente todas y cada una de las posibles variaciones de polinucleótidos que podrían realizarse codificando los polipéptidos descritos en este documento mediante la selección de combinaciones basadas en las posibles opciones de codones, y todas estas variaciones deben considerarse específicamente descritas para cualquier polipéptido descrito en este documento, incluidas las secuencias de aminoácidos presentadas en las **Tablas 2A, 2B, 2C, 2D, 2E, 2F, 2G y 2H** y se describen en la lista de secuencias como SEQ ID NO:8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228.

[0175] En diversas realizaciones, los codones se seleccionan preferiblemente para adaptarse a la célula huésped en la que se produce la proteína. Por ejemplo, los codones preferidos utilizados en bacterias se utilizan para expresar el gen en bacterias; Los codones preferidos usados en levadura se usan para la expresión en levadura; y los codones preferidos usados en mamíferos se usan para la expresión en células de mamíferos. En algunas realizaciones, no es necesario reemplazar todos los codones para optimizar el uso de codones de las hidroxilasas de prolina, ya que la secuencia natural comprenderá los codones preferidos y porque el uso de codones preferidos puede no ser necesario para todos los residuos de aminoácidos. En consecuencia, los polinucleótidos optimizados de codones que codifican las enzimas hidroxilasas de la prolina pueden contener codones preferidos en aproximadamente el 40%, 50%, 60%, 70%, 80% o más del 90% de las posiciones de los codones de la región codificante de longitud completa.

[0176] En algunas realizaciones, el polinucleótido comprende una secuencia de nucleótidos optimizada de codón que codifica la secuencia de aminoácidos del polipéptido de hidroxilasa de prolina que se produce de forma natural, como se representa por la SEQ ID NO:2. En algunas realizaciones, el polinucleótido tiene una secuencia de ácido nucleico que comprende al menos 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más de identidad. para las secuencias de ácido nucleico optimizadas en codón de SEQ ID NO:1, 3 o 5, cada una de las

cuales codifica las secuencias polipeptídicas idénticas de SEQ ID NO:2, 4 o 6, respectivamente. Las secuencias optimizadas de codón de SEQ ID NO:1, 3 o 5 mejoran la expresión de la hidroxilasa de prolina de tipo salvaje codificada, proporcionando preparaciones de enzima capaces de convertir *in vitro* más del 80% del compuesto (2) al compuesto (1) bajo condiciones de ensayo mini DSP, y conversión de más del 45% del compuesto (2) al compuesto (1) en condiciones de ensayo DSP. En algunas realizaciones, la secuencia de polinucleótido optimizada en codón puede mejorar la expresión de la hidroxilasa de prolina en al menos 1,2 veces, 1,5 veces o 2 veces o más en comparación con la secuencia de polinucleótido natural de *Sinorhizobium meliloti*, que se describe en este documento como SEQ ID NO:135.

[0177] En algunas realizaciones, los polinucleótidos son capaces de hibridar en condiciones altamente rigurosas con una secuencia de referencia de SEQ ID NO:1, 3 o 5, o un complemento de los mismos, y codifica un polipéptido que tiene actividad de hidroxilasa de prolina.

[0178] En algunas realizaciones, como se describió anteriormente, el polinucleótido codifica un polipéptido diseñado que tiene actividad de hidroxilasa de prolina con propiedades mejoradas en comparación con la SEQ ID NO:2, donde el polipéptido comprende una secuencia de aminoácidos que tiene al menos el 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más de identidad con una secuencia de referencia seleccionada de SEQ ID NO:8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228, y una o más diferencias de residuos en comparación con la SEQ ID NO:2 seleccionadas de: X2K; X2T; X3S; X4Q; X4L; X4E; X4S; X5I; X5L; X5M; X9I; X13T; X17V; X24R; X24S; X25R; X26R; X26T; X26W; X29A; X30V; X30P; X36T; X42E; X52P; X57T; X57A; X58A; X59G; X62Q; X66Q; X86S; X88R; X92V; X95M; X98F; X98T; X103L; X103Q; X112T; X112V; X113E; X114N; X115E; X115H; X115D; X115G; X115S; X115A; X116L; X121F; X131Y; X131F; X140L; X150S; X151A; X151H; X151S; X166T; X166L; X166Q; X186G; X188G; X205V; X225L; X225Y; X225W; X230V; X270E; X271K; y X271R. En algunas realizaciones, la secuencia de aminoácidos de referencia se selecciona de la SEQ ID NO:10, 24, 104, 106, 108, 110, 132, 164, 222, 224, 226 y 228. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:10. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:24. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:104. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:108. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:110. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:132. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:164. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:222. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:224. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:226. En algunas realizaciones, la secuencia de aminoácidos de referencia es la SEQ ID NO:228.

[0179] En algunas realizaciones, el polinucleótido codifica un polipéptido de hidroxilasa de prolina capaz de convertir el compuesto de sustrato (2) en el compuesto de producto (1) con propiedades mejoradas en comparación con la SEQ ID NO:2, en donde el polipéptido comprende una secuencia de aminoácidos que tiene al menos 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más la identidad de la secuencia a la secuencia de referencia SEQ ID NO: 2 y una o más diferencias de residuos en comparación con la SEQ ID NO: 2 en las posiciones de residuos seleccionadas de: X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X186; X188; X205; X225; X230; X270; y X271.

[0180] En algunas realizaciones, el polinucleótido codifica un polipéptido de hidroxilasa de prolina capaz de convertir el compuesto de sustrato (2) en el compuesto de producto (1) con propiedades mejoradas en comparación con la SEQ ID NO: 2, en donde el polipéptido comprende una secuencia de aminoácidos que tiene en al menos 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más la identidad de la secuencia a la secuencia de referencia SEQ ID NO: 2, y al menos una combinación de diferencias de residuos en comparación con la SEQ ID NO: 2 seleccionada de: (a) X103L y X166Q; (b) X52P y X255Y; (c) X4E/L/S y X115A; (d) X25R y X58A; (e) X29A y X166T/Q/L; (f) X115H/D/G y X121F; (g) X3S, X103L y X166Q; (h) X103L, X131Y/F, y X166T/Q/L; (i) X26T, X103L y X166T/Q/L; (j) X25R, X66Q, X92V y X115E; (k) X25R, X66Q, X92V, X103L, X115E y X166Q; y (l) X3S, X25R, X66Q, X92V, X103L, X115E y X166Q.

[0181] En algunas realizaciones, el polinucleótido codifica un polipéptido de hidroxilasa de prolina capaz de convertir el compuesto de sustrato (2) en el compuesto de producto (1) con propiedades enzimáticas mejoradas en comparación con el polipéptido de referencia de la SEQ ID NO: 2, en donde el polipéptido comprende una secuencia de aminoácidos que tiene al menos 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% de identidad con un polipéptido de referencia seleccionado de una cualquiera de las SEQ ID NO: 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172,

174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228, con la condición de que la secuencia de aminoácidos comprende cualquiera de los conjuntos de diferencias de residuos en comparación con la SEQ ID NO: 2 contenida en una cualquiera de las secuencias polipeptídicas de la SEQ ID NO: 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228, como se indica en las **Tablas 2A, 2B, 2C, 2D, 2E, 2F, 2G y 2H.**

[0182] En algunas realizaciones, el polinucleótido que codifica la hidroxilasa de prolina diseñada comprende una secuencia de polinucleótido seleccionada de la SEQ ID NO: 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177, 179, 181, 183, 185, 187, 189, 191, 193, 195, 197, 199, 201, 203, 205, 207, 209, 211, 213, 215, 217, 219, 221, 223, 225 y 227.

[0183] En algunas realizaciones, los polinucleótidos son capaces de hibridar en condiciones altamente rigurosas con una secuencia de polinucleótidos de referencia seleccionada de la SEQ ID NO: 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177, 179, 181, 183, 185, 187, 189, 191, 193, 195, 197, 199, 201, 203, 205, 207, 209, 211, 213, 215, 217, 219, 221, 223, 225 y 227. o un complemento del mismo, y codifica un polipéptido que tiene actividad hidroxilasa de prolina con una o más de las propiedades mejoradas descritas en el presente documento. En algunas realizaciones, el polinucleótido capaz de hibridar en condiciones altamente rigurosas codifica un polipéptido de hidroxilasa de prolina que tiene una secuencia de aminoácidos que comprende una o más diferencias de residuos en comparación con la SEQ ID NO: 2 en las posiciones de residuos seleccionadas de: X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X186; X188; X205; X225; X230; X270; y X271. En algunas realizaciones, las diferencias de residuos en las posiciones de residuos X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X186; X188; X205; X225; X230; X270; y X271 se seleccionan de X2K; X2T; X3S; X4Q; X4L; X4E; X4S; X5I; X5L; X5M; X9I; X13T; X17V; X24R; X24S; X25R; X26R; X26T; X26W; X29A; X30V; X30P; X36T; X42E; X52P; X57T; X57A; X58A; X59G; X62Q; X66Q; X86S; X88R; X92V; X95M; X98F; X98T; X103L; X103Q; X112T; X112V; X113E; X114N; X115E; X115H; X115D; X115G; X115S; X115A; X116L; X121F; X131F; X140L; X150S; X151A; X151H; X151S; X166T; X166L; X166Q; X186G; X188G; X205V; X225L; X225Y; X225W; X230V; X270E; X271K; y X271R.

[0184] En algunas realizaciones, los polinucleótidos codifican los polipéptidos descritos en este documento pero tienen al menos aproximadamente el 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% o más de identidad de secuencia a nivel de nucleótido a un polinucleótido de referencia que codifica la hidroxilasa de prolina modificada. En algunas realizaciones, la secuencia de polinucleótidos de referencia se selecciona de la SEQ ID NO: 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177, 179, 181, 183, 185, 187, 189, 191, 193, 195, 197, 199, 201, 203, 205, 207, 209, 211, 213, 215, 217, 219, 221, 223, 225 y 227.

[0185] Un polinucleótido aislado que codifica un polipéptido de hidroxilasa de prolina mejorado puede manipularse en una variedad de formas para proporcionar la expresión del polipéptido. En algunas realizaciones, los polinucleótidos que codifican los polipéptidos pueden proporcionarse como vectores de expresión en los que está presente una o más secuencias de control para regular la expresión de los polinucleótidos y/o polipéptidos. La manipulación del polinucleótido aislado antes de su inserción en un vector puede ser deseable o necesaria dependiendo del vector de expresión. Las técnicas para modificar polinucleótidos y secuencias de ácido nucleico que utilizan métodos de ADN recombinante son bien conocidas en la técnica. Se proporciona orientación en Sambrook et al., 2001, *Molecular Cloning: A Laboratory Manual*, 3^a Ed., Cold Spring Harbor Laboratory Press; y *Current Protocols in Molecular Biology*, Ausubel. F. Ed., Greene Pub. Asociados, 1998, actualizaciones a 2006.

[0186] En algunas realizaciones, la secuencia de control incluye, entre otras, un promotor, una secuencia líder, una secuencia de poliadenilación, una secuencia de propéptidos, una secuencia de péptidos señal y un terminador de la transcripción. Los promotores adecuados pueden seleccionarse basándose en las células hospedadoras utilizadas. Para las células hospedadoras bacterianas, los promotores adecuados para dirigir la transcripción de los constructos de ácido nucleico de la presente descripción incluyen los promotores obtenidos del operón lac de *E. coli*, el gen de la agarasa de *Streptomyces coelicolor* (dagA), el gen de *Bacillus subtilis* levansacarasa (sacB), *Bacillus licheniformis* gen de alfa-amilasa (amyL), gen de la amilasa maltogénica de *Bacillus stearothermophilus* (amyM), gen de la alfa-

amilasa de *Bacillus amyloliquefaciens* (amyQ), gen de penicilinas *Bacillus licheniformis* (penP), genes xylA y xylB *Bacillus subtilis*, y gen de betalactamasa procariótica (Villa-Kamaroff et al., 1978, Proc. Natl Acad. SCI. EE.UU. 75: 3727-3731), así como el promotor *tac* (DeBoer et al., 1983, Proc. Natl Acad. SCI. EE.UU. 80: 21-25). Promotores ejemplares para células huésped fúngicas filamentosas, incluyen promotores obtenidos a partir de los genes para amilasa TAKA *Aspergillus oryzae*, proteinasa aspártica *Rhizomucor miehei*, alfa-amilasa neutra *Aspergillus niger*, alfa-amilasa estable de ácido *Aspergillus niger*, glucoamilasa *Aspergillus niger* o *Aspergillus awamori* (glaA), lipasa *Rhizomucor miehei*, proteasa alcalina *Aspergillus oryzae*, isomerasa de fosfato de triosa *Aspergillus oryzae*, acetamidasa *Aspergillus nidulans*, y proteasa de tipo tripsina *Fusarium oxysporum* (WO 96/00787), así como el promotor NA2-tpi (un híbrido de los promotores de los genes de alfa-amilasa neutra *Aspergillus niger* y isomerasa de fosfato de triosa *Aspergillus oryzae*), y promotores mutantes, truncados e híbridos de los mismos. Promotores celulares de levadura ejemplares pueden ser de genes que pueden ser genes de enolasa *Saccharomyces cerevisiae* (eNO-1), galactocinasa *Saccharomyces cerevisiae* (GAL1), alcohol deshidrogenasa/gliceraldehído-3-fosfato dehidrogenasa *Saccharomyces cerevisiae* (ADH2/GAP), y 3-fosfoglicerato quinasa *Saccharomyces cerevisiae*. Otros promotores útiles para células huésped de levadura se describen en Romanos et al., 1992, Yeast 8: 423-488.

[0187] La secuencia de control también puede ser una secuencia de terminación de la transcripción adecuada, una secuencia reconocida por una célula huésped para terminar la transcripción. La secuencia de terminación está unida operativamente al extremo 3' de la secuencia de ácido nucleico que codifica el polipéptido. Cualquier terminador que sea funcional en la célula huésped de elección puede usarse en la presente invención. Por ejemplo, los terminadores de transcripción ejemplares para células bacterianas se describen en Ermolaeva et al., 2001, J. Mol. Biol. 301: 27-33. Ejemplos de terminadores de la transcripción para células huésped fúngicas filamentosas pueden obtenerse a partir de los genes de la amilasa TAKA *Aspergillus oryzae*, glucoamilasa *Aspergillus niger*, antranilato sintasa *Aspergillus nidulans*, alfa-glucosidasa *Aspergillus niger*, proteasa de tipo tripsina *Fusarium oxysporum*. Los terminadores ejemplares para las células huésped de levadura se pueden obtener de los genes para la enolasa *Saccharomyces cerevisiae*, citocromo C *Saccharomyces cerevisiae* (CYC1) y *Saccharomyces cerevisiae* gliceraldehído-3-fosfato deshidrogenasa. Otros terminadores útiles para células huésped de levadura se describen por Romanos et al., 1992, supra.

[0188] La secuencia de control también puede ser una secuencia líder adecuada que contiene una secuencia de inicio de la traducción. La secuencia líder está unida operativamente al extremo 5' de la secuencia de ácido nucleico que codifica el polipéptido. Se puede usar cualquier secuencia líder que sea funcional para iniciar la traducción en la célula huésped elegida. La secuencia de iniciación de la traducción bacteriana ejemplar se puede obtener de cualquier gen bacteriano expresado, tal como de *E. coli*, *Bacillus subtilis*, *Lactococcus lactis* y *Sinorhizobium meliloti* (véase, por ejemplo, Sakai, et al., 2001, J. Mol. Evol. 52: 164-170; Ma et al., 2002, J. Bacteriol. 184 (20): 5733-5745). En algunas realizaciones, se pueden usar secuencias de iniciación de la traducción artificial (por ejemplo, secuencia de Shine-Delgarno) (ver, por ejemplo, Vimberg et al., 2007, BMC Molecular Biology 8: 100). Líderes ejemplares para células huésped fúngicas filamentosas se obtienen a partir de los genes para la amilasa TAKA de *Aspergillus oryzae* y la isomerasa de fosfato triosa de *Aspergillus nidulans*. Los líderes adecuados para las células huésped de levadura se obtienen de los genes para enolasa *Saccharomyces cerevisiae* (ENO-1), 3-fosfoglicerato kinasa *Saccharomyces cerevisiae*, alfa-factor *Saccharomyces cerevisiae*, y alcohol deshidrogenasa/gliceraldehído-3-fosfato deasa *Saccharomyces cerevisiae*.

[0189] La secuencia de control también puede ser una secuencia de poliadenilación, una secuencia operativamente unida al extremo 3' de la secuencia de ácido nucleico y que, cuando se transcribe, es reconocida por la célula huésped como una señal para agregar residuos de poliadenosina al ARNm transcrito. Cualquier secuencia de poliadenilación que sea funcional en la célula huésped elegida puede usarse en la presente invención. Secuencias de poliadenilación ejemplares para células huésped fúngicas filamentosas pueden ser de los genes de amilasa TAKA *Aspergillus oryzae*, glucoamilasa *Aspergillus niger*, antranilato sintasa *Aspergillus nidulans*, proteasa de tipo tripsina *Fusarium oxysporum*, y alfa-glucosidasa *Aspergillus niger*. Las secuencias de poliadenilación para las células huésped de levadura se describen por Guo y Sherman, 1995, Mol Cell Bio 15: 5983-5990.

[0190] La secuencia de control también puede ser una región codificante de péptido señal que codifica una secuencia de aminoácidos unida al extremo amino de un polipéptido y dirige el polipéptido codificado hacia la vía secretora de la célula. El extremo 5' de la secuencia codificante de la secuencia de ácido nucleico puede contener inherentemente una región codificante de péptido señal enlazada naturalmente en el marco de lectura de traducción con el segmento de la región codificante que codifica el polipéptido secretado. Alternativamente, el extremo 5' de la secuencia de codificación puede contener una región de codificación del péptido señal que es ajena a la secuencia de codificación. En la presente invención se puede usar cualquier región codificante de péptido señal que dirige el polipéptido expresado hacia la ruta secretora de una célula huésped elegida. Las regiones codificantes del péptido señal eficaz para las células huésped bacterianas son las regiones codificantes del péptido señal obtenidas de los genes para la amilasa maltogénica de *Bacillus* NCIB 11837, *Bacillus stearothermophilus* alfa, la amígdala de *Bacillus*, la adición de *Bacillus licheniformis* beta-lactamasa, beta-lactamasa, nprM), y *Bacillus subtilis* prsA. Otros péptidos señal son descritos por Simonen y Palva, 1993, Microbiol Rev 57: 109-137. Las regiones codificantes del péptido señal eficaz para las células huésped fúngicas filamentosas pueden ser las regiones codificantes del péptido señal obtenidas de los genes de la amilasa TAKA de *Aspergillus oryzae*, la amilasa neutra de *Aspergillus niger*, la glucoamilasa de *Aspergillus niger*, la proteinasa aspártica *Rhizomucor miehei*, celulasa *Humicola insolens*, y lipasa

Humicola lanuginosa. Los péptidos señales útiles para células huésped de levadura pueden proceder de los genes para alfa-factor *Saccharomyces cerevisiae* e invertasa *Saccharomyces cerevisiae*. Otras regiones de codificación de péptidos señal útiles están descritas por Romanos et al., 1992, supra.

[0191] La secuencia de control también puede ser una región de codificación de propéptido que codifica una secuencia de aminoácidos posicionada en el extremo amino de un polipéptido. El polipéptido resultante se conoce como proenzima o propolipéptido (o zimógeno en algunos casos). Un propolipéptido puede convertirse en un polipéptido activo maduro por escisión catalítica o autocatalítica del propéptido del propolipéptido. La región codificante del propéptido se puede obtener a partir de los genes de la proteasa alcalina de *Bacillus subtilis* (aprE), de la proteasa neutra de *Bacillus subtilis* (nprT), del factor alfa de *Saccharomyces cerevisiae*, de la proteinasa aspártica *Rhizomucor miehei*, y de la lactasa *Myceliophthora thermophila* (WO 95/33836). Donde ambas regiones de péptidos y propéptidos señales están presentes en el extremo amino de un polipéptido, la región propeptídica se coloca al lado del extremo amino de un polipéptido y la región peptídica se coloca al lado del extremo amino de la región propéptido.

[0192] También puede ser deseable añadir secuencias reguladoras, que permitan la regulación de la expresión del polipéptido con respecto al crecimiento de la célula huésped. Ejemplos de sistemas reguladores son aquellos que causan que la expresión del gen se active o desactive en respuesta a un estímulo químico o físico, incluida la presencia de un compuesto regulador. En las células hospedadoras procarióticas, las secuencias reguladoras adecuadas incluyen los sistemas operadores lac, tac y trp. En células huésped de levadura, los sistemas reguladores adecuados incluyen, como ejemplos, el sistema ADH2 o el sistema GAL1. En los hongos filamentosos, las secuencias reguladoras adecuadas incluyen el promotor de alfa-amilasa TAKA, el promotor de glucoamilasa de *Aspergillus niger* y el promotor de glucoamilasa de *Aspergillus oryzae*.

[0193] En otro aspecto, la presente descripción también se dirige a un vector de expresión recombinante que comprende un polinucleótido que codifica un polipéptido de hidroxilasa de prolina diseñado, y una o más regiones reguladoras de la expresión tales como un promotor y un terminador, un origen de replicación, etc. Dependiendo del tipo de huéspedes en los que se van a introducir. Las diversas secuencias de control y ácido nucleico descritas anteriormente pueden unirse para producir un vector de expresión recombinante que puede incluir uno o más sitios de restricción convenientes para permitir la inserción o sustitución de la secuencia de ácido nucleico que codifica el polipéptido en dichos sitios. Alternativamente, la secuencia de ácido nucleico de la presente descripción puede expresarse insertando la secuencia de ácido nucleico o una construcción de ácido nucleico que comprende la secuencia en un vector apropiado para la expresión. Al crear el vector de expresión, la secuencia de codificación se ubica en el vector de modo que la secuencia de codificación esté unida operativamente con las secuencias de control apropiadas para la expresión.

[0194] El vector de expresión recombinante puede ser cualquier vector (por ejemplo, un plásmido o virus), que puede someterse convenientemente a procedimientos de ADN recombinante y puede provocar la expresión de la secuencia polinucleotídica. La elección del vector dependerá típicamente de la compatibilidad del vector con la célula huésped en la que se va a introducir el vector. Los vectores pueden ser plásmidos circulares lineales o cerrados.

[0195] El vector de expresión puede ser un vector de replicación autónoma, es decir, un vector que existe como una entidad extracromosómica, cuya replicación es independiente de la replicación cromosómica, por ejemplo, un plásmido, un elemento extracromosómico, un minicromosoma o cromosoma artificial. El vector puede contener cualquier medio para asegurar la autorreplicación. Alternativamente, el vector puede ser uno que, cuando se introduce en la célula huésped, se integra en el genoma y se replica junto con el (los) cromosoma(s) en donde se ha integrado. Además, se puede usar un solo vector o plásmido o dos o más vectores o plásmidos que juntos contienen el ADN total que se introducirá en el genoma de la célula huésped, o un transposón.

[0196] El vector de expresión contiene preferiblemente uno o más marcadores seleccionables, que permiten una fácil selección de células transformadas. Un marcador seleccionable es un gen cuyo producto proporciona biocida o resistencia viral, resistencia a metales pesados, prototrofia a auxótrofos y similares. Ejemplos de marcadores seleccionables bacterianos son los genes *dal* de *Bacillus subtilis* o *Bacillus licheniformis*, o marcadores, que confieren resistencia a los antibióticos como la ampicilina, kanamicina, cloranfenicol (ejemplo 1) o resistencia a la tetraciclina. Los marcadores adecuados para células huésped de levadura son ADE2, HIS3, LEU2, LYS2, MET3, TRP1 y URA3. Los marcadores seleccionables para usar en una célula huésped fúngica filamentosa incluyen, pero no se limitan a, *amdS* (acetamidasa), *argB* (ornitina carbamoiltransferasa), *barra* (fosfinotricina acetiltransferasa), *hph* (hidro-micina fosfotransferasa), *niaD* (nitrato reductasa), *pyrg* (orotidina-5'-fosfato descarboxilasa), *sC* (sulfato adenilo-transferasa) y *trpC* (antranilato sintetasa), así como sus equivalentes. Las realizaciones para uso en una célula de *Aspergillus* incluyen los genes *amdS* y *pyrG* de *Aspergillus nidulans* o *Aspergillus oryzae* y el gen *bar* de *Streptomyces hygroscopicus*.

[0197] En otro aspecto, la presente descripción proporciona una célula huésped que comprende un polinucleótido que codifica un polipéptido de hidroxilasa de prolina mejorado de la presente descripción, estando el polinucleótido unido operativamente a una o más secuencias de control para la expresión de la enzima hidroxina de prolina en la

célula huésped. Las células hospedadoras para uso en la expresión de los polipéptidos codificados por los vectores de expresión son bien conocidas en la técnica e incluyen, pero no se limitan a, células bacterianas, tales como células de *E. coli*, *Bacillus subtilis*, *Streptomyces* y *Salmonella typhimurium*; células fúngicas, tales como células de levadura (p. ej., *Saccharomyces cerevisiae* o *Pichia pastoris* (Nº de acceso ATCC 201178)); células de insecto tales como células *Drosophila* S2 y *Spodoptera* Sf9; células animales tales como CHO, COS, BHK, 293 y células de melanoma de Bowes; y células vegetales. Las células hospedadoras ejemplares son *Escherichia coli* W3110 (Δ fluA) y BL21.

[0198] Por consiguiente, en otro aspecto, la presente descripción proporciona métodos para fabricar los polipéptidos de hidroxilasa de prolina diseñados, en los que el método puede comprender cultivar una célula huésped capaz de expresar un polinucleótido que codifica el polipéptido de hidroxilasa de prolina en condiciones adecuadas para la expresión del polipéptido. El método puede comprender además aislar o purificar el polipéptido de hidroxilasa de prolina expresado, como se describe en el presente documento.

[0199] Los medios de cultivo apropiados y las condiciones de crecimiento para las células huésped descritas anteriormente son bien conocidas en la técnica. Los polinucleótidos para la expresión de la hidroxilasa de prolina pueden introducirse en las células mediante diversos métodos conocidos en la técnica. Las técnicas incluyen, entre otras, electroporación, bombardeo de partículas biolísticas, transfección mediada por liposomas, transfección con cloruro de calcio y fusión de protoplastos.

[0200] En las realizaciones de este documento, se pueden obtener los polipéptidos mejorados y los polinucleótidos correspondientes utilizando métodos utilizados por los expertos en la técnica. La secuencia de polinucleótidos de origen natural que codifica la hidroxilasa de cis-4-prolina de *Sinorhizobium meliloti* que se produce de manera natural se describe en la publicación de patente de EE.UU. US20110091942 y la publicación de Patente Internacional Nº WO2009139365. Las hidroxilasas de prolina manipuladas descritas en el presente documento pueden obtenerse sometiendo el polinucleótido que codifica la hidroxilasa de prolina de origen natural u otra modificada a métodos de mutagénesis y/o evolución dirigida, como se discute aquí. Una técnica de evolución dirigida ejemplar es la mutagénesis y/o el barajado de ADN como se describe en Stemmer, 1994, Proc Natl Acad Sci EE.UU. 91: 10747-10751; WO 95/22625; WO 97/0078; WO 97/35966; WO 98/27230; WO 00/42651; WO 01/75767 y la patente de EE.UU. 6.537.746. Otros procedimientos de evolución dirigida que pueden usarse incluyen, entre otros, el proceso de extensión escalonada (StEP), la recombinación *in vitro* (Zhao et al., 1998, Nat. Biotechnol. 16: 258-261), la PCR mutagénica (Caldwell et al., 1994, PCR METHODS Appl. 3: S136-S140), y mutagénesis en casete (Black et al., 1996, Proc Natl Acad Sci EE.UU. 93: 3525-3529). Las técnicas de mutagénesis y evolución dirigida útiles para los propósitos en este documento también se describen en las siguientes referencias: Ling, et al., 1997. Anal. Biochem. 254 (2): 157-78; Dale et al., 1996, "Oligonucleotide-directed random mutagenesis using the phosphorothioate method", en Methods Mol. Biol. 57: 369-74; Smith, 1985, Ann. Rev. Genet. 19: 423-462; Botstein et al., 1985, Science 229: 1193-1201; Carter, 1986, Biochem. J. 237: 1-7; Kramer et al., 1984, Cell, 38: 879-887; Wells et al., 1985, Gene 34: 315-323; Minshull et al., 1999, Curr Opin Chem Biol 3: 284-290; Christians et al., 1999, Nature Biotech 17: 259-264; Cramer et al., 1998, Nature 391: 288-291; Cramer et al., 1997, Nature Biotech 15: 436-438; Zhang et al., 1997, Proc Natl Acad Sci EE.UU. 94: 45-4-4509; Cramer et al., 1996, Nature Biotech 14: 315-319; Stemmer, 1994, Nature 370: 389-391; Stemmer, 1994, Proc Natl Acad Sci EE.UU. 91: 10747-10751; WO 95/22625; WO 97/0078; WO 97/35966; WO 98/27230; WO 00/42651; WO 01/75767 y la patente de EE.UU. 6.537.746.

[0201] Los clones obtenidos después del tratamiento de mutagénesis pueden cribarse para detectar hidroxilasas de prolina que tienen una o más propiedades enzimáticas mejoradas deseadas. Por ejemplo, cuando la propiedad enzimática mejorada deseada es la regiospecificidad, la actividad enzimática puede medirse para la producción del compuesto (1) y del compuesto (1a). Los clones que contienen un polinucleótido que codifica una hidroxilasa de prolina con las características deseadas, por ejemplo, una mayor proporción de compuesto (1) sobre el compuesto (1a), se aíslan, se secuencian para identificar los cambios en la secuencia de nucleótidos (si los hay) y se usan para expresar la enzima en una célula huésped. La medición de la actividad enzimática de las bibliotecas de expresión se puede realizar utilizando las técnicas bioquímicas estándar, como el análisis por HPLC y/o la derivación de productos (antes o después de la separación), por ejemplo, con cloruro de dansilo u OPA (ver, por ejemplo, Yaegaki et al., 1986, J. Chromatogr. 356 (1): 163-70).

[0202] En el caso de que se conozca la secuencia del polipéptido modificado, los polinucleótidos que codifican la enzima se pueden preparar mediante métodos estándar en fase sólida, de acuerdo con métodos sintéticos conocidos. En algunas realizaciones, los fragmentos de hasta aproximadamente 100 bases pueden sintetizarse individualmente, luego unirse (por ejemplo, mediante métodos de litigación enzimáticos o químicos, o métodos mediados por polimerasa) para formar cualquier secuencia continua deseada. Por ejemplo, los polinucleótidos y oligonucleótidos que codifican porciones de la hidroxilasa de prolina se pueden preparar por síntesis química utilizando, por ejemplo, el método clásico de fosforamidita descrito por Beaucage et al., 1981, Tet Lett 22: 1859-69, o el método descrito por Matthes et al., 1984, EMBO J. 3: 801-05, por ejemplo, tal como se practica típicamente en métodos sintéticos automatizados. De acuerdo con el método de la fosforamidita, los oligonucleótidos se sintetizan, por ejemplo, en un sintetizador de ADN automático, se purifican, se recocen, se ligan y se clonan en vectores apropiados. Además, esencialmente cualquier ácido nucleico puede obtenerse de cualquiera de una variedad de fuentes comerciales. En algunas realizaciones, se pueden crear variaciones adicionales sintetizando

oligonucleótidos que contienen deleciones, inserciones y/o sustituciones, y combinando los oligonucleótidos en varias permutaciones para crear hidroxilasas de prolina modificadas con propiedades mejoradas.

[0203] Por consiguiente, en algunas realizaciones, un método para preparar el polipéptido de hidroxilasas de prolina modificado puede comprender: (a) sintetizar un polinucleótido que codifica un polipéptido que comprende una secuencia de aminoácidos seleccionada de la SEQ ID NO: 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228, y que tienen una o más diferencias de residuos en comparación con la SEQ ID NO: 2 en las posiciones de residuos seleccionadas de: X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X186; X188; X205; X225; X230; X270; y X271; y (b) expresan el polipéptido de hidroxilasa de prolina codificado por el polinucleótido.

[0204] En algunas realizaciones del método, las diferencias de residuos en las posiciones de residuos X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X186; X188; X205; X225; X230; X270; y X271 se seleccionan de X2K; X2T; X3S; X4Q; X4L; X4E; X4S; X5I; X5L; X5M; X9I; X13T; X17V; X24R; X24S; X25R; X26R; X26T; X26W; X29A; X30V; X30P; X36T; X42E; X52P; X57T; X57A; X58A; X59G; X62Q; X66Q; X86S; X88R; X92V; X95M; X98F; X98T; X103L; X103Q; X112T; X112V; X113E; X114N; X115E; X115H; X115D; X115G; X115S; X115A; X116L; X121F; X131Y; X131F; X140L; X150S; X151A; X151H; X151S; X166T; X166L; X166Q; X186G; X188G; X205V; X225L; X225Y; X225W; X230V; X270E; X271K; y X271R.

[0205] En algunas realizaciones del método, el polinucleótido puede codificar una hidroxilasa de prolina diseñada que tiene opcionalmente una o varias (por ejemplo, hasta 3, 4, 5 o hasta 10) deleciones, inserciones y/o sustituciones de residuos de aminoácidos. En algunas realizaciones, la secuencia de aminoácidos tiene opcionalmente 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-15, 1-20, 1-21, 1-22, 1-23, 1-24, 1-25, 1-30, 1-35, 1-40, 1-45, o 1-50 eliminación de residuos de aminoácidos, inserciones y/o sustituciones. En algunas realizaciones, la secuencia de aminoácidos tiene opcionalmente 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, 40, 45 o 50 deleciones, inserciones y/o sustituciones de residuos de aminoácidos. En algunas realizaciones, la secuencia de aminoácidos tiene opcionalmente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 20, 21, 22, 23, 24 o 25 eliminaciones, inserciones y/o sustituciones de residuos de aminoácidos. En algunas realizaciones, las sustituciones pueden ser sustituciones conservativas o no conservativas.

[0206] En algunas realizaciones, cualquiera de las enzimas hidroxilasas de prolina modificadas expresadas en una célula huésped puede recuperarse de las células y/o el medio de cultivo utilizando una o más de las técnicas bien conocidas para la purificación de proteínas, incluyendo, entre otras, tratamiento con lisozima, sonicación, filtración, lavado, ultra centrifugación y cromatografía. Las soluciones adecuadas para la lisis y la extracción de proteínas de alta eficacia de bacterias, como *E. coli*, están disponibles comercialmente, como CellLytic B™ de Sigma-Aldrich de St. Louis MO.

[0207] Las técnicas cromatográficas para el aislamiento del polipéptido de hidroxilasa de prolina incluyen, entre otras, cromatografía líquida de alta resolución, cromatografía de fase inversa, cromatografía de intercambio iónico, electroforesis en gel y cromatografía de afinidad. Las condiciones para purificar una enzima particular dependerán, en parte, de factores tales como la carga neta, la hidrofobicidad, la hidrofiliidad, el peso molecular, la forma molecular, etc., y serán evidentes para los expertos en la técnica.

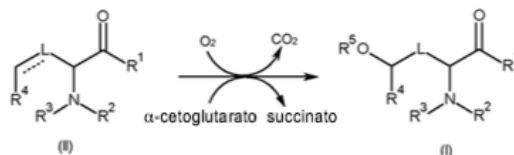
[0208] En algunas realizaciones, pueden usarse técnicas de afinidad para aislar las enzimas hidroxilasas de prolina mejoradas. Para la purificación por cromatografía de afinidad, se puede usar cualquier anticuerpo que se una específicamente al polipéptido de la hidroxilasa de prolina. Para la producción de anticuerpos, pueden inmunizarse diversos animales hospedadores, que incluyen, entre otros, conejos, ratones, ratas, etc., mediante inyección con un polipéptido de hidroxilasa de prolina, o un fragmento del mismo. El polipéptido o fragmento de hidroxilasa de prolina se puede unir a un portador adecuado, tal como BSA, por medio de un grupo funcional de cadena lateral o enlazadores unidos a un grupo funcional de cadena lateral. En algunas realizaciones, la purificación por afinidad puede usar un ligando específico unido a la hidroxilasa de prolina, como la columna de afinidad de tinte o poli(L-prolina) (ver, por ejemplo, EP0641862; Stellwagen, E., 2001, "Dye Affinity Chromatography," In Current Protocols in Protein Science Unit 9.2-9.2.16).

5.7 Métodos de uso de las enzimas de hidroxilasa de prolina diseñadas

[0209] En otro aspecto, las hidroxilasas de prolina descritas en el presente documento pueden usarse en un proceso para convertir un sustrato adecuado en su producto hidroxilado. En general, el proceso para realizar la reacción de hidroxilación comprende poner en contacto o incubar el compuesto de sustrato en presencia de un co-sustrato, como el α -cetoglutarato, con un polipéptido de hidroxilasa de prolina de la descripción en condiciones de reacción adecuadas para la formación del producto hidroxilado. En algunas formas de realización, las hidroxilasas de prolina

se pueden usar en la conversión del compuesto de sustrato (II) en el compuesto de producto (I), como se ilustra en el **Esquema 2**:

Esquema 2



en donde

L se selecciona del grupo que consiste en un enlace, alquileo (C₁-C₄) y alquenileno (C₂-C₄);

R¹ se selecciona del grupo que consiste en hidroxilo, amino, alquilo (C₁-C₆), arilo, alquilo (C₁-C₆) y arilo;

R² y R³ se seleccionan cada uno independientemente del grupo que consiste en hidrógeno y alquilo (C₁-C₆) opcionalmente sustituido, alquilo (C₂-C₆) y alquilo (C₂-C₆);

R⁴ se selecciona del grupo que consiste en alquilo (C₁-C₆), alquilo (C₂-C₆) opcionalmente sustituido, alquilo (C₂-C₆), arilo, heteroarilo, cicloalquilo, heterocicloalquilo; o R⁴ junto con uno de R¹ o R² es un alquilo (C₁-C₅) o alquilo (C₂-C₅) y forma un anillo heterocíclico de 5 a 8 miembros que contiene el átomo de nitrógeno, en donde el anillo está opcionalmente sustituido con 1 a 4 grupos R⁶ seleccionados independientemente;

R⁵ es hidrógeno o un enlace que forma un epóxido con un átomo de carbono de L;

cada aparición de R⁶ se selecciona independientemente del grupo que consiste en halo, alquilo (C₁-C₆) y alquilo (C₁-C₆); y

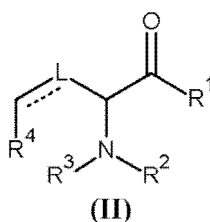
----- representa un enlace opcional a un átomo de carbono de L para formar un doble enlace, con la condición de que

(i) cuando R⁴ no forma un anillo con un R² o R³, o cuando R⁴ forma un anillo heterocíclico de 5 miembros que contiene el átomo de nitrógeno con uno de R² o R³, entonces L es un metileno;

(ii) cuando R⁴ forma un anillo heterocíclico de 6 miembros que contiene el átomo de nitrógeno con uno de R² o R³, entonces L es un enlace o etileno; y

(iii) cuando R⁵ es un enlace a un átomo de carbono de L para formar un epóxido, entonces R⁴ forma el anillo heterocíclico que contiene el átomo de nitrógeno con uno de R² o R³ y L es un alquilo (C₁-C₄) o alquilo (C₂-C₄).

[0211] Por consiguiente, en algunas realizaciones, un proceso para preparar el compuesto de producto (I) puede comprender ponerse en contacto con el compuesto de sustrato de fórmula (II)



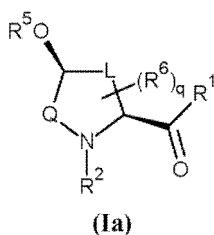
en donde

L, R¹, R², R³, R⁴ y R⁶ son como se definieron anteriormente,

----- representa un enlace opcional a un átomo de carbono de L para formar un doble enlace

con un polipéptido manipulado descrito aquí en presencia de un co-sustrato en condiciones de reacción adecuadas.

[0212] En algunas realizaciones del proceso, el compuesto de fórmula (I) comprende el compuesto de fórmula (Ia),

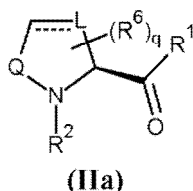


en donde

Q se selecciona del grupo que consiste en un alquileo (C_1-C_5) y alquenileno (C_2-C_5);
 L se selecciona del grupo que consiste en un enlace, alquileo (C_1-C_4) y alquenileno (C_2-C_4);
 R^1 se selecciona del grupo que consiste en hidroxilo, amino, alquiloxi (C_1-C_6), arilo, alquiltio (C_1-C_6) y ariltio;
 R^2 se selecciona del grupo que consiste en hidrógeno y alquilo (C_1-C_6) opcionalmente sustituido, alquenilo (C_2-C_6) y alquinilo (C_2-C_6);
 R^5 es hidrógeno, o un enlace a un átomo de carbono de L para formar un epóxido;
 cada aparición de R^6 se selecciona del grupo que consiste en halo, alquilo (C_1-C_6) y alquiloxi (C_1-C_6); y
 q es un número entero de 0 a 4;
 en donde la suma de los átomos de carbono del anillo para Q + L es un número entero de 2 a 5;
 con la condición de que

- (i) cuando la suma de los átomos de carbono del anillo para Q + L es 2, entonces L es un metileno; y
- (ii) cuando la suma de los átomos de carbono del anillo para Q + L es 3, entonces L es un enlace o etileno.

[0213] Por consiguiente, un proceso para preparar el compuesto de fórmula **(Ia)** comprende poner en contacto el compuesto de fórmula **(IIa)**,

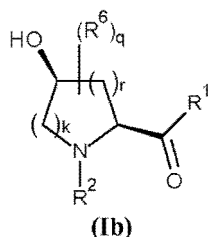


en donde

L, Q, R^1 , R^2 , R^6 y q son como se definieron anteriormente para el compuesto de fórmula **(Ia)**; y
 ---- representa un enlace opcional a un átomo de carbono de L para formar un doble enlace;

con un polipéptido diseñado de la descripción en presencia de un co-sustrato en condiciones de reacción adecuadas.

[0214] En algunas realizaciones del proceso, el compuesto de fórmula **(Ia)** comprende el compuesto de fórmula **(Ib)**,

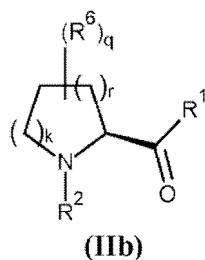


en donde

R^1 se selecciona del grupo que consiste en hidroxilo, amino, alquiloxi (C_1-C_6), arilo, alquiltio (C_1-C_6) y ariltio;
 R^2 se selecciona del grupo que consiste en hidrógeno y alquilo (C_1-C_6) opcionalmente sustituido, alquenilo (C_2-C_6) y alquinilo (C_2-C_6);
 cada aparición de R^6 se selecciona independientemente del grupo que consiste en halo, alquilo (C_1-C_6) y alquiloxi (C_1-C_6); y
 k es un número entero de 1 a 5;
 r es un número entero de 0 a 4;
 en donde k + r es 3, 4 o 5; y
 q es un número entero de 0 a 4;
 con la condición de que cuando k + r es 3, entonces k es 1 o 3.

[0215] Por consiguiente, un proceso para preparar el compuesto de fórmula **(Ib)** comprende poner en contacto el

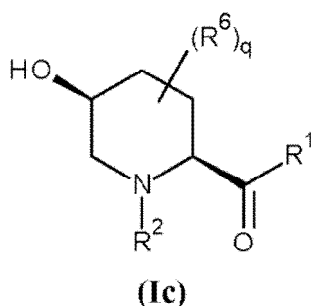
compuesto de fórmula (IIb),



en donde

R¹, R², R⁶, k, r y q son los definidos anteriormente para el compuesto de fórmula (Ib); con un polipéptido diseñado de la descripción en presencia de un co-sustrato en condiciones de reacción adecuadas. En algunas realizaciones, k es 1 y r es 2, 3 o 4.

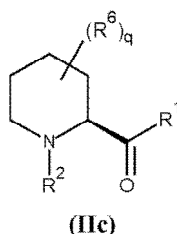
[0216] En algunas realizaciones del proceso, el compuesto de fórmula (Ia) comprende el compuesto de fórmula (Ic),



en donde

R¹ se selecciona del grupo que consiste en hidroxilo, amino, alquilo (C₁-C₆), alilo, alquilo (C₁-C₆) y arilo; R² se selecciona del grupo que consiste en hidrógeno y alquilo (C₁-C₆) opcionalmente sustituido, alquilo (C₂-C₆) y alquilo (C₂-C₆); cada aparición de R⁶ se selecciona independientemente del grupo que consiste en hidrógeno, halo, alquilo (C₁-C₆) y alquilo (C₁-C₆); y q es un número entero de 0 a 4.

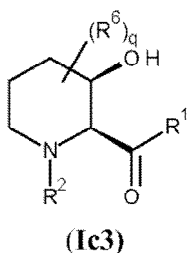
[0217] Por consiguiente, un proceso para preparar el compuesto de fórmula (Ic) comprende poner en contacto el compuesto de fórmula (Ib),



en donde

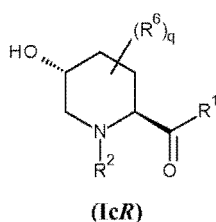
R¹, R², R⁶ y q son como se definieron anteriormente para el compuesto de fórmula (Ic); con un polipéptido diseñado de la descripción en presencia de un co-sustrato en condiciones de reacción adecuadas.

[0218] En algunas realizaciones del proceso, el compuesto de fórmula (Ic) se forma en exceso del compuesto del producto de fórmula (Ic3),



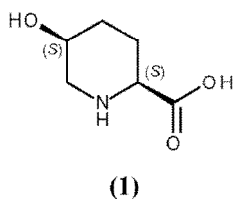
[0219] Por consiguiente, un proceso para preparar el compuesto de fórmula **(Ic)** en exceso del compuesto de fórmula **(Ic3)** comprende poner en contacto el compuesto de fórmula **(Ic)** con un polipéptido modificado descrito en el presente documento que tiene una capacidad de reacción para el compuesto del producto **(1)** sobre el compuesto del producto **(1a)** en presencia de un co-sustrato en condiciones de reacción adecuadas. En algunas realizaciones del proceso, el compuesto producto de fórmula **(Ic)** se forma en exceso del compuesto producto de fórmula **(Ic3)**, donde la proporción de compuesto **(Ic)** formado sobre el compuesto **(Ic3)** es de al menos 1,5, 2, 3, 4, 5 o 6 o más.

[0220] En algunas realizaciones del proceso, el compuesto de producto **(Ic)** se forma en un exceso diastereomérico de compuesto **(IcR)**,

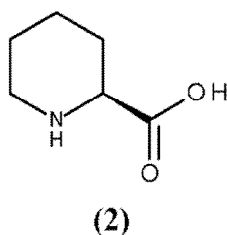


[0221] En algunas realizaciones del proceso, el compuesto del producto **(Ic)** se forma en al menos un 90%, 95%, 96%, 97%, 98%, 99% o mayor exceso diastereomérico del compuesto **(IcR)**. En algunas realizaciones, no se forma una cantidad detectable de compuesto **(IcR)** en el proceso.

[0222] En algunas realizaciones del proceso para la preparación de los compuestos de producto de fórmula **(Ic)**, R^1 es hidroxilo, R^2 es hidrógeno y q es 0. Como tal, en algunas realizaciones del proceso, el compuesto de fórmula **(I)** comprende el compuesto de fórmula **(1)**,



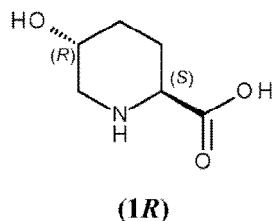
[0223] Por consiguiente, un procedimiento para preparar el compuesto de fórmula **(1)** comprende poner en contacto el compuesto de fórmula **(2)**,



con un polipéptido diseñado de la descripción en presencia de un co-sustrato en condiciones de reacción adecuadas.

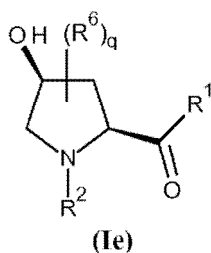
[0224] En algunas realizaciones del proceso, el compuesto del producto (1) se forma en exceso del compuesto del producto (1a). En algunas realizaciones, el compuesto del producto (1) se forma en exceso sobre el compuesto (1a) en una relación de al menos 1,5, 2, 3, 4, 5 o 6 o más. En algunas realizaciones del proceso, el polipéptido diseñado útil para preparar el compuesto (1) en exceso del compuesto de fórmula (1a) comprende poner en contacto el compuesto de fórmula (2) con un polipéptido diseñado descrito en el presente documento que tiene regioselectividad para el compuesto del producto (1) sobre el compuesto (1a) en condiciones de reacción adecuadas.

[0225] En algunas realizaciones del procedimiento, el compuesto de producto (1) se forma en un exceso diastereomérico del compuesto (1R),



[0226] En algunas realizaciones del proceso, el compuesto del producto (1) se forma en al menos 90%, 95%, 96%, 97%, 98%, 99% o mayor exceso diastereomérico del compuesto (1R). En algunas realizaciones, no se forma una cantidad detectable de compuesto (1R) en el proceso.

[0227] En algunas realizaciones, el compuesto de fórmula (1a) comprende el compuesto de fórmula (Ie),



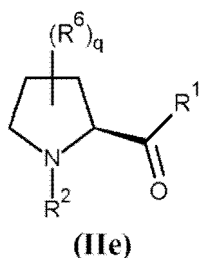
en donde

R¹ se selecciona del grupo que consiste en hidroxilo, amino, alquilo (C₁-C₆), arilo, alquilo (C₁-C₆) y arilo; R² se selecciona del grupo que consiste en hidrógeno y alquilo (C₁-C₆) opcionalmente sustituido, alquilo (C₂-C₆) y alquilo (C₂-C₆);

cada aparición de R⁶ se selecciona independientemente del grupo que consiste en hidrógeno, halo, alquilo (C₁-C₆) y alquilo (C₁-C₆); y

q es un número entero de 0 a 3.

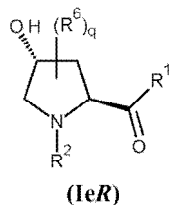
[0228] Por consiguiente, en algunas realizaciones, un proceso para preparar el compuesto de producto de fórmula (Ie) comprende poner en contacto el compuesto de fórmula (IIe),



en donde R¹, R², R⁶ y q son como se definen para el compuesto de fórmula (Ie),

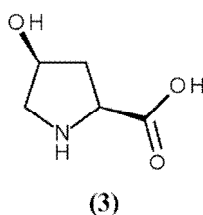
con un polipéptido diseñado de la descripción en presencia de un co-sustrato en condiciones de reacción adecuadas.

[0229] En algunas realizaciones del proceso, el compuesto de fórmula (**1e**) se forma en exceso diastereomérico del compuesto de fórmula (**1eR**),

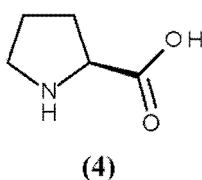


[0230] En algunas realizaciones del proceso, el compuesto producto de fórmula (**1e**) se forma en al menos un 90%, 95%, 96%, 97%, 98%, 99% o mayor exceso diastereomérico del compuesto de fórmula (**1eR**). En algunas realizaciones, no se forma una cantidad detectable del compuesto de fórmula (**1eR**).

[0231] En algunas realizaciones del procedimiento para preparar el producto compuesto de fórmula (**1e**), R¹ es hidroxilo, R² es hidrógeno y q es 0. Como tal, en algunas realizaciones, el compuesto de fórmula (**1e**) comprende el compuesto (**3**),

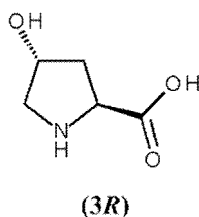


en donde el proceso para preparar el compuesto (**3**) comprende poner en contacto el compuesto (**4**),



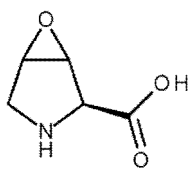
con un polipéptido diseñado de la descripción en presencia de un co-sustrato en condiciones de reacción adecuadas.

[0232] En algunas realizaciones del proceso, el compuesto (**3**) se forma en exceso diastereomérico del compuesto (**3R**),



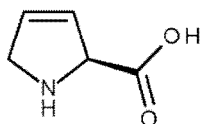
[0233] En algunas realizaciones del proceso, el compuesto del producto (**3**) se forma en al menos 90%, 95%, 96%, 97%, 98%, 99% o mayor exceso diastereomérico del compuesto (**3R**). En algunas realizaciones, no se forma una cantidad detectable del compuesto (**3R**).

[0234] En algunas realizaciones, el compuesto de fórmula (**1a**) comprende el compuesto (**5**),



(5)

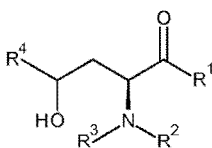
[0235] Por consiguiente, un procedimiento para preparar el compuesto (5) comprende poner en contacto el compuesto (6);



(6)

con un polipéptido diseñado de la descripción en presencia de un co-sustrato en condiciones de reacción adecuadas.

[0236] En algunas realizaciones, el compuesto de fórmula (I) comprende el compuesto de fórmula (III),



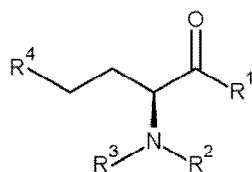
(III)

en donde,

R¹ se selecciona del grupo que consiste en hidroxilo, amino, alquilo (C₁-C₆), arilo, alquilo (C₁-C₆) y arilo;
R² y R³ se seleccionan cada uno independientemente del grupo que consiste en hidrógeno y alquilo (C₁-C₆) opcionalmente sustituido, alqueno (C₂-C₆) y alquino (C₂-C₆); y
R⁴ se selecciona del grupo que consiste en alquilo, alqueno, alquino, arilo, heteroarilo, cicloalquilo y heterocicloalquilo opcionalmente sustituidos;

[0237] En algunas realizaciones, el alquilo opcionalmente sustituido se selecciona del grupo que consiste en alquilo (C₁-C₆), aminocarbonilo (C₁-C₆)alquilo, amino (C₁-C₆)alquilo, tior (C₁-C₆)alquilo, y alquilo (C₁-C₆)alquilo, alquilsulfonilo (C₁-C₆)alquilo, arilo (C₁-C₆)alquilo, heteroarilo (C₁-C₆)alquilo, cicloalquilo (C₁-C₆) alquilo y heterocicloalquilo (C₁-C₆) alquilo.

[0238] Por consiguiente, en algunas realizaciones, un proceso para preparar el compuesto de fórmula (III) comprende poner en contacto el compuesto de fórmula (IV),



(IV)

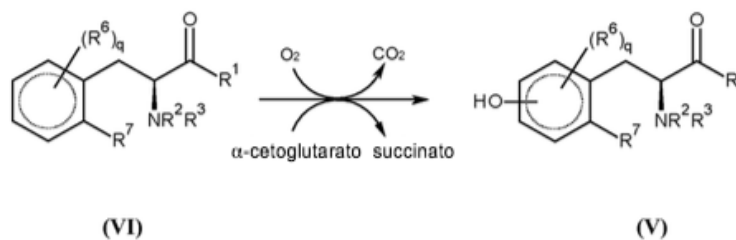
en donde,

R^1 , R^2 , R^3 y R^4 son como se definieron anteriormente para el compuesto de fórmula (III),

con un polipéptido diseñado de la descripción en presencia de un co-sustrato en condiciones de reacción adecuadas.

[0239] En algunas realizaciones, los polipéptidos diseñados pueden usarse en un proceso para llevar a cabo la conversión mostrada en el **Esquema 3**,

Esquema 3



en donde,

R^1 se selecciona del grupo que consiste en hidroxilo, amino, alquilo (C₁-C₆), arilo, alquilo (C₁-C₆) y arilo;

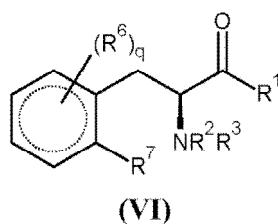
R^2 y R^3 se seleccionan independientemente del grupo que consiste en hidrógeno y alquilo (C₁-C₆) opcionalmente sustituido, alquilo (C₂-C₆) y alquilo (C₂-C₆);

cada aparición de R^6 se selecciona independientemente del grupo que consiste en halo, alquilo (C₁-C₆) y alquilo (C₁-C₆);

R^7 se selecciona del grupo que consiste en hidrógeno, halo y alquilo (C₁-C₆) y alquilo (C₁-C₆) opcionalmente sustituido o R^7 junto con uno de R^2 o R^3 forma un anillo heterocíclico de 5 a 7 miembros que contiene el átomo de nitrógeno; q es un número entero de 0 a 4; y

----- representan dobles enlaces opcionales que forman un anillo arilo.

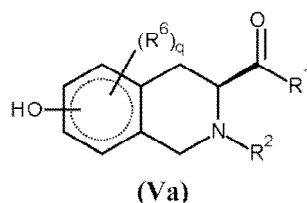
[0240] Por consiguiente, en algunas realizaciones, un procedimiento para preparar el compuesto de fórmula (V) comprende poner en contacto el compuesto de fórmula (VI),



en donde R^1 , R^2 , R^3 , R^6 , R^7 y q son como se definen para el compuesto de fórmula (V),

con un polipéptido diseñado de la descripción en presencia de un co-sustrato en condiciones de reacción adecuadas.

[0241] En algunas realizaciones, el compuesto de fórmula (V) comprende el compuesto de fórmula (Va),

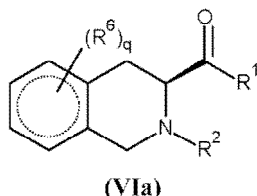


en donde,

R^1 se selecciona del grupo que consiste en hidroxilo, amino, alquilo (C₁-C₆), arilo, alquilo (C₁-C₆) y arilo;

R^2 y R^3 se seleccionan independientemente del grupo que consiste en hidrógeno y alquilo (C_1-C_6) opcionalmente sustituido, alqueno (C_2-C_6) y alquino (C_2-C_6);
 cada aparición de R^6 se selecciona independientemente del grupo que consiste en halo, alquilo (C_1-C_6) y alquilo (C_1-C_6); y
 q es un número entero de 0 a 4.

[0242] Por consiguiente, en algunas realizaciones, un proceso para preparar el compuesto de fórmula **(Va)** comprende poner en contacto el compuesto de fórmula **(Via)**,



en donde R^1 , R^2 , R^3 , R^6 y q son como se definieron anteriormente para el compuesto de fórmula **(Va)**, con un polipéptido diseñado de la descripción en presencia de un co-sustrato en condiciones de reacción adecuadas.

[0243] Para los procesos anteriores, se puede usar cualquiera de las hidroxilasas de prolina diseñadas aquí. A modo de ejemplo y sin limitación, en algunas realizaciones, los procesos pueden usar un polipéptido de hidroxilasa de prolina diseñado que comprende una secuencia de aminoácidos que tiene al menos 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más de identidad con una secuencia de referencia seleccionada de la SEQ ID NO: 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226, y 228.

[0244] En algunas realizaciones de los procesos, el polipéptido de hidroxilasa de prolina diseñado comprende una secuencia de aminoácidos que tiene una o más diferencias de residuos en comparación con la SEQ ID NO: 2 en las posiciones de residuos X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X186; X205; X225; X230; X270; y X271.

[0245] En algunas realizaciones de los procesos, las diferencias de residuos en las posiciones de residuos X2; X3; X4; X5; X9; X13; X17; X24; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X62; X66; X86; X88; X92; X95; X98; X103; X112; X113; X114; X115; X116; X121; X131; X140; X150; X151; X166; X186; X205; X225; X230; X270; y X271 se seleccionan de X2K; X2T; X3S; X4Q; X4L; X4E; X4S; X5I; X5L; X5M; X9I; X13T; X17V; X24R; X24S; X25R; X26R; X26T; X26W; X29A; X30V; X30P; X36T; X42E; X52P; X57T; X57A; X58A; X59G; X62Q; X66Q; X86S; X88R; X92V; X95M; X98F; X98T; X103L; X103Q; X112T; X112V; X113E; X114N; X115E; X115H; X115D; X115G; X115S; X115A; X116L; X121F; X131Y; X131F; X140L; X150S; X151A; X151H; X151S; X166T; X166L; X166Q; X186G; X188G; X205V; X225L; X225Y; X225W; X230V; X270E; X271K; y X271R.

[0246] Como se indicó anteriormente, en algunas realizaciones para preparar un compuesto de producto de fórmula **(Ic)** en exceso sobre el compuesto de fórmula **(Ic3)**, o para preparar compuesto **(1)** en exceso sobre el compuesto **(1a)**, el polipéptido diseñado puede comprender una secuencia de aminoácidos que tiene una o más características seleccionadas de: X103L; X115E; X131Y y X166Q. Polipéptidos ejemplares diseñados con la regioselectividad relevante pueden comprender una secuencia de aminoácidos seleccionada del grupo que consiste en las SEQ ID NO: 10, 24, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226, y 228.

[0247] En algunas realizaciones, los polipéptidos de hidroxilasas de prolina ejemplares capaces de llevar a cabo los procedimientos en el presente documento pueden tener una secuencia de aminoácidos que comprende una secuencia seleccionada de la SEQ ID NO: 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228. En las descripciones de este documento se proporciona orientación sobre la elección y el uso de las hidroxilasas de prolina modificadas, por ejemplo, las **Tablas**

2A, 2B, 2C, 2D, 2E, 2F, 2G y 2H y los Ejemplos.

[0248] En las realizaciones de este documento e ilustradas en los Ejemplos, diversos intervalos de condiciones de reacción adecuadas que pueden usarse en los procesos incluyen, entre otros, carga de sustrato, carga de co-sustrato, reductor, metal de transición divalente, pH, temperatura, tampón, sistema de disolventes, carga de polipéptidos y tiempo de reacción. Otras condiciones de reacción adecuadas para llevar a cabo el proceso para la conversión biocatalítica de compuestos de sustrato a compuestos de producto que utilizan un polipéptido de hidroxilasa de prolina descrito en este documento pueden optimizarse fácilmente en vista de la guía proporcionada en este documento por experimentación rutinaria que incluye, pero no se limita a, poner en contacto el polipéptido de hidroxilasa de prolina y el compuesto de sustrato bajo condiciones de reacción experimentales de concentración, pH, temperatura y condiciones de solvente, y la detección del compuesto del producto.

[0249] Las condiciones de reacción adecuadas que usan los polipéptidos de hidroxilasa de prolina modificadas comprenden típicamente un co-sustrato, que se utiliza estequiométricamente en la reacción de hidroxilación. Generalmente, el co-sustrato para las hidroxilasas de prolina es α -cetoglutarato, también denominado ácido α -cetoglutarico y ácido 2-oxoglutarico. Se pueden usar otros análogos del α -cetoglutarato que son capaces de servir como co-sustratos para las hidroxilasas de prolina. Un análogo ejemplar que puede servir como un co-sustrato es el 2-oxoadipato. Debido a que el co-sustrato se utiliza estequiométricamente, el co-sustrato está presente en una cantidad equimolar o mayor que la del compuesto de sustrato, es decir, la concentración molar de co-sustrato es equivalente o superior a la concentración molar del compuesto de sustrato. En algunas realizaciones, las condiciones de reacción adecuadas pueden comprender una concentración molar de co-sustrato de al menos 1 vez, 1,5 veces, 2 veces, 3 veces 4 veces o 5 veces o más que la concentración molar del compuesto de sustrato. En algunas realizaciones, las condiciones de reacción adecuadas pueden comprender una concentración de co-sustrato, particularmente α -cetoglutarato, de aproximadamente 0,001 M a aproximadamente 2 M, 0,01 M a aproximadamente 2 M, 0,1 M a aproximadamente 2 M, 0,2 M a aproximadamente 2 M de aproximadamente 0,5 M a aproximadamente 2 M, o de aproximadamente 1 M a aproximadamente 2 M. En algunas realizaciones, las condiciones de reacción comprenden una concentración de co-sustrato de aproximadamente 0,001, 0,01, 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 1, 1,5 o 2 M. En algunas realizaciones, se puede agregar un co-sustrato adicional durante la reacción.

[0250] El compuesto de sustrato en las mezclas de reacción se puede variar, teniendo en cuenta, por ejemplo, la cantidad deseada de compuesto de producto, el efecto de la concentración de sustrato sobre la actividad enzimática, la estabilidad de la enzima en condiciones de reacción y el porcentaje de conversión de sustrato a producto. En algunas realizaciones, las condiciones de reacción adecuadas comprenden una carga de compuesto de sustrato de al menos aproximadamente 0,5 a aproximadamente 200 g/L, 1 a aproximadamente 200 g/L, 5 a aproximadamente 150 g/L, aproximadamente 10 a aproximadamente 100 g/L, 20 a aproximadamente 100 g/L o aproximadamente 50 a aproximadamente 100 g/L. En algunas realizaciones, las condiciones de reacción adecuadas comprenden una carga de compuesto de sustrato de al menos aproximadamente 0,5 g/L, al menos aproximadamente 1 g/L, al menos aproximadamente 5 g/L, al menos aproximadamente 10 g/L, al menos aproximadamente 15 g/L g/L, al menos aproximadamente 20 g/L, al menos aproximadamente 30 g/L, al menos aproximadamente 50 g/L, al menos aproximadamente 75 g/L, al menos aproximadamente 100 g/L, al menos aproximadamente 150 g/L L o al menos unos 200 g/L, o incluso mayor. Los valores para cargas de sustratos que se proporcionan aquí se basan en el peso molecular del compuesto (2), sin embargo, también contempla que las cantidades molares equivalentes de varios hidratos y sales del compuesto (2) también se pueden usar en el proceso. Además, los compuestos de sustrato cubiertos por las fórmulas (II) y (VI), incluidos los compuestos de fórmula (IIa), (IVa) y (VIa) también pueden usarse en cantidades apropiadas, a la luz de las cantidades utilizadas para el compuesto (2).

[0251] Al llevar a cabo los procesos mediados por la hidroxilasa de prolina descritos en el presente documento, el polipéptido diseñado se puede agregar a la mezcla de reacción en forma de una enzima purificada, una enzima parcialmente purificada, células completas transformadas con el gen o los genes que codifican la enzima, como extractos celulares y/o lisados de dichas células, y/o como una enzima inmovilizada sobre un soporte sólido. Se pueden emplear células completas transformadas con el gen o los genes que codifican la enzima hidroxilasa de prolina o los extractos celulares diseñados, sus lisados y enzimas aisladas en una variedad de formas diferentes, incluido el sólido (por ejemplo, liofilizado, secado por pulverización y similares) o semisólido (por ejemplo, una pasta cruda). Los extractos celulares o lisados celulares pueden purificarse parcialmente mediante precipitación (sulfato de amonio, polietilenimina, tratamiento térmico o similar, seguido de un procedimiento de desalinización antes de la liofilización (por ejemplo, ultrafiltración, diálisis y similares). Cualquiera de las preparaciones enzimáticas (incluyendo preparaciones de células completas) se pueden estabilizar mediante reticulación utilizando agentes de reticulación conocidos, como, por ejemplo, glutaraldehído o inmovilización en una fase sólida (por ejemplo, Eupergit C, y similares).

[0252] El gen o los genes que codifican los polipéptidos de hidroxilasa de prolina modificados se pueden transformar en la célula huésped por separado o juntos en la misma célula huésped. Por ejemplo, en algunas realizaciones, un conjunto de células huésped puede transformarse con uno o varios genes que codifican un polipéptido de hidroxilasa de prolina diseñado y otro conjunto puede transformarse con uno o más genes que codifican otro polipéptido de hidroxilasa de prolina. Ambos conjuntos de células transformadas se pueden utilizar juntos en la mezcla de reacción

en forma de células completas, o en forma de lisados o extractos derivados de ellos. En otras realizaciones, una célula huésped puede transformarse con uno o varios genes que codifican polipéptidos de hidroxilasa de prolina de ingeniería múltiple. En algunas realizaciones, los polipéptidos modificados pueden expresarse en forma de polipéptidos secretados y el medio de cultivo que contiene los polipéptidos secretados puede usarse para la reacción de la hidroxilasa de prolina.

[0253] La actividad mejorada y/o la estereoselectividad de los polipéptidos de hidroxilasa de prolina modificados proporcionan procesos en los que se puede lograr un mayor porcentaje de conversión con concentraciones más bajas del polipéptido modificado. En algunas realizaciones del proceso, las condiciones de reacción adecuadas comprenden una cantidad de polipéptido diseñado de aproximadamente 1% (p/p), 2% (p/p), 5% (p/p), 10% (p/p), 20% (p/p), 30% (p/p), 40% (p/p), 50% (p/p), 75% (p/p), 100% (p/p) o más de carga de compuestos de sustrato.

[0254] En algunas realizaciones, el polipéptido modificado está presente a aproximadamente 0,01 g/L hasta aproximadamente 50 g/L; aproximadamente 0,05 g/L a aproximadamente 50 g/L; aproximadamente 0,1 g/L a aproximadamente 40 g/L; aproximadamente 1 g/L a aproximadamente 40 g/L; aproximadamente 2 g/L a aproximadamente 40 g/L; aproximadamente 5 g/L a aproximadamente 40 g/L; aproximadamente 5 g/L a aproximadamente 30 g/L; aproximadamente 0,1 g/L a aproximadamente 10 g/L; aproximadamente 0,5 g/L a aproximadamente 10 g/L; aproximadamente 1 g/L a aproximadamente 10 g/L; aproximadamente 0,1 g/L hasta aproximadamente 5 g/L; aproximadamente 0,5 g/L a aproximadamente 5 g/L; o aproximadamente 0,1 g/L hasta aproximadamente 2 g/L. En algunas realizaciones, el polipéptido de hidroxilasa de prolina está presente a aproximadamente 0,01 g/L, 0,05 g/L, 0,1 g/L, 0,2 g/L, 0,5 g/L, 1, 2 g/L, 5 g/L, 10 g/L, 15 g/L, 20 g/L, 25 g/L, 30 g/L, 35 g/L, 40 g/L, o 50 g/L.

[0255] En algunas realizaciones, las condiciones de reacción también comprenden un metal de transición divalente capaz de servir como un cofactor en la reacción de oxidación. En general, el cofactor del metal de transición divalente es un ion ferroso, es decir, Fe^{+2} . El ion ferroso puede proporcionarse en diversas formas, tales como sulfato ferroso (FeSO_4), cloruro ferroso (FeCl_2), carbonato ferroso (FeCO_3) y las sales de ácidos orgánicos tales como citratos, lactatos y fumaratos. Una fuente ejemplar de sulfato ferroso es la sal de Mohr, que es sulfato de amonio ferroso $(\text{NH}_4)_2\text{Fe}(\text{SO}_4)_2$ y está disponible en formas anhidras e hidratadas (es decir, hexahidrato). Si bien el ión ferroso es el cofactor del metal de transición que se encuentra en la hidroxilasa de prolina natural y funciona de manera eficiente en las enzimas diseñadas, debe entenderse que otros metales de transición divalentes capaces de actuar como cofactor pueden usarse en los procesos. En algunas realizaciones, el cofactor de metal de transición divalente puede comprender Mn^{+2} y Cr^{+2} . En algunas realizaciones, las condiciones de reacción pueden comprender un cofactor de metal de transición divalente, particularmente Fe^{+2} , a una concentración de aproximadamente 0,1 mM a 10 mM, 0,1 mM a aproximadamente 5 mM, 0,5 mM a aproximadamente 5 mM, aproximadamente 0,5 mM a aproximadamente 3 mM o aproximadamente 1 mM a aproximadamente 2 mM. En algunas realizaciones, las condiciones de reacción comprenden una concentración de cofactor de metal de transición divalente de aproximadamente 0,1 mM, 0,2 μM , 0,5 mM, 1 mM, 1,5 mM, 2 mM, 3 mM, 5 mM, 7,5 mM o 10 mM. En algunas realizaciones, se pueden usar concentraciones más altas de cofactor de metal de transición divalente, por ejemplo hasta 50 mM o hasta 100 mM.

[0256] En algunas realizaciones, las condiciones de reacción pueden comprender además un reductor capaz de reducir el ión férrico, Fe^{+3} a ión ferroso, Fe^{+2} . En algunas realizaciones, el reductor comprende ácido ascórbico, típicamente ácido L-ascórbico. Si bien no se requiere ácido ascórbico para la reacción de hidroxilación, la actividad enzimática aumenta en su presencia. Sin pretender imponer ninguna teoría, se cree que el ascorbato mantiene o regenera la forma de la enzima- Fe^{+2} , que es la forma activa que media en la reacción de hidroxilación. En general, las condiciones de reacción pueden comprender una concentración de ácido ascórbico que corresponde proporcionalmente a la carga del sustrato. En algunas realizaciones, el ácido ascórbico está presente en al menos aproximadamente 0,1 veces, 0,2 veces, 0,3 veces, 0,5 veces, 0,75 veces, 1 vez, 1,5 veces o al menos 2 veces la cantidad molar de sustrato. En algunas realizaciones, el reductor, particularmente el ácido ascórbico L, está en una concentración de aproximadamente 0,001 M a aproximadamente 0,5 M, aproximadamente 0,01 M a aproximadamente 0,5 M, aproximadamente 0,01 M a aproximadamente 0,4 M, aproximadamente 0,1 a aproximadamente 0,4 M, o aproximadamente 0,1 a aproximadamente 0,3 M. En algunas realizaciones, el reductor, particularmente ácido ascórbico, está a una concentración de aproximadamente 0,001 M, 0,005 M, 0,01 M, 0,02 M, 0,03 M, 0,05 M, 0,1 M, 0,15 M, 0,2 M, 0,3 M, 0,4 M o 0,5 M.

[0257] En algunas realizaciones, las condiciones de reacción comprenden oxígeno molecular, es decir, O_2 . Sin pretender imponer ninguna teoría, se incorpora un átomo de oxígeno del oxígeno molecular al compuesto del sustrato para formar el compuesto producto hidroxilado. El O_2 puede estar presente de forma natural en la solución de reacción, o introducirse y/o complementarse en la reacción artificialmente. En algunas realizaciones, las condiciones de reacción pueden comprender una aireación forzada (por ejemplo, burbujeo) con aire, gas O_2 u otros gases que contienen O_2 . En algunas realizaciones, el O_2 en la reacción se puede aumentar al aumentar la presión de la reacción con O_2 o un gas que contiene O_2 . Esto se puede hacer llevando a cabo la reacción en un recipiente que puede presurizarse con gas O_2 . En algunas realizaciones, el gas O_2 se puede rociar a través de la solución de reacción a una velocidad de al menos 1 litro por hora (L/h), al menos 2 L/h, al menos 3 L/h, al menos 4 L/h, al menos 5 L/h, o más. En algunas realizaciones, el gas O_2 se puede rociar a través de la solución de reacción a una velocidad

de entre aproximadamente 1 L/h y 10 L/h, entre aproximadamente 2 L/h y 7 L/h, o entre aproximadamente 3 L/h y 5 L/h.

[0258] Durante el curso de la reacción, el pH de la mezcla de reacción puede cambiar. El pH de la mezcla de reacción se puede mantener a un pH deseado o dentro de un rango de pH deseado. Esto se puede hacer mediante la adición de un ácido o una base, antes y/o durante el curso de la reacción. Alternativamente, el pH puede controlarse utilizando un tampón. Por consiguiente, en algunas realizaciones, la condición de reacción comprende un tampón. Los tampones adecuados para mantener los intervalos de pH deseados son conocidos en la técnica e incluyen, a modo de ejemplo y sin limitación, borato, fosfato, ácido 2-(N-morfolino) etanosulfónico (MES), ácido 3-(N-morfolino)propanosulfónico (MOPS), acetato, trietanolamina y 2-amino-2-hidroximetilo-propano-1,3-diol (Tris), y similares. En algunas realizaciones, el tampón es fosfato. En algunas realizaciones del proceso, las condiciones de reacción adecuadas comprenden una concentración de tampón (por ejemplo, fosfato) de aproximadamente 0,01 a aproximadamente 0,4 M, de 0,05 a aproximadamente 0,4 M, de 0,1 a aproximadamente 0,3 M, o de aproximadamente 0,1 a aproximadamente 0,2 M. En algunas realizaciones, la condición de reacción comprende una concentración de tampón (por ejemplo, fosfato) de aproximadamente 0,01, 0,02, 0,03, 0,04, 0,05, 0,07, 0,1, 0,12, 0,14, 0,16, 0,18, 0,2, 0,3 o 0,4 M. En algunas realizaciones, las condiciones de reacción comprenden agua como un disolvente adecuado sin tampón presente.

[0259] En las realizaciones del proceso, las condiciones de reacción pueden comprender un pH adecuado. El pH deseado o el rango de pH deseado se pueden mantener mediante el uso de un ácido o base, un tampón apropiado, o una combinación de tamponamiento y adición de ácido o base. El pH de la mezcla de reacción se puede controlar antes y/o durante el curso de la reacción. En algunas realizaciones, las condiciones de reacción adecuadas comprenden una solución de pH de aproximadamente 4 a aproximadamente 10, pH de aproximadamente 5 a aproximadamente 10, pH de aproximadamente 5 a aproximadamente 9, pH de aproximadamente 6 a aproximadamente 9, pH de aproximadamente 6 a aproximadamente 8. En algunas realizaciones, las condiciones de reacción comprenden un pH de solución de aproximadamente 4, 4,5, 5, 5,5, 6, 6,5, 7, 7,5, 8, 8,5, 9, 9,5 o 10.

[0260] En las realizaciones de los procesos de este documento, se puede usar una temperatura adecuada para las condiciones de reacción, por ejemplo, teniendo en cuenta el aumento en la velocidad de reacción a temperaturas más altas y la actividad de la enzima durante el período de tiempo de reacción. Por consiguiente, en algunas realizaciones, las condiciones de reacción adecuadas comprenden una temperatura de aproximadamente 10°C a aproximadamente 60°C, aproximadamente 10°C a aproximadamente 55°C, aproximadamente 15°C a aproximadamente 60°C, aproximadamente 20°C a aproximadamente 60°C, aproximadamente 20°C a aproximadamente 55°C, aproximadamente 25°C a aproximadamente 55°C, o aproximadamente 30°C a aproximadamente 50°C. En algunas realizaciones, las condiciones de reacción adecuadas comprenden una temperatura de aproximadamente 10°C, 15°C, 20°C, 25°C, 30°C, 35°C, 40°C, 45°C, 50°C, 55°C, o 60°C. En algunas realizaciones, la temperatura durante la reacción enzimática se puede mantener a una temperatura específica a lo largo del curso de la reacción. En algunas realizaciones, la temperatura durante la reacción enzimática se puede ajustar sobre un perfil de temperatura durante el curso de la reacción.

[0261] Los procesos de la descripción se llevan a cabo generalmente en un disolvente. Los disolventes adecuados incluyen agua, soluciones tampón acuosas, disolventes orgánicos, disolventes poliméricos y/o sistemas de codisolventes, que generalmente comprenden disolventes acuosos, disolventes orgánicos y/o disolventes poliméricos. El solvente acuoso (agua o sistema de co-solvente acuoso) puede estar tamponado con pH o sin amortiguar. En algunas realizaciones, los procesos que utilizan los polipéptidos de hidroxilasa de prolina diseñados se pueden llevar a cabo en un sistema acuoso de co-solvente que comprende un solvente orgánico (por ejemplo, etanol, isopropanol (IPA), dimetilsulfóxido (DMSO), dimetilformamida (DMF) acetato etílico, acetato de butilo, 1-octanol, heptano, octano, metilo t butilo éter (MTBE), tolueno y similares), disolventes iónicos o polares (por ejemplo, tetrafluoroborato de 1-etilo 4 metilimidazolio, tetrafluoroborato de 1-butilo-3-metilimidazolio, hexafluorofosfato de 1-butilo-3-metilimidazolio, glicerol, polietilenglicol y similares). En algunas realizaciones, el co-disolvente puede ser un disolvente polar, tal como un poliol, dimetilsulfóxido (DMSO) o alcohol inferior. El componente de co-disolvente no acuoso de un sistema de co-disolvente acuoso puede ser miscible con el componente acuoso, proporcionando una sola fase líquida, o puede ser parcialmente miscible o inmiscible con el componente acuoso, proporcionando dos fases líquidas. Los sistemas de co-solvente acuosos ejemplares pueden comprender agua y uno o más co-solventes seleccionados de entre un solvente orgánico, solvente polar y solvente de poliol. En general, el componente co-solvente de un sistema acuoso co-solvente se elige de tal manera que no inactive adversamente la enzima hidroxilasa de prolina bajo las condiciones de reacción. Los sistemas de co-solventes apropiados pueden identificarse fácilmente midiendo la actividad enzimática de la enzima hidroxilasa de prolina diseñada específica con un sustrato definido de interés en el sistema de solvente candidato, utilizando un ensayo de actividad enzimática, como los descritos en este documento.

[0262] En algunas realizaciones del proceso, las condiciones de reacción adecuadas comprenden un co-solvente acuoso, donde el co-solvente comprende DMSO en aproximadamente 1% a aproximadamente 50% (v/v), aproximadamente 1 a aproximadamente 40% (v/v), de aproximadamente 2% a aproximadamente 40% (v/v), de aproximadamente 5% a aproximadamente 30% (v/v), de aproximadamente 10% a aproximadamente 30% (v/v) o de aproximadamente 10% a aproximadamente 20% (v/v). En algunas realizaciones del proceso, las condiciones de

reacción adecuadas pueden comprender un co-disolvente acuoso que comprende DMSO a aproximadamente el 1% (v/v), aproximadamente el 5% (v/v), aproximadamente el 10% (v/v), aproximadamente el 15% (v/v), aproximadamente el 20% (v/v), aproximadamente el 25% (v/v), aproximadamente el 30% (v/v), aproximadamente el 35% (v/v), aproximadamente el 40% (v/v), aproximadamente el 45% (v/v), o aproximadamente el 50% (v/v).

[0263] En algunas realizaciones, las condiciones de reacción pueden comprender un tensioactivo para estabilizar o mejorar la reacción. Los tensioactivos pueden comprender tensioactivos no iónicos, catiónicos, aniónicos y/o anfifílicos. Los tensioactivos ejemplares incluyen, a modo de ejemplo y sin limitación, nonilo fenoxipoliétoxietanol (NP40), Triton X-100, polioxietileno-estearilamina, bromuro de cetiltrimetilamonio, oleilamidossulfato de sodio, polioxietileno-sorbitanmonostearato, hexadecildimetilamina, etc. Puede emplearse cualquier tensioactivo que podría estabilizar o potenciar la reacción. La concentración del tensioactivo a emplear en la reacción puede ser generalmente de 0,1 a 50 mg/ml, particularmente de 1 a 20 mg/ml.

[0264] En algunas realizaciones, las condiciones de reacción pueden incluir un agente antiespumante, que ayuda a reducir o evitar la formación de espuma en la solución de reacción, tal como cuando las soluciones de reacción se mezclan o burbujan. Los agentes antiespumantes incluyen aceites no polares (por ejemplo, minerales, siliconas, etc.), aceites polares (por ejemplo, ácidos grasos, alquilaminas, alquilo amidas, alquilo sulfatos, etc.) e hidrófobos (por ejemplo, sílice tratada, polipropileno), etc.), algunos de los cuales también funcionan como tensioactivos. Los agentes antiespumantes ejemplares incluyen, Y-30® (Dow Corning), copolímeros de poliglicol, alcoholes oxi/etoxilados y polidimetilsiloxanos. En algunas realizaciones, el antiespumante puede estar presente en aproximadamente 0,001% (v/v) a aproximadamente 5% (v/v), aproximadamente 0,01% (v/v) a aproximadamente 5% (v/v), aproximadamente 0,1 % (v/v) a aproximadamente 5% (v/v), o aproximadamente 0,1% (v/v) a aproximadamente 2% (v/v). En algunas realizaciones, el agente antiespumante puede estar presente a aproximadamente el 0,001% (v/v), aproximadamente el 0,01% (v/v), aproximadamente el 0,1% (v/v), aproximadamente el 0,5% (v/v), aproximadamente 1% (v/v), aproximadamente 2% (v/v), aproximadamente 3% (v/v), aproximadamente 4% (v/v), o aproximadamente 5% (v/v) o más como deseable para promover la reacción.

[0265] Las cantidades de reactivos usados en la reacción de hidroxilasa generalmente variarán dependiendo de las cantidades de producto deseadas, y concomitantemente la cantidad de sustrato de hidroxilasa de prolina empleada. Los expertos en la técnica entenderán fácilmente cómo variar estas cantidades para adaptarlas al nivel deseado de productividad y escala de producción.

[0266] En algunas realizaciones, el orden de adición de los reactivos no es crítico. Los reactivos pueden agregarse juntos al mismo tiempo a un solvente (por ejemplo, solvente monofásico, sistema de co-solvente acuoso bifásico y similares), o alternativamente, algunos de los reactivos pueden agregarse por separado, y algunos juntos en diferentes momentos. Por ejemplo, el cofactor, el co-sustrato, la hidroxilasa de prolina y el sustrato se pueden agregar primero al solvente.

[0267] Los reactivos sólidos (por ejemplo, enzimas, sales, etc.) pueden proporcionarse a la reacción en una variedad de formas diferentes, que incluyen polvo (por ejemplo, liofilizado, secado por aspersión y similares), solución, emulsión, suspensión, y similares. Los reactivos se pueden liofilizar o secar por pulverización fácilmente utilizando métodos y equipos que son conocidos por los expertos en la técnica. Por ejemplo, la solución de proteína se puede congelar a -80°C en pequeñas alícuotas, y luego agregarse a una cámara de liofilización previamente enfriada, seguido de la aplicación de un vacío.

[0268] Para mejorar la eficacia de la mezcla cuando se usa un sistema de co-disolvente acuoso, la hidroxilasa de prolina y el cofactor se pueden agregar y mezclar primero en la fase acuosa. La fase orgánica puede luego agregarse y mezclarse, seguido de la adición del sustrato de hidroxilasa de prolina y el co-sustrato. Alternativamente, el sustrato de hidroxilasa de prolina se puede premezclar en la fase orgánica, antes de la adición a la fase acuosa.

[0269] El proceso de hidroxilación generalmente se deja avanzar hasta que la conversión adicional del sustrato a producto hidroxilado no cambie significativamente con el tiempo de reacción, por ejemplo, menos del 10% del sustrato que se está convirtiendo, o menos del 5% del sustrato que se está convirtiendo). En algunas realizaciones, se permite que la reacción continúe hasta que haya una conversión completa o casi completa del sustrato en producto. La transformación del sustrato en producto se puede monitorear usando métodos conocidos mediante la detección de sustrato y/o producto, con o sin derivatización. Los métodos analíticos adecuados incluyen cromatografía de gases, HPLC, MS y similares.

[0270] En algunas realizaciones del proceso, las condiciones de reacción adecuadas comprenden una carga de sustrato de al menos aproximadamente 5 g/L, 10 g/L, 20 g/L, 30 g/L, 40 g/L, 50 g/L, 60 g/L, 70 g/L, 100 g/L, o más, y en donde el método resulta en al menos aproximadamente 50%, 60%, 70%, 80%, 90%, 95% o más de conversión de sustrato de compuesto a producto compuesto en aproximadamente 48 h o menos, en aproximadamente 36 h o menos, o en aproximadamente 24 h o menos.

[0271] Los polipéptidos de hidroxilasa de prolina modificados de la presente descripción cuando se usan en el

proceso en condiciones de reacción adecuadas dan como resultado un exceso del producto hidroxilado en *cis* en al menos el 90%, 95%, 96%, 97%, 98%, 99%, o mayor exceso diastereomérico sobre el producto *trans*-hidroxilado. En algunas realizaciones, no se forma una cantidad detectable de compuesto *trans*-hidroxilado.

[0272] En realizaciones adicionales de los procesos para convertir el compuesto de sustrato en un compuesto de producto hidroxilado usando los polipéptidos de hidroxilasa de prolina diseñados, las condiciones de reacción adecuadas pueden comprender una carga de sustrato inicial en la solución de reacción que luego se pone en contacto con el polipéptido. Esta solución de reacción se complementa luego con un compuesto de sustrato adicional como una adición continua o discontinua a lo largo del tiempo a una velocidad de al menos aproximadamente 1 g/L/h, al menos aproximadamente 2 g/L/h, al menos aproximadamente 4 g/L/h, al menos aproximadamente 6 g/L/h, o superior. Por lo tanto, de acuerdo con estas condiciones de reacción adecuadas, el polipéptido se agrega a una solución que tiene una carga de sustrato inicial de al menos aproximadamente 20 g/L, 30 g/L o 40 g/L. A continuación, a esta adición de polipéptido le sigue una adición continua de sustrato adicional a la solución a una velocidad de aproximadamente 2 g/L/h, 4 g/L/h, o 6 g/L/h hasta una carga final de sustrato mucho mayor. se alcanza al menos aproximadamente 30 g/L, 40 g/L, 50 g/L, 60 g/L, 70 g/L, 100 g/L, 150 g/L, 200 g/L o más. En consecuencia, en algunas realizaciones del proceso, las condiciones de reacción adecuadas comprenden la adición del polipéptido a una solución que tiene una carga de sustrato inicial de al menos aproximadamente 20 g/L, 30 g/L, o 40 g/L seguido de la adición de un sustrato adicional a la solución a una velocidad de aproximadamente 2 g/L/h, 4 g/L/h, o 6 g/L/h hasta que se alcanzan una carga de sustrato final de al menos aproximadamente 30 g/L, 40 g/L, 50 g/L, 60 g/L, 70 g/L, 100 g/L o más. Esta condición de reacción de suplementación de sustrato permite lograr mayores cargas de sustrato mientras que se mantienen altas tasas de conversión de sustrato a producto hidroxilado de al menos aproximadamente 50%, 60%, 70%, 80%, 90% o más de conversión de sustrato. En algunas realizaciones de este proceso, el sustrato agregado se encuentra en una solución que comprende α -cetoglutarato en una cantidad equimolar o superior del sustrato adicional agregado.

[0273] En algunas realizaciones de los procesos, la reacción que usa un polipéptido de hidroxilasa de prolina diseñada puede comprender las siguientes condiciones de reacción adecuadas: (a) carga de sustrato de aproximadamente 5 g/L a 30 g/L; (b) aproximadamente 0,1 g/L a 10 g/L del polipéptido diseñado; (c) aproximadamente 19 g/L (0,13 M) a 57 g/L (0,39 M) de α -cetoglutarato; (d) aproximadamente 14 g/L (0,08 M) a 63 g/L (0,36 M) de ácido ascórbico; (e) aproximadamente 1,5 g/L (3,8 mM) a 4,5 g/L (11,5 mM) de FeSO_4 ; (f) un pH de aproximadamente 6 a 7; (g) temperatura de aproximadamente 20° a 40°C; y (h) tiempo de reacción de 2-24 h.

[0274] En algunas realizaciones de los procesos, la reacción que usa un polipéptido de hidroxilasa de prolina diseñada puede comprender las siguientes condiciones de reacción adecuadas: (a) carga de sustrato de aproximadamente 10 g/L a 100 g/L; (b) de aproximadamente 1 g/L a aproximadamente 50 g/L de polipéptido diseñado; (c) α -cetoglutarato en aproximadamente 1 a 2 equivalentes molares de compuesto de sustrato; (d) ácido ascórbico en aproximadamente 0,25 a 0,75 equivalentes molares de compuesto de sustrato; (e) de aproximadamente 0,5 mM a aproximadamente 12 mM de FeSO_4 ; (f) pH de aproximadamente 6 a 8; (g) temperatura de aproximadamente 20° a 40°C; y (h) tiempo de reacción de 6 a 120 h.

[0275] En algunas realizaciones, se llevan a cabo componentes de reacción adicionales o técnicas adicionales para complementar las condiciones de reacción. Estos pueden incluir tomar medidas para estabilizar o prevenir la inactivación de la enzima, reducir la inhibición del producto, cambiar el equilibrio de la reacción a la formación del producto hidroxilado.

[0276] En otras realizaciones, cualquiera de los procesos descritos anteriormente para la conversión de compuesto de sustrato en compuesto de producto puede comprender además una o más etapas seleccionadas de: extracción; aislamiento; purificación; y cristalización del producto compuesto. Métodos, técnicas y protocolos para extraer, aislar, purificar y/o cristalizar el producto hidroxilado a partir de mezclas de reacciones biocatalíticas producidas por los procesos descritos anteriormente son conocidos por los expertos en la técnica y/o se accede a ellos mediante experimentación rutinaria. Además, los métodos ilustrativos se proporcionan en los ejemplos a continuación.

[0277] Se ilustran varias características y realizaciones de la descripción en los siguientes ejemplos representativos, que pretenden ser ilustrativos, y no limitativos.

6. EJEMPLOS

Ejemplo 1: Síntesis, optimización y cribado de polipéptidos de hidroxilasa de prolina diseñados

[0278] Síntesis y optimización de genes: la secuencia de polinucleótidos que codifica el polipéptido de *cis*-4-hidroxilasa de prolina de tipo salvaje informado de *Sinorhizobium meliloti*, como se representa por la SEQ ID NO: 2, se sintetizó como el gen de la SEQ ID NO: 1, el gen sintético de la SEQ ID NO: 1 se clonó en un sistema de vector pCK110900 (véase, por ejemplo, la publicación de solicitud de patente estadounidense 20060195947) y se expresó posteriormente en *E. coli* W3110fhua. La *E. coli* W3110 expresa los polipéptidos de hidroxilasa de prolina bajo el control del promotor lac. Sobre la base de las comparaciones de secuencia con otras hidroxilasas de prolina y el modelado computarizado de la estructura enzimática acoplada al sustrato prolina, se identificaron las posiciones de

residuos asociadas con el sitio activo, los bucles peptídicos, la interfaz solución/sustrato y las posibles posiciones de estabilidad y se sometieron a mutagénesis. Estas variantes de la primera ronda se seleccionaron en condiciones de ensayo HTP con ácido (2S)-piperidina-2-carboxílico como sustrato. Se identificaron variantes con mayor actividad enzimática y/o expresión. También se generaron dos polinucleótidos optimizados en codones adicionales que codifican la secuencia de aminoácidos de la enzima natural (es decir, la SEQ ID NO: 3 y 5) con fines de comparación. Los polinucleótidos 3 y 5 optimizados de codón que expresan la hidroxilasa de cis-4-prolina de origen natural mostraron una expresión aumentada con respecto al polinucleótido de la SEQ ID NO: 1. Las diferencias de residuos de la primera ronda de selección se combinaron en varias permutaciones y se seleccionaron para obtener propiedades mejoradas en las condiciones del Ensayo HTP, el Ensayo SFP y el Ensayo DSP. Las secuencias polipeptídicas de hidroxilasa de la prolina y las mutaciones específicas y las actividades relativas obtenidas de las pantallas se enumeran en la **Tabla 2A, 2B, 2C, 2D, 2E, 2F, 2G y 2H**.

Ejemplo 2: Producción de hidroxilasas de prolina modificadas

[0279] Los polipéptidos de hidroxilasa de prolina modificados se produjeron en *E. coli* W3110 bajo el control del promotor lac. Las preparaciones de enzimas para los ensayos de HTP, DSP y SFP se realizaron de la siguiente manera.

[0280] Preparación de crecimiento, expresión y de lisado de alto rendimiento. Las células se recogieron y se cultivaron durante la noche en medio LB que contenía 1% de glucosa y 30 µg/ml de cloranfenicol (CAM), 30°C, 200 rpm, 85% de humedad. Una alícuota de 20 µl de crecimiento durante la noche se transfirió a una placa de pocillos profundos que contenía 380 µl de medio de cultivo 2xTB que contenía 30 µg/ml de CAM, 1 mM de IPTG y se incubó durante aproximadamente 18 horas a 30°C, 200 rpm, 85% de humedad. Los cultivos celulares se centrifugaron a 4000 rpm, 4°C durante 10 min, y los medios desechados. Los sedimentos celulares se resuspendieron en 100 µl de tampón de lisis (50 mM de tampón fosfato, pH 6,3, que contenía sal de Mohr 100 µM (es decir, (NH₄)₂Fe(SO₄)₂), 0,5 mg/ml de PMBS (sulfato de polimixina B) y 1 mg/mL de Lisozima). El tampón de lisis se preparó recién añadiendo a 60 ml de tampón fosfato 50 mM, pH 6,3, 60 mg de lisozima y 30 mg de PMBS. Después de mezclar la solución de lisozima, se agregaron 0,6 ml de solución de sal de Mohr 10 mM (en H₂O).

[0281] Producción de polvos de matraz de agitación (SFP): Se usó un procedimiento de matraz de agitación para generar polvos polipeptídicos de polipropilasa usados en ensayos de cribado secundarios o en los procesos biocatalíticos descritos en el presente documento. El polvo del matraz de agitación (SFP) proporciona una preparación más purificada (por ejemplo, hasta el 30% de la proteína total) de la enzima diseñada en comparación con el lisado celular utilizado en los ensayos de HTP. Una sola colonia de *E. coli* que contiene un plásmido que codifica un polipéptido de interés se inocula en 50 ml de caldo Luria Bertani que contiene 30 µg/ml de cloranfenicol y 1% de glucosa. Las células se cultivan durante la noche (al menos 16 horas) en una incubadora a 30°C con agitación a 250 rpm. El cultivo se diluye en 250 ml de Terrific Broth (12 g/L de bacto-triptona, 24 g/L de extracto de levadura, 4 ml/l de glicerol, fosfato de potasio 65 mM, pH 7, MgSO₄ 1 mM) que contiene 30 µg/ml de cloranfenicol en un matraz de 1 litro a una densidad óptica de 600 nm (OD₆₀₀) de 0,2 y se deja crecer a 30°C. La expresión del gen de la hidroxilasa de prolina se induce mediante la adición de isopropilo-β-D-tiogalactósido ("IPTG") a una concentración final de 1 mM cuando la OD₆₀₀ del cultivo es de 0,6 a 0,8. La incubación se continúa durante la noche (al menos 16 horas). Las células se recogen por centrifugación (5.000 rpm, 15 min, 4°C) y el sobrenadante se descarta. El sedimento celular se resuspende con un volumen igual de tampón de fosfato de potasio 50 mM frío (4°C), pH 6,3, y se recoge por centrifugación como anteriormente. Las células lavadas se resuspenden en dos volúmenes del tampón de fosfato de potasio 50 mM frío, pH 6,3 y se pasan a través de un French Press dos veces a 12.000 psi mientras que se mantienen a 4°C. Los residuos celulares se eliminan por centrifugación (9.000 rpm, 45 minutos, 4°C). El sobrenadante de lisado transparente se recoge y almacena a -20°C. La liofilización del lisado claro congelado proporciona un polvo de matraz de agitación seco de polipéptido en bruto. Alternativamente, el sedimento celular (antes o después del lavado) se puede almacenar a 4°C o -80°C.

[0282] Producción de polvos de proceso descendente (DSP): Los polvos DSP proporcionan una preparación más purificada de la enzima hidroxilasa de prolina modificada en comparación con el lisado celular utilizado en los ensayos de HTP o SFP. La fermentación a gran escala de la hidroxilasa de prolina diseñada para la producción de polvos DSP (~100 - 120 g a partir de 10 L) se puede llevar a cabo como un lote corto seguido de un proceso de alimentación por lotes de acuerdo con los métodos estándar de bioprocesos. En resumen, la expresión de hidroxilasa de prolina se induce mediante la adición de IPTG a una concentración final de 1 mM. Después de la fermentación, las células se recolectan y se resuspenden en tampón fosfato 50 mM, luego se rompen mecánicamente por homogeneización. Los residuos celulares y el ácido nucleico se flocculan con polietilénimina (PEI) y la suspensión se clarifica por centrifugación. El sobrenadante transparente resultante se concentra utilizando una membrana de ultrafiltración de flujo cruzado tangencial para eliminar las sales y el agua. El concentrado de enzimas parcialmente purificado y concentrado puede luego secarse en un liofilizador y envasarse (por ejemplo, en recipientes de polietileno).

Ejemplo 3: Procedimientos analíticos

[0283] Método 1 - Análisis por HPLC de las reacciones de ensayo de HTP: En un bloque de ensayo de 96 pocillos

profundos, se diluyeron 10 µl de solución de reacción con 230 µl de solución de bicarbonato de sodio al 5% seguido de 160 µl de solución de cloruro de dansilo (6 mg/mL de cloruro de dansilo en MeCN). La placa se selló térmicamente, se centrifugó y se colocó en una incubadora a 55°C durante 45 minutos. La solución de reacción cambia de amarillo a amarillo claro cuando se completa la derivación con cloruro de dansilo. En los casos en que la solución permaneció amarilla, la placa se calentó durante otros 15 minutos. Después de la incubación, la placa se centrifugó durante 5 minutos a 4.000 rpm. Una alícuota de 200 µl de sobrenadante se transfirió a una placa de 96 Corning para el análisis de HPLC. La concentración final del sustrato fue inferior a 0,25 g/L.

[0284] La reacción inactivada se sometió a análisis por HPLC en las siguientes condiciones.

Columna	Agilent Poroshell120 SB-C18 50 x 4,6mm (2,7 11m) con columna protectora			
Temperatura	30°C			
Fase móvil	Solución A: 2 mM Acetato de amonio pH 3,3 Solución B: Acetonitrilo Perfil de fase móvil			
Postiempo	Tiempo: min	A%	B%	Vel. de flujo: mL/min
	0,00	80	20	2,0
	2,00	30	70	2,0
	3,00	30	70	2,0
	1 min			
Detección longitud de onda	250 nm			
Temperatura de columna	25°C			
Volumen de inyección	10 µL			
Tiempo de ejecución total	4 min			
Factor de respuesta (área de sustrato/área de producto)	N.A.			

[0285] La conversión del compuesto (2) en el compuesto (1) se determinó a partir de los cromatogramas resultantes de la siguiente manera:

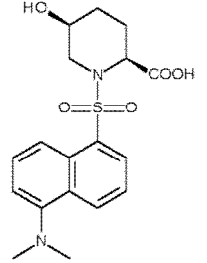
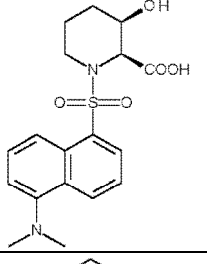
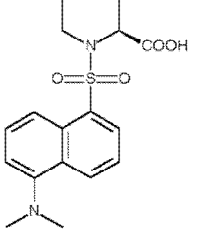
$$\% \text{ Conversión} = \{(\text{RF} \times \text{Área de producto}) / [(\text{RF} \times \text{Área de producto}) + \text{Área de sustrato}]\} \times 100$$

dónde

$$\text{Factor de respuesta (RF)} = \text{Área de sustrato} / \text{Área de producto}.$$

[0286] Este método se utilizó para la identificación rápida para la conversión del ácido (2S)-piperidina-2-carboxílico en ácido hidroxipiperidina-2-carboxílico.

[0287] Los perfiles cromatográficos de elución, indicados como "Tiempo de respuesta", son los siguientes:

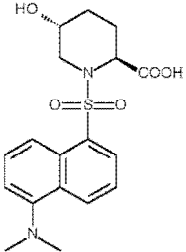
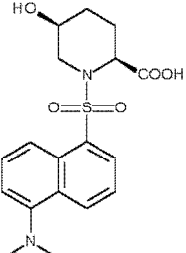
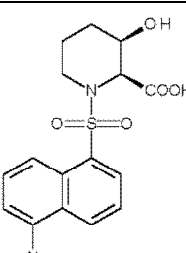
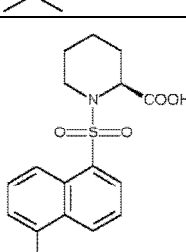
5	Ácido (2S,5S)-1-(5-(dimetilamino)naftaleno-1-ilsulfonilo)-5-hidroxipiperidina-2-carboxílico		Tiempo de respuesta: 1,59 min
10			
15	Ácido (2S,3R)-1-(5-(dimetilamino)naftaleno-1-ilsulfonilo)-3-hidroxipiperidina-2-carboxílico		Tiempo de respuesta: 1,76 min
20			
25	Ácido (5)-1-(5-(dimetilamino)naftaleno-1-ilsulfonilo)piperidina-2-carboxílico		Tiempo de respuesta: 2,23 min
30			

[0288] Método 2 - Análisis por HPLC de las reacciones de DSP y SFP: 10 µl de solución de reacción de una reacción de DSP o SFP se pipeteó en un tubo Eppendorf de 1,5 ml y se diluyó con 230 µl de bicarbonato de sodio al 5%. Luego se añadió una alícuota de 160 µl de solución de cloruro de dansilo (6 mg/ml de cloruro de dansilo en MeCN). El tubo se calentó con una cubierta abierta a 55°C durante al menos 30 minutos en un bloque de calentamiento para asegurar la derivación completa, como lo indica el cambio en el color de la mezcla de derivación de un color amarillo a un amarillo claro. El tubo se agitó en vórtex y luego se centrifugó a 12.000 rpm durante 5 minutos. Se transfirió una parte alícuota de 200 µl de sobrenadante a un vial de HPLC de 2 ml con el inserto. El vial se sometió a fase inversa HPLC-UV para su análisis, como se describe a continuación. La concentración final del sustrato fue inferior a 0,25 g/l.

[0289] La reacción inactivada se sometió a análisis por HPLC en las siguientes condiciones.

Columna	Supelco Ascentis Express C18 100 x 4.6mm (2,7 11m), adjunto			
Fase móvil	Solución A: 2 mM Acetato de amonio, pH 3,3 Solución B: Acetonitrilo Perfil de fase móvil			
	Tiempo: min	A%	B%	Vel. de flujo: mL/min
	0	80	20	1,0
	9,0	27,5	72,5	1,0
	9,1	0	100	1,0
	12,0	0	100	1,0
	12,1	80	20	1,0
Postiempo	2,90 min			
Longitud de onda de detección	250 nm			
Temperatura de columna	25°C			
Volumen de inyección	10 mL			
Tiempo de ejecución total	15,0 min			
Factor de respuesta (RF) (Área de sustrato/Área de producto)	N.A.			

[0290] Los perfiles cromatográficos de elución son los siguientes:

Ácido (2S, 5R)-1-(5-(dimetilamino)naftaleno-1-ilsulfonilo)-5-hidroxipiperidina-2-carboxílico		Tiempo de respuesta: 5,16 min
Ácido (2S, 5S)-1-(5-(dimetilamino)naftaleno-1-ilsulfonilo)-5-hidroxipiperidina-2-carboxílico		Tiempo de respuesta: 5,57 min
Ácido (2S,3R)-1-(5-(dimetilamino)naftaleno-1-ilsulfonilo)-3-hidroxipiperidina-2-carboxílico		Tiempo de respuesta: 6,02 min
Ácido (5)-1-(5-(dimetilamino)naftaleno-1-ilsulfonilo)piperidina-2-carboxílico		Tiempo de respuesta: 8,17 min

Ejemplo 4: Cribado de alto rendimiento (HTP) de hidroxilasas de prolina para la conversión del compuesto (2) en el compuesto (1)

[0291] Ensayos de cribado de HTP: el cribado de alto rendimiento utilizado para guiar la selección primaria de variantes se llevó a cabo en placas de 96 pocillos utilizando lisados celulares. Se utilizaron dos condiciones, la condición A y la condición B.

[0292] Las reacciones de la condición A se llevaron a cabo como sigue. Las células se cultivaron en placas de 96 pocillos como se describe anteriormente y los lisados se prepararon dispensando 100 µl de tampón de lisis en cada pocillo. El tampón de lisis se preparó disolviendo 30 mg de lisozima y 15 mg de PMBS en 30 ml de tampón fosfato 50 mM, pH 6,3. Se añadió un volumen de 600 µl de sal de Mohr 10 mM recién preparada en agua estéril a la solución de lisozima. La placa se selló térmicamente y luego se agitó en un agitador de placa de titulación a velocidad n° 8 durante 2 horas a temperatura ambiente. Posteriormente, la placa se centrifugó rápidamente para asentar el lisado en el fondo de la placa. Este lisado crudo se usaría para la reacción.

[0293] Las reacciones de la condición A en una escala de 200 µl se llevaron a cabo en placas de 98 pocillos. Se preparó una solución madre de premezcla disolviendo 1,33 g de ácido α-cetoglutarico y 1,47 g de ácido L-ascórbico en 31,5 ml de tampón fosfato 50 mM, pH 6,3 (pH ajustado usando KOH). Después de mezclarse, el pH se ajustó a 6,3 con KOH. A la premezcla de pH ajustado, se agregaron 41,16 mg de sal de Mohr. La solución se vuelve turbia

debido a la baja solubilidad de la sal de Mohr en disolventes acuosos.

[0294] A cada 100 µl de lisados crudos preparados anteriormente, se agregaron 90 µl de la solución madre de premezcla a cada pocillo, seguido inmediatamente por 10 µL/pocillo de una solución madre de sustrato de 200 g/L, es decir, el compuesto (2) preparado en tampón fosfato 50 mM, pH 6,3. La placa se selló con un sello AirPore (Qiagen) y la reacción se dejó proceder durante toda la noche en un agitador de títulos, con una velocidad de agitación de N° 2,5 a temperatura ambiente.

[0295] La condición A tiene los siguientes parámetros de reacción finales: (a) 10 g/L de carga de sustrato; (b) 19 g/L de ácido α-cetoglutarico; (c) 21 g/L de ácido ascórbico; (d) sal de Mohr 1,5 mM; (e) tampón fosfato 50 mM, pH 6,3 (pH ajustado con KOH); (f) temperatura ambiente (20°C a 25°C); y (g) tiempo de reacción de aprox. 24 h.

[0296] Después de la incubación durante la noche, las placas se centrifugaron a 4.000 rpm durante 5 minutos a temperatura ambiente. Las muestras de reacción se derivaron y se detuvieron mediante la alícuota de 10 µl de la mezcla de reacción transparente en una placa de 96 pocillos profundos que contenía 230 µl/pocillo de bicarbonato de sodio al 5% (ac.). Se añadió un volumen de 160 µl de 6 mg/ml de cloruro de dansilo en MeCN a cada pocillo, la placa se selló por calor y luego se centrifugó rápidamente para depositar la solución de reacción en el fondo del pozo. La placa se calentó luego a 55°C durante al menos 45 minutos sin agitación, y se centrifugó a 4.000 rpm durante 10 minutos a temperatura ambiente. Un volumen de 200 µl de la solución derivada se transfirió a una placa de fondo redondo 96 y se envió para análisis de HPLC.

[0297] Las reacciones de la condición B se llevaron a cabo de la siguiente manera. Las células se cultivaron en placas de 96 pocillos como se describió anteriormente y los lisados se prepararon dispensando 100 µl de tampón de lisis en cada pocillo. El tampón de lisis se preparó recientemente disolviendo 30 mg de lisozima y 15 mg de PMBS en 30 ml de tampón fosfato 50 mM, pH 6,3, seguido de un volumen de 600 µl de sal de Mohr 10 mM, recién preparado en agua estéril. La placa de lisis se selló térmicamente y luego se agitó en un agitador de placa de titulación a velocidad N° 8 durante 2 horas a temperatura ambiente. Posteriormente, la placa se centrifugó rápidamente para asentar el lisado en el fondo de la placa. Este lisado crudo de 100 µl debía usarse para la reacción.

[0298] Las reacciones de la condición B en una escala de 200 µl se llevaron a cabo en placas de 96 pocillos. Una solución de reserva de premezcla fue preparada disolviendo 1,33 g de ácido α-cetoglutarico y 1,47 g de ácido L-ascórbico en 31,5 ml de tampón fosfato 50 mM, pH 6,3 (pH ajustado utilizando KOH). Después de mezclarse, el pH se ajustó a 6,3 con KOH. A la premezcla de pH ajustado, se agregaron 41,16 mg de sal de Mohr.

[0299] A cada 100 µl de lisados brutos preparados anteriormente, se agregaron 90 µl de la solución madre de premezcla a cada pocillo, seguido inmediatamente por 10 µL/pocillo de una solución madre de sustrato de 200 g/L, es decir, el compuesto (2) preparado en tampón fosfato 50 mM, pH 6,3. La placa se selló con un sello AirPore® (Qiagen) y la reacción se dejó proceder durante la noche en un agitador de títulos, con una velocidad de agitación de N° 2,5 a temperatura ambiente.

[0300] La condición B tiene los siguientes parámetros de reacción finales: (a) 10 g/L de carga de sustrato; (b) 19 g/L de ácido α-cetoglutarico; (c) 21 g/L de ácido ascórbico; (d) sal de Mohr 1,5 mM; (e) tampón fosfato 50 mM, pH 6,3 (pH ajustado con KOH); (f) temperatura ambiente (20°C a 25°C); y (g) tiempo de reacción de ca 24 h.

[0301] Después de la incubación general, las placas se centrifugaron a 4.000 rpm durante 5 minutos a temperatura ambiente. Las muestras de reacción se derivaron y se detuvieron mediante la alícuota de 10 µl de la mezcla de reacción transparente en una placa de 96 pocillos profundos que contenía 230 µl/pocillo de bicarbonato de sodio al 5% (ac.). Se añadió un volumen de 160 µl de 6 mg/ml de cloruro de dansilo en MeCN a cada pocillo, la placa se selló por calor y luego se centrifugó rápidamente para depositar la solución de reacción en el fondo del pozo. La placa se calentó luego a 55°C durante al menos 45 minutos con agitación en un Infors HT Microtron con una velocidad de agitación de 500 rpm. Las placas se centrifugaron a 4.000 rpm durante 10 minutos a temperatura ambiente. Un volumen de 200 µl de la solución derivada se transfirió a una placa de fondo redondo 96 y se envió para HPLC para su análisis.

Ejemplo 5: Proceso para la conversión del Compuesto (2) en el Compuesto (1) usando preparaciones en polvo de matraz de agitación (SFP)

[0302] Se llevó a cabo una reacción a escala de 200 µl usando polvo de enzima SFP como sigue. Se preparó recientemente una solución madre de premezcla disolviendo 1,05 g de ácido α-cetoglutarico, 420 mg de ácido L-ascórbico y 600 mg de sustrato (2S)-piperidina-2-carboxílico en 10 ml de tampón fosfato 50 mM, pH 6,3 (pH ajustado usando KOH). Después de mezclar bien la solución, el pH se ajustó a 6,3 con KOH. A la solución de premezcla de pH ajustado, se agregaron 45 mg de sal de Mohr.

[0303] Se preparó una solución madre de enzima disolviendo 20 mg de polvo de enzima SFP en 2 ml de tampón fosfato 50 mM, pH 6,3. Para iniciar la reacción, se agregaron 100 µl de solución de enzima a una placa, seguido de

100 µl de solución madre de premezcla para un volumen de reacción final de 200 µl. La placa se selló con un sello AirPore® (Qiagen) y se permitió que la reacción transcurriera durante la noche (aprox. 24 h) con agitación en un mezclador de placas de titulación (velocidad N° 2.5) a temperatura ambiente.

[0304] La condición del Ensayo de SFP (es decir, la Condición C) tiene los siguientes parámetros finales: (a) 5 g/L de carga de enzima en polvo; (b) 30 g/L de carga de sustrato; (c) 52,5 g/L de ácido α-cetoglutarico; (d) 21 g/L de ácido L-ascórbico; (e) sal de Mohr 2,25 mM; (f) tampón de fosfato de potasio 50 mM, pH 6,3 (pH ajustado con KOH), (g) temperatura de reacción a temperatura ambiente; y (h) tiempo de reacción de aprox. 24 h. En algunas reacciones, las condiciones de reacción contenían además un 1% (v/v) de antiespumante Y-30® (Dow Corning), y la solución de reacción se roció con gas O₂ a 2 l/h.

[0305] Las reacciones se inactivaron con 400 µl de MeCN al 75% y H₂O al 25%. Las placas se agitaron durante 10 minutos a temperatura ambiente y se centrifugaron a 4.000 rpm. La derivación se llevó a cabo transfiriendo 20 µl de reacción inactivada a una placa de 96 pocillos que contenía 230 µl/pocillo de bicarbonato de sodio al 5% (ac). Se añadió a cada pocillo una alícuota de 150 µl de cloruro de dansilo 21 mg/ml en MeCN. La placa se selló térmicamente y se centrifugó rápidamente, y las placas se incubaron a 65°C durante al menos 1 h con agitación en un Infors HT Microtron a una velocidad de agitación de 500 rpm. Las placas se centrifugaron a 4.000 rpm durante 10 minutos a temperatura ambiente. Un volumen de 200 µl de la solución derivada se transfirió a una placa de fondo redondo 96 y se envió para análisis por HPLC.

Ejemplo 6: Proceso para la conversión del compuesto (2) en el compuesto (1) usando preparaciones de polvo de proceso aguas abajo (DSP)

[0306] Se usaron dos condiciones de reacción para preparaciones de polvo de proceso aguas abajo (DSP). Las primeras condiciones de reacción, denominadas condiciones "mini-DSP" (es decir, condición D) se llevaron a cabo en una escala de 1 ml de la siguiente manera. Se preparó recientemente una solución madre de premezcla disolviendo 120 mg de ácido (2S)-piperidina-2-carboxílico (es decir, ácido L-pipecólico), 228 mg de ácido α-cetoglutarico y 252 mg de ácido L-ascórbico en 11,88 ml de tampón fosfato 50 mM, pH 6,3. El pH de la solución de premezcla se ajustó luego a 6,3 utilizando KOH. Se añadió un volumen de 120 µl de sal de Mohr 150 mM (en agua estéril) para formar la solución madre de premezcla.

[0307] Las reacciones se llevaron a cabo pesando 20 mg del polvo de enzima DSP en un vial seguido de 1 ml de solución madre de premezcla. La solución se mezcló completamente y el vial se dejó abierto durante la noche (~ 24 h) a temperatura ambiente. La solución de reacción se agitó a 1.200 rpm durante el curso de la reacción.

[0308] Las condiciones de reacción del "mini DSP" tienen los siguientes parámetros: (a) 20 g/L de carga de sustrato; (b) 34 g/L de ácido α-cetoglutarico (1,5 equivalentes de sustrato); (c) 13,6 g/L de ácido ascórbico (0,5 equivalentes de sustrato); (d) 1,5 mM de sal de mohr; (e) 20 g/L de proteína de preparación de enzima DSP; (f) tampón fosfato 50 mM, pH 6,3 (pH ajustado con KOH); (g) temperatura ambiente; y (g) tiempo de reacción de ~ 24 h. En algunas reacciones, la solución de reacción contenía además un 1% (v/v) de antiespumante Y-30® (Dow Corning) y la solución de reacción se roció con gas O₂ a 2 l/h durante el curso de la reacción.

[0309] Para seguir el curso de la reacción, se tomaron muestras de 10 µl y se disolvieron en 230 µl de bicarbonato de sodio al 5% (acuoso). Luego se agregó un volumen de 160 µl de 6 mg/ml de cloruro de dansilo en MeCN a la mezcla, los tubos se mezclaron bien y luego se calentaron, se destaparon a 50°C durante 30 minutos. Las muestras se centrifugaron y el sobrenadante transparente se analizó por HPLC como se describe en el Ejemplo 2.

[0310] Las segundas condiciones de reacción, denominadas condiciones "DSP completas", se llevaron a cabo de la siguiente manera. Se preparó una solución madre de premezcla recién preparada para reacciones a escala de 1 ml disolviendo 240 mg de ácido (2S)-piperidina-2-carboxílico (ácido L-pipecólico), 228 mg de ácido α-cetoglutarico y 252 mg de ácido L-ascórbico en 11,88 ml de tampón fosfato 50 mM, pH 6,3. El pH de la solución de premezcla se ajustó a 6,3 con KOH. Se añadió un volumen de 120 µl de sal de Mohr 150 mM (en agua estéril) para formar la solución madre de premezcla.

[0311] Las reacciones se realizaron pesando 10 mg del polvo de enzima DSP y añadiendo 1 ml de solución madre de premezcla. Después de mezclarse, el vial se dejó abierto durante la noche (~ 24 h) a temperatura ambiente. La solución de reacción se agitó a 1.200 rpm durante el curso de la reacción.

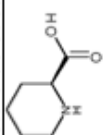
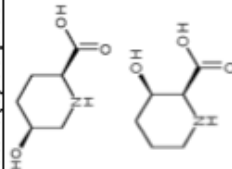
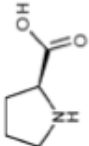
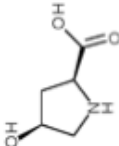
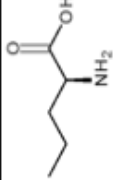
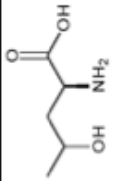
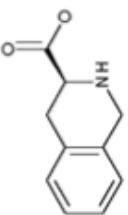
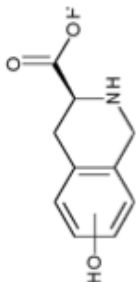
[0312] Las condiciones de reacción de "DSP completo" tienen los siguientes parámetros; (a) 10 g/L de carga de sustrato; (b) 38 g/L de ácido α-cetoglutarico; (c) 21 g/L de ácido ascórbico; (d) sal de Mohr 1,5 mM; (e) 10 g/L de preparación de enzima DSP; (f) tampón fosfato 50 mM, pH 6,3 (pH ajustado con KOH); (f) temperatura de reacción de 25°C; y (g) tiempo de reacción de ca 24 h. En algunas reacciones, la solución de reacción contenía un 1% (v/v) de antiespumante Y-30® (Dow Corning) y la solución de reacción se roció con gas O₂ a 2 l/h.

[0313] Para seguir el curso de la reacción, se retiraron muestras de 10 µl y se mezclaron con 230 µl de bicarbonato de sodio al 5% (ac). Luego se añadió a la mezcla un volumen de 160 µl de 6 mg/ml de cloruro de dansilo en MeCN.

Los tubos se mezclaron completamente y luego se calentaron, se destaparon a 50°C durante 30 minutos. Las muestras se centrifugaron y el sobrenadante transparente se analizó por HPLC como se describe en el Ejemplo 2.

Ejemplo 7: Proceso para la conversión de compuestos de fórmula (II) en compuestos de fórmula (I) usando polvos DSP de polipéptidos de hidroxilasa de prolina modificados

[0314] Se examinó la capacidad de las hidroxilasas de prolina modificadas para reconocer sustratos distintos de la prolina o el ácido pipecólico. Las condiciones de reacción comprendieron (a) 20 g/L de carga de sustrato; (b) 35 g/L de ácido α -cetoglutarico; (c) 14 g/L de ácido ascórbico; (d) sal de Mohr 1,5 mM; (e) 10 g/L de proteína de preparación de enzima DSP de la SEQ ID NO: 108; (f) tampón fosfato 50 mM, p H 6,3 (pH ajustado con KOH); (f) temperatura de reacción de 25°C; y (g) tiempo de reacción de ~ 24 h. Los controles negativos utilizaron preparaciones de enzimas obtenidas de células transformadas con un vector de expresión que no tenía un gen que codificara una hidroxilasa de prolina.

Sustrato	Estructura de sustrato	Control Neg	Rxn	Producto(s)	Estructura(s) de producto
Ácido L-pipecólico		-	+++++	2 regioisómeros	
L-prolina		-	++++	1 isómero	
L-norvalina		+	++	2 isómeros	
Ácido S-(1,2,3,4)-tetrahidro-3-isoquinolina carboxílico		+	++	3 productos significativos	

[0315] Las reacciones se inactivaron diluyendo 2.000 veces en 50:50 acetonitrilo:H₂O, y los productos de la reacción se analizaron por LC/MS/MS.

- 5 **[0316]** El análisis de LC/MS/MS para ácido pipecólico, prolina y norvalina se llevó a cabo en las siguientes condiciones:

Columna	ChiroBiotic TAG 250x4,6mm, 5µm
Fase móvil	Solución A: 0,1% ácido fórmico Solución B: 0,1% ácido fórmico en acetonitrilo A:B=50:50
Postiempo	5,0 min
Condiciones MS	Parámetros dependientes de fuente: CUR 30, IS 5500, TEM 590°C, GS1 60, GS2 60, DP30, EP10, CE 20 MRM: 130/84 (ácido pipecólico TA 4,2min), 146/100 y 146/82 (ácido pipecólico hidroxilado 3,2min y 3,7min) MRM: 116/70 (prolina TA 3,4min), 132/86 y 132/68 (hidroxilprolina TA 2,7min)
	MRM: 118/72 (norvalina TA 2,7min), 134/88, 134/74 y 134/70 (norvalina hidroxilada TA 2,6 y 2,7min)
Temperatura de columna	No controlada
Volumen de inyección	2 µL

[0317] La reacción inactivada para el ácido tetrahydroisoquinolina carboxílico se sometió a análisis de LC/MS/MS en las siguientes condiciones:

Columna	Poroshell EC C18 100x4,6mm, 2,7µm			
Fase móvil	Solución A: 0,5mM ácido perfluoroheptanoico Solución B: Acetonitrilo			
	Tiempo (min)	Flujo (ml/min)	A	B
	0-1,5	0,8	97	3
	9	0,8	70	30
	12	0,8	70	30
	13	0,8	97	3
	20	0,8	97	3
Postiempo	20 min			
Condiciones MS	Parámetros dependientes de fuente: CUR 30, IS 5500, TEM 600°C, GS1 60, GS2 60, DP30, EP10, CE 30 MRM: 194/148 y 194/146 (tetrahydroisoquinolina hidroxilada ácido carboxílico TA 5,7min, 7,26min y 9,73min)			
Temperatura de columna	No controlada			
Volumen de inyección	2 µL			

Ejemplo 8: Proceso para la conversión de ácido L-pipecólico (compuesto (2)) en ácido (2S, 5S)-5-hidroxipiperidina-2-carboxílico (compuesto (1)) seguido de la etapa de protección con Boc

[0318] Reacción enzimática: una solución de ácido L-pipecólico (15 g) disuelta en 138 ml de tampón de fosfato de potasio 50 mM, pH 6,3, se cargó en una solución premezclada que contenía: (i) Preparación de DSP del polipéptido de la SEQ. ID NO: 132 (5 g); (ii) emulsión antiespumante Y-30 (5 ml); (iii) sal de Mohr (1,08 g); (iv) ácido α-cetoglutarico (25,5 g); y (iv) ácido ascórbico (10,2 g); todo disuelto en 250 ml de tampón de fosfato de potasio 50 mM, pH 6.3. La mezcla resultante se agitó y se roció con oxígeno a una velocidad de 3 L/h y 25°C. El progreso de la reacción enzimática se controló mediante HPLC utilizando el Método 2 del Ejemplo 3. La conversión del ácido L-pipecólico a (2S, 5S)-5-hidroxipiperidina-2-carboxílico siguió un curso de la reacción de ~ conversión 78% en 25 h, ~ 92% de conversión en 45 h, y ~ 94% de conversión en 52 h, pero no alcanzó una mayor conversión a las 75 h. La pureza de la región del producto ácido (2S, 5S)-5-hidroxipiperidina-2-carboxílico después de 52 h de reacción fue de 6: 1.

[0319] Protección de Boc: la mezcla cruda de la reacción enzimática se ajustó a pH 9,5 con KOH (50% p/p), se calentó a 60°C durante 1 h, y luego se enfrió a temperatura ambiente. Posteriormente, se añadió adyuvante de filtración de Celite (15 g) con agitación (10 minutos) y la mezcla se filtró a través de una almohadilla de Celite 545 de 1 cm de espesor. La torta del filtro se lavó con agua (60 ml) y el filtrado se cargó con NaOH. (38,5 ml a 10 M) y

dicarbonato de di-*terc*-butilo ($\text{BoC}_2\text{-O}$) (42,2 g en 75 ml de THF). Al finalizar la reacción (~96% de conversión en dos días) la fase acuosa se lavó dos veces con heptanos (2 x 125 ml). Los lavados con heptano se descartaron y la fase acuosa se ajustó a pH 3,5 con 5 M HCl y se trató con NaCl (40 g). La fase acuosa se extrajo con *t*-butilo metilo éter (TBME) (3 X 125 ml), y los extractos de la fase orgánica resultante se combinaron, se secaron sobre MgSO_4 , se filtraron y se concentraron para dar ácido (2*S*,5*S*)-1-(*terc*-butoxicarbonilo)-5-hidroxipiridina-2-carboxílico.

[0320] Aislamiento: Una solución del producto bruto de ácido (2*S*,5*S*)-1-(*terc*-butilxicarbonil)-5-hidroxipiridina-2-carboxílico (25 g) en TBME (250 ml) se agitó con una barra agitadora magnética, los productos insolubles se filtraron y se enjuagaron con TBME (150 ml). El filtrado concentrado se secó a vacío durante la noche y se disolvió en acetato de isopropilo (100 ml) y heptano (100 ml). La mezcla resultante se calentó a 80°C durante 20-25 minutos, los insolubles se eliminaron mediante filtración en caliente, el filtrado se enfrió y se agitó durante 24 horas a temperatura ambiente. El sólido se filtró y la torta se lavó con una mezcla de heptanos-acetato de isopropilo enfriada (0 a 5°C) (1:1 de 50 ml). El producto sólido se recogió y se secó a vacío durante una noche para proporcionar ácido (2*S*,5*S*)-1-(*terc*-butilxicarbonil)-5-hidroxipiridina-2-carboxílico purificado (8,6 g, 30% de rendimiento).

LISTADO DE SECUENCIAS

[0321]

<110> Codexis, Inc.
CHEN, Haibin
BONG, Yong Koy
CABIROL, Fabien
COLLIER, Steven J.
GOHEL, Anupam
LI, Tao
MOORE, Jeffrey C.
QUINTANAR-AUDELO, Martina
SMITH, Derek
YANG, Hong

<120> BIOCATALIZADORES Y MÉTODOS PARA LA HIDROXILACIÓN DE COMPUESTOS QUÍMICOS

<130> CX2-095WO2

<150> 61/644.135
<151> 2012-05-08

<160> 228

<170> PatentIn versión 3.5

<210> 1
<211> 843
<212> ADN
<213> Secuencia artificial

<220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon *Sinorhizobium meliloti*

<400> 1

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg      60
      totacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg      120
      gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac      180
      aatgcgcccg cattagcgac ccctttatct aaatccactgc cgtacttgaa tgagcttggt      240
      gaaacacact tcgactgcag tgcgggttaga tatacgcgta tcgtgcgtgt ctccgagaac      300
      gcctgcatta tccctcatag tgattatttg gagctcgatg aaacgttcac ccggtgcac      360
10     ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg      420
      ggtctgggcg agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct      480
      aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc      540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg      600
      acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac      660
15     cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtc      720
      atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag      780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat      840
      taa

```

20

```

<210> 2
<211> 280
<212> PRT
25    <213> Sinorhizobium meliloti

<400> 2

```

```

30      Met Ser Thr His Phe Leu Gly Lys Val Lys Phe Asp Glu Ala Arg Leu
      1              5              10              15
      Ala Glu Asp Leu Ser Thr Leu Glu Val Ala Glu Phe Ser Ser Ala Tyr

```

35

40

45

50

55

60

65

				20					25					30					
5	Ser	Asp	Phe	Ala	Cys	Gly	Lys	Trp	Glu	Ala	Cys	Val	Leu	Arg	Asn	Arg			
			35					40					45						
	Thr	Gly	Met	Gln	Glu	Glu	Asp	Ile	Val	Val	Ser	His	Asn	Ala	Pro	Ala			
		50					55					60							
10	Leu	Ala	Thr	Pro	Leu	Ser	Lys	Ser	Leu	Pro	Tyr	Leu	Asn	Glu	Leu	Val			
		65				70					75					80			
	Glu	Thr	His	Phe	Asp	Cys	Ser	Ala	Val	Arg	Tyr	Thr	Arg	Ile	Val	Arg			
				85					90					95					
	Val	Ser	Glu	Asn	Ala	Cys	Ile	Ile	Pro	His	Ser	Asp	Tyr	Leu	Glu	Leu			
			100						105					110					
15	Asp	Glu	Thr	Phe	Thr	Arg	Leu	His	Leu	Val	Leu	Asp	Thr	Asn	Ser	Gly			
		115						120					125						
	Cys	Ala	Asn	Thr	Glu	Glu	Asp	Lys	Ile	Phe	His	Met	Gly	Leu	Gly	Glu			
		130					135					140							
20	Ile	Trp	Phe	Leu	Asp	Ala	Met	Leu	Pro	His	Ser	Ala	Ala	Cys	Phe	Ser			
		145			150						155					160			
	Lys	Thr	Pro	Arg	Leu	His	Leu	Met	Ile	Asp	Phe	Glu	Ala	Thr	Ala	Phe			
				165					170						175				
	Pro	Glu	Ser	Phe	Leu	Arg	Asn	Val	Glu	Gln	Pro	Val	Thr	Thr	Arg	Asp			
25				180					185						190				
	Met	Val	Asp	Pro	Arg	Lys	Glu	Leu	Thr	Asp	Glu	Val	Ile	Glu	Gly	Ile			
		195						200					205						
	Leu	Gly	Phe	Ser	Ile	Ile	Ile	Ser	Glu	Ala	Asn	Tyr	Arg	Glu	Ile	Val			
		210					215					220							
30	Ser	Ile	Leu	Ala	Lys	Leu	His	Phe	Phe	Tyr	Lys	Ala	Asp	Cys	Arg	Ser			
		225				230					235					240			
	Met	Tyr	Asp	Trp	Leu	Lys	Glu	Ile	Cys	Lys	Arg	Arg	Gly	Asp	Pro	Ala			
				245					250						255				
35	Leu	Ile	Glu	Lys	Thr	Ala	Ser	Leu	Glu	Arg	Phe	Phe	Leu	Gly	His	Arg			
			260						265					270					
	Glu	Arg	Gly	Glu	Val	Met	Thr	Tyr											
			275					280											
40	<210> 3																		
	<211> 843																		
	<212> ADN																		
	<213> Secuencia artificial																		
45	<220>																		
	<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti																		
	<400> 3																		
50	atgagtagccc	acttttctcgg	caaagtga	aa	ttcgatgaag	cacggcttgc	agaagacttg												60
	tctacgctgg	aagtggcgga	attttctctcc	gcgtatagt	attttgcctg	cggaaaatgg													120
	gaagcatgtg	ttcttagaaa	ccgtacaggt	atgcaagaag	aggatattgt	ggtttcccat													180
	aacgcacggg	ctcttgcaac	cccgtgagt	aaatcattgc	cttaccttaa	tgaattagtt													240
55	gagacacatt	tcgactgttc	ggctgtccgt	tacacgcgta	ttgtaagagt	gagcgagaac													300
	gcgtgcatta	taccacattc	agattacctc	gagcttgacg	agacatttac	cagattgcac													360
	ctggtcttgg	ataccaactc	aggttgtgcc	aataccgaag	aggacaagat	attccatatg													420
	ggtctcggtg	aaatatggtt	cttggtatgcg	atgctgcctc	atagtgcagc	atgtttttct													480
	aagacgccta	gactgcacat	gatgattgac	ttcgaagcta	cagcgtttcc	cgaatcattt													540
60	ttacggaacg	tcgaacagcc	ggttaccacc	cgtgatattg	tcgaccctag	aaaagaatta													600
	acggacgaag	tgatcgaagg	tatactcgcc	ttttctatta	tcatatcgga	ggcgaactac													660
	cgtgaaatcg	tgtctatttt	agcgaagctg	cacttctttt	ataaagccga	ttgccggtcg													720
	atgtatgatt	ggctgaaaga	aatttgcaag	agacggggag	acccggcgct	tattgaaaag													780
	acggcttcat	tggaacgctt	cttttttaggc	cacagagaac	gcggggaagt	aatgacctac													840
65	taa																		843

<210> 4
 <211> 280
 <212> PRT
 <213> Sinorhizobium meliloti

<400> 4

```

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

<210> 5
 <211> 843
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 5

ES 2 712 682 T3

5	atgagcacc	atttcttggg	caaggtcaag	ttcgaatgaag	cgcgattggc	agaagatcta	60									
	tctaccttgg	aagttgccga	gttctcgagt	gcatactcgg	acttcgcgtg	cggtaaatgg	120									
	gaggcatg	tgctacgcaa	tcggaccgga	atgcaggagg	aagatatcgt	cgtaagtcac	180									
	aacgctcctg	cactggccac	gocgctgagc	aagtcgctgc	cgtatctgaa	cgaacttggt	240									
	gaaacccact	tcgattgcag	cgtctgttcg	tatacaagaa	ttgtccgtgt	atcagaaaac	300									
	gcatgtataa	tcccccatag	tgattaccta	gaactagatg	agaccttcac	aaggttacac	360									
	ctggtgttag	acactaattc	aggatgcgct	aatactgagg	aagataaaat	atttcataatg	420									
10	ggactgggag	agatttggtt	ccttgacgct	atgttaccgc	atagcgctgc	ttgtttttcc	480									
	aaaactccgc	gcctgcatct	gatgatcgac	tttgaggcta	ccgcttttcc	cgaatctttt	540									
	ctgcgaaatg	tcgaacaacc	agtgacaaca	cgagacatgg	ttgatcctcg	gaaggaacta	600									
15	accgatgagg	ttatcgaagg	tattctgggg	ttttcaataa	ttattagcga	agccaattac	660									
	cgggaaattg	tttctattct	ggcgaagcta	cacttcttct	acaaggcaga	ctgtcgatca	720									
	atgtacgact	ggctgaagga	aatctgcaaa	cgtcgagggg	atcctgcact	tattgaaaag	780									
	accgcctcgc	tcgagcgatt	ttttctaggg	caccgtgaac	gtggcgaggt	gatgacatac	840									
	taa						843									
20	<210> 6															
	<211> 280															
	<212> PRT															
	<213> Sinorhizobium meliloti															
25	<400> 6															
30	Met	Ser	Thr	His	Phe	Leu	Gly	Lys	Val	Lys	Phe	Asp	Glu	Ala	Arg	Leu
	1				5					10					15	
	Ala	Glu	Asp	Leu	Ser	Thr	Leu	Glu	Val	Ala	Glu	Phe	Ser	Ser	Ala	Tyr
				20					25					30		
	Ser	Asp	Phe	Ala	Cys	Gly	Lys	Trp	Glu	Ala	Cys	Val	Leu	Arg	Asn	Arg
35			35					40					45			
	Thr	Gly	Met	Gln	Glu	Glu	Asp	Ile	Val	Val	Ser	His	Asn	Ala	Pro	Ala
	50						55					60				
	Leu	Ala	Thr	Pro	Leu	Ser	Lys	Ser	Leu	Pro	Tyr	Leu	Asn	Glu	Leu	Val
	65				70						75					80
40	Glu	Thr	His	Phe	Asp	Cys	Ser	Ala	Val	Arg	Tyr	Thr	Arg	Ile	Val	Arg
				85						90					95	
	Val	Ser	Glu	Asn	Ala	Cys	Ile	Ile	Pro	His	Ser	Asp	Tyr	Leu	Glu	Leu
				100					105					110		
	Asp	Glu	Thr	Phe	Thr	Arg	Leu	His	Leu	Val	Leu	Asp	Thr	Asn	Ser	Gly
45			115					120					125			
	Cys	Ala	Asn	Thr	Glu	Glu	Asp	Lys	Ile	Phe	His	Met	Gly	Leu	Gly	Glu
	130						135					140				
	Ile	Trp	Phe	Leu	Asp	Ala	Met	Leu	Pro	His	Ser	Ala	Ala	Cys	Phe	Ser
	145				150						155					160
50	Lys	Thr	Pro	Arg	Leu	His	Leu	Met	Ile	Asp	Phe	Glu	Ala	Thr	Ala	Phe
				165						170					175	
	Pro	Glu	Ser	Phe	Leu	Arg	Asn	Val	Glu	Gln	Pro	Val	Thr	Thr	Arg	Asp
			180						185					190		
	Met	Val	Asp	Pro	Arg	Lys	Glu	Leu	Thr	Asp	Glu	Val	Ile	Glu	Gly	Ile
55			195					200					205			
	Leu	Gly	Phe	Ser	Ile	Ile	Ile	Ser	Glu	Ala	Asn	Tyr	Arg	Glu	Ile	Val
	210						215					220				
	Ser	Ile	Leu	Ala	Lys	Leu	His	Phe	Phe	Tyr	Lys	Ala	Asp	Cys	Arg	Ser
	225				230						235					240
60	Met	Tyr	Asp	Trp	Leu	Lys	Glu	Ile	Cys	Lys	Arg	Arg	Gly	Asp	Pro	Ala
				245						250					255	
	Leu	Ile	Glu	Lys	Thr	Ala	Ser	Leu	Glu	Arg	Phe	Phe	Leu	Gly	His	Arg
			260						265					270		
65	Glu	Arg	Gly	Glu	Val	Met	Thr	Tyr								
		275						280								

ES 2 712 682 T3

5	<210> 7 <211> 843 <212> ADN <213> Secuencia artificial	
10	<220> <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti <400> 7	
15		
	atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60 tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg 120	
20		
	gaagagtgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180 aatgcgcccg cattagcgac ccotttatct aaatcactgc cgtacttgaa tgagcttggt 240 gaaacacact tgcactgcag tgcgggttaga tatacgcgta tcgtgcgtgt ctccgagAAC 300 gcctgcatta tccctcatag tgattatttg gagctcgatg aaacgttcac ccggctgcac 360 ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420 ggtctgggag agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480	
25		
	aaaactccta gactgcattc catgatagat tttgaagcta ccgccttccc ggaatccttc 540 ttacgcaatg tggAACagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600 acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660 cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcggg ctgtcgggtc 720 atgtatgatt ggcttaaaga aatatgtaaa cgcgggggag atcccgcggt gattgagaag 780 acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840	
30		
	taa 843	
35	<210> 8 <211> 280 <212> PRT <213> Secuencia artificial	
40	<220> <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti <400> 8	
45		
50		
55		
60		
65		

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

45 <210> 9
 <211> 843
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 9

55

60

65

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg      60
      tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg      120
      gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac      180
      aatgcgcccg cattagcgac ccctttatct aaatcacctgc cgtacttgaa tgagcttggt      240
      gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac      300
      gcctgcttga tcctcatag tgattatttg gagctcgatg aaacgttcac ccgctgcac      360
10     ctgctgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg      420
      ggtctggggc agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct      480
      aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc      540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg      600
      acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac      660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc      720
15     atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag      780
      acagcgagtc tcgaacggtt ctttttgggg catcgtgaac gcggagaagt aatgacgtat      840
      taa
      843

```

```

20     <210> 10
      <211> 280
      <212> PRT
      <213> Secuencia artificial

25     <220>
      <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

      <400> 10

```

```

30

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

```

60

65

5 Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val
210 215 220
Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
225 230 235 240
Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
245 250 255
10 Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
260 265 270
Glu Arg Gly Glu Val Met Thr Tyr
275 280

15

<210> 11

<211> 843

<212> ADN

20 <213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

25 <400> 11

30 atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg 120
gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggttttctcac 180
aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac 300
gcctgccaga tccctcatag tgattatttg gagctcgatg aaacggtcac ccggctgcac 360
ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
ggtctgggag agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct 480
35 aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
cgcgagatcg tctctatcct ggccaaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag 780
40 acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa 843

45 <210> 12

<211> 280

<212> PRT

<213> Secuencia artificial

<220>

50 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 12

55

60

65

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

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

<210> 13

<211> 843

<212> ADN

<213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 13

ES 2 712 682 T3

	atgtctacgc	atttccttgg	caaggtcaaa	ttcgacgagg	ccagattagc	cgaagatctg	60
	tctacttttag	aagtcgctga	gttttctagt	gcatattcgg	acttcgcttg	tggaaagtgg	120
5	gaagcatgtg	ttttacgcaa	cagaacaggc	atgcaagaag	aagacattgt	ggtttctcac	180
	aatgcgcccg	cattagcgac	ccctttatct	aaatcaactgc	cgtacttgaa	tgagcttggt	240
	gaaacacact	tcgactgcag	tgcggttaga	tatacgcgta	tcgtgcgtgt	ctccgagaac	300
	gcctgcatta	tcctcatag	tgattatttg	gagctcgatg	aaacgctgac	ccggctgcac	360
	ctcgtgttgg	atactaacag	tggttggtgc	aacactgagg	aagataagat	attccacatg	420
10	ggtctgggcg	agatctggtt	tctggacgca	atgttacctc	attctgctgc	ctgcttttct	480
	aaaactccta	gactgcatct	catgatagat	tttgaagcta	ccgccttccc	ggaatccttc	540
	ttacgcaatg	tggaacagcc	tgtaactact	cgcgacatgg	tagaccctcg	taaagaattg	600
	acggacgaag	ttatagaagg	catttttaggg	ttctccataa	tcatcagtga	agcgaattac	660
	cgcgagatcg	tctctatcct	ggccaaatta	cacttctttt	ataaagcgga	ctgtcgggtc	720
	atgtatgatt	ggcttaaaga	aatatgtaaa	cgcgggggag	atcccgcgtt	gattgagaag	780
15	acagcgagtc	tcgaacgttt	ctttttgggg	catcgtgaac	gcggagaagt	aatgacgtat	840
	ttaa						843

	<210> 14
20	<211> 280
	<212> PRT
	<213> Secuencia artificial
	<220>
25	<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
	<400> 14

30

35

40

45

50

55

60

65

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

<210> 15

<211> 843

45 <212> ADN

<213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

50

<400> 15

55

60

65

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg      60
      tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaagtg      120
      gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac      180
      aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt      240
      gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac      300
      gcctgcatta tccctcatag tgattatttg gagctcgatg aaacgttcac ccggctgcac      360
      ctcggtgttg atactaacag tggttgtgcc tatactgagg aagataagat attccacatg      420
10     ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct      480
      aaaactccta gactgcatct catgatatag tttgaagcta ccgccttccc ggaatccttc      540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg      600
      aoggaagcga ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac      660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcggtcc      720
      atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag      780
15     acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat      840
      taa

```

<210> 16

<211> 280

20 <212> PRT

<213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

25

<400> 16

```

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

```

ES 2 712 682 T3

5	<210> 17 <211> 843 <212> ADN <213> Secuencia artificial	
10	<220> <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti	
15	<400> 17	
	atgtctacgc atttcottgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg	60 120
20	gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt gaaacacact tcgactgcag tgcgggttaga tatacgcgta tcgtgcgtgt ctccgagAAC	180 240 300
25	gcctgcatta tccctcatag tgattatttg gagctcgatg aaacgttcac ccggctgcac ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg ggctctggcg agatctggtt tctggacagt atgttacctc attctgctgc ctgcttttct aaaactccta gactgcatct catgatagat ttgaaagcta ccgccttccc ggaatccttc ttacgcaatg tggAACagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg	360 420 480 540 600
30	acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcggA ctgtcgggtcc atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat taa	660 720 780 840 843
35	<210> 18 <211> 280 <212> PRT <213> Secuencia artificial	
40	<220> <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti	
45	<400> 18	
50		
55		
60		
65		

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

<210> 19

<211> 843

<212> ADN

45 <213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

50 <400> 19

55 atgtctacgc atttccttgg caaggtcaaa ttcgaacgag ccagattagc cgaagatctg 60
tctacttttag aagtcgctga gtttgctagt gcatattcgg acttcgcttg tggaaagtgg 120
gaagcatgcg ttttacgcaa cagaacaggc atgcaagaag aagacattgt gggttctcac 180
aatgcgcccc cattagcgac ccctttatct aaactactgc cgtacttgaa tgagcttggt 240
gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac 300
gcctgcatta tccctcatag tgattatttg gagctcgatg aaacggtcac ccggtgcac 360
ctcgtgttgg atactaacag tgggtgtgcc aactactgagg aagataagat attccacatg 420
ggctctggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
60 aaaactccta gactgactct catgatagat ttggaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggaaacagc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgctt gattgagaag 780
65 acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa 843

<210> 20
 <211> 280
 <212> PRT
 5 <213> Secuencia artificial

 <220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

 10 <400> 20

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

 195 200 205
 45 Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val
 210 215 220
 Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
 225 230 235 240
 Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
 245 250 255
 50 Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
 260 265 270
 Glu Arg Gly Glu Val Met Thr Tyr
 275 280

 55

 <210> 21
 <211> 843
 <212> ADN
 60 <213> Secuencia artificial

 <220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

 65 <400> 21

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg      60
      tctacttttag aagtcgctga gtttttctagt gcatattcgg acttcgcttg tggaaagtgg      120
      gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac      180
      aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt      240
      gaaacacact tcgactgcag tgcgggttaga tatacgcgta tcgtgcgtgt ctccgagaac      300
      gcctgcatta tcctcatag tgattatttg gagctcgatg aaacggtcac ccggctgcac      360
10     ctctgtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg      420
      ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct      480
      aaaactocta gactgacgct catgatagat tttgaagcta ccgccttccc ggaatccttc      540
      ttacgcaatg tggaaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg      600
      acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac      660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc      720
15     atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag      780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat      840
      taa
      843

```

20 <210> 22
 <211> 280
 <212> PRT
 <213> Secuencia artificial

25 <220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
 <400> 22

```

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

```

45

50

55

60

65

ES 2 712 682 T3

5 Val Ser Glu Asn Ala Cys Ile Ile Pro His Ser Asp Tyr Leu Glu Leu
100 105 110
Asp Glu Thr Phe Thr Arg Leu His Leu Val Leu Asp Thr Asn Ser Gly
115 120 125
Cys Ala Asn Thr Glu Glu Asp Lys Ile Phe His Met Gly Leu Gly Glu
130 135 140
10 Ile Trp Phe Leu Asp Ala Met Leu Pro His Ser Ala Ala Cys Phe Ser
145 150 155 160
Lys Thr Pro Arg Leu Thr Leu Met Ile Asp Phe Glu Ala Thr Ala Phe
165 170 175
15 Pro Glu Ser Phe Leu Arg Asn Val Glu Gln Pro Val Thr Thr Arg Asp
180 185 190
Met Val Asp Pro Arg Lys Glu Leu Thr Asp Glu Val Ile Glu Gly Ile
195 200 205
Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val
210 215 220
20 Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
225 230 235 240
Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
245 250 255
25 Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
260 265 270
Glu Arg Gly Glu Val Met Thr Tyr
275 280

30

<210> 23
<211> 843
35 <212> ADN
<213> Secuencia artificial

<220>
40 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
<400> 23

45 atgtctacgc acttccttgg caagggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaaagtgg 120
gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggttttctcac 180
aatgcgcccc cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
gaaacacact tcgactgcag tgcgggttaga tatacgcgta tcgtgcgtgt ctccgagaac 300
gcctgcatta tccctcatag tgattatttg gagctcgatg aaacgttcac ccggctgcac 360
ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
50 ggtctgggag agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct 480
aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
55 atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag 780
acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa 843

60 <210> 24
<211> 280
<212> PRT
<213> Secuencia artificial

<220>
65 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium melilo

<400> 24

```

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

```

45 <210> 25
 <211> 843
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 25

55

60

65

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg caaggtcaaa ttgcgacgagg ccagattagc cgaagatctg      60
      tctacttttag aagtcgctga gttttctgtt gcatattcgg acttcgcttg tggaaagtgg      120
      gaagcatgtg ttttacgcaa cagaacaggg atgcaagaag aagacattgt ggtttctcac      180
      aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt      240
      gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac      300
      gcctgcatta tccctcatag tgattatttg gagctcgatg aaacgttcac ccgggtgcac      360
      ctcggtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg      420
10     ggtctggggg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct      480
      aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc      540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg      600
      acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac      660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc      720
15     atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag      780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat      840

```

taa 843

```

20     <210> 26
      <211> 280
      <212> PRT
      <213> Secuencia artificial

25     <220>
      <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

      <400> 26

```

```

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

```

ES 2 712 682 T3

5	<210> 27 <211> 843 <212> ADN <213> Secuencia artificial	
10	<220> <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti	
	<400> 27	
15	atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg	60
20	tctacttttag aagtcgctga gttttctccg gcatattcgg acttcgcttg tggaaagtgg	120
	gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac	180
	aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt	240
	gaaacacact tcgactgcag tgcgggttaga tatacgcgta tcgtgcgtgt ctccgagaac	300
	gcctgcatta tccctcatag tgattatttg gagctcgatg aaacgttcac ccggtgcac	360
	ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg	420
	ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct	480
	aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc	540
25	ttacgcaatg tggacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg	600
	acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac	660
	cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaaagcgga ctgtcgggtcc	720
	atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag	780
	acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat	840
30	taa	843
35	<210> 28 <211> 280 <212> PRT <213> Secuencia artificial	
40	<220> <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti	
	<400> 28	
45		
50		
55		
60		
65		

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

45

<210> 29

<211> 843

<212> ADN

50

<213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

55

<400> 29

60

65

5 atgtctacgc atttccttgg caagggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
tctacttttag aagtcgctga gtttttctagt gcatattcgg acttcacttg tggaaagtgg 120
gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt gggtttctcac 180
aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
gaaacacact tcgactgcag tgcgggttaga tatacgcgta tcgtgcgtgt ctccgagAAC 300
gcctgcatta tccctcatag tgattatttg gagctcgatg aaacggtcac ccggctgcac 360
10 ctctgtgttg atactaacag tgggtgtgcc aacactgagg aagataagat attccacatg 420
gggtctgggag agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct 480
aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggAACagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
15 cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcggg ctgtcgggtcc 720
atgtatgatt ggcttaaaga aatatgtaaa cgcgggggag atcccgcggt gattgagaag 780
acagcgagtc tcgaacggtt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa 843

20 <210> 30
<211> 280
<212> PRT
<213> Secuencia artificial
25 <220>
<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
30 <400> 30

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

60

65

5 Met Val Asp Pro Arg Lys Glu Leu Thr Asp Glu Val Ile Glu Gly Ile
195 200 205
Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val
210 215 220
10 Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
225 230 235 240
Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
245 250 255
15 Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
260 265 270
Glu Arg Gly Glu Val Met Thr Tyr
275 280

20 <210> 31
<211> 843
<212> ADN
<213> Secuencia artificial

25 <220>
<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
<400> 31

30 atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg 120
gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggttggtcac 180
aatgcgcccc cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
35 gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac 300
gcctgcatta tcctcatag tgattatttg gagctcgatg aaacggtcac ccgctgcac 360
ctcgtgttgg atactaacag tgggtgtgcc aacactgagg aagataagat attccacatg 420
ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
aaaactocta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggaaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
40 acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag 780
acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa 843

45 <210> 32
<211> 280
<212> PRT
50 <213> Secuencia artificial

<220>
<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
55 <400> 32

60

65

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

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

<210> 33

<211> 843

<212> ADN

<213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 33

ES 2 712 682 T3

5	atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg	60
	tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg	120
	gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattac ggtttctcac	180
	aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt	240
	gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac	300
	gcctgcatta tcctcatag tgattatttg gagctcgatg aaacgttcac ccggctgcac	360
10	ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg	420
	ggtctggggc agatctggtt tctggacgca atgttacctc actctgctgc ctgcttttct	480
	aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc	540
	ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg	600
	acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac	660
15	cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc	720
	atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag	780
	acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat	840
	taa	843

20	<210> 34
	<211> 280
	<212> PRT
	<213> Secuencia artificial
25	<220>
	<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
30	<400> 34

35

40

45

50

55

60

65

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

<210> 35

<211> 843

<212> ADN

45 <213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

50 <400> 35

55

60

65

ES 2 712 682 T3

5	atgtctacgc	atctccttgg	caaggtcaaa	ttcgacgagg	ccagattagc	cgaagatctg	60
	tctactttag	aagtcgctga	gttttctagt	gcatattcgg	acttcgcttg	tggaaagtgg	120
	gaagcatgtg	ttttacgcaa	cagaacaggg	atgcctgaag	aagacattgt	ggtttctcac	180
	aatgcgcccg	cattagcgac	ccctttatct	aaatcactgc	cgtacttgaa	tgagcttggt	240
	gaaacacact	tcgactgcag	tcggttaga	tatacgcgta	tcgtgcgtgt	ctccgagAAC	300
10	gcctgcatta	tccctcatag	tgattatgtg	gagctcgatg	aaacgttcac	ccggctgcac	360
	ctcgtgttgg	atactaacag	tggttgtgcc	aacactgagg	aagataagat	attccacatg	420
	ggctctggcg	agatctgggt	tctggacgca	atgttacctc	attctgctgc	ctgcttttct	480
	aaaactccta	gactgcatct	catgatagat	tttgaagcta	ccgccttccc	ggaatccttc	540
	ttacgcaatg	tggaacagcc	tgtaactact	cgcgacatgg	tagaccctcg	taaagaattg	600
15	acggacgaag	ttatagaagg	catttttaggg	ttctccataa	tcacacagtga	agcgaattac	660
	cgcgagatcg	tctatatact	ggccaaatta	cacttccttt	ataaagcggg	ctgtcgggtcc	720
	atgtatgatt	ggcttaaaga	aatatgtaaa	cgcgggggag	atcccgcggt	gattgagaag	780

20	acagcgagtc	tcgaacgttt	ctttttgggg	catcgtgaac	gcggagaagt	aatgacgtat	840
	taa						843

25 <210> 36
 <211> 280
 <212> PRT
 <213> Secuencia artificial

30 <220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
 <400> 36

35

40

45

50

55

60

65

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

45

<210> 37
 <211> 843
 <212> ADN
 <213> Secuencia artificial

50

<220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

55

<400> 37

60

65

ES 2 712 682 T3

5	atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg	60
	tctacttttag aagtcgctga gttttctagt gcataattcg acttcgcttg tggaaagtgg	120
	gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgc tgtttctcac	180
	aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt	240
10	gaaacacact tgcactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac	300
	gcctgcatta tcoctcatag tgattatttg gagctcgatg aaacgttcac ccggtgcac	360
	ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg	420
	ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct	480
	aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc	540
15	ttacgcaatg tggaaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg	600
	acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac	660
	cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc	720
	atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag	780
	acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat	840
20	taa	843

<210> 38

<211> 280

25 <212> PRT

<213> Secuencia artificial

<220>

30 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 38

35

40

45

50

55

60

65

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

45		275	280
----	--	-----	-----

50	<210> 39
	<211> 843
	<212> ADN
	<213> Secuencia artificial
55	<220>
	<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
	<400> 39

60

65

5 atgtctacgc atttccttgg caagggtcaaa ttgcacgagg ccagattagc cgaagatctg 60
tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg 120
gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
gagacacact tcgactgcag tcgggttaga tatacgcgta tcgtgcgtgt ctccgagaac 300
gcctgcatta tccctcatag tgattatttg gagctcgatg aacatttcac ccggctgcac 360
ctcgtgttgg ataactaacag tgggtgtgcc aacactgagg aagataagat attccacatg 420
10 ggtctgggcg agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct 480
aaaactccta gactgcctct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
15 atgtatgatt ggcttaaaga aatatgtaa cgcgggggag atcccgcggt gattgagaag 780
acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa 843

20

<210> 40

<211> 280

<212> PRT

25 <213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

30 <400> 40

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

60

65

5 Met Val Asp Pro Arg Lys Glu Leu Thr Asp Glu Val Ile Glu Gly Ile
180 185 190
195 200 205
Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val
210 215 220
10 Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
225 230 235 240
Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
245 250 255
15 Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
260 265 270
Glu Arg Gly Glu Val Met Thr Tyr
275 280

20

<210> 41
<211> 843
<212> ADN
25 <213> Secuencia artificial

<220>
<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

30

<400> 41

35

atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg 120
gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac 300
gcctgcatta tccctcatag tgattatttg gagacggatg aaacgttcac ccggctgcac 360
ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
ggctctggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
40 aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcggtcc 720
atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgctt gattgagaag 780
45 acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa 843

50

<210> 42
<211> 280
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

55

<400> 42

60

65

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

<210> 43

<211> 843

50 <212> ADN

<213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

55

<400> 43

60

65

ES 2 712 682 T3

5	atgtctacgc	atctccttgg	caaggtcaaa	ttcgacgagg	ccagattagc	cgaagatctg	60
	tctacttttag	aagtcgctga	gttttctagt	gcatattcgg	acttcgcttg	tggaagtgg	120
	gaagcatgtg	ttttacgcaa	cagaacaggc	atgcaagaag	aagacattgt	ggtttctcac	180
	aatgcgcccg	cattacagac	ccctttatct	aaatcactgc	cgtacttgaa	tgagcttggt	240
	gaaacacact	tcgactgcag	tgcggttaga	tatacgcgta	tcgtgcgtgt	ctccgagaac	300
	gcctgcatta	tccctcatag	tgattatitg	gagctcgatg	aaacgttcac	ccggctgcac	360
	ctcgtgttgg	atactaacag	tggttgtgcc	aacactgagg	aagataagat	attccacatg	420
10	ggctctggcg	agatctggtt	tctggacgca	atgttacctc	attctgctgc	ctgcttttct	480
	aaaactccta	gactgcatct	catgatagat	tttgaagcta	ccgccttccc	ggaatccttc	540
	ttacgcaatg	tggaacagcc	tgtaactact	cgcgacatgg	tagaccctcg	taaagaattg	600
	acggacgaag	ttatagaagg	catttttaggg	ttctccataa	tcatcagtga	agcgaattac	660
	cgcgagatcg	tctctatcct	ggccaaatta	cacttctttt	ataaagcgga	ctgtcggtcc	720
15	atgtatgatt	ggcttaaaga	aatatgtaaa	cgcgggggag	atcccgcggt	gattgagaag	780
	acagcgagtc	tcgaacgttt	ctttttgggg	catcgtgaac	gcggagaagt	aatgacgtat	840
	taa						843

20	<210> 44
	<211> 280
	<212> PRT
	<213> Secuencia artificial
25	<220>
	<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
30	<400> 44

35

40

45

50

55

60

65

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

45 <210> 45
 <211> 843
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 45

55	atgtctacgc	atttccttgg	caaggtcaaa	ttcgacgagg	ccagattagc	cgaagatctg	60
	tctacttttag	aagtcgctga	gttttctagt	gcatattcgg	acttcgcttg	tggaaagtgg	120
	gaagcatgtg	ttttacgcaa	cagaacaggc	atgcaagaag	aagacattgt	ggtttctcac	180
	aatgcgcccc	cattagcgac	ccctttatct	aaatcactgc	cgtacttgaa	tgagcttggt	240
	gaaacacact	tcgactgcag	tcggttaga	tatacgcgta	tcgtgcgtgt	ctccgagaac	300
60	gcctgcatta	tcctcatag	tgattatttg	gagctcgatg	aaacgttcac	ccggctgcac	360
	ctcgtgttgg	atactaacag	tggttgtgcc	aacactgagg	aagataagat	attccacatg	420
	ggtctggggc	agatctggtt	tctggacgca	atgttacctc	attctgctgc	ctgcttttct	480
	aaaactccta	gactgcatct	catgatagat	tttgaagcta	ccgccttccc	ggaatccttc	540
	ttacgcaatg	tggaacagcc	tgtaactact	cgcgacatgg	tagaccctcg	taaagaattg	600
	acggacgaag	ttatagaagg	catttttaggg	ttctccataa	tcatcagtga	agcgaattac	660
65	cgcgagatcg	tctctatcct	ggccaaatta	cacttctttt	ataaagcgga	ctgtcggtcc	720

ES 2 712 682 T3

5 atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggtt gattgagaag 780
acagcgagtc tcgaacgttt ctttttgggg aagcgtgaac gcggagaagt aatgacgtat 840
taa 843

10 <210> 46
<211> 280
<212> PRT
<213> Secuencia artificial

<220>
15 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 46

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

60 <210> 47
<211> 843
<212> ADN
<213> Secuencia artificial

<220>
65 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

ES 2 712 682 T3

<400> 47

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg      60
      tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg      120
      gaagcatgtg ttttacgcaa cagaacaggg atgcaagaag aagacattgt ggtttctcac      180
      aatgcgcccgc cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt      240
      gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac      300
      gcctgcatta tccctcatag tgattatttg gagctcgatg aagatttcac ccggctgcac      360
      ctcggtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg      420
10     ggtctgggcy agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct      480
      aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc      540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg      600
      acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac      660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc      720
15     atgtatgatt ggcttaaga aatatgtaa cgccggggag atcccgcggt gattgagaag      780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat      840
      taa
      843

```

<210> 48

20 <211> 280

<212> PRT

<213> Secuencia artificial

<220>

25 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 48

```

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

```

ES 2 712 682 T3

Glu Arg Gly Glu Val Met Thr Tyr
275 280

5

<210> 49

<211> 843

<212> ADN

10 <213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

15 <400> 49

20	atgtctacgc	atcttccttgg	caagggtcaaa	ttcgcacgagg	ccagattagc	cgaagatctg	60
	tctacttttag	aagtcgctga	gtttttctagt	gcataattcgg	acttcgcttg	tggaaaagtgg	120
	gaagcatgtg	ttttacgcaa	cagaacaggc	atgcaagaag	aagacattgt	ggttttctcac	180
	aatgcgccc	cattagcgac	ccctttatct	aaatcaactgc	cgtacttgaa	tgagcttggt	240
	gaaacacact	tcgactgcag	tgccggttaga	tatacgcgta	tcgtgcgtgt	ctccgagaac	300
	gcctgcatta	tcctcatag	tgattatttg	gagctcgatg	aagggtttcac	ccggctgcac	360
	ttcgtgttgg	atactaaccag	tggttggtgcc	aacactgagg	aagataagat	attccacatg	420
25	gggtctggg	agatctgggt	tctggacgca	atgttacctc	attctgctgc	ctgcttttct	480
	aaaactccta	gactgcattc	catgatagat	tttgaagcta	ccgccttccc	ggaatccttc	540
	ttacgcaatg	tggaaacagcc	tgttaactact	cgcgacatgg	tagaccctcg	taaagaattg	600
	acggacgaag	ttatagaagg	catttttaggg	ttctccataa	tcatcagtga	agcgaattac	660
	cgcgagatcg	tctctatcct	ggccaaatta	cacttctttt	ataaagcgga	ctgtcgggtcc	720
30	atgtatgatt	ggcttaaaga	aatatgtaaa	cgcgggggag	atcccgcggt	gattgagaag	780
	acagcgagtc	tcgaacgttt	ctttttgggg	catcgtgaac	gcggagaagt	aatgacgtat	840
	taa						843

35 <210> 50

<211> 280

<212> PRT

<213> Secuencia artificial

40 <220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 50

45

50

55

60

65

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

45

<210> 51
 <211> 843
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 51

60

65

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg caagggtcaaa ttccgacgagg ccagattagc cgaagatctg      60
      tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg      120
      gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac      180
      aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt      240
      gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac      300
      gcctgcatta tccctcatag tgattatttg gagctcgatg aaacgttcac ccggctgcac      360
10     ctctgtttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg      420
      ggtctgggcg agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct      480
      aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc      540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg      600
      acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac      660
15     cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc      720
      atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag      780
      acagcgagtc tcgaacgttt ctttttgggg cggcgtgaac gcggagaagt aatgacgtat      840
      taa

```

```

20
    <210> 52
    <211> 280
    <212> PRT
    <213> Secuencia artificial
25
    <220>
    <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
30
    <400> 52

```

```

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

```

```

45
50
55
60
65

```

ES 2 712 682 T3

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

<210> 53

35 <211> 843

<212> ADN

<213> Secuencia artificial

<220>

40 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 53

45	atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg	60
	tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg	120
	gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac	180
	aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt	240
	gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac	300
	gcctgcatta tccctcatag tgattatttg gaggttgatg aaacgttcac ccggctgcac	360
50	ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg	420
	ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct	480
	aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc	540
	ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg	600
	acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac	660
55	cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc	720
	atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag	780
	acagcgagtc tcgaacggtt ctttttgagg catcgtgaac ggggagaagt aatgacgtat	840
	ttaa	843

60 <210> 54

<211> 280

<212> PRT

<213> Secuencia artificial

ES 2 712 682 T3

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 54

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

<210> 55

<211> 843

45 <212> ADN

<213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

50

<400> 55

55

60

65

ES 2 712 682 T3

5	atgtctacgc atttccttgg caaggcctaaa ttcgacgagg ccagattagc cgaagatctg	60
	tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg	120
	gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac	180
	aatgcgccc cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt	240
	gaaacacact tgcactgcag tgcggttaga tatacgcgta tctgctgctgt ctccgagaac	300
10	gcctgcatta tccctcatag tgattatttg gagctcgatg aaagtttcac ccggctgcac	360
	ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg	420
	ggctctggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct	480
	aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc	540
	ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg	600
	acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac	660
15		
	cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcggg ctgtcggtcc	720
	atgtatgatt ggcttaaaga aatatgtaaa cgcgggggag atcccgcggt gattgagaag	780
	acagcgagtc tcgaacgttt ctttttgagg catcgtgaac gcggagaagt aatgacgtat	840
	taa	843

20

25

<210> 56
 <211> 280
 <212> PRT
 <213> Secuencia artificial

30

<220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
 <400> 56

35

40

45

50

55

60

65

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

45

<210> 57

<211> 843

<212> ADN

<213> Secuencia artificial

50

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

55

<400> 57

60

65

5	atgtctacgc atttccttgg caaggtcaaa ttccgacgagg ccagattagc cgaagatctg 60
	tctacttttag aacgtgctga gtttttctagt gcatattcgg acttcgcttg tggaaagtgg 120
	gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggcttctcac 180
	aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
	gaaacacact tcgactgcag tgcgggttaga tatacgcgta tcgtgcgtgt ctccgagAAC 300
10	gcctgcatta tccctcatag tgattatttg gagctcgatg aaacgttcac ccggctgcac 360
	ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
	ggctctggggc agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
	aaaactccta gactgcactc catgatagat tttgaagcta ccgccttccc ggaatccttc 540
	ttacgcaatg tggAACagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
	acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
15	cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
	atgtatgatt ggcttaaaga aatatgtaaa cgcgcggggag atcccgcggtt gattgagaag 780
	acagcgagtc tcgaacggtt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
	tAA 843
20	<210> 58
	<211> 280
	<212> PRT
	<213> Secuencia artificial
25	<220>
	<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
	<400> 58
30	
	Met Ser Thr His Phe Leu Gly Lys Val Lys Phe Asp Glu Ala Arg Leu
	1 5 10 15
35	Ala Glu Asp Leu Ser Thr Leu Glu Arg Ala Glu Phe Ser Ser Ala Tyr
	20 25 30
	Ser Asp Phe Ala Cys Gly Lys Trp Glu Ala Cys Val Leu Arg Asn Arg
	35 40 45
	Thr Gly Met Gln Glu Glu Asp Ile Val Ala Ser His Asn Ala Pro Ala
	50 55 60
40	Leu Ala Thr Pro Leu Ser Lys Ser Leu Pro Tyr Leu Asn Glu Leu Val
	65 70 75 80
	Glu Thr His Phe Asp Cys Ser Ala Val Arg Tyr Thr Arg Ile Val Arg
	85 90 95
45	Val Ser Glu Asn Ala Cys Ile Ile Pro His Ser Asp Tyr Leu Glu Leu
	100 105 110
	Asp Glu Thr Phe Thr Arg Leu His Leu Val Leu Asp Thr Asn Ser Gly
	115 120 125
50	Cys Ala Asn Thr Glu Glu Asp Lys Ile Phe His Met Gly Leu Gly Glu
	130 135 140
	Ile Trp Phe Leu Asp Ala Met Leu Pro His Ser Ala Ala Cys Phe Ser
	145 150 155 160
	Lys Thr Pro Arg Leu His Leu Met Ile Asp Phe Glu Ala Thr Ala Phe
	165 170 175
55	Pro Glu Ser Phe Leu Arg Asn Val Glu Gln Pro Val Thr Thr Arg Asp
	180 185 190
	Met Val Asp Pro Arg Lys Glu Leu Thr Asp Glu Val Ile Glu Gly Ile
	195 200 205
60	Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val
	210 215 220
	Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
	225 230 235 240
65	Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
	245 250 255
	Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg

ES 2 712 682 T3

[illegible]

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

<210> 61

<211> 843

50 <212> ADN

<213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

55

<400> 61

60

65

ES 2 712 682 T3

```

5      atgactacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg      60
      tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg      120
      gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac      180
      aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt      240
      gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac      300
      gcctgcatta tcctcatag tgattatttg gagctcgatg aaacggtcac cgggctgcac      360
10     ctctgtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg      420
      ggtctggggc agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct      480
      aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc      540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg      600
      acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac      660
15     cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc      720
      atgtatgatt ggcttaaaga aatatgtaaa cgcgggggag atcccgcggt gattgagaag      780
      acagcgagtc tcgaacggtt ctttttgggg catcgtgaac gcggagaagt aatgacgtat      840
      taa

```

```

20     <210> 62
      <211> 280
      <212> PRT
      <213> Secuencia artificial
25     <220>
      <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
30     <400> 62

```

```

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

```

45

50

55

60

65

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

<210> 63

<211> 843

40 <212> ADN

<213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

45

<400> 63

50 atgtctacgc atttccttgg caaggtcaaa ttcgacactg ccagattagc cgaagatctg 60
tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaaagtgg 120
gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
aatgcgcccc cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaaac 300
gcctgcatta tccctcatag tgattatattg gagctcgatg aaacgttcac ccggctgcac 360
ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
55 ggtctgggag agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
60 atgtatgatt ggcttaaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag 780
acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa 843

<210> 64

65 <211> 280

<212> PRT

<213> Secuencia artificial

<220>

5 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 64

```

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

```

50

<210> 65

<211> 843

55 <212> ADN

<213> Secuencia artificial

<220>

60 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 65

65

ES 2 712 682 T3

5	atgtctacgc agttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg	60
	tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg	120
	gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac	180
	aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt	240
	gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagAAC	300
	gcctgcatta tccctcatag tgattatttg gagctcgatg aaacgttcac ccggtgcac	360
10	ctcgtgttgg atactaacag tgggtgtgcc aacactgagg aagataagat attccacatg	420
	ggtctgggCG agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct	480
	aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc	540
	ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg	600
15	acggacgaag ttatagaagg catTTTTaggg ttctccataa tcatcagtga agcgaattac	660
	cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc	720
	atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag	780
	acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat	840
20	taa	843

	<210> 66
	<211> 280
25	<212> PRT
	<213> Secuencia artificial
	<220>
30	<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
	<400> 66

35

40

45

50

55

60

65

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

45

50 <210> 67
<211> 843
<212> ADN
<213> Secuencia artificial

<220>
55 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 67

60

65

ES 2 712 682 T3

5	atgtctacgt	tggttccttg	caagggtcaaa	ttcgacgagg	ccagattagc	cgaagatctg	60									
	tctacttttag	aagtcgctga	gttttctagt	gcatattcgg	acttcgcttg	tggaagtgg	120									
	gaagcatgtg	ttttacgcaa	cagaacaggc	atgcaagaag	aagacattgt	ggtttctcac	180									
	aatgcgcccg	cattagcgac	ccctttatct	aaatcactgc	cgtacttgaa	tgagcttggt	240									
	gaaacacact	tcgactgcag	tgcggttaga	tatacgcgta	tcgtgcgtgt	ctccgagaac	300									
	gcctgcatta	tccctcatag	tgattatttg	gagctcgatg	aaacgttcac	ccggctgcac	360									
10	ctcgtgttgg	atactaacag	tggttgtgcc	aacactgagg	aagataagat	attccacatg	420									
	ggctctgggg	agatctggtt	tctggacgca	atgttacctc	attctgctgc	ctgcttttct	480									
	aaaactccta	gactgcactc	catgatagat	tttgaagcta	ccgccttccc	ggaatccttc	540									
	ttacgcaatg	tggaacagcc	tgtaactact	cgcgacatgg	tagaccctcg	taaagaattg	600									
	acggacgaag	ttatagaagg	catttttaggg	ttctccataa	tcatcagtga	agcgaattac	660									
	cgcgagatcg	tctctatcct	ggccaaatta	cacttctttt	ataaagcgga	ctgtcggtoc	720									
15	atgtatgatt	ggcttaaaga	aatatgtaaa	cgcgggggag	atcccgcggt	gattgagaag	780									
	acagcgagtc	tcgaacgttt	ctttttgggg	catcgtgaac	gcggagaagt	aatgacgtat	840									
	taa						843									
20																
	<210> 68															
	<211> 280															
	<212> PRT															
	<213> Secuencia artificial															
25																
	<220>															
	<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti															
30																
	Met	Ser	Thr	Leu	Phe	Leu	Gly	Lys	Val	Lys	Phe	Asp	Glu	Ala	Arg	Leu
	1				5					10					15	
35	Ala	Glu	Asp	Leu	Ser	Thr	Leu	Glu	Val	Ala	Glu	Phe	Ser	Ser	Ala	Tyr
				20					25					30		
	Ser	Asp	Phe	Ala	Cys	Gly	Lys	Trp	Glu	Ala	Cys	Val	Leu	Arg	Asn	Arg
			35					40					45			
	Thr	Gly	Met	Gln	Glu	Glu	Asp	Ile	Val	Val	Ser	His	Asn	Ala	Pro	Ala
		50					55					60				
40	Leu	Ala	Thr	Pro	Leu	Ser	Lys	Ser	Leu	Pro	Tyr	Leu	Asn	Glu	Leu	Val
	65					70					75					80
	Glu	Thr	His	Phe	Asp	Cys	Ser	Ala	Val	Arg	Tyr	Thr	Arg	Ile	Val	Arg
					85					90					95	
45	Val	Ser	Glu	Asn	Ala	Cys	Ile	Ile	Pro	His	Ser	Asp	Tyr	Leu	Glu	Leu
				100					105					110		
	Asp	Glu	Thr	Phe	Thr	Arg	Leu	His	Leu	Val	Leu	Asp	Thr	Asn	Ser	Gly
				115				120					125			
50	Cys	Ala	Asn	Thr	Glu	Glu	Asp	Lys	Ile	Phe	His	Met	Gly	Leu	Gly	Glu
		130					135					140				
	Ile	Trp	Phe	Leu	Asp	Ala	Met	Leu	Pro	His	Ser	Ala	Ala	Cys	Phe	Ser
	145					150					155					160
	Lys	Thr	Pro	Arg	Leu	His	Leu	Met	Ile	Asp	Phe	Glu	Ala	Thr	Ala	Phe
					165					170					175	
55	Pro	Glu	Ser	Phe	Leu	Arg	Asn	Val	Glu	Gln	Pro	Val	Thr	Thr	Arg	Asp
				180					185					190		
	Met	Val	Asp	Pro	Arg	Lys	Glu	Leu	Thr	Asp	Glu	Val	Ile	Glu	Gly	Ile
			195					200					205			
60	Leu	Gly	Phe	Ser	Ile	Ile	Ile	Ser	Glu	Ala	Asn	Tyr	Arg	Glu	Ile	Val
		210					215					220				
	Ser	Ile	Leu	Ala	Lys	Leu	His	Phe	Phe	Tyr	Lys	Ala	Asp	Cys	Arg	Ser
	225					230					235					240
65	Met	Tyr	Asp	Trp	Leu	Lys	Glu	Ile	Cys	Lys	Arg	Arg	Gly	Asp	Pro	Ala
					245					250					255	

ES 2 712 682 T3

5	Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg	260	265	270
	Glu Arg Gly Glu Val Met Thr Tyr	275	280	
10	<210> 69 <211> 843 <212> ADN <213> Secuencia artificial			
15	<220> <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti			
20	<400> 69			
25	atgtctacgc atattcttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg	60		
	tctacttttag aagtcgctga gttttctagt gcataattcgg acttcgcttg tggaaaagtgg	120		
	gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt gggtttctcac	180		
	aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttgtt	240		
	gaaacacact tcgactgcag tgcggtaga tatacgcgta tcgtgcgtgt ctccgagaac	300		
	gcctgcatta tccctcatag tgattatttg gagctcgatg aaacggttcac ccggctgcac	360		
	ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg	420		
30	gggtctggcg agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct	480		
	aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc	540		
	ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg	600		
	acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac	660		
	cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgtcc	720		
	atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag	780		
35	acagcgagtc tcgaacgttt ctttttgagg catcgtgaac gcggagaagt aatgacgtat	840		
	taa	843		
40	<210> 70 <211> 280 <212> PRT <213> Secuencia artificial			
45	<220> <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti			
	<400> 70			
50				
55				
60				
65				

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

Lys Thr Pro Arg Leu His Leu Met Ile Asp Phe Glu Ala Thr Ala Phe
165 170 175
30 Pro Glu Ser Phe Leu Arg Asn Val Glu Gln Pro Val Thr Thr Arg Asp
180 185 190
Met Val Asp Pro Arg Lys Glu Leu Thr Asp Glu Val Ile Glu Gly Ile
195 200 205
Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val
210 215 220
35 Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
225 230 235 240
Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
245 250 255
40 Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
260 265 270
Glu Arg Gly Glu Val Met Thr Tyr
275 280

45

50 <210> 71
<211> 843
<212> ADN
<213> Secuencia artificial

55 <220>
<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 71

60

65

ES 2 712 682 T3

```

5      atgtctacgc atttgcttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg      60
      tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg      120
      gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac      180
      aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttgtt      240
      gaaacacact tcgactgcag tgcgggttaga tatacgcgta tcgtgcgtgt ctccgagaac      300
      gcctgcatta tccctcatag tgattatttg gagctcgatg aaacggtcac ccggctgcac      360
10     ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg      420
      ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgottttct      480
      aaaactccta gactgcatct catgatatag tttgaagcta ccgccttccc ggaatccttc      540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg      600
      acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agogaattac      660
      cgcgagatcg tctctatcct gcccaaatta caattctttt ataaagcgga ctgtcgggtcc      720
15     atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag      780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat      840
      taa

```

```

20     <210> 72
      <211> 280
      <212> PRT
      <213> Secuencia artificial
25
      <220>
      <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
30
      <400> 72

```

```

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

```

40

45

50

55

60

65

5	50	55	60	
	Leu Ala Thr Pro Leu Ser Lys Ser Leu Pro Tyr Leu Asn Glu Leu Val			
	65	70	75	80
	Glu Thr His Phe Asp Cys Ser Ala Val Arg Tyr Thr Arg Ile Val Arg			
		85	90	95
10	Val Ser Glu Asn Ala Cys Ile Ile Pro His Ser Asp Tyr Leu Glu Leu			
		100	105	110
	Asp Glu Thr Phe Thr Arg Leu His Leu Val Leu Asp Thr Asn Ser Gly			
		115	120	125
	Cys Ala Asn Thr Glu Glu Asp Lys Ile Phe His Met Gly Leu Gly Glu			
15		130	135	140
	Ile Trp Phe Leu Asp Ala Met Leu Pro His Ser Ala Ala Cys Phe Ser			
	145	150	155	160
	Lys Thr Pro Arg Leu His Leu Met Ile Asp Phe Glu Ala Thr Ala Phe			
		165	170	175
20	Pro Glu Ser Phe Leu Arg Asn Val Glu Gln Pro Val Thr Thr Arg Asp			
		180	185	190
	Met Val Asp Pro Arg Lys Glu Leu Thr Asp Glu Val Ile Glu Gly Ile			
		195	200	205
25	Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val			
		210	215	220
	Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser			
	225	230	235	240
	Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala			
30		245	250	255
	Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg			
		260	265	270
	Glu Arg Gly Glu Val Met Thr Tyr			
35		275	280	
	<210> 73			
	<211> 843			
	<212> ADN			
40	<213> Secuencia artificial			
	<220>			
	<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti			
45	<400> 73			
	atgtctacgc atagtcttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg			60
	tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg			120
50	gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac			180
	aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtactgaa tgagcttggt			240
	gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac			300
	gcctgcatta tccctcatag tgattatttg gagctcgatg aaacgttcac ccggtgcac			360
	ctcgtgttgg atactaacag tgggttggtcc aacactgagg aagataagat attccacatg			420
55	ggtctgggag agatctggtt tctggacgca atgtttacctc attctgctgc ctgcttttct			480
	aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc			540
	ttacgcaatg tggaaacagc tgtaactact cgcgacatgg tagaccctcg taaagaattg			600
	acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac			660
	cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcggtcc			720
60	atgtatgatt ggcttaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag			780
	acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat			840
	taa			843
65	<210> 74			

<211> 280
<212> PRT
<213> Secuencia artificial

5 <220>
<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 74

10

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

50

55

<210> 75
<211> 843
<212> ADN
60 <213> Secuencia artificial

<220>
<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

65 <400> 75

ES 2 712 682 T3

5	atgtctacgc	atttccttgg	caagattaaa	ttcgacgagg	ccagattagc	cgaagatctg	60
	tctacttttag	aagtcgctga	gttttctagt	gcatattcgg	acttcgcttg	tggaaagtgg	120
	gaagcatgtg	ttttacgcaa	cagaacaggc	atgcaagaag	aagacattgt	ggtttctcac	180
	aatgcgccc	cattagcgac	ccctttatct	aaatcactgc	cgtacttgaa	tgagcttggt	240
	gaaacacact	tcgactgcag	tgogggttaga	tatacgcgt	tcgtgcgtgt	ctccgagAAC	300
	gcctgcatta	tcctcatag	tgattatttg	gagctcgatg	aaacgttcac	ccggctgcac	360
10	ctcgtgttg	atactaacag	tggttgtgcc	aacactgagg	aagataagat	attccacatg	420
	ggtctggg	agatctggt	tctggacgca	atgttacctc	attctgctgc	ctgcttttct	480
	aaaactccta	gactgcatct	catgatagat	tttgaagcta	ccgccttccc	ggaatccttc	540
15	ttacgcaatg	tggaacagcc	tgtactact	cgcgacatgg	tagaccctcg	taaagaattg	600
	acggacgaag	ttatagaagg	catttttaggg	ttctccataa	tcatcagtga	agcgaattac	660
	cgcgagatcg	tctctatcct	ggccaaatta	cacttctttt	ataaagcgga	ctgtcggtcc	720
	atgtatgatt	ggcttaaaga	aatatgtaaa	cgccggggag	atcccgcggt	gattgagaag	780
	acagcgagtc	tcgaacgttt	ctttttgggg	catcgtgaac	gcggagaagt	aatgacgtat	840
	ttaa						843
20							
	<210> 76						
	<211> 280						
	<212> PRT						
25	<213> Secuencia artificial						
	<220>						
	<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti						
30	<400> 76						
35							
40							
45							
50							
55							
60							
65							

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

<210> 77

<211> 843

<212> ADN

<213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 77

5 atgtctacgg aattccttgg caagggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg 120
gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
aatgcgccc cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
gaaacacact tcgactgcag tgcgggttaga tatacgcgta tcgtgcgtgt ctccgagaac 300
gcctgcatta tccctcatag tgattatttg gagctcgatg aaacgttcac ccggctgcac 360
10 ctctgtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
ggctctggcg agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct 480
aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
15 atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag 780
acagcgagtc tcgaacggtt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa 843

20 <210> 78
<211> 280
<212> PRT
<213> Secuencia artificial
25 <220>
<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
30 <400> 78

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

ES 2 712 682 T3

5				245					250					255		
	Leu	Ile	Glu	Lys	Thr	Ala	Ser	Leu	Glu	Arg	Phe	Phe	Leu	Gly	His	Arg
				260					265					270		
	Glu	Arg	Gly	Glu	Val	Met	Thr	Tyr								
			275					280								

<210> 79
<211> 843
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 79

25	atgtctacgc	tcttccttgg	caaggtcaaa	ttcgacgagg	ccagattatgc	cgaagatctg	60
	tctacttttag	aagtoctga	gttttctagt	gcatattcgg	acttcgcttg	tggaaagtgg	120
	gaagcatgtg	ttttacgcaa	cagaacaggc	atgcaagaag	aagacattgt	ggtttctcac	180
	aatgcgccc	cattagcgac	ccctttatct	aaatcactgc	cgtacttgaa	tgagcttgtt	240
	gaacacact	tcgactgcag	tcgggttaga	tatcgcgcta	tcgtgcgttg	ctccgagaac	300
30	gcctgcatta	tccctcatag	tgattatttg	gagctcgatg	aagcgttcac	ccggctgcac	360
	ctcgtgttgg	atactaacag	tggttgtgcc	aacactgagg	aagataagat	attccacatg	420
	ggctctggcg	agatctggtt	tctggacgca	atgttacctc	attctgctgc	ctgcttttct	480
	aaaactccta	gactgcatct	catgatagat	tttgaaagta	ccgcctctccc	ggaactccttc	540
	ttacgcaatg	tggaaacgcc	tgtaactact	cgcgacatgg	tagaccctcg	taaagaattg	600
35	acggacgaag	ttatagaagg	catttttaggg	ttctccataa	tcatcagtga	agcgaattac	660
	cgcgagatcg	tctctatcct	ggccaaatta	cacttctttt	ataaagcgga	ctgtcgcgtc	720
	atgatatgatt	ggcttaaaga	aatatgtaa	cgccggggag	atcccgcggt	gattgagaag	780
	acagcgagtc	tcgaacgttt	ctttttgggg	catcgtgaac	gcggagaagt	aatgacgtat	840
	taa						843

<210> 80
<211> 280
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 80

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

50 <210> 81
 <211> 843
 <212> ADN
 <213> Secuencia artificial

55 <220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 81

60

65

ES 2 712 682 T3

```

5      atgtctacgt cattccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg      60
      tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg      120
      gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggttttctcac      180
      aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt      240
      gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac      300
10     gcctgcatta tcccccatag tgattatgtg gagctcgatg aaacgttcac ccggtgcac      360
      ctgctgttgg atactaacag tgggtgtgcc aacactgagg aagataagat attccacatg      420
      ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct      480
      aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc      540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg      600
15     acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac      660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcggtcc      720
      atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag      780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat      840
      taa

```

20

<210> 82

<211> 280

<212> PRT

25 <213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

30 <400> 82

```

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

```

40

45

50

55

60

65

5	Thr Gly Met Gln Glu Glu Asp Ile Val Val Ser His Asn Ala Pro Ala	
	50 55 60	
	Leu Ala Thr Pro Leu Ser Lys Ser Leu Pro Tyr Leu Asn Glu Leu Val	
	65 70 75 80	
10	Glu Thr His Phe Asp Cys Ser Ala Val Arg Tyr Thr Arg Ile Val Arg	
	85 90 95	
	Val Ser Glu Asn Ala Cys Ile Ile Pro His Ser Asp Tyr Leu Glu Leu	
	100 105 110	
	Asp Glu Thr Phe Thr Arg Leu His Leu Val Leu Asp Thr Asn Ser Gly	
15	115 120 125	
	Cys Ala Asn Thr Glu Glu Asp Lys Ile Phe His Met Gly Leu Gly Glu	
	130 135 140	
	Ile Trp Phe Leu Asp Ala Met Leu Pro His Ser Ala Ala Cys Phe Ser	
20	145 150 155 160	
	Lys Thr Pro Arg Leu His Leu Met Ile Asp Phe Glu Ala Thr Ala Phe	
	165 170 175	
	Pro Glu Ser Phe Leu Arg Asn Val Glu Gln Pro Val Thr Thr Arg Asp	
	180 185 190	
25	Met Val Asp Pro Arg Lys Glu Leu Thr Asp Glu Val Ile Glu Gly Ile	
	195 200 205	
	Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val	
	210 215 220	
30	Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser	
	225 230 235 240	
	Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala	
	245 250 255	
	Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg	
	260 265 270	
35	Glu Arg Gly Glu Val Met Thr Tyr	
	275 280	
40	<210> 83	
	<211> 843	
	<212> ADN	
	<213> Secuencia artificial	
45	<220>	
	<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti	
	<400> 83	
50	atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg	60
	tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaaagtgg	120
	gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac	180
	aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt	240
55	gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac	300
	gcctgcatta tccctcatag tgattatttg gagctcgatg aaacgttcac ccggctgcac	360
	ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg	420
	ggtctggggc agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct	480
	aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc	540
	ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg	600
60	acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac	660
	cgcgagatcg tctggatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc	720
	atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag	780
	acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat	840
	taa	843
65		

<210> 84
 <211> 280
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
 <400> 84

```

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

<210> 85
 <211> 843
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

ES 2 712 682 T3

<400> 85

5	atgtctacgc	atctccttgg	caaggtcaaa	ttcgacgagg	ccagattagc	cgaagatctg	60
	tctacttttag	aagtcgctga	gttttctagt	gcataattcgg	acttcgcttg	tggaagtgg	120
	gaagcatgtg	ttttacgcaa	cagaacaggc	atgcaagaag	aagacattgt	ggtttctcac	180
	aatgcgccc	cattagcgac	ccctttatct	aaatcactgc	cgtacttgaa	tgagcttggt	240
	gaaacacact	tcgactgcag	tgcggttaga	tatacgcgta	tcgtgcgtgt	ctccgagaac	300
10	gcctgcatta	tccctcatag	tgattatttg	gagctcgatg	aaacgttcac	ccggctgcac	360
	ctcgtgttgg	atactaacag	tggttggtgc	aacactgagg	aagataagat	attccacatg	420
	ggctctggg	agatctggtt	tctggacgca	atgttacctc	attctgctgc	ctgcttttct	480
15	aaaactccta	gactgcatct	catgatagat	tttgaagcta	ccgccttccc	ggaatccttc	540
	ttacgcaatg	tggaacagcc	tgtaactact	cgcgacatgg	tagaccctcg	taaagaattg	600
	acggacgaag	ttatagaagg	cattttaggg	ttctccataa	tcatcagtga	agcgaattac	660
	cgcgagatcg	tccttatcct	ggccaaatta	cacttctttt	ataaagcgga	ctgtcggtcc	720
	atgtatgatt	ggcttaaaga	aatatgtaaa	cgcgggggag	atcccgcggt	gattgagaag	780
	acagcgagtc	tcgaacgttt	ctttttgggg	catcgtgaac	gcggagaagt	aatgacgtat	840
20	taa						843

<210> 86

25 <211> 280

<212> PRT

<213> Secuencia artificial

<220>

30 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 86

35

40

45

50

55

60

65

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

45

<210> 87

<211> 843

50

<212> ADN

<213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

55

<400> 87

60

65

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg caagggtcaaa ttcgacgagg ccagattagc cgaagatctg      60
      tctacttttag aagtcgctga gtttttctagt gcatattcgg acttcgcttg tggaaaagtgg      120
      gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac      180
      aatgcgcccg cattagcgac ccctttatct aaactactgc cgtacttgaa tgagcttggt      240
      gaaacacact tcgactgcag tgcgggttaga tatacgcgta tcgtgcgtgt ctccgagaac      300
      gcctgcatta tccctcatag tgattatttg gagctcgatg aaacgttcac ccggctgcac      360
10     ctcggttggg atactaacag tgggttggtcc aacactgagg aagataagat attccacatg      420
      ggtctggggc agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct      480
      aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc      540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg      600
      acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac      660
15     cgcgagatcg tctctatcct ggccaaagtg cacttctttt ataaagcggg ctgtcgggtcc      720
      atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag      780
      acagcgagtc tcgaacggtt ctttttgggg catcgtgaac gcggagaagt aatgacgtat      840
      taa

```

```

20
<210> 88
<211> 280
<212> PRT
<213> Secuencia artificial
25
<220>
<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
30
<400> 88

```

```

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

```

ES 2 712 682 T3

5 Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
245 250 255
Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
260 265 270
10 Glu Arg Gly Glu Val Met Thr Tyr
275 280

15 <210> 89
<211> 843
<212> ADN
<213> Secuencia artificial
20 <220>
<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
<400> 89

25
30 atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg 120
gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggttttctcac 180
aatgcgcccg cattagcgac ccttttatct aaatcactgc cgtacttgaa tgagcttggt 240
gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagAAC 300
gcctgcatta tccctcatag tgattatttg gagctcgatg aaacgttcac ccggtgcac 360
ctcgtgttgg atactaacag tggttgtgcc ttactgagg aagataagat attccacatg 420
ggctctgggag agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
35 aaaactccta gactgcatct catgatagat ttggaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggAACagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag 780
40 acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa 843

45 <210> 90
<211> 280
<212> PRT
<213> Secuencia artificial
50 <220>
<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
<400> 90

55

60

65

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

45

<210> 91

<211> 843

<212> ADN

50

<213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

55

<400> 91

60

65

5 Met Ser Thr His Phe Leu Gly Lys Val Lys Phe Asp Glu Ala Arg Leu
1 5 10 15
Ala Glu Asp Leu Ser Thr Leu Glu Val Ala Glu Phe Ser Ser Ala Tyr
20 25 30
Ser Asp Phe Ala Cys Gly Lys Trp Glu Ala Cys Val Leu Arg Asn Arg
35 40 45
10 Thr Gly Met Gln Glu Glu Asp Ile Val Val Ser His Asn Ala Pro Ala
50 55 60
Leu Ala Thr Pro Leu Ser Lys Ser Leu Pro Tyr Leu Asn Glu Leu Val
65 70 75 80
15 Glu Thr His Phe Asp Cys Ser Ala Val Arg Tyr Thr Arg Ile Val Arg
85 90 95
Val Ser Glu Asn Ala Cys Ile Ile Pro His Ser Asp Tyr Leu Glu Leu
100 105 110
20 Asp Glu Thr Phe Thr Arg Leu His Leu Val Leu Asp Thr Asn Ser Gly
115 120 125
Cys Ala Phe Thr Glu Glu Asp Lys Ile Phe His Met Gly Leu Gly Glu
130 135 140
25 Ile Trp Phe Leu Asp Ala Met Leu Pro His Ser Ala Ala Cys Phe Ser
145 150 155 160
Lys Thr Pro Arg Leu His Leu Met Ile Asp Phe Glu Ala Thr Ala Phe
165 170 175
30 Pro Glu Ser Phe Leu Arg Asn Val Glu Gln Pro Val Thr Thr Arg Asp
180 185 190
Met Val Asp Pro Arg Lys Glu Leu Thr Asp Glu Val Ile Glu Gly Ile
195 200 205
Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val
210 215 220
35 Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
225 230 235 240
Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
245 250 255
40 Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
260 265 270
Glu Arg Gly Glu Val Met Thr Tyr
275 280
45
50 <210> 92
<211> 280
<212> PRT
<213> Secuencia artificial
55 <220>
<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
60 <400> 92
Met Ser Thr His Phe Leu Gly Lys Val Lys Phe Asp Glu Ala Arg Leu
1 5 10 15
Ala Glu Asp Leu Ser Thr Leu Glu Val Ala Glu Phe Ser Ser Ala Tyr
20 25 30
Ser Asp Phe Ala Cys Gly Lys Trp Glu Ala Cys Val Leu Arg Asn Arg
65

		35		40		45	
5	Thr	Gly	Met	Gln	Glu	Glu	Asp
	50					55	
	Leu	Ala	Thr	Pro	Leu	Ser	Lys
	65					70	
10	Glu	Thr	His	Phe	Asp	Cys	Ser
				85			
	Val	Ser	Glu	Asn	Ala	Cys	Ile
				100			
	Asp	Glu	Thr	Phe	Thr	Arg	Leu
				115			
15	Cys	Ala	Asn	Thr	Glu	Glu	Asp
	Ile	Trp	Phe	Leu	Asp	Ala	Met
	145					150	
20	Lys	Thr	Pro	Arg	Leu	His	Leu
	Pro	Glu	Ser	Phe	Leu	Arg	Asn
	Met	Val	Asp	Pro	Arg	Lys	Glu
25	Leu	Gly	Phe	Ser	Ile	Ile	Ile
	Ser	Ile	Leu	Ala	Lys	Leu	His
	225					230	
30	Met	Tyr	Asp	Trp	Leu	Lys	Glu
	Leu	Ile	Glu	Lys	Thr	Ala	Ser
35	Glu	Arg	Gly	Glu	Val	Met	Thr

<210> 93
 <211> 843
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 93

50	atgtctacgc	atttccttgg	caaggtcaaa	ttcgacgagg	ccagattagc	cgaagatctg	60
	tctacttttag	aagtcgctga	gttttctagt	gcatattcgg	acttcgcttg	tggaaagtgg	120
	gaagcatgtg	ttttacgcaa	cagaacaggc	atgcaagaag	aagacattgt	ggtttctcac	180
	aatgcgccc	cattagcgac	ccctttatct	aaatcactgc	cgtacttgaa	tgagcttggt	240
55	gaaacacact	tcgactgcag	tgcggttaga	tatacgcgta	tcatgcgtgt	ctccgagAAC	300
	gcctgcatta	tcctcatag	tgattatttg	gagctcgatg	aaacggtcac	ccggctgcac	360
	ctcgtgttgg	atactaacag	tggttgtgcc	aacactgagg	aagataagat	attccacatg	420
	ggctctggcg	agatctggtt	tctggacgca	atgttacctc	attctgctgc	ctgcttttct	480
	aaaactccta	gactgcatct	catgatagat	tttgaagcta	ccgccttccc	ggaatccttc	540
	ttacgcaatg	tggaacagcc	tgtaactact	cgcgacatgg	tagaccctcg	taaagaattg	600
60	acggacgaag	ttatagaagg	catttttaggg	ttctccataa	tcatcagtga	agcgaattac	660
	cgcgagatcg	tctctatcct	ggccaaatta	cacttctttt	ataaagcgga	ctgtcgtgtc	720
	atgtatgatt	ggcttaaaga	aatatgtaaa	cgccggggag	atcccgcggt	gattgagaag	780
	acagcgagtc	tcgaacgttt	ctttttgggg	catcgtgaac	gcggaagaag	aatgacgtat	840
	taa						843

ES 2 712 682 T3

5 <210> 94
 <211> 280
 <212> PRT
 <213> Secuencia artificial

 10 <220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

 <400> 94

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

 55 <210> 95
 <211> 843
 <212> ADN
 60 <213> Secuencia artificial

 <220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

 <400> 95

ES 2 712 682 T3

5	atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
	tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg 120
	gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
	aatgcgcccg cattagcgac ccccttatct aaatcactgc cgtacttgaa tgagcttggt 240
	gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac 300
10	gcctgcatta tcctcatag tgattatttg gagctcgatg aaggcttcac ccgctgcac 360
	ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
	ggctctggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
	aaaactccta gactgcatct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
15	ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
	acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
	cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcggg ctgtcggtcc 720
	atgtatgatt ggcttaaaga aatatgtaaa cgcgggggag atcccgcggt gattgagaag 780
	acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
20	taa 843
	<210> 96
	<211> 280
25	<212> PRT
	<213> Secuencia artificial
	<220>
30	<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
	<400> 96
35	
40	
45	
50	
55	
60	
65	

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

<210> 97

<211> 843

<212> ADN

<213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 97

```

5      atgtctagcc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg      60
      tctacttttag aagtcgctga gtttttctagt gcatattcgg acttcgcttg tggaaaagtgg      120
      gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac      180
      aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt      240
      gaaacacact tcgactgcag tgcgggttaga tatacgcgta tcgtgcgtgt ctccgagaac      300
      gcctgcatta tccctcatag tgattatttg gagctcgatg aaacgttcac ccggctgcac      360
10     ctctgtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg      420
      ggtctgggag agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct      480
      aaaactccta gactgcctct catgatagat tttgaagcta ccgccttccc ggaatccttc      540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg      600
      acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac      660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc      720
15     atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag      780
      acagcgagtc tcgaacggtt ctttttgggg catcgtgaac gcggagaagt aatgacgtat      840
      taa

```

```

20
<210> 98
<211> 280
<212> PRT
<213> Secuencia artificial
25
<220>
<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
30
<400> 98

```

```

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

```

ES 2 712 682 T3

5	225	230	235	240
	Met Tyr Asp Trp	Leu Lys Glu Ile Cys	Lys Arg Arg Gly Asp	Pro Ala
		245	250	255
	Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg			
		260	265	270
10	Glu Arg Gly Glu Val Met Thr Tyr			
		275	280	

15

<210> 99
<211> 843
<212> ADN
<213> Secuencia artificial

20

<220>
<223> Variante diseñada de prolina 4-hidroxilasa de *Sinorhizobium meliloti*

<400> 99

30	atgagtaccc	actttctcgg	caaagtga	ttcgatgaag	cacggcttgc	agaagacttg	60
	tctacgctgg	aagtggcgga	attttcctcc	gcgtatagtg	attttgcctg	cggaaatgg	120
	gaagcatgtg	ttcttagaaa	ccgtacaggt	atgcaagaag	aggatatattg	ggtttcccat	180
	aacgcaccgg	ctcttgcaac	ccgctgagt	aatcattgc	cttaccttaa	tgaattagtt	240
	gagacacatt	tcgactgttc	ggctgtccgt	tacacgcgta	ttgtaaagat	gagcgagaac	300
	gcgtgcatta	taccacattc	agattacctc	gagcttgacg	agacatttac	cagattgcat	360
35	ctggctcttg	ataccaactc	aggttgtgcc	aatacgaag	aggacaagat	attccatatg	420
	ggctctcggtg	aaatatgggt	cttggtatgcg	atgctgcctc	atagtgacgc	atgtttttct	480
	aagacgccta	gactgcagct	gatgattgac	ttcgaagcta	cagcgctttcc	cgaattcattt	540
	ttacggaacg	tcgaacagcc	ggttaccacc	cgtgatattg	tcgaccctga	aaaagaatta	600
	acggacgaag	tgatcgaagg	tatactcggc	ttttctatta	tcatatcgga	ggcgaactac	660
	cgtgaaatcg	tgcttatttt	agcgaagctg	cacttctttt	ataaagccga	ttgccggtcg	720
40	atgatatgatt	ggctgaaaga	aatttgcaag	agacggggag	acccggcgct	tattgaaaag	780
	acggcttcatt	tggaacgctt	ctttttaggc	cacagagaac	gcgggggaagt	aatgacctac	840
	ttaa						843

45	<210>	100
	<211>	280
	<212>	PRT
	<213>	Secuencia artificial
50	<220>	
	<223>	Variante diseñada de prolina 4-hidroxilasa de <i>Sinorhizobium meliloti</i>
	<400>	100

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

45

<210> 101

<211> 843

50 <212> ADN

<213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

55

<400> 101

60

65

ES 2 712 682 T3

```

5      atgagtaccc actttctcgg caaagtgaaa ttcgatgaag cacggcttgc agaagacttg      60
      tctacgctgg aagtggcgga attttcctcc gcgtatagtg attttgcctg cggaaaaatgg      120
      gaagcatgtg ttcttagaaa ccgtacaggt atgcaagaag aggatattgt ggtttcccat      180
      aacgcaccgg ctcttgcaac cccgctgagt aaatcattgc cttaccttaa tgaattagtt      240
      gagacacatt tcgactgttc ggctgtccgt tacacgcgta ttgtaagagt gagcgagaac      300
      gcgtgcctga taccacattc agattacctc gagcttgacg agacatttac cagattgcat      360
10     ctggctcttg ataccaactc aggttgtgcc aataccgaag aggacaagat attccatatg      420
      ggtctcggtg aaatatgggt cttggatgcg atgctgcctc atagtgcagc atgtttttct      480
      aagacgccta gactgcatct gatgattgac ttcgaagcta cagcgtttcc cgaatcattt      540
      ttacggaacg tcgaacagcc ggttaccacc cgtgatatgg tcgaccctag aaaagaatta      600
      acggacgaag tgatcgaagg tatactcggc ttttctatta tcatatcgga ggcgaaactac      660
      cgtgaaatcg tgtctatttt agcgaagctg cacttctttt ataaagccga ttgccggtcg      720
15     atgtatgatt ggctgaaaga aatttgcaag agacggggag acccgggcgt tattgaaaag      780
      acggcttcat tggaacgctt ctttttaggc cacagagaac gcggggaagt aatgacctac      840
      taa

```

```

20     <210> 102
      <211> 280
      <212> PRT
      <213> Secuencia artificial
25     <220>
      <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
30     <400> 102

```

```

      Met Ser Thr His Phe Leu Gly Lys Val Lys Phe Asp Glu Ala Arg Leu
      1          5          10          15
35     Ala Glu Asp Leu Ser Thr Leu Glu Val Ala Glu Phe Ser Ser Ala Tyr
           20           25           30

```

40

45

50

55

60

65

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

40 <210> 103
<211> 843
<212> ADN
<213> Secuencia artificial

45 <220>
<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
400> 103

50 atgagtaccc actttctcgg caaagtgaag ttccgatgaag cacggcttgc agaagacttg 60
tctacgctgg aagtggcgga attttcctcc gcgtatagtg attttgcctg cggaaaatgg 120
gaagcatgtg ttcttagaaa ccgtacaggt atgcaagaag aggatattgt ggtttcccat 180
55 aacgcaccgg ctcttgcaac cccgctgagt aaatcattgc cttaccttaa tgaattagtt 240
gagacacatt tcgactgttc ggctgtccgt tacacgcgta ttgtaagagt gagcgagaac 300
gcgtgcctga taccacattc agattacctc gagcttgacg agacatttac cagattgcat 360
ctggtcttgg ataccaactc aggttgtgcc aataccgaag aggacaagat attccatag 420
ggtctcgggtg aaatatgggt cttggatgag atgctgcctc atagtgcagc atgtttttct 480
aagacgccta gactgcagct gatgattgac ttccaagcta cagcgtttcc cgaatcattt 540
60 ttacggaaac tcgaacagcc gggtaccacc cgtgatatgg tcgaccctag aaaagaatta 600
acggacgaag tgatcgaagg tatactcggc ttttctatta tcatatcgga ggccgaactac 660
cgtgaaatcg tgtctatatt agcgaagctg cacttctttt ataaagccga ttgccggctc 720
atgtatgatt ggctgaaaga aatttgcaag agacggggag acccggcgct tattgaaaag 780
acggcttcat tggaacgctt ctttttaggc cacagagaac gcggggaagt aatgacctac 840
65 taa 843

<210> 104
 <211> 280
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 104

```

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

```

<210> 105
 <211> 843
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 105

ES 2 712 682 T3

5	atgagtaccc	acttttctcgg	caaagtgaag	ttcgatgaag	cacggcttgc	agaagacttg	60									
	tctacgctgg	aagtggcgga	attttcctcc	gcgtatagtg	attttgcctg	cggaaaatgg	120									
	gaagcatgtg	ttcttagaaa	ccgtacaggt	atgcaagaag	aggatattgt	ggtttcccat	180									
	aacgcaccgg	ctcttgcaac	cccgtgagt	aaatcattgc	cttaccttaa	tgaattagtt	240									
	gagacacatt	tcgactgttc	ggctgtccgt	tacacgcgta	ttgtaagagt	gagcgagaac	300									
	gcgtgcctga	taccacattc	agattacctc	gagcttgacg	agacattttac	cagattgcat	360									
10	ctggtcttgg	ataccaactc	aggttgtgcc	tataccgaag	aggacaagat	attccatatg	420									
	gggtctcggg	aaatatggtt	cttgatgacg	atgctgcctc	atagtgcagc	atgtttttct	480									
	aagacgccta	gactgcagct	gatgattgac	ttcgaagcta	cagcgtttcc	cgaatcattt	540									
	ttacggaacg	tcgaacagcc	ggttaccacc	cgtgatattg	tcgaccctag	aaaagaatta	600									
	acggacgaag	tgatcgaagg	tatactcggc	ttttctatta	tcatatcgga	ggcgaactac	660									
15	cgtgaaatcg	tgtctatttt	agcgaagctg	cacttctttt	ataaagccga	ttgccggctg	720									
	atgtatgatt	ggctgaaaga	aatttgcaag	agacggggag	acccggcgct	tattgaaaag	780									
	acggcttcat	tggaacgctt	ctttttaggc	cacagagaac	gcggggaagt	aatgacctac	840									
	taa						843									
20	<210>	106														
	<211>	280														
	<212>	PRT														
	<213>	Secuencia artificial														
25	<220>															
	<223>	Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti														
	<400>	106														
30	Met	Ser	Thr	His	Phe	Leu	Gly	Lys	Val	Lys	Phe	Asp	Glu	Ala	Arg	Leu
	1				5					10					15	
	Ala	Glu	Asp	Leu	Ser	Thr	Leu	Glu	Val	Ala	Glu	Phe	Ser	Ser	Ala	Tyr
				20					25					30		
	Ser	Asp	Phe	Ala	Cys	Gly	Lys	Trp	Glu	Ala	Cys	Val	Leu	Arg	Asn	Arg
35			35				40						45			
	Thr	Gly	Met	Gln	Glu	Glu	Asp	Ile	Val	Val	Ser	His	Asn	Ala	Pro	Ala
	50					55						60				
	Leu	Ala	Thr	Pro	Leu	Ser	Lys	Ser	Leu	Pro	Tyr	Leu	Asn	Glu	Leu	Val
	65				70						75				80	
40	Glu	Thr	His	Phe	Asp	Cys	Ser	Ala	Val	Arg	Tyr	Thr	Arg	Ile	Val	Arg
				85						90				95		
	Val	Ser	Glu	Asn	Ala	Cys	Leu	Ile	Pro	His	Ser	Asp	Tyr	Leu	Glu	Leu
				100					105					110		
	Asp	Glu	Thr	Phe	Thr	Arg	Leu	His	Leu	Val	Leu	Asp	Thr	Asn	Ser	Gly
45			115				120					125				
	Cys	Ala	Tyr	Thr	Glu	Glu	Asp	Lys	Ile	Phe	His	Met	Gly	Leu	Gly	Glu
		130				135					140					
	Ile	Trp	Phe	Leu	Asp	Ala	Met	Leu	Pro	His	Ser	Ala	Ala	Cys	Phe	Ser
	145				150						155				160	
50	Lys	Thr	Pro	Arg	Leu	Gln	Leu	Met	Ile	Asp	Phe	Glu	Ala	Thr	Ala	Phe
				165						170					175	
	Pro	Glu	Ser	Phe	Leu	Arg	Asn	Val	Glu	Gln	Pro	Val	Thr	Thr	Arg	Asp
				180					185					190		
	Met	Val	Asp	Pro	Arg	Lys	Glu	Leu	Thr	Asp	Glu	Val	Ile	Glu	Gly	Ile
55			195				200						205			
	Leu	Gly	Phe	Ser	Ile	Ile	Ile	Ser	Glu	Ala	Asn	Tyr	Arg	Glu	Ile	Val
		210				215						220				
	Ser	Ile	Leu	Ala	Lys	Leu	His	Phe	Phe	Tyr	Lys	Ala	Asp	Cys	Arg	Ser
	225				230						235				240	
60	Met	Tyr	Asp	Trp	Leu	Lys	Glu	Ile	Cys	Lys	Arg	Arg	Gly	Asp	Pro	Ala
				245						250				255		
	Leu	Ile	Glu	Lys	Thr	Ala	Ser	Leu	Glu	Arg	Phe	Phe	Leu	Gly	His	Arg
			260						265					270		
	Glu	Arg	Gly	Glu	Val	Met	Thr	Tyr								
65		275						280								

	<210> 107	
	<211> 843	
	<212> ADN	
5	<213> Secuencia artificial	
	<220>	
	<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti	
10	<400> 107	
	atgagtagcc actttctcgg caaagtgaag ttcgatgaag cacggcttgc agaagacttg 60	
	tctacgctgg aagtggcgga attttcctcc gcgtatagtg attttgcctg cggaaaatgg 120	
15	gaagcatgtg ttcttagaaa ccgtacaggt atgcaagaag aggatattgt ggtttcccat 180	
	aacgcaccgg ctcttgcaac ccgctgagt aaatcattgc cttaccttaa tgaattagtt 240	
	gagacacatt tcgactgttc ggctgtccgt tacacgcgta ttgtaagagt gagcgagaac 300	
	gcgtgcctga taccacattc agattacctc gagcttgacg agacatttac cagattgcat 360	
	ctggtcttgg ataccaactc aggttgtgcc aataccgaag aggacaagat attccatatg 420	
20	ggtctcgggtg aaatatggtt cttggatgcg atgctgcctc atagtgcagc atgtttttct 480	
	aagacgccta gactgcagct gatgattgac ttcgaagcta cagcgtttcc cgaatcattt 540	
	ttacggaacg tcgaacagcc ggttaccacc cgtgatatgg tcgaccctag aaaagaatta 600	
	acggaogaag tgatogaagg tatactcggc ttttctatta tcatatcgga ggcgaactac 660	
	cgtgaaatcg tgtctatattt agcgaagctg cacttctttt ataaagccga ttgccggtcg 720	
25	atgtatgatt ggctgaaaga aatttgcaag agacggggag acccggcgct tattgaaaag 780	
	acggcttcat tggaacgctt ctttttaggc cacagagaac gcggggaagt aatgacctac 840	
	taa 843	
30	<210> 108	
	<211> 280	
	<212> PRT	
	<213> Secuencia artificial	
35	<220>	
	<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti	
	<400> 108	
40		
45		
50		
55		
60		
65		

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

<210> 109

<211> 843

<212> ADN

<213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 109

ES 2 712 682 T3

5
10
15

```

atgagtaccc actttctcgg caaagtgaag ttcgatgaag cacggcttgc agaagacttg      60
tctacgctgg aagtgcagga attttcctcc gcgtatagtg attttgcttg cggaaaatgg      120
gaagcatgtg ttcttagaaa ccgtacaggt atgcaagaag aggatattgt ggtttcccat      180
aacgcaccgg ctcttgcaac cccgctgagt aaatcattgc cttaccttaa tgaattagtt      240
gagacacatt tcgactgttc ggctgtccgt tacacgcgta ttgtaagagt gagcgagAAC      300
gcgtgcctga taccacattc agattacctc gagcttgacg agacattttac cagattgcat      360
ctggtcttgg ataccaactc aggttgtgcc aataccgaag aggacaagat attccatatg      420
ggctctgggtg aaatatgggt cttggatgcg atgctgcctc atagtgcagc atgtttttct      480
aagacgccta gactgcagct gatgattgac ttcgaagcta cagcgtttcc cgaatcattt      540
ttacggaacg tcgaacagcc ggttaccacc cgtgatattg tcgaccctag aaaagaatta      600
acggacgaag tgatcgaagg tatactcggc ttttctatta tcatatcgga gccgaactac      660
cgtgaaatcg tgtctatttt agcgaagctg cacttctttt ataaagccga ttgccgggtcg      720
atgtatgatt ggctgaaaaga aatttgcaag agacggggag acccggcgct tattgaaaag      780
acggcttcat tggaacgctt ctttttaggc cacagagaac gcggggaagt aatgacctac      840
taa

```

20
25

```

<210> 110
<211> 280
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 110

```

30
35
40
45
50
55
60
65

```

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

Cys Ala Asn Thr Glu Glu Asp Lys Ile Phe His Met Gly Leu Gly Glu
130      135      140
Ile Trp Phe Leu Asp Ala Met Leu Pro His Ser Ala Ala Cys Phe Ser
145      150      155      160
Lys Thr Pro Arg Leu Gln Leu Met Ile Asp Phe Glu Ala Thr Ala Phe
165      170      175
Pro Glu Ser Phe Leu Arg Asn Val Glu Gln Pro Val Thr Thr Arg Asp
180      185      190
Met Val Asp Pro Arg Lys Glu Leu Thr Asp Glu Val Ile Glu Gly Ile
195      200      205
Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val
210      215      220
Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
225      230      235      240
Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
245      250      255
Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
260      265      270
Glu Arg Gly Glu Val Met Thr Tyr
275      280

```

ES 2 712 682 T3

<210> 111
 <211> 843
 <212> ADN
 5 <213> Secuencia artificial

 <220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

 10 <400> 111

 atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
 tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg 120
 15 gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
 aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
 gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac 300
 gcctgcctga tccctcatag tgattatttg gagctcgatg aaacgttcac ccggctgcac 360
 ctcggtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
 20 ggtctgggag agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
 aaaactccta gactgcagct catgatagat tttgaagcta ccgccttocc ggaatccttc 540
 ttacgcaatg tgggaacagcc tgttaactact cgcgacatgg tagaccctcg taaagaattg 600
 acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
 cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtc 720
 25 atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag 780
 acagcgagtc tgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
 taa 843

 <210> 112
 30 <211> 280
 <212> PRT
 <213> Secuencia artificial

 <220>
 35 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

 <400> 112

 40 Met Ser Thr His Phe Leu Gly Lys Val Lys Phe Asp Glu Ala Arg Leu
 1 5 10 15
 Ala Glu Asp Leu Ser Thr Leu Glu Val Ala Glu Phe Ser Ser Ala Tyr

 45

 50

 55

 60

 65

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

40 <210> 113
 <211> 843
 <212> ADN
 <213> Secuencia artificial

45 <220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 113

50	atgtctacgc	atttccttgg	caaggtcaaa	ttcgacgagg	ccagattagc	cgaagatctg	60
	tctacttttag	aagtcgctga	gttttctagt	gcatattcgg	acttcgcttg	tggaaagtgg	120
	gaagcatgtg	ttttacgcaa	cagaacaggg	atgcaagaag	aagacattgt	ggttttctcac	180
	aatgcgcccc	cattagcgac	ccctttatct	aaatcactgc	cgtacttgaa	tgagcttggt	240
55	gaaacacact	tcgacagcag	tgcggttaga	tatgtccgta	tcgtgcgtgt	ctccgagaac	300
	gcctgcctga	tcctcatag	tgattatttg	gagctcgatg	aaacgttcac	ccggctgcac	360
	ctcgtgttgg	atactaacag	tggttgtgcc	aacactgagg	aagataagat	attccacatg	420
	ggctcgggag	agatctggtt	tctggacgca	tccttacctc	attctgctgc	ctgcttttct	480
	aaaactccta	gactgcagct	catgatagat	tttgaagcta	ccgccttccc	ggaatccttc	540
	ttacgcaatg	tggaacagcc	tgtaactact	cgcgacatgg	tagaccctcg	taaagaattg	600
60	acggacgaag	ttatagaagg	cattttaggg	ttctccataa	tcacagtgta	agcgaattac	660
	cgcgagatcg	tctctatcct	ggccaaatta	cacttctttt	ataaagcgga	ctgtcgggtc	720
	atgtatgatt	ggcttaaaga	aatatgtaaa	cgccggggag	atcccgcggt	gattgagaag	780
	acagcgagtc	tcgaacgttt	ctttttggag	catcgtgaac	gcggagaagt	aatgacgtat	840
	taa						843

65

<210> 114

<211> 280

<212> PRT

5 <213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

10 <400> 114

```

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

```

55

60

65

ES 2 712 682 T3

	<210> 115	
	<211> 843	
	<212> ADN	
5	<213> Secuencia artificial	
	<220>	
	<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti	
10	<400> 115	
15	atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg tctacttttag aagtcgctga gttttctagt gcatattcgg acttcgcttg tggaaagtgg gaagcatgtg ttttacgcaa cagaacaggg atgcaagaag aagacattgt ggtttctcac aatgcgcccg cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac	60 120 180 240 300
20	gcctgccaga tccctcatag tgattatttg gagctcgatg aaacgttaac ccggtgcgat ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg ggtctgggag agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac	360 420 480 540 600 660
25	cgcgagatcg tcctgatcct ggccaaatta caattctttt ataaagcgga ctgtcggtcc atgtatgatt ggcttaaaga aatatgtaaa cgcgggggag atcccgcgtt gattgagaag acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat taa	720 780 840 843
30	<210> 116	
	<211> 280	
	<212> PRT	
	<213> Secuencia artificial	
35	<220>	
	<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti	
40	<400> 116	
45		
50		
55		
60		
65		

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

45 <210> 117
 <211> 843
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 117

55

60

65

5 atgtctacgc atttccttgg caagggtcaaa ttcgacgagg ccagatttagc cgaagatctg 60
tctacttttag aagtcgctga gttttcttagt gcatattcgg acttcgcttg tggaaaagtgg 120
gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
aatgcgcccc cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
gaaacacact tcgactgcag tgcggttaga tatacgcgta tcgtgcgtgt ctccgagaac 300
gcctgcctga tccctcatag tgattatttg gagctcgatg aaacggtcac ccggctgcac 360
ctcgtgttgg atactaacag tgggttggtcc aacactgagg aagataagat attccacatg 420
10 ggtctgggag agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggaaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
15 atgtatgatt ggcttaaaga aatatgtaa cgcgaggag atcccgcggt gattgagaag 780
acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa 843

20 <210> 118
<211> 280
<212> PRT
<213> Secuencia artificial
25 <220>
<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
30 <400> 118

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

65

ES 2 712 682 T3

	210		215		220	
5	Ser Ile Leu Ala Lys	Leu His Phe Phe Tyr	Lys Ala Asp Cys Arg Ser			
	225	230	235		240	
	Met Tyr Asp Trp Leu	Lys Glu Ile Cys Lys	Arg Arg Gly Asp Pro Ala			
		245	250		255	
10	Leu Ile Glu Lys Thr	Ala Ser Leu Glu Arg	Phe Phe Leu Gly His Arg			
		260	265		270	
	Glu Arg Gly Glu Val	Met Thr Tyr				
		275	280			

15
 <210> 119
 <211> 843
 <212> ADN
 <213> Secuencia artificial

20
 <220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

25
 <400> 119

	atgtctacgc	atttccttgg	caaggtcaaa	ttcgacgagg	ccagattagc	cgaagatctg	60
	tctactttag	aagttgctga	gttttcttcc	gcatattcgg	acttcgcatg	tggaaaagtgg	120
	gaagcatgtg	ttttacgcaa	cagaacaggc	atgcaagaag	aagacattgt	ggtttctcac	180
30	aatgcgccc	cattacagac	ccctttatct	aaatcactgc	cgtacttgaa	tgagcttggt	240
	gaaacacact	tcgactgcag	tgccggttaga	tatacgcgta	tcgtgcgtgt	ctccgagaac	300
	gcctgcctga	tcctcatag	tgattatttg	gagctcgatg	aaacggtcac	ccggctgcac	360
	ctcgtgttgg	atactaacag	tggttgtgcc	aacactgagg	aagataagat	attccacatg	420
	gggtctggcg	agatctgggt	tctggacgca	atgttaacctc	attctgctgc	ctgcttttct	480
	aaaactccta	gactgcagct	catgatagat	tttgaagcta	ccgccttccc	ggaatccttc	540
35	ttacgcaatg	tggaacagcc	tgtaactact	cgcgacatgg	tagaccctcg	ttaaagaattg	600
	acggacgaag	ttatagaagg	catttttaggg	ttctccataa	tcatcagtga	agcgaattac	660
	cgcgagatcg	tctctatcct	ggccaaatta	cacttctttt	ataaagcgga	ctgtcgggtcc	720
	atgtatgatt	ggcttaaaga	aatatgtaaa	cgccggggag	atcccgcggt	gattgagaag	780
	acagcgagtc	tcgaacgttt	ctttttgggg	catcgtgaac	gcggagaagt	aatgacgtat	840
40	taa						843

45
 <210> 120
 <211> 280
 <212> PRT
 <213> Secuencia artificial

50
 <220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

55
 <400> 120

55

60

65

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

<210> 121

<211> 843

<212> ADN

<213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 121

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg      60
      totacttttag aagtcgctga gttttcttcc gcatattcgg acttcgcttg tggaaaagtgg      120
      gaagcatgtg ttttaacgaa cagaacaggc atgcaagaag aagacattgt gggtttctcac      180
      aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt      240
      gaaacacact tcgactgcag tgcgggttaga tatgtacgta tcgtgcgtgt ctccgagaac      300
      gcctgcctga tccctcatag tgattatttg gagctcgatg aaacgttcac ccggctgcac      360
10     ctctgtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg      420
      ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct      480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc      540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg      600
      acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac      660
15     cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcggg ctgtcgggtcc      720
      atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag      780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat      840
      taa

```

20 <210> 122
 <211> 280
 <212> PRT
 <213> Secuencia artificial

25 <220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
 <400> 122

```

30      Met Ser Thr His Phe Leu Gly Lys Val Lys Phe Asp Glu Ala Arg Leu
          1              5              10              15

```

35

40

45

50

55

60

65

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

<210> 123

<211> 843

<212> ADN

45 <213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

50 <400> 123

55 atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
tctacttttag aagttgctga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
aatgcgcccgc cattagcgac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
gaaacacact tcgactgcag tgcggttaga tatgtacgta tcgtgcgtgt ctccgagaac 300
gcctgcctga tccctcatag tgattatttg gagctcgatg aaacgttcac ccggctgcac 360
ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
60 ggtctgggag agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
cgcgagatcg tctctatcct ggccaaatta caactctttt ataaagcgga ctgtcggttc 720
atgtatgatt ggcttaaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag 780
65 acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa 843

<210> 124

<211> 280

<212> PRT

5 <213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

10 <400> 124

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

<210> 125

<211> 843

<212> ADN

55 <213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

60 <400> 125

65

ES 2 712 682 T3

5	atgtctacgc atttcottgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg	60
	tctactttag aagttgctga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg	120
	gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac	180
	aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt	240
10	gaaacacact tcgactgcag tgcgggttaga tatgtacgta tcgtgcgtgt ctccgagaac	300
	gcctgcctga tccctcatag tgattatttg gagttggagg aatcattcac ccggctgcac	360
	ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg	420
	gggtctgggag agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct	480
15	aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc	540
	ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg	600
	acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac	660
	cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc	720
20	atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag	780
	acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat	840
	taa	843
25	<210> 126	
	<211> 280	
	<212> PRT	
	<213> Secuencia artificial	
30	<220>	
	<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti	
	<400> 126	
35		
40		
45		
50		
55		
60		
65		

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

<210> 127

45 <211> 843

<212> ADN

<213> Secuencia artificial

<220>

50 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 127

55

60

65

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg      60
      tctacttttag aacgtgctga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg      120
      gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac      180
      aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt      240
      gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt ctccgagaac      300
      gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggctgcac      360
10     ctctgtgttg atactaacag tgggtgtgcc aacactgagg aagataagat attccacatg      420
      ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct      480
      aaaactccta gactgcagct catgatagat ttggaagcta ccgccttccc ggaatccttc      540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg      600
      acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac      660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc      720
15     atgtatgatt ggcttaaaga aatatgtaaa cgcgggggag atcccgcgtt gattgagaag      780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat      840
      taa
      843

```

```

20     <210> 128
      <211> 280
      <212> PRT
      <213> Secuencia artificial

25     <220>
      <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

      <400> 128

```

```

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

```

60

65

5 Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val
210 215 220
Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
225 230 235 240
Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
245 250 255
10 Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
260 265 270
Glu Arg Gly Glu Val Met Thr Tyr
275 280

15

<210> 129

<211> 843

<212> ADN

20 <213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

25 <400> 129

atgagtagcc actttctcgg caaagtgaag ttcgatgaag cacggcttgc agaagacttg 60
tctacgctgg aagtggcgga attttcctcc gcgtatagtg attttgcctg cggaaaatgg 120
30 gaagcatgtg ttcttagaaa ccgtacaggt atgcaagaag aggatattgt ggtttcccat 180
aacgcaccgg ctcttcagac cccgctgagt aaatcattgc cttaccttaa tgaattagtt 240
gagacacatt tcgactgttc ggctgtccgt tacgttcgta ttgtaagagt gacgcgagaac 300
gcgtgcctga taccacattc agattacctc gagcttgacg agacatttac cagattgcat 360
ctggtcttgg ataccaactc aggttgtgcc aataccgaag aggacaagat attccatag 420
ggctctcggg aaatatggtt cttggatgag atgctgcctc atagtgcagc atgtttttct 480
35 aagacgccta gactgcagct gatgattgac ttcgaagcta cagcgtttcc cgaatcattt 540
ttacggaacg tcgaacagcc ggttaccacc cgtgatatgg tcgaccctag aaaagaatta 600
acggacgaag tgatcgaagg tatactcggc ttttctatta tcatatcgga gccgaactac 660
cgtgaaatcg tgtctatatt agcgaagctg cacttctttt ataaagccga ttgccgggtcg 720
atgtatgatt ggctgaaaga aatttgcaag agacggggag acccggcgct tattgaaaag 780
40 acggcttcat tggaacgctt ctttttaggc cacagagaac gcggggaagt aatgacctac 840
taa 843

<210> 130

45 <211> 280

<212> PRT

<213> Secuencia artificial

<220>

50 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 130

55

60

65

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

<210> 131

45 <211> 843

<212> ADN

<213> Secuencia artificial

<220>

50 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 131

55

60

65

ES 2 712 682 T3

5	atgagtagcc	actttctcgg	caaagtgaaa	ttcgatgaag	cacggcttgc	agaagacttg	60
	tctacgctgg	aacgtgcgga	attttcctcc	gcgtatagt	attttgcctg	cgaaaaatgg	120
	gaagcatgtg	ttcttagaaa	ccgtacaggt	atgcaagaag	aggatattgt	ggtttcccat	180
	aacgcaccgg	ctcttcagac	cccgtcgagt	aaatcattgc	cttaccttaa	tgaattagtt	240
	gagacacatt	tcgactgttc	ggctgtccgt	tacgttcgta	ttgtaagagt	gagcgagAAC	300
	gcgtgcctga	taccacattc	agattacctc	gagcttgacg	aggagtttac	cagattgcat	360
	ctggtccttg	ataccaactc	aggttgtgcc	aataccgaag	aggacaagat	attccatatg	420
10	ggtctcgggtg	aaatatggtt	cttggatgcg	atgctgcctc	atagtgcagc	atgtttttct	480
	aagacgccta	gactgcagct	gatgattgac	ttcgaagcta	cagcgtttcc	cgaatcattt	540
	ttacggaacg	tcgaacagcc	ggttaccacc	cgtgatatgg	tcgaccctag	aaaagaatta	600
	acggacgaag	tgatcgaagg	tatactcggc	ttttctatta	tcatatcgga	ggcgaactac	660
	cgtgaaatcg	tgtctatttt	agcgaagctg	cacttctttt	ataaagccga	ttgccggtcg	720
15	atgtatgatt	ggctgaaaga	aatttgcaag	agacggggag	acccggcgct	tattgaaaag	780
	acggcttcat	tggaacgctt	cttttttaggc	cacagagaac	gcggggaagt	aatgacctac	840
	taa						843

20 <210> 132
 <211> 280
 <212> PRT
 <213> Secuencia artificial

25 <220>
 <223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
 <400> 132

30 Met Ser Ser His Phe Leu Gly Lys Val Lys Phe Asp Glu Ala Arg Leu

35

40

45

50

55

60

65

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

<210> 133

<211> 843

<212> ADN

<213> Secuencia artificial

<220>

<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti

<400> 133

ES 2 712 682 T3

5	atgagtagcc	actttctcgg	caaagtgaaa	ttcgatgaag	cacggcttgc	agaagacttg	60
	tctacgctgg	aacgtgcgga	attttcctcc	gcgtatagtg	attttgccctg	cggaaaatgg	120
	gaagcatgtg	ttcttagaaa	ccgtacaggt	atgcaagaag	aggatattgt	ggtttcccat	180
	aacgcaccgg	ctcttcagac	cccgctgagt	aaatcattgc	cttaccttaa	tgaattagtt	240
	gagacacatt	tcgactgttc	ggctgtccgt	tacgttcgta	ttgtaagagt	gagcgagaac	300
	gcgtgcctga	taccacattc	agattacctc	gagcttgacg	agacatttac	cagattgcat	360
	ctggctcttg	ataccaactc	aggttgtgcc	aataccgaag	aggacaagat	attccatattg	420
10	ggctctcggtg	aaatatgggt	cttggtatgag	atgctgcctc	atagtgcagc	atgtttttct	480
	aagacgccta	gactgcagct	gatgattgac	ttcgaagcta	cagcgtttcc	cgaatcattt	540
	ttacggaacg	tcgaacagcc	ggttaccacc	cgtgatattg	tcgaccctag	aaaagaatta	600
	acggacgaag	tgatcgaagg	tatactcggc	ttttctatta	tcatatcgga	ggcgaactac	660
	cgtgaaatcg	tgtctatttt	agcgaagctg	cacttctttt	ataaagccga	ttgccggtcg	720
	atgtatgatt	ggctgaaaaga	aatttgcaag	agacggggag	acccggcgct	tattgaaaag	780
15	acggcttcat	tggaacgctt	cttttttaggc	cacagagaac	gcggggaagt	aatgacctac	840
	ttaa						843

20	<210> 134
	<211> 280
	<212> PRT
	<213> Secuencia artificial
25	<220>
	<223> Variante diseñada de prolina 4-hidroxilasa de Sinorhizobium meliloti
	<400> 134

30

35

40

45

50

55

60

65

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

<210> 135
<211> 843
<212> ADN
<213> Sinorhizobium meliloti

<400> 135

50 atgagcacc c atttcttggg caaggtcaag ttcgatgaag cgcgattggc agaagatcta 60
tctaccttgg aagttgccga gttctcgagt gcatactcgg acttcgcgtg cggtaaatgg 120
gaggcatcgc tgctacgcaa tcggaccgga atgcaggagg aagatatcgt cgtaagtcac 180
aacgctcctg cactggccac gccgctgagc aagtcgctgc cgtatctgaa cgaacttggt 240
55 gaaacccact tcgattgcag cgctgttcgg tatacaagaa ttgtccgtgt atcagaaaac 300
gcatgtataa tcccccatag tgattacct gaactagatg agaccttcac aaggttacac 360
ctggtgtagg acactaattc aggatgcgct aatactgagg aagataaaat atttcatatg 420

60 ggactgggag agatttggtt ccttgacgct atgttacgcg atagcgotgc ttgtttttcc 480
aaaactccgc gcctgcatct gatgatcgac tttgaggcta ccgcttttcc cgaatctttt 540
ctgcgaaatg tcgaacaacc agtgacaaca cgagacatgg ttgatoctcg gaaggaacta 600
accgatgagg ttatcgaagg tattctgggg ttttcaataa ttattagcga agccaattac 660
cgggaaattg tttctattct ggcgaaagcta cacttcttct acaaggcaga ctgtcgatca 720
atgtacgact ggctgaagga aatctgcaaa cgctcgaggg atcctgcact tattgaaaag 780
65 accgcctcgc tcgagcgatt ttttctaggg caccgtgaac gtggcgaggt gatgacatac 840
taa 843

<210> 136
 <211> 280
 <212> PRT
 <213> Sinorhizobium meliloti

<400> 136

```

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

```

<210> 137
 <211> 843
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 137

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
      tctacttttag aacgtgctga gttttcttcc gcatattcgg acttcgcttg tggaaaagtgg 120
      gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
      aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
      gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt ctccgagaac 300
      gcctgcctga tccctcatag tgattatttg gagttggata atgaattcac ccggctgcac 360
10     ctcgtggttg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
      ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
15     cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
      atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa
      843

```

20

<210> 138
 <211> 280
 <212> PRT
 <213> Secuencia artificial

25

<220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

30

<400> 138

35

40

45

50

55

60

65

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

<210> 139

45 <211> 843

<212> ADN

<213> Secuencia artificial

<220>

50 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 139

55

60

65

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
      tctacttttag aacgtgctga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
      gaagcatgtg ttttaacgaa cagaacaggc atgcaagaag aagacattgt gggtttctcac 180
      aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
      gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt ctccgagaac 300
      gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggctgcac 360
      ctcggtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacttg 420
10     ggtctgggcg agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
      cgcgagatcg tctctatcct ggccaaatta caottctttt ataaagcgga ctgtcggtcc 720
15     atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa

```

20 <210> 140
 <211> 280
 <212> PRT
 <213> Secuencia artificial

25 <220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti
 <400> 140

```

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

```

60

65

ES 2 712 682 T3

5 Met Val Asp Pro Arg Lys Glu Leu Thr Asp Glu Val Ile Glu Gly Ile
195 200 205
Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val
210 215 220
Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
225 230 235 240
10 Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
245 250 255
Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
260 265 270
15 Glu Arg Gly Glu Val Met Thr Tyr
275 280

<210> 141

<211> 843

20 <212> ADN

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

25

<400> 141

30 atgtctacgc atttcctcgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
tctacttttag aacgtgctga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt cactgagaac 300
gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggctgcac 360
ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
35 ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtoggtcc 720
40 atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag 780
acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa 843

<210> 142

45 <211> 280

<212> PRT

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

50

<400> 142

55

60

65

ES 2 712 682 T3

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

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

<210> 143

<211> 843

45 <212> ADN

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

50

<400> 143

55

60

65

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
      tctacttttag aacgtgctga gttttcttcc gcataattcgg acttcgcttg tggaaagtgg 120
      gaactttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
      aatgcgcccg cattacagac ccctttatct aaatcaactgc cgtacttgaa tgagcttggt 240
      gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt ctccgagAAC 300
      gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggctgcac 360
10     ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
      ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tggAACagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
15     cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
      atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggaagaat aatgacgtat 840
      taa

```

20

<210> 144
 <211> 280
 <212> PRT
 <213> Secuencia artificial

25

<220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

30

<400> 144

35

40

45

50

55

60

65

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

45 <210> 145
 <211> 843
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

55 <400> 145

60

65

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagt cgaagatctg 60
      tctacttttag aacggttgga gttttcttcc gcatattcgg acttcgcttg tggaaaagtgg 120
      gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
      aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
      gaaacacact togactgcag tgcgggttaga tatgtacgga tcgtgcgtgt ctccgagAAC 300
      gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggctgcac 360
10     ctcgtggttg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
      ggtctgggCG agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat ttgaaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tggAACagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcggA ctgtcgggtcc 720
15     atgtatgatt ggcttaaaga aatatgtaaa cgcCGgggag atcccgcggt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa
      843

```

20

<210> 146
 <211> 280
 <212> PRT
 <213> Secuencia artificial

25

<220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

30

<400> 146

35

40

45

50

55

60

65

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

40

<210> 147

<211> 843

45 <212> ADN

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

50

<400> 147

55 atgtctacgc atttccttgg caaggtcaaaa ttcgacgagg ccagattagc cgaagatctg 60
tctacttttag aacgtgctga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180

60

65

ES 2 712 682 T3

```

5      aatcagcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
      gaaacacact tgcactgcag tgcggttaga tatgtacgga tgcgtcgtgt ctccgagaac 300
      gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggctgcac 360
      ctctgtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
      ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
10     acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggcc 720
      atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa

```

15

<210> 148

<211> 280

<212> PRT

20

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

25

<400> 148

```

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

```

<210> 149
 <211> 843
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 149

```

atgtctacgc atttccttgg caaggtcaaa ttcgacgagg cdagattagc cgaagatctg 60
tctacttttag aacgtagggg gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt ctccgagaac 300
gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggctgcac 360
ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
ggtctgggag agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct 480
aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag 780
acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa
    
```

<210> 150
 <211> 280
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 150

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

210 215 220
35 Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
225 230 235 240
Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
245 250 255
40 Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
260 265 270
Glu Arg Gly Glu Val Met Thr Tyr
275 280

<210> 151

<211> 843

<212> ADN

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 151

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
      tctacttttag aacgtgctga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
      gaagcatgtg ttttacgcaa cagaacaggg atgcaagaag aagacattgt ggttttctcac 180
      aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggtt 240
      gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt ctccgagAAC 300
      gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggctgcac 360
      ctogtggttg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
10     ggtctgggCG agatctggtt tctggacgca gcgttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tggAACagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
15     atgtatgatt ggcttaaaga aatatgtaaa cgcCGggag atcccgctt gattgagaag 780
      acagcgagtc tcgaacggtt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa
      843

```

20

<210> 152

<211> 280

<212> PRT

25 <213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

30 <400> 152

```

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

```

50

55

60

65

5 Cys Ala Asn Thr Glu Glu Asp Lys Ile Phe His Met Gly Leu Gly Glu
130 135 140
Ile Trp Phe Leu Asp Ala Ala Leu Pro His Ser Ala Ala Cys Phe Ser
145 150 155 160
Lys Thr Pro Arg Leu Gln Leu Met Ile Asp Phe Glu Ala Thr Ala Phe
165 170 175
10 Pro Glu Ser Phe Leu Arg Asn Val Glu Gln Pro Val Thr Thr Arg Asp
180 185 190
Met Val Asp Pro Arg Lys Glu Leu Thr Asp Glu Val Ile Glu Gly Ile
195 200 205
15 Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val
210 215 220
Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
225 230 235 240
Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
245 250 255
20 Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
260 265 270
Glu Arg Gly Glu Val Met Thr Tyr
275 280

25
30 <210> 153
<211> 843
<212> ADN
<213> Secuencia artificial

<220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti
35
<400> 153

40 atgtctaagc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
tctacttttag aacgtgctga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
aatgcgcocg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
gaaacacact tcgactgcag tcgtgttaga tatgtacgga tcgtgcgtgt ctccgagaac 300
gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggtgcac 360
45 ctctgtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
ggtctgggag agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct 480
aaaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
ttacgcgaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
50 atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag 780
acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa 843

55 <210> 154
<211> 280
<212> PRT
<213> Secuencia artificial

60 <220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti
<400> 154
65

ES 2 712 682 T3

5 Met Ser Thr His Phe Leu Gly Lys Val Lys Phe Asp Glu Ala Arg Leu
 1 5 10 15
 Ala Glu Asp Leu Ser Thr Leu Glu Arg Ala Glu Phe Ser Ser Ala Tyr

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

45

<210> 155
 <211> 843
 50 <212> ADN
 <213> Secuencia artificial

55

<220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti
 <400> 155

60

65

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg taagggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
      tctacttttag aacgtgctga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
      gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
      aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
      gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt ctccgagAAC 300
      gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggctgcac 360
      ctctgtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
10     ggtctgggCG agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tggaacagcc tgggactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcggtcc 720
15     atgtatgatt ggcttaaaga aatatgtaa cgcgggggag atcccgcggt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa
      843

```

20

<210> 156
 <211> 280
 <212> PRT
 <213> Secuencia artificial

25

<220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

30

<400> 156

35

40

45

50

55

60

65

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

45 <210> 157
<211> 843
<212> ADN
<213> Secuencia artificial

50 <220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 157

55 atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
tctactttac gtcgtgctga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
gaagcatgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt gggttctcac 180
aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
60 gaaacacact tcgactgcag tgcgggttaga tatgtacgga tcgtgcgtgt ctccgagaac 300
gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggctgcac 360

65

ES 2 712 682 T3

```

5      ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
      ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
10     atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa
      843

```

15 <210> 158
 <211> 280
 <212> PRT
 <213> Secuencia artificial

20 <220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 158

```

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

```

ES 2 712 682 T3

<210> 159
 <211> 843
 <212> ADN
 5 <213> Secuencia artificial

<220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

10 <400> 159

```

    atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
    tctacttttaa gtcgtgctga gttttcttcc gcatattcgg acttcgcttg tggaagtgg 120
    15 gaagcatgtg ttttacgcaa cagaacaggg atgcaagaag aagacattgt ggtttctcac 180
    aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
    gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt ctccgagaac 300
    gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggctgcac 360
    ctcggtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
    20 ggtctgggcg agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct 480
    aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
    ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
    acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
    cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
    25 atgtatgatt ggcttaaaga aatatgtaaa cgcgggggag atcccgcgtt gattgagaag 780
    acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
    taa
  
```

30 <210> 160
 <211> 280
 <212> PRT
 <213> Secuencia artificial

35 <220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 160

40

45

50

55

60

65

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

<210> 161

<211> 843

<212> ADN

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 161

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg caaggtcaaa ttgcacgagg ccagattagc cgaagatctg 60
      tctacttttag aacgtgctga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
      gaagcatgtg ttttacgcaa cagaacaggg atgcaagaag aagacattgt ggtttctcac 180
      aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
      gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt ctttgagaac 300
      gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccgctgcac 360
10     ctgctgttgg atactaacag tggttggtgcc aacactgagg aagataagat attccacatg 420
      ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatatag tttgaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tgggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
15     cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
      atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atccgcggtt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa
      843

```

```

20     <210> 162
      <211> 280
      <212> PRT
      <213> Secuencia artificial

```

```

25     <220>
      <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

```

```

30     <400> 162

```

```

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

```

```

55

```

```

60

```

```

65

```

5 Ile Trp Phe Leu Asp Ala Met Leu Pro His Ser Ala Ala Cys Phe Ser
145 150 155 160
Lys Thr Pro Arg Leu Gln Leu Met Ile Asp Phe Glu Ala Thr Ala Phe
165 170 175
Pro Glu Ser Phe Leu Arg Asn Val Glu Gln Pro Val Thr Thr Arg Asp
180 185 190
10 Met Val Asp Pro Arg Lys Glu Leu Thr Asp Glu Val Ile Glu Gly Ile
195 200 205
Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val
210 215 220
15 Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
225 230 235 240
Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
245 250 255
Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
260 265 270
20 Glu Arg Gly Glu Val Met Thr Tyr
275 280

25 <210> 163
<211> 843
<212> ADN
<213> Secuencia artificial

30 <220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

35 <400> 163

atgagtagcc actttctcgg caaagtgaag ttcgatgaag cacggcttgc agaagacttg 60
tctacgctgg aacgtgcgga attttcctcc gcgtatagtg attttgcctg cggaaaatgg 120
gaactgtgtg ttcttagaaa ccgtacaggt atgcaagaag aggatattgt gggttcccat 180
aacgcaccgg ctcttcagac cccgctgagt aaatcattgc cttaccttaa tgaattagtt 240
gagacacatt tcgactgttc ggctgtccgt tacgttcgta ttgtaagagt gagcgagaac 300
gcgtgcctga taccacattc agattacctc gagcttgacg aggagtttac cagattgcat 360
ctggctcttg ataccaactc aggttgtgccc aataccgaag aggacaagat attccatatg 420
ggctctcggtg aaatatgggt cttggatgcg atgctgcctc atagtgcagc atgtttttct 480
45 aagacgccta gactgcagct gatgattgac ttccaagcta cagcgtttcc cgaatcattt 540
ttacggaacg tcgaacagcc gggtaccacc cgtgatatgg tcgaccctag aaaagaatta 600
acggacgaag tgatcgaagg tatactcggc ttttctatta tcatatcgga gccgaactac 660
cgtgaaatcg tgtctatatt agcgaagctg cacttctttt ataaagccga ttgccggtcg 720
atgtatgatt ggctgaaaga aatttgcaag agacggggag acccggcgct tattgaaaag 780
50 acggcttcat tggaacgctt ctttttaggc cacagagaac gcggggaagt aatgacctac 840
taa 843

55 <210> 164
<211> 280
<212> PRT
<213> Secuencia artificial

60 <220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 164

65

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

10

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

35 Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
225 230 235 240
Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
245 250 255
Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
260 265 270
Glu Arg Gly Glu Val Met Thr Tyr
275 280

40

45

<210> 165

<211> 843

<212> ADN

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 165

ES 2 712 682 T3

```

5      atgagtagcc actttctcgg caaagtgaaa ttcgatgaag cacggccttg agaagacttg 60
      tctacgctgg aacgtgcgga attttcctcc gcgtatagtg attttgccctg cggaaaaatgg 120
      gaactgtgtg ttcttagaaa ccgtacaggt atgcaagaag aggatattgt ggtttcccat 180
      aacgcaccgg ctcttcagac cccgctgagt aaatcattgc cttaccttaa tgaattagtt 240
      gagacacatt tcgactgttc ggctgtccgt tacgttcgta ttgtaagagt gaccgagaac 300
      gcgtgcctga taccacattc agattacctc gagcttgacg aggagtttac cagattgcat 360
10     ctggtcttgg ataccaactc aggttggtgc aataccgaag aggacaagat attccatatg 420
      ggtctcgggtg aaatatggtt cttggatgcg atgctgcctc atagtgcagc atgtttttct 480
      aagacgccta gactgcagct gatgattgac ttcgaagcta cagcgtttcc cgaatcattt 540
      ttacggaacg tcgaacagcc ggttaccacc cgtgatatgg tcgaccctag aaaagaatta 600
      acggacgaag tgatcgaagg tatactcggc ttttctatta tcatatcgga ggcgaactac 660
      cgtgaaatcg tgtctatattt agcgaagctg cacttccttt ataaagccga ttgccgggtcg 720
15     atgtatgatt ggctgaaaga aatttgcaag agacggggag acccggcgct tattgaaaag 780
      acggccttcatt tggaacgctt ctttttaggc cacagagaac gcggggaagt aatgacctac 840
      taa

```

20

<210> 166

<211> 280

<212> PRT

25 <213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

30

<400> 166

35

40

45

50

55

60

65

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

<210> 167

<211> 843

45 <212> ADN

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

50

<400> 167

55 atgagtagcc actttctcgg caaagtgaag ttcgatgaag cacggcttgc agaagacttg 60
tctacgctgg aacgtgcgga attttcctcc gcgtatagtg attttgcctg cggaaaatgg 120
gaactgtgtg ttcttagaaa ccgtacaggt atgcaagaag aggatattgt ggtttcccat 180
aacgcaccgg ctcttcagac cccgctgagt aaatcattgc cttaccttaa tgaattagtt 240
gagacacatt tcgactgttc ggctgtccgt tacgttcgta ttgtaagagt gaccgagaac 300
gcgtgcctga taccacattc agattacctc gagcttgacg aggagtttac cagattgcat 360
ctggtcttgg ataccaactc aggttggtgc aataccgaag aggacaagat attocatatg 420
ggtctcgggt aaatatggtt cttggatgac atgctgcctc atagtgcagc atgtttttct 480
aagacgccta gactgcagct gatgattgac ttogaagcta cagcgtttcc cgaatcattt 540

65

```

ttacggaacg tcgaacagcc gggcaccacc cgtgatatgg tcgaccctag aaaagaatta 600
acggacgaag tgatcgaagg tatactcggc ttttctatta tcatatcgga ggcgaactac 660
cgtgaaatcg tgtctatattt agcgaagctg cacttctttt ataaagccga ttgccggtcg 720
atgtatgatt ggctgaaaga aatttgcaag agacggggag acccggcgct tattgaaaag 780
acgggttcat tggaacgctt ctttttaggc cacagagaac gcggggaagt aatgacctac 840
taa 843

```

<210> 168

<211> 280

<212> PRT

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 168

```

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

```

ES 2 712 682 T3

5 <210> 169
 <211> 843
 <212> ADN
 <213> Secuencia artificial

 10 <220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti
 <400> 169

 15 atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
 tctacttta gtcgtgcaga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
 gaactttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
 aatgcgccc cttacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
 20 gaaacacact tgcactgcag tgcggttaga tatgtacgga tcgtgcgtgt ctccgagAAC 300
 gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac cgggctgcac 360
 ctgctggttg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
 ggtctggggc agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
 aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
 25 ttacgcaatg tggAACagcc tggactact cgcgacatgg tagaccctcg taaagaattg 600
 acggacgaag ttatagaagg ctttttaggg ttctccataa tcatcagtga agcgaattac 660
 cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcggA ctgtcggtcc 720
 atgtatgatt ggcttaaaga aatatgtaaa cgcgggggag atcccgcgtt gattgagaag 780
 acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
 30 taa 843

35 <210> 170
 <211> 280
 <212> PRT
 <213> Secuencia artificial

 40 <220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti
 <400> 170

45

50

55

60

65

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

<210> 171

<211> 843

50 <212> ADN

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

55 <400> 171

60

65

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
      tctacttttaa gtogtgcaga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
      gaactttgtg ttttacgcaa cagaacaggg atgcaagaag aagacattgt ggtttctcac 180
      aatcagcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
      gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt cacggagAAC 300
10     gcctgcctga tccctcatag tgattatttg gagttggata acgaattcac ccggctgcac 360
      ctctgtgttg atactaacag tgggtgtgcc aacactgagg aagataagat attccacctg 420
      ggtctggggc agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tggAACagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
15     acggacgaag ttatagaagg cattttaggg ttctccataa tcatacgtga agcgaattac 660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcggA ctgtcgggtcc 720
      atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa 843

```

20

<210> 172

<211> 280

25

<212> PRT

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

30

<400> 172

```

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

```

60

65

5 Pro Glu Ser Phe Leu Arg Asn Val Glu Gln Pro Val Thr Thr Arg Asp
165 170 175
180 185 190
Met Val Asp Pro Arg Lys Glu Leu Thr Asp Glu Val Ile Glu Gly Ile
195 200 205
Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val
210 215 220
10 Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
225 230 235 240
Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
245 250 255
15 Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
260 265 270
Glu Arg Gly Glu Val Met Thr Tyr
275 280

20

25 <210> 173
<211> 843
<212> ADN
<213> Secuencia artificial

30 <220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 173

35 atgtctacgc atttccttgg caaggtcaaa ttgcacgagg ccagattagc cgaagatctg 60
tctacttttaa gtctgtgcaga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
gaactttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
aatgcgcccgc cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggtt 240
gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt cacggagAAC 300
40 gcctgcctga tcctcatag tgattatttg gagttggata acgaattcac ccggctgcac 360
ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
ggtctgggCG agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggAACagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
45 acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcggtcc 720
atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag 780
acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa 843

50

<210> 174
<211> 280
<212> PRT
55 <213> Secuencia artificial

<220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

60 <400> 174

65

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

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

45

<210> 175
<211> 843
50 <212> ADN
<213> Secuencia artificial

55

<220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti
60
65 <400> 175

60

65

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
      tctacttttaa gtcgtcgtga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
      gaacttttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
      aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
      gaaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt ctccgagaac 300
      gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggctgcac 360
10     ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
      ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
15     cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
      atgtatgatt ggcttaaaga aatatgtaaa cgcgggggag atcccgcggt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa
      843

```

20

<210> 176

<211> 280

<212> PRT

25 <213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

30 <400> 176

35

40

45

50

55

60

65

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

<210> 177

<211> 843

45 <212> ADN

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

50

<400> 177

55	atgtctacgc	atttccttgg	caaggtcaaa	ttcgacgagg	ccagattagc	cgaagatctg	60
	tctacttttag	aacgtgcaga	gttttcttcc	gcatattcgg	acttcgcttg	tggaaagtgg	120
	gaacttttgtg	ttttacgcaa	cagaacaggc	atgcaagaag	aagacattgt	ggtttctcac	180
	aatcagcccg	cattacagac	ccctttatct	aatcactgc	cgtacttgaa	tgagcttggt	240
	gaaacacact	tcgactgcag	tcggttaga	tatgtacgga	tcgtgcgtgt	ctccgagaac	300
	gcctgcctga	tccctcatag	tgattatttg	gagttggatg	aagaattcac	ccggtgcac	360
60	ctcgtgttg	atactaacag	tggttgtgcc	aacactgagg	aagataagat	attccacctg	420
	ggctctgggcg	agatctggtt	tctggacgca	atgttacctc	attctgctgc	ctgcttttct	480
	aaaactccta	gactgcagct	catgatagat	tttgaagcta	ccgccttccc	ggaatccttc	540
	ttacgcaatg	tggaacagcc	tggtactact	cgcgacatgg	tagaccctcg	taaagaattg	600
	acggacgaag	ttatagaagg	cattttaggg	ttctccataa	tcatcagtga	agcgaattac	660
65	cgcgagatcg	tctctatcct	ggccaaatta	cacttctttt	ataaagcgga	ctgtcgggtcc	720

ES 2 712 682 T3

5 atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggtt gattgagaag 780
acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa 843

10 <210> 178
<211> 280
<212> PRT
<213> Secuencia artificial

15 <220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 178

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

60 <210> 179
<211> 843
<212> ADN
65 <213> Secuencia artificial

ES 2 712 682 T3

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

5 <400> 179

```

10      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
      tctacttttaa gtcgtcgtga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
      gaactttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
      aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
      gaaacacact tcgactgcag tcgtgttaga tatgtacgga tcgtgcgtgt ctccgagaac 300
      gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac cgggctgcac 360
      ctcggtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacctg 420
15      ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
20      atgtatgatt ggcttaaaga aatatgtaaa cgcgggggag atcccgcggt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa
      843

```

25 <210> 180
<211> 280

<212> PRT

<213> Secuencia artificial

30

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 180

35

40

45

50

55

60

65

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

<210> 181

45 <211> 843

<212> ADN

<213> Secuencia artificial

<220>

50 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 181

55

60

65

```

5      atgtctacgc atttcottgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
      totacttttaa gtcgtogtga gttttcttcc gcatattcgg acttgcgttg tggaaagtgg 120
      gaacttttggt ttttaacgaa cagaacaggc atgcaagaag aagacattgt ggttttctcac 180
      aatgcgcccgc cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
      gaaacacact tcgactgcag tcgtgttaga tatgtacgga tcgtgcgtgt cacggagaat 300
      gcctgcctga tccctcatag tgattatttg gagttggata acgaattcac ccggctgcac 360
10     ctctgtgttg atactaacag tgggtgtgcc aacactgagg aagataagat attccacctg 420
      ggtctggggcg agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tggaacagcc tggtagtact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
15     cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
      atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag 780
      acagcgagtc tcgaacggtt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa                                     843

```

20

<210> 182
 <211> 280
 <212> PRT
 <213> Secuencia artificial

25

<220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

30

<400> 182

```

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

```

60

65

5 Met Val Asp Pro Arg Lys Glu Leu Thr Asp Glu Val Ile Glu Gly Ile
195 200 205
Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val
210 215 220
Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
225 230 235 240
10 Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
245 250 255
Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
260 265 270
15 Glu Arg Gly Glu Val Met Thr Tyr
275 280

<210> 183
<211> 843
20 <212> ADN
<213> Secuencia artificial

<220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti
25 <400> 183

30 atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
tctacttttaa gtcgtgcaga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
gaactttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
aatcagcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt ctccgagaac 300
gcctgcctga tccctcatag tgattatttg gagttggata acgaattcac ccgctgcac 360
35 ctctgtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacctg 420
ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
40 cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcggtcc 720
atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag 780
acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa 843

45 <210> 184
<211> 280
<212> PRT
50 <213> Secuencia artificial

<220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti
55 <400> 184

60

65

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

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

<210> 185

<211> 843

50 <212> ADN

<213> Secuencia artificial

<220>

55 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 185

60

65

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
      tctactttaaa gtcgtcgtga gttttcttcc gcataattcgg acttcgcttg tggaaagtgg 120
      gaactttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggttttctcac 180
      aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
      gaaacacact tcgactgcag tcgtgttaga tatgtacgga tcgtgcgtgt ctccgagAAC 300
10     gcctgcctga tccctcatag tgattatTTg gagttggata acgaattcac ccggctgcac 360
      ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
      ggtctgggCG agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tggAACagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
15     acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcggtcc 720
      atgtatgatt ggcttaaaga aatatgtaaa cgcgggggag atcccgcggt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa
      843

```

20

<210> 186

<211> 280

<212> PRT

25 <213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

30 <400> 186

35

40

45

50

55

60

65

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

45 <210> 187
<211> 843
<212> ADN
<213> Secuencia artificial

50 <220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 187

55

60

65

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
      tctactttaa gtcgtgcaga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
      gaactttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
      aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
      gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt ctccgagAAC 300
10     gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggtgcac 360
      ctctgtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacctg 420
      ggtctgggag agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat ttggaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tggAACagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
15     cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
      atgtatgatt ggcttaaaga aatatgtaaa cgcgggggag atcccgcggt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggaagaat aatgacgtat 840
      taa 843

```

20

<210> 188

<211> 280

<212> PRT

25 <213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

30 <400> 188

35

40

45

50

55

60

65

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

45 <210> 189
<211> 843
<212> ADN
<213> Secuencia artificial

50 <220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

55 <400> 189

atgtctacgc atttccttgg caagggtcaaa ttcgacgagg ccagatttagc cgaagatctg 60
tctacttttat ctcggtgcaga gttttcttcc gcatattcgg acttogettg tggaaagtgg 120
gaacttttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180

60

65

ES 2 712 682 T3

```

5      aatgcgccccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
      gaaacacact tgcactgcag tgcgggttaga tatgtacgga tgcgtgcgtgt cacggagaac 300
      gcctgcctga tccctcatag tgattatattg gagttggata atgaattcac ccggctgcac 360
      ctctgtgttg atactaacag tggttgcgcc aacactgagg aagataagat attccacatg 420
      ggtctggggc agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
10     aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tggaacaacc tgtgactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
      atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
15     taa
      843

```

<210> 190

<211> 280

20 <212> PRT

<213> Secuencia artificial

<220>

25 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 190

```

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

```

ES 2 712 682 T3

<210> 191
 <211> 843
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 191

```

atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
tctacttttag aacgtcgtga gttttcttcc gcatattcgg acttgcgttg tggaaagtgg 120
gaactttgtg ttttaacgaa cagaacaggc atgcaagaag aagacattgt gggtttctcac 180
aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt ctccgagAAC 300
gcctgcctga tccctcatag tgattatttg gagttggata atgaattcac ccggtgcac 360
ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
gggtctggcg agatctgggt tctggacgca catttacctc attctgctgc ctgcttttct 480
aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggacaacc tgtgactact cgcgacatgg tagaccctcg taaagaattg 600
acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag 780
acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa

```

<210> 192
 <211> 280
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 192

ES 2 712 682 T3

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

210 215 220
 35 Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
 225 230 235 240
 Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
 245 250 255
 40 Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
 260 265 270
 Glu Arg Gly Glu Val Met Thr Tyr
 275 280

<210> 193

<211> 843

<212> ADN

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 193

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
      tctactttat ctctgtgcaga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
      gaactttgtg ttttacgcaa cagaacaggg atgcaagaag aagacattgt ggtttctcac 180
      aatcagcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
      gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt cacggagAAC 300
      gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggtgcac 360
      ctctgtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacctg 420
10     ggtctgggCG agatctggtt tctggacgca catttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tggAACagcc tgtgactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
15     atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccggtt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa
      843

```

20

<210> 194

<211> 280

<212> PRT

25 <213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

30 <400> 194

```

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

```

50

55

60

65

ES 2 712 682 T3

[illegible]

<210> 195
<211> 843
<212> ADN
<213> Secuencia artificial

<220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 195

40	atgtctacgc	atcttccttg	caaggtcaaa	tctgcacgag	ccagattagc	cgaagatctg	60
	tctacttttag	aacgtgcaga	gttttcttcc	gcataattcg	acttcgcttg	tggaaagtgg	120
	gaacttttgt	ttttacgcaa	cagaacaggc	atgcaagaag	aagacattgt	ggtttctcac	180
	aatgcgccc	cattacagac	ccctttatct	aaatcactgc	cgtacttgaa	tgagcttggt	240
	gaaacacact	tgcactgcag	tgcggttaga	tatgtacgga	tcggtcgtgt	ctccgagaac	300
45	gcctgcctga	tcctcatag	tgattatttg	gagttggata	atgaattcac	ccggctgcac	360
	ctcgtgttgg	atactaacag	tggttgtgcc	aacactgagg	aagcataagt	attccacatg	420
	ggtctggggc	agatctgggt	tctggacgca	atgttacctc	attctgctgc	ctgcttttct	480
	aaaactccta	gactgcagct	catgatagat	tttgaagcta	ccgccttccc	ggaatccttc	540
	ttacgcaatg	tggaacagcc	tgtaactact	cgcgacatgg	tagaccctcg	taaagaattg	600
50	acggacgaag	ttatagaagg	cattttaggg	ttctccataa	tcatcagtga	agcgaattac	660
	cgcgagatcg	tctctatcct	ggccaaatta	cacttctttt	ataaagcgga	ctgtcggtcc	720
	atgtatgatt	ggcttaaaga	aatatgtaaa	cgcgggggag	atcccgcgtt	gattgagaag	780
	acagcgagtc	tccaacgttt	ctttttgggg	catcgtgaac	gcggagaagt	aatgacgtat	840
	taa						843

<210> 196
<211> 280
<212> PRT
<213> Secuencia artificial

<220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de *Sinorhizobium meliloti*

<400> 196

5 Met Ser Thr His Phe Leu Gly Lys Val Lys Phe Asp Glu Ala Arg Leu
1 5 10 15
Ala Glu Asp Leu Ser Thr Leu Glu Arg Ala Glu Phe Ser Ser Ala Tyr

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

<210> 197

<211> 843

<212> ADN

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 197

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
      tctacttttag aacgtgctga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
      gaactttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggttttctcac 180
      aatcagcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
      gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt ctccgagaac 300
      gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggctgcac 360
10     ctogtggttg atactaacag tggttgtgcc aacactgagg aagataagat attccacctg 420
      ggtctgggcg agatctggtt tctggacgca catttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttgtagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcggtcc 720
15     atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa

```

20

<210> 198

<211> 280

<212> PRT

25 <213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

30 <400> 198

35

40

45

50

55

60

65

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

<210> 199

45 <211> 843

<212> ADN

<213> Secuencia artificial

<220>

50 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 199

55 atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
tctactttat ctctgtgcaga gttttcttcc gcatattcgg acttgcgttg tggaaagtgg 120
gaactttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
aatgcgcccc cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
gaaacacact tcgactgcag tcgtgttaga tatgtacgga tcgtgcgtgt ctccgagaac 300
60 gcctgcctga tccctcatag tgattatttg gagttggata atgaattcac ccggtgcac 360

65

ES 2 712 682 T3

```

5      ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
      ggtctgggcg agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta cgccttcccc ggaatccttc 540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
10     atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa
                                     843

```

15 <210> 200
 <211> 280
 <212> PRT
 <213> Secuencia artificial

20 <220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 200

```

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

```

65 <210> 201

ES 2 712 682 T3

<211> 843
<212> ADN
<213> Secuencia artificial

5 <220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 201

```

10      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
      tctactttat ctcgctcgtga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
      gaactttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
      aatcagcccc cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
15      gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt ctccgagAAC 300
      gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggctgcac 360
      ctcggtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
      ggtctgggcg agatctggtt tctggacgca cattacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
20      ttacgcaatg tggAACagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
      atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggaagaat aatgacgtat 840
25      taa 843

```

<210> 202
<211> 280
<212> PRT
30 <213> Secuencia artificial

<220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

35 <400> 202

40

45

50

55

60

65

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

<210> 203

<211> 843

<212> ADN

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 203

5 atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
 tctactttat ctcgctgtga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
 gaactttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggttttctcac 180
 aatgcgcccgc cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
 gaaacacact tcgactgcag tctgtttaga tatgtacgga tctgctgtgt ctccgagAAC 300
 gcctgcctga tccctcatag tgattatttg gagttggata atgaattcac ccggctgcac 360
 10 ctcggtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
 ggtctggggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
 aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
 ttacgcaatg tggAACagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
 acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
 15 cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
 atgtatgatt ggcttaaaga aatatgtaaa cgcgggggag atcccgctt gattgagaag 780
 acagcgagtc tcgaacggtt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
 taa 843

20

<210> 204
 <211> 280
 <212> PRT
 25 <213> Secuencia artificial

<220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

30 <400> 204

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

55

60

65

5 Ile Trp Phe Leu Asp Ala Met Leu Pro His Ser Ala Ala Cys Phe Ser
145 150 155 160
Lys Thr Pro Arg Leu Gln Leu Met Ile Asp Phe Glu Ala Thr Ala Phe
165 170 175
Pro Glu Ser Phe Leu Arg Asn Val Glu Gln Pro Val Thr Thr Arg Asp
10 180 185 190
Met Val Asp Pro Arg Lys Glu Leu Thr Asp Glu Val Ile Glu Gly Ile
195 200 205
Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val
210 215 220
15 Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
225 230 235 240
Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
245 250 255
20 Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
260 265 270
Glu Arg Gly Glu Val Met Thr Tyr
275 280

25
<210> 205
<211> 843
<212> ADN
<213> Secuencia artificial

30
<220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

35
<400> 205

atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
tctactttat ctcgctgtga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
gaactttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
40 aatcagcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
gaaacacact tgcactgcag tgcggttaga tatgtacgga togtgcgtgt ctccgagAAC 300
gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggtgcac 360
ctcgtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacctg 420
ggctctggcg agatctggtt tctggacgca catttacctc attctgctgc ctgcttttct 480
45 aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggAACagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
acggacgaag ttatagaagg catTTtaggg ttctccataa tcatcagtga agcgaattac 660
cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcggtcc 720
atgtatgatt ggcttaaaga aatatgtaaa cgcgggggag atcccgcggt gattgagaag 780
50 acagcgagtc tcgaacggtt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa 843

55
<210> 206
<211> 280
<212> PRT
<213> Secuencia artificial

60
<220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

65
<400> 206

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

45

<210> 207

<211> 843

<212> ADN

50

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

55

<400> 207

60

65

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
      tctactttat ctcgtcgtga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
      gaactttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
      aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
      gaaacacact tcgactgcag tcgtgttaga tatgtacgga tcgtgcgtgt ctccgagAAC 300
      gcctgcctga tccctcatag tgattatttg gagttggata atgaattcac ccggctgcac 360
10     ctcggtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacatg 420
      ggtctgggCG agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
      ttaCGcaatg tggAaggtcc tgtgactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
15     cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcggtcc 720
      atgtatgatt ggcttaaaga aatatgtaaa cgcgggggag atcccgcggt gattgagaag 780
      acagcgagtc tcgaacggtt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa
      843

```

20

<210> 208

<211> 280

<212> PRT

25 <213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

30

<400> 208

35

40

45

50

55

60

65

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

<210> 209

45 <211> 843

<212> ADN

<213> Secuencia artificial

<220>

50 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 209

55 atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
tctactttat ctcggttggga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
gaactttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
aatgcgccc cttacagac ccctttatct aaatcactgc cgtacttgaa tgagcttgtt 240
gaaacacact tcgactgcag tcgtgttaga tatgtacgga tcgtgcgtgt ctccgagaac 300
gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggctgcac 360
60 ctcggtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacctg 420
ggtctgggcg agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct 480
aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540

65

5	ttacgcaatg	tggaaacagcc	tgtaactact	cgcgacatgg	tagaccctcg	taaagaattg	600
	acggacgaag	ttatagaagg	cattttaggg	ttctccataa	tcatcagtga	agcgaattac	660
	cgcgagatcg	tctctatcct	ggccaaatta	cacttctttt	ataaagcgga	ctgtcggtcc	720
	atgtatgatt	ggcttaaaga	aatatgtaa	cgccggggag	atcccgcggt	gattgagaag	780
	acagcgagtc	tcgaaacgttt	ctttttgggg	catcgtgaac	gcggagaagt	aatgacgtat	840
	taa						843

	<210>	210
	<211>	280
15	<212>	PRT
	<213>	Secuencia artificial

<220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 210

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

65 <210> 211

ES 2 712 682 T3

<211> 843
<212> ADN
<213> Secuencia artificial

5 <220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 211

```

10      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
      tctactttat ctcgtcgtga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
      gaactttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
      aatcagcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
15      gaaacacact ttgactgcag tcgtgttaga tatgtacgga tcgtgogtgt ctccgagAAC 300
      gcctgcctga tccctcatag tgattatttg gagttggata atgaattcac ccggctgcac 360
      ctogtggttg atactaacag tggttgtgcc aacactgagg aagataagat attccacctg 420
      ggtctgggCG agatctggtt tctggatgca catttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
20      ttacgcaatg tggaaagtcc tgtgactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
      atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag 780
      acagcgagtc tcgaacggtt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
25      taa
      843
  
```

<210> 212
<211> 280
30 <212> PRT
<213> Secuencia artificial

<220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti
35

<400> 212

40

45

50

55

60

65

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

40 260 265 270
Glu Arg Gly Glu Val Met Thr Tyr
275 280

<210> 213

<211> 843

<212> ADN

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 213

5 atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
 tctacttttaa gtcgtgcaga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
 gaactttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt gggtttctcac 180
 aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggtt 240
 gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt ctttgagaac 300
 gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggtgcac 360
 ctctgttgg atactaacag tgggtgtgcc aacactgagg aagataagat attccacctg 420
 10 ggtctgggcg agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
 aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
 ttacgcaatg tggaacagcc tggactact cgcgacatgg tagaccctcg taaagaattg 600
 acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
 cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
 15 atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcggt gattgagaag 780
 acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggaagaat aatgaogtat 840
 taa 843

20 <210> 214
 <211> 280
 <212> PRT
 <213> Secuencia artificial

25 <220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti
 <400> 214

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

55

60

65

5 Pro Glu Ser Phe Leu Arg Asn Val Glu Gln Pro Gly Thr Thr Arg Asp
180 170 175
Met Val Asp Pro Arg Lys Glu Leu Thr Asp Glu Val Ile Glu Gly Ile
195 200 205
10 Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val
210 215 220
Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
225 230 235 240
15 Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
245 250 255
Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
260 265 270
Glu Arg Gly Glu Val Met Thr Tyr
275 280

20

<210> 215

<211> 843

<212> ADN

25

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

30

<400> 215

atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
tctactttat ctcgctgtga gttttcttcc gcatattcgg acttogettg tggaaagtgg 120
35 gaactttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt gggtttctcac 180
aatcagcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
gaaacacact tcgactgcag tcgtgttaga tatgtacgga tcgtgcgtgt ctttgagaac 300
gcctgcctga tccctcatag tgattatttg gagttggata gtgaattcac ccggctgcac 360
40 ctcgtgttgg atactaacag tggttgtgcc aacactgagg aagataagat attccacctg 420
ggctctggcg agatctggtt tctggacgca catttacctc attctgctgc ctgcttttct 480
aaaactccta gactgcagct catgtagat tttgaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggaacaacc tgtaactact cgcgacatgg tagacctcg taaagaattg 600
acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggcc 720
45 atgtatgatt ggcttaaaaga aatatgtaaa cgcgggggag atcccgcggt gattgagaag 780
acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
taa 843

50

<210> 216

<211> 280

<212> PRT

<213> Secuencia artificial

55

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 216

60

65

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

45

<210> 217

<211> 843

<212> ADN

50

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

55

<400> 217

60

65

ES 2 712 682 T3

```

5      atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
      tctactttat ctcgtcgtga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
      gaactttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
      aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
      gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt ctttgagaac 300
      gcctgtctga tccctcatag tgattatttg gagttggatg aagaattcac ccggctgcac 360
10     ctcggtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacctg 420
      ggtctggggc agatctggtt tctggatgca catttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg catttttaggg ttctccataa tcatcagtga agcgaattac 660
15     cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtcc 720
      atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa
      843

```

20

<210> 218
 <211> 280
 <212> PRT
 <213> Secuencia artificial

25

<220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

30

<400> 218

35

40

45

50

55

60

65

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

<210> 219

<211> 843

45 <212> ADN

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

50

<400> 219

55 atgtctacgc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
tctactttat ctggtgcaga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
gaactttgtg ttttacgcaa cagaacaggg atgcaagaag aagacattgt ggtttctcac 180
aatgcgcccg cattacagac ccttttatct aaatcactgc cgtacttgaa tgagcttggt 240
gaaacacact tcgactgcag tgcggttaga tatgtacgga tcgtgcgtgt ctttgagaac 300
gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggtgcac 360
60 ctggtgttgg atactaacag tggttggtgcc aacactgagg aagataagat attccacatg 420
ggtctgggag agatctggtt tctggacgca atgttacctc attctgctgc ctgcttttct 480
aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
ttacgcaatg tggacaacc tgtgactact cgcgacatgg tagaccctcg taaagaattg 600
acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
65 cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcgggtc 720

ES 2 712 682 T3

```

5      atgtatgatt ggcttaaaga aatatgtaaa cgccggggag atcccgcgtt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa                                     843

<210> 220
10    <211> 280
      <212> PRT
      <213> Secuencia artificial

<220>
15    <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 220

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

60

65

```

ES 2 712 682 T3

<210> 221
 <211> 843
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti
 <400> 221

```

atgagtagcc actttctcgg caaagtgaag ttcgatgaag cacggcttgc agaagacttg 60
tctacgctga gtctgtcaga attttccctcc gcgtatagtg attttgcctg cggaaaatgg 120
gaactgtgtg ttcttagaaa ccgtacaggt atgcaagaag aggatattgt ggtttcccat 180
aacgcaccgg ctcttcagac cccgctgagt aaatcattgc cttaccttaa tgaattagtt 240
gagacacatt tcgactgttc ggctgtccgt tacgttcgta ttgtaagagt gaccgagAAC 300
gcgtgcctga taccacattc agattacctc gagcttgaca acgagtttac cagattgcat 360
ctggctcttg ataccaactc aggttgtgcc aataccgaag aggacaagat attccatatg 420
ggctctcgtg aaatatggtt cttggatgcg atgctgcctc atagtgcagc atgtttttct 480
aagacgccta gactgcagct gatgattgac ttcgaagcta cagcgtttcc cgaatcattt 540
ttacggaacg tcgaacagcc ggttaccacc cgtgatatgg tcgaccctag aaaagaatta 600
acggacgaag tgatcgaagg tatactcggc ttttctatta tcatatcgga ggcgaaactac 660
cgtgaaatcg tgtctatttt agcgaagctg cacttctttt ataaagccga ttgccgggtcg 720
atgtatgatt ggctgaaaga aatttgcaag agacggggag acccggcgct tattgaaaag 780
acggcttcat tggaacgctt ctttttaggc cacagagAAC gcggggaagt aatgacctac 840
taa

```

<210> 222
 <211> 280
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti
 <400> 222

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

45
 <210> 223
 <211> 843
 <212> ADN
 <213> Secuencia artificial

50
 <220>
 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

55
 <400> 223

60

65

5
 10
 15
 20

```

atgagtagcc actttctcgg caaagtgaag ttcgatgaag cacggcttgc agaagacttg 60
tctacgctga gtcgtcgtga attttcctcc gcgtatagtg attttgcctg cggaaaatgg 120
gaactgtgtg ttcttagaaa ccgtacaggt atgcaagaag aggatattgt ggtttcccat 180
aacgcaccgg ctcttcagac cccgctgagt aaatcattgc cttaccttaa tgaattagtt 240
gagacacatt tcgactgttc gcgtgtccgt tacgttcgta ttgtaagagt gagcgagaac 300
gcgtgcctga taccacattc agattacctc gagcttgacg aggagtttac cagattgcat 360
ctggtcttgg ataccaactc aggttgtgcc aataccgaag aggacaagat attccatctg 420
ggtctcggtg aaatatggtt cttggatgcg atgctgcctc atagtgcagc atgtttttct 480
aagacgccta gactgcagct gatgattgac ttcgaagcta cagcgtttcc cgaatcattt 540
ttacggaacg tcgaacagcc ggttaccacc cgtgatatgg tcgaccctag aaaagaatta 600
acggacgaag tgatcgaagg tatactcggc ttttctatta tcatatcgga gccgaactac 660
cgtgaaatcg tgtctatttt agcgaagctg cacttctttt ataaagccga ttgccggtcg 720
atgtatgatt ggctgaaaga aatttgcaag agacggggag acccggcgct tattgaaaag 780
acggcttcat tggaacgctt ctttttaggc cacagagaac gcggggaagt aatgacctac 840
taa 843

```

<210> 224

<211> 280

<212> PRT

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 224

35
 40
 45
 50
 55
 60
 65

```

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

```

5 Met Val Asp Pro Arg Lys Glu Leu Thr Asp Glu Val Ile Glu Gly Ile
195 200 205
Leu Gly Phe Ser Ile Ile Ile Ser Glu Ala Asn Tyr Arg Glu Ile Val
210 215 220
Ser Ile Leu Ala Lys Leu His Phe Phe Tyr Lys Ala Asp Cys Arg Ser
225 230 235 240
10 Met Tyr Asp Trp Leu Lys Glu Ile Cys Lys Arg Arg Gly Asp Pro Ala
245 250 255
Leu Ile Glu Lys Thr Ala Ser Leu Glu Arg Phe Phe Leu Gly His Arg
260 265 270
15 Glu Arg Gly Glu Val Met Thr Tyr
275 280

20 <210> 225
<211> 843
<212> ADN
<213> Secuencia artificial

<220>
25 <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 225

30 atgagtagcc acttttctcgg caaagtgaag ttcgatgaag cacggcttgc agaagacttg 60
totacgctga gtcgtgcaga attttctctcc gcgtatagtg attttgcctg cggaaaatgg 120
gaactgtgtg ttcttagaaa ccgtacaggt atgcaagaag aggatattgt ggtttcccat 180
aaccagccgg ctcttcagac cccgctgagt aaatcattgc cttaccttaa tgaattagtt 240
gagacacatt tcgactgttc ggctgtccgt tacgttcgta ttgtaagagt gagcgagAAC 300
35 gcgtgcctga taccacattc agattacctc gagcttgaca acgagtttac cagattgcat 360
ctggctcttg ataccaactc aggttgtgcc aataccgaag aggacaagat attccatctg 420
ggtctcgggt aaatatgggt cttggatgag atgctgcctc atagtgcagc atgtttttct 480
aagacgccta gactgcagct gatgattgac ttggaagcta cagcgtttcc cgaatcattt 540
ttacggaacg tcgaacagcc gggtaccacc cgtgatatgg tcgaccctag aaaagaatta 600
40 acggacgaag tgatcgaagg tatactcggc ttttctatta tcatatcgga ggcgaactac 660
cgtgaaatcg tgtctatatt agcgaagctg cacttctttt ataaagccga ttgccgggtcg 720
atgtatgatt ggctgaaaga aatttgcaag agacggggag acccgggcgt tattgaaaag 780
acggcttcat tggaaacgct ctttttaggc cacagagaac gcggggaagt aatgacctac 840
taa 843

45

50 <210> 226
<211> 280
<212> PRT
<213> Secuencia artificial

<220>
<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

55 <400> 226

60

65

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

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

<210> 227

<211> 843

<212> ADN

<213> Secuencia artificial

<220>

<223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

<400> 227

ES 2 712 682 T3

```

5      atgtctagcc atttccttgg caaggtcaaa ttcgacgagg ccagattagc cgaagatctg 60
      tctactttaa gtcgtcgtga gttttcttcc gcatattcgg acttcgcttg tggaaagtgg 120
      gaactttgtg ttttacgcaa cagaacaggc atgcaagaag aagacattgt ggtttctcac 180
      aatgcgcccg cattacagac ccctttatct aaatcactgc cgtacttgaa tgagcttggt 240
      gaaacacact tcgactgcag tcgtgttaga tatgtacgga tcgtgcgtgt ctccgagaac 300
      gcctgcctga tccctcatag tgattatttg gagttggatg aagaattcac ccggtgcac 360
      ctctgtgttg atactaacag tggttgtgcc aacactgagg aagataagat attccacctg 420
10     ggtctgggcg agatctgggt tctggacgca atgttacctc attctgctgc ctgcttttct 480
      aaaactccta gactgcagct catgatagat tttgaagcta ccgccttccc ggaatccttc 540
      ttacgcaatg tggaacagcc tgtaactact cgcgacatgg tagaccctcg taaagaattg 600
      acggacgaag ttatagaagg cattttaggg ttctccataa tcatcagtga agcgaattac 660
      cgcgagatcg tctctatcct ggccaaatta cacttctttt ataaagcgga ctgtcggtcc 720
15     atgtatgatt ggcttaaaga aatatgtaa cgccggggag atcccgcgtt gattgagaag 780
      acagcgagtc tcgaacgttt ctttttgggg catcgtgaac gcggagaagt aatgacgtat 840
      taa                                     843

```

```

20     <210> 228
      <211> 280
      <212> PRT
      <213> Secuencia artificial

```

```

25     <220>
      <223> Gen de 4-hidroxilasa de prolina optimizada por codon de Sinorhizobium meliloti

      <400> 228

```

30

35

40

45

50

55

60

65

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

45

50

55

60

65

REIVINDICACIONES

1. Un polipéptido diseñado que tiene actividad de prolina hidroxilasa, que comprende una secuencia de aminoácidos que tiene al menos un 80% de identidad de secuencia con la SEQ ID NO: 2 en la que la prolina hidroxilasa tiene al menos 1,2 veces mayor actividad en comparación con el polipéptido de la SEQ ID NO: 2 y en el que la secuencia de aminoácidos comprende una diferencia de residuos en comparación con la secuencia de la SEQ ID NO: 2 en la posición del residuo X166.

2. El polipéptido diseñado según la reivindicación 1:

(i) en el que las diferencias de residuos en las posiciones de residuos X166 se seleccionan de X166Q; X166L; X166T y/o

(ii) que comprenden además una o más diferencias de residuos en comparación con la secuencia de la SEQ ID NO: 2 en las posiciones de residuos; X2; X3; X4; X5; X9; X13; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X66; X86; X92; X95; X103; X112; X113; X115; X116; X121; X131; X150; X151; X225; X230; X270; y X271.

3. El polipéptido diseñado según la reivindicación 2, en el que las diferencias de residuos en las posiciones de residuos X2; X3; X4; X5; X9; X13; X25; X26; X29; X30; X36; X42; X52; X57; X58; X59; X66; X92; X95; X103; X112; X115; X116; X121; X131; X150; X151; X225; X230; y X271 se seleccionan de X2K; X2T; X3S; X4Q; X4L; X4E; X4S; X5I; X5L; X5M; X9I; X13T; X25R; X26T; X29A; X30V; X30P; X36T; X42E; X52P; X57T; X57A; X58A; X59G; X66Q; X86S; X92V; X95M; X103L; X103Q; X112T; X112V; X113E; X115E; X115H; X115D; X115G; X115S; X115A; X116L; X121F; X131Y; X131F; X150S; X151S; X225L; X225Y; X225W; X230V; X270E; X271K; y X271R.

4. El polipéptido diseñado de acuerdo con una cualquiera de las reivindicaciones 1 a 3, en el que la secuencia de aminoácidos comprende al menos una combinación de características seleccionadas de:

(a) X103L y X166Q;

(b) X52P y X255Y;

(c) X4E/L/S y X115A;

(d) X25R y X58A;

(e) X29A y X166T/Q/L;

(f) X115H/D/G y X121F;

(g) X3S, X103L y X166Q;

(h) X103L, X131Y/F, y X166T/Q/L;

(i) X26T, X103L y X166T/Q/L;

(j) X25R, X66Q, X92V y X115E;

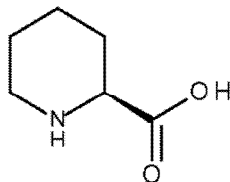
(k) X25R, X66Q, X92V, X103L, X115E y X166Q; y

(l) X3S, X25R, X66Q, X92V, X103L, X115E y X166Q.

5. El polipéptido diseñado de acuerdo con una cualquiera de las reivindicaciones 1 a 4, que comprende además una o más diferencias de residuos en comparación con la secuencia de la SEQ ID NO: 2 en las posiciones de residuos seleccionadas de: X17, X24, X26, X62, X88, X98, X114, X140, X151, X186, X188 y X205.

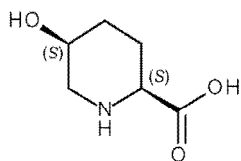
6. El polipéptido diseñado según la reivindicación 5, en el que las diferencias de residuos en las posiciones de residuos X17, X24, X26, X62, X88, X114, X140, X151, X186, X188 y X205 se seleccionan entre X17V, X24R, X24S, X26R, X26W, X62Q, X88R, X98F, X98T, X114N, X140L, X151A, X151H, X186G, X188G y X205V.

7. El polipéptido diseñado de acuerdo con una cualquiera de las reivindicaciones 1 a 6, que es capaz de convertir el compuesto de sustrato (2), ácido (2S)-piperidina-2-carboxílico,



(2)

al producto compuesto (1), ácido (2S, 5S)-5-hidroxipiperidina-2-carboxílico,



(1)

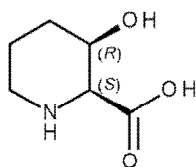
bajo condiciones de reacción adecuadas.

8. El polipéptido diseñado según la reivindicación 7, en el que el polipéptido es capaz de convertir el compuesto de sustrato (2) en el compuesto del producto (1) con al menos 2 veces la actividad de la SEQ ID NO: 2.

9. El polipéptido diseñado según la reivindicación 8, en el que la secuencia de aminoácidos comprende una o más características seleccionadas de: X3S; X4Q; X4L; X5I; X5L; X24S; X25R; X30P; X66Q; X86S; X92V; X103L; X103Q; X113E; X115E; X150S; X166Q; X151S; X225L; y X270E.

10. El polipéptido diseñado de acuerdo con una cualquiera de las reivindicaciones 1 a 9, que es capaz de:

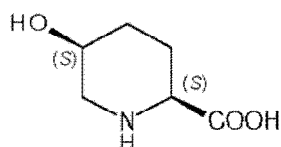
i) convertir el compuesto de sustrato (2) en el compuesto de producto (1) en exceso del producto compuesto (1a), ácido (2S, 3R)-3-hidroxipiperidina-2-carboxílico,



(1a)

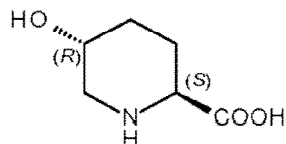
en donde la secuencia de aminoácidos comprende una o más características seleccionadas de: X103L; X115E; X131Y y X166Q; o

ii) compuesto de producto formador (1),



(1)

en exceso diastereomérico del producto compuesto (1R), ácido (2S, 5R)-5-hidroxipiperidina-2-carboxílico,



(1R)

11. El polipéptido diseñado según la reivindicación 1 que tiene actividad de prolina hidroxilasa en la que la secuencia de aminoácidos comprende una secuencia seleccionada del grupo que consiste en la SEQ ID NO: 24, 20, 22, 100, 104, 106, 108, 110, 112, , 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164 , 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196,

198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226 y 228.

12. Un polinucleótido que codifica el polipéptido diseñado de acuerdo con una cualquiera de las reivindicaciones 1 a 11.

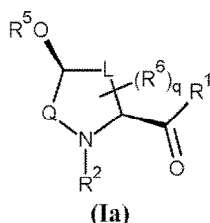
13. Un polinucleótido que codifica el polipéptido de la reivindicación 1-11 que incluye una secuencia de ácido nucleico optimizada para la expresión en *E. coli*.

14. Un vector de expresión que comprende el polinucleótido de la reivindicación 12 o 13.

15. Una célula huésped que comprende el polinucleótido de la reivindicación 12 o 13.

16. Un método para preparar el polipéptido modificado de una cualquiera de las reivindicaciones 1 a 11, que comprende cultivar la célula huésped de la reivindicación 15 en condiciones adecuadas para la expresión del polipéptido.

17. Un proceso para preparar un producto compuesto de fórmula **(Ia)**,



en donde

Q se selecciona del grupo que consiste en un alquileo (C₁-C₅) y alquilenilo (C₂-C₅);

L se selecciona del grupo que consiste en un enlace, alquileo (C₁-C₄) y alquilenilo (C₂-C₄);

R¹ se selecciona del grupo que consiste en hidróxido, amino, alquilo (C₁-C₆), alilo, alquilo (C₁-C₆) y arilo; R² se selecciona del grupo que consiste en hidrógeno y alquilo (C₁-C₆) opcionalmente sustituido, alquilenilo (C₂-C₆) y alquilenilo (C₂-C₆);

R⁵ es hidrógeno, o un enlace directo a un átomo de carbono de L para formar un epóxido;

cada aparición de R⁶ se selecciona del grupo que consiste en halo, alquilo (C₁-C₆) y alquilo (C₁-C₆); y

q es un número entero de 0 a 4;

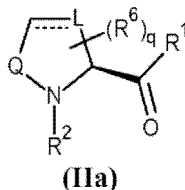
en donde la suma de los átomos de carbono del anillo para Q + L es un número entero de 2 a 5;

con la condición de que

(i) cuando la suma de los átomos de carbono del anillo para Q + L es 2, entonces L es un metileno; y

(ii) cuando la suma de los átomos de carbono del anillo para Q + L es 3, entonces L es un enlace o etileno;

el proceso que comprende poner en contacto el compuesto de sustrato de fórmula **(IIa)**,

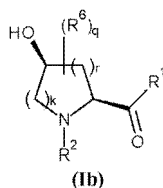


en donde L, Q, R¹, R², R⁶ y q son como se definieron anteriormente para el compuesto de fórmula **(Ia)**; y ----- representa un enlace opcional a un átomo de carbono de L para formar un doble enlace;

con un polipéptido diseñado de acuerdo con una cualquiera de las reivindicaciones 1 a 11, en presencia de un co-sustrato en condiciones de reacción adecuadas.

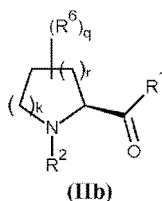
18. El proceso de la reivindicación 17 en el que:

i) el compuesto de fórmula **(Ia)** comprende el compuesto de fórmula **(Ib)**,



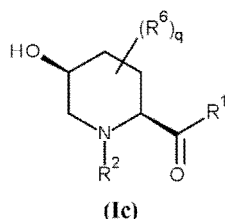
en donde

R^1 se selecciona del grupo que consiste en hidroxilo, amino, alquilo (C₁-C₆), arilo, alquiltio (C₁-C₆) y ariltio;
 R^2 se selecciona del grupo que consiste en hidrógeno y alquilo (C₁-C₆) opcionalmente sustituido, alqueno (C₂-C₆) y alquino (C₂-C₆);
 cada aparición de R^6 se selecciona independientemente del grupo que consiste en halo, alquilo (C₁-C₆) y alquilo (C₁-C₆); y
 k es un número entero de 1 a 5;
 r es un número entero de 0 a 4;
 en donde $k + r$ es 3, 4 o 5; y
 q es un número entero de 0 a 4;
 con la condición de que cuando $k + r$ es 3, entonces k es 1 o 3;
 el proceso que comprende poner en contacto el compuesto de sustrato de fórmula **(Ib)**,



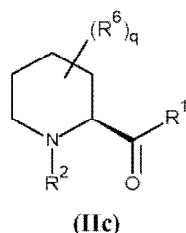
en la que R^1 , R^2 , R^6 , k , r y q son como se definieron anteriormente para el compuesto de fórmula **(Ib)**, con un polipéptido diseñado de una cualquiera de las reivindicaciones 1 a 11 en presencia de un co-sustrato en condiciones de reacción adecuadas;

ii) el compuesto de fórmula **(Ia)** comprende el compuesto de fórmula **(Ic)**,



en donde

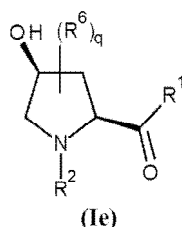
R^1 se selecciona del grupo que consiste en hidroxilo, amino, alquilo (C₁-C₆), arilo, alquiltio (C₁-C₆) y ariltio;
 R^2 se selecciona del grupo que consiste en hidrógeno y alquilo (C₁-C₆) opcionalmente sustituido, alqueno (C₂-C₆) y alquino (C₂-C₆);
 cada aparición de R^6 se selecciona independientemente del grupo que consiste en hidrógeno, halo, alquilo (C₁-C₆) y alquilo (C₁-C₆); y
 q es un número entero de 0 a 4;
 que comprende poner en contacto el compuesto de sustrato de fórmula **(Ic)**,



en donde

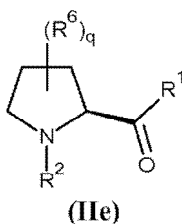
R¹, R², R⁶ y q son como se definieron anteriormente para el compuesto de fórmula **(Ic)**, con un polipéptido diseñado de acuerdo con una cualquiera de las reivindicaciones 1 a 11 en presencia de un co-sustrato en condiciones de reacción adecuadas;

iii) el compuesto de fórmula **(Ia)** comprende el compuesto de fórmula **(Ie)**,



en donde

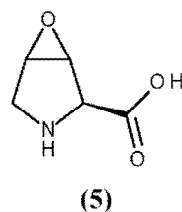
R¹ se selecciona del grupo que consiste en hidroxilo, amino, alquilo (C₁-C₆), arilo, alquilo (C₁-C₆) y arilo; R² se selecciona del grupo que consiste en hidrógeno y alquilo (C₁-C₆) opcionalmente sustituido, alquilo (C₂-C₆) y alquilo (C₂-C₆); cada aparición de R⁶ se selecciona independientemente del grupo que consiste en hidrógeno, halo, alquilo (C₁-C₆) y alquilo (C₁-C₆); y q es un número entero de 0 a 3; el proceso que comprende poner en contacto el compuesto de sustrato de fórmula **(Ie)**,



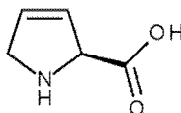
en donde

R¹, R², R⁶ y q son los definidos para el compuesto de fórmula **(Ie)**; con un polipéptido diseñado de acuerdo con una cualquiera de las reivindicaciones 1 a 11 en presencia de un co-sustrato en condiciones de reacción adecuadas; o

iv) el compuesto de fórmula **(Ia)** comprende el compuesto **(5)**,



el proceso que comprende poner en contacto el compuesto de sustrato (6),



(6)

con un polipéptido diseñado de acuerdo con una cualquiera de las reivindicaciones 1 a 11, en presencia de un co-sustrato en condiciones de reacción adecuadas.

19. El proceso de cualquiera de las reivindicaciones 17 y 18 en el que:

- i) el co-sustrato comprende ácido α -cetoglutarico, opcionalmente a una concentración equimolar o superior a la de la concentración del compuesto de sustrato;
- ii) las condiciones de reacción comprenden Fe^{+2} , opcionalmente en el que el Fe^{+2} está en forma de $(\text{NH}_4)_2(\text{FeSO}_4)_2$;
- iii) las condiciones de reacción comprenden un reductor capaz de reducir Fe^{+3} a Fe^{+2} , opcionalmente en el que el reductor comprende ácido ascórbico presente en al menos aproximadamente 0,1 veces, 0,2 veces, 0,3 veces, 0,5 veces, 0,75 veces, 1 vez, 1,5 veces, o al menos 2 veces la concentración molar del compuesto de sustrato;
- iv) las condiciones de reacción comprenden O_2 , opcionalmente en el que el O_2 se proporciona por aireación forzada; o
- v) las condiciones de reacción adecuadas comprenden (a) carga de sustrato de aproximadamente 10 g/L a 100 g/L; (b) de aproximadamente 1 g/l a aproximadamente 50 g/l de polipéptido diseñado; (c) α -cetoglutarato en aproximadamente 1 a 2 equivalentes molares de compuesto de sustrato; (d) ácido ascórbico en aproximadamente 0,25 a 0,75 equivalentes molares de compuesto de sustrato; (e) de aproximadamente 0,5 mM a aproximadamente 12 mM de FeSO_4 ; (f) pH de aproximadamente 6 a 8; (g) temperatura de aproximadamente 20° a 40°C; y (h) tiempo de reacción de 6 a 120 h.