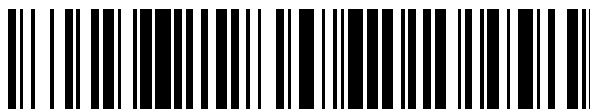


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 773 889**

51 Int. Cl.:

C12N 9/00 (2006.01)

C12N 1/21 (2006.01)

C12N 15/09 (2006.01)

C12P 13/04 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **27.01.2015 PCT/JP2015/052126**

87 Fecha y número de publicación internacional: **06.08.2015 WO15115398**

96 Fecha de presentación y número de la solicitud europea: **27.01.2015 E 15742961 (4)**

97 Fecha y número de publicación de la concesión europea: **11.12.2019 EP 3101123**

54 Título: **Hidroxilasa en la posición 4 de ácido pipecolínico y método para producir 4-hidroxiaminoácido usando la misma**

30 Prioridad:

31.01.2014 JP 2014017716

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
15.07.2020

73 Titular/es:

**API CORPORATION (100.0%)
13-4, Uchikanda 1-chome, Chiyoda-ku
Tokyo 1010047, JP**

72 Inventor/es:

**HIBI, MAKOTO;
OGAWA, JUN;
MIYAKE, RYOMA y
KAWABATA, HIROSHI**

74 Agente/Representante:

ARIAS SANZ, Juan

ES 2 773 889 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

Hidroxilasa en la posición 4 de ácido pipecolínico y método para producir 4-hidroxiaminoácido usando la misma

Campo técnico

La presente invención se refiere a un método para producir 4-hidroxiaminoácido, método que utiliza una 4-hidroxilasa de ácido pipecólico novedosa.

Técnica anterior

Las hidroxilasas de aminoácidos son enzimas útiles para la producción de productos intermedios de productos farmacéuticos y similares, y se ha informado de la presencia de prolina 4-hidroxilasa (documento no de patente 1), L-isoleucina dioxigenasa (documento no de patente 2), y similares. En cuanto a enzimas que tienen la capacidad de hidroxilar el ácido pipecólico, se ha informado que varios tipos de prolina hidroxilasas tienen la capacidad de hidroxilar la posición 3 o la posición 5 del ácido L-pipecólico (documento no de patente 3). Sin embargo, hasta el momento no se ha informado que ninguna enzima hidroxile la posición 4 del ácido pipecólico.

Las secuencias de aminoácidos de SEQ ID NO: 2, 4, 6, 8, 10 y 12 son las mismas que los números de registro de GenBank. EGU81245, XP_001827566, XP_002558179, XP_383389, ELA34460 y XP_659994, respectivamente, que son las secuencias de aminoácidos traducidas a partir de secuencias de ADN que codifican para proteínas supuestas basándose en la información de la secuencia genómica de la cepa *Fusarium oxysporum* Fo5176, cepa *Aspergillus oryzae* RIB40, cepa *Penicillium chrysogenum* Wisconsin 54-1255, cepa *Gibberella zeae* (otro nombre, *Fusarium graminearum*) PH-1, cepa *Colletotrichum gloeosporioides* Nara gc5 y cepa *Aspergillus nidulans* (otro nombre, *Emericella nidulans*) FGSC A4, respectivamente. Dado que todas estas proteínas se derivan de hongos, las proteínas, si existen, es probable que estén en un estado en el que han experimentado glicosilación después de su expresión. Sin embargo, no hay ningún informe que apoye su existencia basándose en el aislamiento real o similar, y las funciones de estas proteínas han sido desconocidas en absoluto.

Los 4-hidroxiaminoácidos ópticamente activos son sustancias útiles como productos intermedios de productos farmacéuticos y similares. Por ejemplo, el ácido (4S)-hidroxi-L-pipecólico puede usarse como precursor de un inhibidor de Rho cinasa (documento de patente 1) y el ácido (4S)-hidroxi-D-pipecólico puede usarse como precursor de CGS-20281, que es un inhibidor del receptor NMDA (documento no de patente 4). La (4R)-hidroxi-L-prolina puede usarse como precursor de acetato de icatibant, que es un inhibidor del receptor de bradicinina B2 (documento no de patente 5) y (4R)-hidroxi-D-prolina puede usarse como precursor de un inhibidor del factor Xa (documento de patente 2). Puede usarse L-homoserina como precursor de Omapatrilat, que es un inhibidor de la ECA (documento de patente 3) y puede usarse 4-hidroxi-L-leucina como precursor de SCY-635, que es un inhibidor de ciclofilina (documento no de patente 6).

Los ejemplos de métodos notificados previamente para la síntesis de un 4-hidroxiaminoácido ópticamente activo incluyen un método en el que se produce ácido (4S)-hidroxi-L-pipecólico o ácido (4S)-hidroxi-D-pipecólico mediante ciclación estereoselectiva de 3-butenol con una C-aminocarbonil nitrona o C-alcoxycarbonil nitrona ópticamente selectiva se ha notificado hasta ahora (documento no de patente 7). También se han notificado un método en el que se sintetiza (4R)-hidroxi-L-prolina a partir de L-prolina usando prolina 4-hidroxilasa derivada de la cepa de *Dactilopora RH1* (documento no de patente 1) y un método en el que se sintetiza (4R)-hidroxi-D-prolina por medio de un α,β -dideshidroaminoácido (documento no de patente 8). También se han notificado un método en el que se sintetiza L-homoserina mediante fermentación usando una *E. coli* recombinante (documento de patente 4) y un método en el que se sintetiza 4-hidroxi-L-leucina a partir de L-leucina usando L-isoleucina dioxigenasa derivada de la cepa de *Bacillus thuringiensis* 2e2 (documento no de patente 2).

Sin embargo, se han demandado métodos de síntesis más eficientes dado que estos métodos presentan problemas tales como materiales costosos, grandes números de etapas, pequeños números de los tipos de compuestos a los que son aplicables los métodos y altas cargas de purificación.

Documentos de la técnica anterior

[Documentos de patente]

Documento de patente 1: solicitud de patente japonesa PCT traducida abierta a consulta por el público 2010-514720

Documento de patente 2: solicitud de patente japonesa PCT traducida abierta a consulta por el público 2009-526813

Documento de patente 3: JP 7-48259 A

Documento de patente 4: WO 2013/134625

[Documentos no de patente]

Documento no de patente 1: Shibasaki *et al.*, Appl. Environ. Microbiol., 1999, 65, 4028

Documento no de patente 2: Hibi *et al.*, Appl. Environ. Microbiol., 2011, 77, 6926

Documento no de patente 3: Klein *et al.*, Adv. Synth. Catal., 2011, 353, 1375

Documento no de patente 4: Occhiato *et al.*, SYNTHESIS, 2009, 3611

Documento no de patente 5: Bork *et al.*, Nat. Rev. Drug Discov., 2008, 7, 801

5 Documento no de patente 6: Hopkins *et al.*, Antimicrobial. Agents Chemother., 2010, 54, 660

Documento no de patente 7: Cordero *et al.*, Eur. J. Org. Chem., 2006, 3235

Documento no de patente 8: Kimura *et al.*, Bull. Chem. Soc. Jpn., 2002, 75, 2517

Sumario de la invención

Problemas que van a resolverse mediante la invención

10 Un objeto de la presente invención es proporcionar una 4-hidroxilasa de ácido pipecólico novedosa, y un método novedoso para producir de manera económica y sencilla un 4-hidroxi-aminoácido, especialmente un 4-hidroxi-aminoácido ópticamente activo.

Medios para resolver los problemas

15 Con el fin de resolver los problemas descritos anteriormente, los presentes inventores estudiaron intensamente un método para producir un 4-hidroxi-aminoácido ópticamente activo. Como resultado, los presentes inventores descubrieron que la proteína FOXB_08233 derivada de la cepa de *Fusarium oxysporum* Fo5176, cuyo aislamiento como proteína no se ha notificado y cuya función se ha desconocido, y sus proteínas homólogas, tienen una actividad de 4-hidroxilasa de ácido pipecólico dependiente de ácido 2-oxoglutarico. Los presentes inventores descubrieron entonces que, al preparar transformantes usando ADN que codifican para estas proteínas y haciendo
20 reaccionar las células transformantes preparadas, los productos tratados de las mismas y/o los líquidos de cultivo de las mismas con diversos aminoácidos, pueden obtenerse diversos 4-hidroxi-aminoácidos ópticamente activos a una alta pureza óptica. La presente invención se logró basándose en estos descubrimientos.

Es decir, la presente invención puede resumirse de la siguiente manera.

25 [1] Una proteína 4-hidroxilasa de ácido pipecólico que tiene actividad para reaccionar con ácido L-pipecólico en presencia de ácido 2-oxoglutarico e iones de hierro (II), para producir ácido trans-4-hidroxi-L-pipecólico.

[2] La proteína 4-hidroxilasa de ácido pipecólico según [1], seleccionada del grupo que consiste en los siguientes (A), (B) y (C):

(A) un polipéptido que comprende la secuencia de aminoácidos representada por SEQ ID NO:2, 4, 6, 8, 10, 12, 16 ó 18;

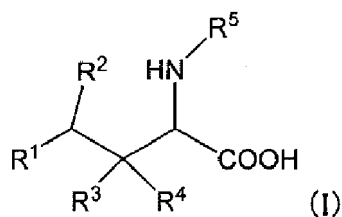
30 (B) un polipéptido que comprende la secuencia de aminoácidos representada por SEQ ID NO:2, 4, 6, 8, 10, 12, 16 ó 18 excepto que se deletan, sustituyen y/o añaden uno o varios aminoácidos, y que tiene actividad de 4-hidroxilasa de ácido pipecólico; y

35 (C) un polipéptido que comprende una secuencia de aminoácidos que no es menos del 80% idéntica a la secuencia de aminoácidos representada por SEQ ID NO:2, 4, 6, 8, 10, 12, 16 ó 18, y que tiene actividad de 4-hidroxilasa de ácido pipecólico.

[3] La proteína 4-hidroxilasa de ácido pipecólico según [2], que se produce como proteína recombinante por un huésped que no tiene capacidad de glicosilación.

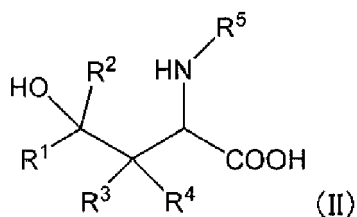
40 [4] Un método para producir 4-hidroxi-aminoácido, comprendiendo el método hacer reaccionar la proteína 4-hidroxilasa de ácido pipecólico según uno cualquiera de [1] a [3], una(s) célula(s) que comprende(n) la proteína, un producto tratado de la(s) célula(s) y/o un líquido de cultivo obtenido cultivando la(s) célula(s), con α -aminoácido para producir 4-hidroxi-aminoácido.

[5] El método para producir 4-hidroxi-aminoácido según [4], en el que el α -aminoácido se representa mediante la siguiente fórmula general (I):



(en la que cada uno de R^1 , R^2 , R^3 , R^4 y R^5 representa un átomo de hidrógeno o alquilo C_1 - C_3 , y alternativamente R^2 puede unirse a R^5 o al átomo de nitrógeno del grupo amino para formar una estructura de anillo),

y el 4-hidroxi-aminoácido se representa mediante la siguiente fórmula general (II):



(en la que cada uno de R^1 , R^2 , R^3 , R^4 y R^5 representa un átomo de hidrógeno o alquilo C_1 - C_3 , y alternativamente R^2 puede unirse a R^5 o al átomo de nitrógeno del grupo amino para formar una estructura de anillo).

[6] El método para producir 4-hidroxi-aminoácido según [4], en el que el α -aminoácido es ácido pipercolico y el 4-hidroxi-aminoácido es ácido 4-hidroxi-L-pipercolico.

10 [7] El método para producir 4-hidroxi-aminoácido según uno cualquiera de [4] a [6], en el que la célula que comprende 4-hidroxilasa de ácido pipercolico es una célula transformada con un ADN que codifica para proteína 4-hidroxilasa de ácido pipercolico.

[8] El método para producir 4-hidroxi-aminoácido según [7], en el que el ADN que codifica para proteína 4-hidroxilasa de ácido pipercolico se selecciona del grupo que consiste en los siguientes (D), (E) y (F):

15 (D) un ADN que comprende la secuencia de nucleótidos representada por SEQ ID NO:1, 3, 5, 7, 9, 11, 15 ó 17;

(E) un ADN que comprende la secuencia de nucleótidos representada por SEQ ID NO:1, 3, 5, 7, 9, 11, 15 ó 17 excepto que se sustituyen, delecionan y/o añaden una o varias bases, y que codifica para un polipéptido que tiene actividad de 4-hidroxilasa de ácido pipercolico; y

20 (F) un ADN que comprende una secuencia de nucleótidos que se hibrida con la cadena complementaria de la secuencia de nucleótidos representada por SEQ ID NO:1, 3, 5, 7, 9, 11, 15 ó 17 en condiciones rigurosas, y que codifica para un polipéptido que tiene actividad de 4-hidroxilasa de ácido pipercolico.

[9] Un microorganismo transformado con un ADN que codifica para una proteína 4-hidroxilasa de ácido pipercolico seleccionada del grupo que consiste en los siguientes (D), (E) y (F):

(D) un ADN que comprende la secuencia de nucleótidos representada por SEQ ID NO:1, 3, 5, 7, 9, 11, 15 ó 17;

25 (E) un ADN que comprende la secuencia de nucleótidos representada por SEQ ID NO:1, 3, 5, 7, 9, 11, 15 ó 17 excepto que se sustituyen, delecionan y/o añaden una o varias bases, y que codifica para un polipéptido que tiene actividad de 4-hidroxilasa de ácido pipercolico; y

30 (F) un ADN que comprende una secuencia de nucleótidos que se hibrida con la cadena complementaria de la secuencia de nucleótidos representada por SEQ ID NO:1, 3, 5, 7, 9, 11, 15 ó 17 en condiciones rigurosas, y que codifica para un polipéptido que tiene actividad de 4-hidroxilasa de ácido pipercolico.

[10] El microorganismo según [9], que se selecciona del grupo que consiste en los géneros *Escherichia*, *Bacillus*, *Pseudomonas* y *Corynebacterium*.

Efecto de la invención

35 Según la presente invención, pueden producirse de manera eficiente diversos 4-hidroxi-aminoácidos. En particular, pueden producirse de manera eficiente a altas purzas ópticas 4-hidroxi-aminoácidos ópticamente activos tales como ácido (4S)-hidroxi-L-pipercolico y ácido (4S)-hidroxi-D-pipercolico, que son útiles como productos intermedios de productos farmacéuticos.

Breve descripción de los dibujos

La figura 1 es un diagrama que muestra el resultado de análisis de CL-EM de un líquido de reacción obtenido haciendo reaccionar L-prolina con *E. coli* transformada con pFoPA4H.

La figura 2 es un diagrama que muestra el resultado de análisis de CL-EM de un líquido de reacción obtenido haciendo reaccionar L-leucina con *E. coli* transformada con pFoPA4H.

- 5 La figura 3 es un diagrama que muestra el resultado de análisis de CL-EM de un líquido de reacción obtenido haciendo reaccionar ácido L-pipecólico con *E. coli* transformada con pFoPA4H.

La figura 4 es un diagrama que muestra el resultado de análisis de CL-EM de un líquido de reacción obtenido haciendo reaccionar D-prolina con *E. coli* transformada con pEnPA4H.

- 10 La figura 5 es un diagrama que muestra el resultado de análisis de CL-EM de un líquido de reacción obtenido haciendo reaccionar ácido (S)-2-aminobutírico con *E. coli* transformada con pEnPA4H.

La figura 6 es un diagrama que muestra el resultado de análisis de ¹H-RMN de un producto de reacción obtenido haciendo reaccionar ácido L-pipecólico con *E. coli* transformada con pFoPA4H (imagen de medio tono).

La figura 7 es un diagrama que muestra el resultado de análisis de ¹³C-RMN de un producto de reacción obtenido haciendo reaccionar ácido L-pipecólico con *E. coli* transformada con pFoPA4H (imagen de medio tono).

- 15 La figura 8 es un diagrama que muestra el resultado de análisis de CL-EM de un líquido de reacción obtenido haciendo reaccionar ácido (5S)-hidroxi-L-pipecólico con un extracto libre de células obtenido de *E. coli* transformada con pEnPA4H.

- 20 La figura 9 es un diagrama que muestra el resultado de análisis de CL-EM de un líquido de reacción obtenido haciendo reaccionar (3R)-hidroxi-L-prolina con un extracto libre de células obtenido de *E. coli* transformada con pAoPA4H (superior) y con un extracto libre de células obtenido de *E. coli* transformada con pEnPA4H (inferior), respectivamente.

Realizaciones para llevar a cabo la invención

La presente invención se describe a continuación en detalle.

<4-Hidroxilasa de ácido pipecólico y método para producir 4-hidroxiaminoácido usando la misma>

- 25 La 4-hidroxilasa de ácido pipecólico de la presente invención comprende la secuencia de aminoácidos de SEQ ID NO:2, 4, 6, 8, 10, 12, 16 ó 18, o una secuencia homóloga de la misma que tiene actividad de 4-hidroxilasa de ácido pipecólico.

- 30 Las secuencias de aminoácidos de SEQ ID NO:2, 4, 6, 8, 10 y 12 se derivan de la cepa de *Fusarium oxysporum* c8D, cepa de *Aspergillus oryzae* RIB40, cepa de *Penicillium chrysogenum* Wisconsin 54-1255, cepa de *Gibberella zeae* (otro nombre, *Fusarium graminearum*) PH-1, cepa de *Colletotrichum gloeosporioides* Nara gc5 y cepa de *Aspergillus nidulans* (otro nombre, *Emericella nidulans*) FGSC A4, respectivamente. Estas son las secuencias de aminoácidos de proteínas identificadas como 4-hidroxilasa de ácido pipecólico en la presente invención. Las secuencias de aminoácidos de SEQ ID NO:16 y 18 son secuencias obtenidas mediante análisis de ADN recogidos directamente de muestras de suelos en Japón, y son las secuencias de aminoácidos de proteínas identificadas como 4-hidroxilasa de ácido pipecólico en la presente invención. Aunque la secuencia de aminoácidos de SEQ ID NO:16 muestra una homología del 97% con una proteína que se deriva de *Variovorax paradoxus* EPS y cuya función se desconoce (n.º de registro de GenBank YP_004156498), no hay ningún informe que describa su función. Además, aunque la secuencia de aminoácidos de SEQ ID NO:18 muestra una homología del 50% con una proteína que se deriva de *Burkholderia* sp A1 y cuya función se desconoce (n.º de registro de GenBank WP_029951026), no hay ningún informe que describa su función.

Ninguna de estas secuencias puede asumirse fácilmente que sea una hidroxilasa para ácido pipecólico basándose sólo en su información de secuencia. La presente invención las identificó como 4-hidroxilasas de ácido pipecólico por primera vez.

Puede usarse una pluralidad de tipos de 4-hidroxilasas de ácido pipecólico.

- 45 En la presente descripción, la actividad de 4-hidroxilasa de ácido pipecólico significa una actividad que añade un grupo hidroxilo al átomo de carbono en la posición 4 del ácido L-pipecólico en presencia de ácido 2-oxoglutarico e iones de hierro (II), para producir ácido trans-4-hidroxi-L-pipecólico. Una actividad de este tipo puede confirmarse permitiendo que, como enzima, la proteína de interés, una(s) célula(s) que expresa(n) la proteína o un producto tratado de la misma, actúen en un sistema de reacción que contiene ácido L-pipecólico como sustrato y ácido 2-oxoglutarico e iones de hierro (II) como cofactores, y después midiendo la producción de ácido trans-4-hidroxi-L-pipecólico tal como se describe en los ejemplos descritos más adelante.

Los ejemplos del homólogo de la 4-hidroxilasa de ácido pipecólico que comprende la secuencia de aminoácidos de

SEQ ID NO:2, 4, 6, 8, 10, 12, 16 ó 18 en la presente invención incluyen aquellos que comprenden la secuencia de aminoácidos de SEQ ID NO:2, 4, 6, 8, 10, 12, 16 ó 18 excepto que se delecionan, sustituyen y/o añaden uno o varios aminoácidos, siempre que los homólogos tengan actividad de 4-hidroxilasa de ácido pipecólico. En este caso, el término "uno o varios aminoácidos" significa, por ejemplo, de 1 a 100, preferiblemente de 1 a 50, más preferiblemente de 1 a 20, todavía más preferiblemente de 1 a 10, especialmente preferible de 1 a 5 aminoácidos.

El homólogo puede ser una proteína que tiene una secuencia que no es menos del 80%, preferiblemente no menos del 90%, más preferiblemente no menos del 95%, idéntica a la secuencia de aminoácidos de longitud completa de SEQ ID NO: 2, 4, 6, 8, 10, 12, 16 ó 18 siempre que el homólogo mantenga actividad de 4-hidroxilasa de ácido pipecólico.

La 4-hidroxilasa de ácido pipecólico que puede usarse en la presente invención puede obtenerse mediante purificación de los microorganismos descritos anteriormente, o puede obtenerse clonando un ADN que codifica para la 4-hidroxilasa de ácido pipecólico mediante un método conocido tal como PCR o hibridación, y permitiendo la expresión de la 4-hidroxilasa de ácido pipecólico en un huésped apropiado.

Los ejemplos del ADN que codifica para 4-hidroxilasa de ácido pipecólico incluyen ADN que comprenden la secuencia de nucleótidos representada por SEQ ID NO:1, 3, 5, 7, 9, 11, 15 ó 17.

El ADN que codifica para 4-hidroxilasa de ácido pipecólico también puede ser un homólogo de un ADN que comprende la secuencia de nucleótidos de SEQ ID NO:1, 3, 5, 7, 9, 11, 15 ó 17 siempre que el ADN codifique para una proteína que tiene actividad de 4-hidroxilasa de ácido pipecólico. Los ejemplos de un homólogo de este tipo incluyen aquellos que comprenden la secuencia de nucleótidos de SEQ ID NO:1, 3, 5, 7, 9, 11, 15 ó 17 excepto que se sustituyen, delecionan y/o añaden uno o varios nucleótidos.

En este caso, el término "uno o varios nucleótidos" significa, por ejemplo, de 1 a 300, preferiblemente de 1 a 150, más preferiblemente de 1 a 60, todavía más preferiblemente de 1 a 30, especialmente preferible de 1 a 15 nucleótidos.

El homólogo de ADN puede ser un ADN que se hibrida con la cadena complementaria de la secuencia de nucleótidos de SEQ ID NO:1, 3, 5, 7, 9, 11, 15 ó 17 en condiciones rigurosas, siempre que el homólogo de ADN codifique para una proteína que tiene actividad de 4-hidroxilasa de ácido pipecólico. Los ejemplos de las "condiciones rigurosas" en el presente documento incluyen condiciones en las que el lavado se lleva a cabo en presencia de 0,1xSSC (solución salina-citrato de sodio) y SDS al 0,1% (dodecil sulfato de sodio) a 60°C.

Los expertos en la técnica pueden obtener un homólogo de ADN de este tipo introduciendo de manera apropiada una(s) mutación/mutaciones de sustitución, delección, inserción y/o adición en el ADN de SEQ ID NO:1, 3, 5, 7, 9, 11, 15 ó 17 mediante mutagénesis dirigida al sitio (Nucleic Acids Res. 10, pág. 6487 (1982); Methods in Enzymol. 100, pág. 448 (1983); Molecular Cloning, PCR A Practical Approach IRL Press pág. 200 (1991)) o similares.

La secuencia de aminoácidos de SEQ ID NO:2, 4, 6, 8, 10, 12, 16 ó 18 o una parte de la misma, o la secuencia de nucleótidos representada por SEQ ID NO:1, 3, 5, 7, 9, 11, 15 ó 17 o una parte de la misma, puede someterse a búsqueda de homología frente a una base de datos proporcionada por, por ejemplo, el banco de datos de Japón (DNA Databank of JAPAN, DDBJ), para obtener información de aminoácidos para la actividad de 4-hidroxilasa de ácido pipecólico, o información de secuencia de nucleótidos de un ADN que lo codifica.

En el método para producir 4-hidroxiaminoácido de la presente invención, puede usarse directamente 4-hidroxilasa de ácido pipecólico en la reacción. Sin embargo, se prefiere usar células que contienen 4-hidroxilasa de ácido pipecólico, un producto tratado de las células o un líquido de cultivo obtenido cultivando las células.

Las células que contienen 4-hidroxilasa de ácido pipecólico pueden ser células de un microorganismo que tienen intrínsecamente 4-hidroxilasa de ácido pipecólico. Sin embargo, se prefiere usar células transformadas con un gen que codifica para 4-hidroxilasa de ácido pipecólico.

Los ejemplos del producto tratado de las células que contienen 4-hidroxilasa de ácido pipecólico incluyen productos celulares tratados tales como aquellos obtenidos mediante tratamiento con un disolvente orgánico tal como acetona, dimetilsulfóxido (DMSO) o tolueno, o con un tensioactivo, aquellos obtenidos mediante liofilización y aquellos obtenidos mediante destrucción física o enzimática; fracciones enzimáticas de las células obtenidas como producto en bruto o producto purificado; y productos preparados inmovilizándolos en un portador tal como gel de poliacrilamida y gel de carragenano.

Al insertar el ADN aislado de ese modo que codifica para 4-hidroxilasa de ácido pipecólico en un vector de expresión conocido de una manera que permite la expresión de la 4-hidroxilasa de ácido pipecólico, se proporciona un vector que expresa 4-hidroxilasa de ácido pipecólico. Al transformar células huésped con este vector de expresión, puede obtenerse un transformante al que se le introduce un ADN que codifica para 4-hidroxilasa de ácido pipecólico. El transformante también puede obtenerse incorporando un ADN que codifica para 4-hidroxilasa de ácido pipecólico en el ADN cromosómico de un huésped mediante un método tal como recombinación homóloga de una manera que permite la expresión de la 4-hidroxilasa de ácido pipecólico.

Los ejemplos específicos del método para preparar el transformante incluyen un método en el que un ADN que codifica para 4-hidroxilasa de ácido pipécólico se introduce en un vector plasmídico, vector de fago o vector de virus que puede estar presente de manera estable en una célula huésped tal como una célula de microorganismos, y el vector de expresión construido se introduce en la célula huésped, o el ADN está directamente en el genoma del huésped, seguido por permitir la transcripción y traducción de la información genética. En este procedimiento, un promotor adecuado se liga preferiblemente en el sentido de 5' del ADN, y, además, se liga preferiblemente un terminador en el sentido de 3' del ADN. Un promotor y un terminador de este tipo no están limitados siempre que sean un promotor y un terminador que se sabe que actúan en la célula usada como huésped. Por ejemplo, "Fundamental Microbiology 8: Genetic Engineering, KYORITSU SHUPPAN CO., LTD." describe detalles de vectores, promotores y terminadores que pueden usarse en microorganismos huésped.

El microorganismo huésped que va a transformarse para la expresión de 4-hidroxilasa de ácido pipécólico no está limitado siempre que el huésped en sí no afecte de manera negativa a la reacción de α -aminoácido, y los ejemplos específicos de los microorganismos huésped incluyen los siguientes microorganismos:

Bacterias que pertenecen a los géneros *Escherichia*, *Bacillus*, *Pseudomonas*, *Serratia*, *Brevibacterium*, *Corynebacterium*, *Streptococcus*, *Lactobacillus*, y similares cuyos sistemas de vectores de huésped se han establecido.

Actinomicetos que pertenecen a los géneros *Rhodococcus*, *Streptomyces*, y similares cuyos sistemas de vectores de huésped se han establecido.

Levaduras que pertenecen a los géneros *Saccharomyces*, *Kluyveromyces*, *Schizosaccharomyces*, *Zygosaccharomyces*, *Yarrowia*, *Trichosporon*, *Rhodospiridium*, *Hansenula*, *Pichia*, *Candida*, y similares cuyos sistemas de vectores de huésped se han establecido.

Hongos que pertenecen a los géneros *Neurospora*, *Aspergillus*, *Cephalosporium*, *Trichoderma*, y similares cuyos sistemas de vectores de huésped se han establecido.

El procedimiento para la construcción del transformante, el método para la construcción de un vector recombinante adecuado para el huésped y el método para cultivar el huésped pueden llevarse a cabo según técnicas usadas comúnmente en los campos de biología molecular, ingeniería biológica e ingeniería genética (por ejemplo, métodos descritos en Molecular Cloning).

Lo siguiente son ejemplos específicos de microorganismos huésped preferidos, y métodos de transformación, vectores, promotores, terminadores, y similares preferidos para los microorganismos. La presente invención no se limita a estos ejemplos.

Para el género *Escherichia*, especialmente *Escherichia coli*, los ejemplos del vector plasmídico incluyen plásmidos pBR y pUC, y los ejemplos del promotor incluyen promotores derivados de lac (β -galactosidasa), trp (operón de triptófano), tac, trc (fusión de lac y trp) y fago λ PL y PR. Los ejemplos del terminador incluyen terminadores derivados de trpA, fagos y ARN ribosómico rrnB.

Para el género *Bacillus*, los ejemplos del vector incluyen plásmidos pUB110 y plásmidos pC194. La integración en el cromosoma también es posible. Los ejemplos del promotor y el terminador incluyen aquellos de genes de enzimas tales como proteasa alcalina, proteasa neutra y α -amilasa.

Para el género *Pseudomonas*, los ejemplos del vector incluyen sistemas de vectores de huésped comunes establecidos en *Pseudomonas putida*, *Pseudomonas cepacia*, y similares; y un vector de amplia variedad de huéspedes (que contiene genes requeridos para replicación autónoma derivados de RSF1010 y similares) pKT240, que se basa en un plásmido implicado en la degradación de compuestos de tolueno, plásmido TOL (Gene, 26, 273-82 (1983)).

Para el género *Brevibacterium*, especialmente *Brevibacterium lactofermentum*, los ejemplos del vector incluyen vectores plasmídicos tales como pAJ43 (Gene 39, 281 (1985)). Los ejemplos del promotor y el terminador incluyen promotores y terminadores usados en *E. coli*.

Para el género *Corynebacterium*, especialmente *Corynebacterium glutamicum*, los ejemplos del vector incluyen vectores plasmídicos tales como pCS11 (JP 57-183799 A) y pCB101 (Mol. Gen. Genet. 196, 175 (1984)).

Para el género *Saccharomyces*, especialmente *Saccharomyces cerevisiae*, los ejemplos del vector incluyen plásmidos YRp, YEp, YCp y YIp. Los ejemplos de promotores y terminadores que pueden usarse incluyen aquellos de los genes de enzimas tales como alcohol deshidrogenasa, gliceraldehído-3-fosfato deshidrogenasa, fosfatasa ácida, β -galactosidasa, fosfoglicerato cinasa y enolasa.

Para el género *Schizosaccharomyces*, los ejemplos del vector incluyen el vector plasmídico derivado de *Schizosaccharomyces pombe* descrito en Mol. Cell. Biol. 6, 80 (1986). En particular, pAUR224 está disponible comercialmente de Takara Shuzo Co., Ltd., y puede usarse fácilmente.

En cuanto al género *Aspergillus*, *Aspergillus niger*, *Aspergillus oryzae*, y similares son las especies mejor estudiadas entre hongos. Los plásmidos y la integración en el cromosoma son aplicables a estas especies, y pueden usarse promotores derivados de proteasa y amilasa extracelulares (Trends in Biotechnology 7, 283-287 (1989)).

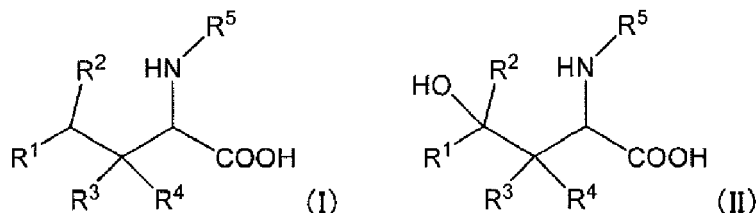
- 5 Entre los microorganismos descritos anteriormente, los ejemplos de microorganismos especialmente preferidos incluyen microorganismos que pertenecen a los géneros *Escherichia*, *Bacillus*, *Pseudomonas* y *Corynebacterium* que no tienen capacidad de glicosilación.

También se han establecido sistemas de huésped-vector distintos de los sistemas descritos anteriormente para diversos microorganismos, y esos sistemas pueden usarse según sea apropiado.

- 10 Se han establecido diversos sistemas de huésped-vector para plantas y animales además de microorganismos. En particular, sistemas para permitir la expresión de una gran cantidad de proteína foránea en un animal tal como un insecto incluyendo el gusano de seda (Nature 315, 592-594 (1985)), o en una planta tal como colza, maíz o patata; y se han establecido sistemas que usan sistemas de síntesis de proteínas libres de células tales como extractos libres de células de *E. coli* y gérmenes de trigo; y pueden usarse preferiblemente.

- 15 Al hacer reaccionar 4-hidroxilasa de ácido pipecólico, células que contienen la enzima, un producto tratado de las células o un líquido de cultivo de las células, con el sustrato de reacción α -aminoácido en presencia de ácido 2-oxoglutarico e iones de hierro (II), se produce 4-hidroxiaminoácido.

- 20 En este caso, el α -aminoácido como sustrato de reacción no está limitado siempre que tenga un átomo de hidrógeno que pueda sustituirse con un grupo hidroxilo en la posición 4. Los ejemplos del α -aminoácido incluyen compuestos representados por la siguiente fórmula general (I). Los ejemplos del 4-hidroxiaminoácido incluyen compuestos representados por la siguiente fórmula general (II). Cada uno del α -aminoácido y el 4-hidroxiaminoácido es preferiblemente un isómero L o un isómero D, pero también puede ser una mezcla racémica.



- 25 En las fórmulas generales (I) y (II), cada uno de R¹, R² y R⁵ representa un átomo de hidrógeno o alquilo C₁-C₃, y cada uno de R³ y R⁴ representa un átomo de hidrógeno, alquilo C₁-C₃ o hidroxilo. Alternativamente, R² puede unirse a R⁵ o al átomo de nitrógeno del grupo amino para formar una estructura de anillo.

Los ejemplos específicos del tipo del α -aminoácido incluyen valina, leucina, isoleucina, prolina (en la que R² se une al átomo de nitrógeno del grupo amino para formar un anillo de cinco miembros), ácido α -aminobutírico, norvalina, norleucina, ácido pipecólico (en el que R² se une a R⁵ para formar un anillo de seis miembros), 3-hidroxiprolina, ácido 3-hidroxipipecólico y ácido 5-hidroxipipecólico. El α -aminoácido es especialmente preferible ácido pipecólico.

- 30 La reacción se lleva a cabo en un medio acuoso que contiene: el α -aminoácido; ácido 2-oxoglutarico; iones de hierro (II); y la 4-hidroxilasa de ácido pipecólico o las células que la contienen, el producto tratado de las células o el producto de cultivo de las células; o se lleva a cabo en una mezcla del medio acuoso y un disolvente orgánico.

- 35 Los ejemplos del medio acuoso incluyen agua y tampones. En este caso, los tampones no están limitados siempre que la actividad de la 4-hidroxilasa de ácido pipecólico no se inhiba de ese modo. Los ejemplos de los tampones incluyen tampón fosfato y tampón MES (ácido 2-morfolinoetanosulfónico). Los ejemplos del disolvente orgánico incluyen aquellos en los que el sustrato de reacción es altamente soluble, tales como metanol, etanol, 1-propanol, 2-propanol, 1-butanol, *tert*-butanol, acetona y dimetilsulfóxido. Otros ejemplos del disolvente orgánico incluyen acetato de etilo, acetato de butilo, tolueno, cloroformo y *n*-hexano, que son eficaces para la retirada de subproductos de reacción, y similares.

- 40 El sustrato de reacción, α -aminoácido, se usa habitualmente a una concentración de sustrato dentro del intervalo del 0,01% p/v al 90% p/v, preferiblemente del 0,1% p/v al 30% p/v. El sustrato de reacción puede añadirse de una vez cuando se inicia la reacción, pero se añade preferiblemente de manera continua o intermitente desde el punto de vista de reducir la influencia de la inhibición de sustrato de la enzima, si la hubiera, y aumentar la concentración del producto acumulado.

- 45 El número de moles del ácido 2-oxoglutarico requerido para la reacción es normalmente equivalente a, o mayor que, el del sustrato, preferiblemente equivalente a, o hasta 1,2 veces mayor que, el del sustrato. El ácido 2-oxoglutarico puede añadirse de una vez cuando se inicia la reacción, pero se añade preferiblemente de manera continua o intermitente desde el punto de vista de reducir la acción inhibitoria en la enzima, si la hubiera, y aumentar la concentración del producto acumulado. Alternativamente, puede añadirse un compuesto económico que puede

metabolizarse por el huésped, tal como glucosa o ácido L-glutámico, en lugar de ácido 2-oxoglutarico, para permitir el metabolismo del compuesto por el huésped, y puede usarse ácido 2-oxoglutarico producido durante este procedimiento para la reacción.

Los iones de hierro (II) requeridos para la reacción se usan dentro del intervalo de habitualmente 0,01 mmol/l a 100 mmol/l, preferiblemente de 0,1 mmol/l a 10 mmol/l. Los iones de hierro (II) se añaden habitualmente de una vez cuando se inicia la reacción. La adición adicional de iones de hierro (II) también es eficaz en casos en los que los iones de hierro (II) disminuyen durante la reacción debido a la oxidación a iones de hierro (III) o la formación de un precipitado. En casos en los que una cantidad suficiente de iones de hierro (II) ya están contenidos en la 4-hidroxilasa de ácido pipécólico, las células que contienen la enzima, el producto tratado de las células o el líquido de cultivo de las células, no se requiere la adición de iones de hierro (II).

La reacción se lleva a cabo a una temperatura de reacción de habitualmente 4°C a 60°C, preferiblemente de 10°C a 45°C, a un pH de habitualmente 3 a 11, preferiblemente de 5 a 8. El tiempo de reacción es habitualmente de aproximadamente 1 hora a aproximadamente 72 horas.

La cantidad de las células y/o el producto celular tratado que van a añadirse al líquido de reacción es de la siguiente manera. En casos de adición de las células, las células se añaden al líquido de reacción de modo que la concentración celular es habitualmente de aproximadamente el 0,1% p/v a aproximadamente el 50% p/v, preferiblemente del 1% p/v al 20% p/v, en cuanto al peso celular húmedo. En los casos de usar el producto celular tratado, se determina la actividad específica de la enzima en el producto celular tratado, y se añade el producto celular tratado de modo que se logra la concentración celular descrita anteriormente.

Después de la finalización de la reacción, el 4-hidroxiaminoácido producido por el método de la presente invención puede someterse a separación de células microbianas, proteínas y/o similares en el líquido de reacción mediante centrifugación, tratamiento de membrana y/o similar, y luego a purificación mediante una combinación apropiada de, por ejemplo, extracción con un disolvente orgánico tal como 1-butanol o *tert*-butanol, destilación, cromatografía en columna usando una resina de intercambio iónico, gel de sílice, o similar, cristalización en el punto isoelectrico, y/o cristalización con monoclóhidrato, diclorhidrato, sal de calcio, o similar.

Ejemplos

La presente invención se describe a continuación en más detalle por medio de ejemplos, pero la presente invención no se limita a los ejemplos.

Ejemplo 1

<Construcción de plásmido que expresa el gen de 4-hidroxilasa de ácido pipécólico>

(1) Clonación de genes

Basándose en una secuencia génica que codifica para una proteína hipotética FOXB_08233 derivada de *Fusarium oxysporum* Fo5176, se diseñaron y sintetizaron cebadores BOF1 y BOR1 para amplificación de genes homólogos de longitud completa para la proteína hipotética FOXB_08233. Las secuencias de nucleótidos de los cebadores se muestran en SEQ ID NO:13 y 14.

Se cultivó *Fusarium oxysporum* c8D durante la noche en un medio líquido de dextrosa de patata (fabricado por Becton, Dickinson and Company, Japón). A partir de las células microbianas obtenidas, se preparó ADN cromosómico usando un kit DNeasy Blood & Tissue (fabricado por QIAGEN).

Usando ADN cromosómico preparado de ese modo derivado de cada cepa microbiana como molde, y los oligonucleótidos de SEQ ID NO:13 y 14 como cebadores, se llevó a cabo reacción en cadena de la polimerasa (PCR) para amplificar un fragmento de ADN de aproximadamente 1 kpb. Se llevó a cabo la PCR usando ADN polimerasa Tks Gflex (fabricada por Takara Bio Inc.) según las condiciones descritas en la instrucción del fabricante. El programa de temperatura fue de la siguiente manera: se mantuvo la temperatura a 95°C durante 1 minuto, y luego se llevaron a cabo 35 ciclos de (98°C, 10 segundos; 56,5°C, 15 segundos; y 68°C, 40 segundos), seguido por mantener la temperatura a 72°C durante 3 minutos para completar la reacción. El resultado del análisis de la secuencia de ADN obtenida se muestra en SEQ ID NO:1, y la secuencia de aminoácidos codificada por esta secuencia de ADN se muestra en SEQ ID NO:2.

(2) Preparación del plásmido de expresión

Se digirió el fragmento de ADN obtenido en (1) con las enzimas de restricción *Bam*HI y *Hind*III, y se introdujo en un vector plasmídico pQE80L (fabricado por QIAGEN) digerido con las enzimas de restricción *Bam*HI y *Hind*III, usando un kit de conveniencia de ligamiento (fabricado por Nippon Gene Co., Ltd.). El plásmido obtenido se denomina a continuación en el presente documento pFoPA4H. Además, las secuencias de ADN de SEQ ID NO:3, 5, 7, 9 y 11 se sintetizaron mediante DNA2.0, Inc., y se insertaron en un plásmido de expresión pJexpress411 por el mismo fabricante. Los plásmidos obtenidos se denominan a continuación en el presente documento pAoPA4H, pPcPA4H,

pGzPA4H, pCgPA4H y pEnPA4H.

Ejemplo 2

<Evaluación de la actividad de 4-hidroxilasa de ácido pipecólico>

5 Los seis tipos de plásmidos obtenidos en el ejemplo 1, pFoPA4H, pAoPA4H, pPcPA4H, pGzPA4H, pCgPA4H y pEnPA4H, se usaron para la transformación de *E. coli* (*Escherichia coli*) Rosetta 2 (DE3) (fabricada por Merck Millipore Corporation) según un método común. Se cultivó cada *E. coli* recombinante obtenida con agitación a 30°C usando medio LB líquido complementado con kanamicina 50 mg/l, cloranfenicol 15 mg/l e IPTG 1 mmol/l (isopropil β-D-1-tiogalactopiranosido) durante 20 horas, y se recogieron las células bacterianas cultivadas.

10 Las células bacterianas obtenidas se sometieron a reacción a 20°C a pH 6 durante 21 horas en tampón MES 50 mmol/l (ácido 2-morfolinoetanosulfónico) complementado con ácido 2-oxoglutarico 15 mmol/l, ácido ascórbico 5 mmol/l, clorhidrato de tris(2-carboxietil)fosfina 5 mmol/l, FeSO₄·7H₂O 0,5 mmol/l, ácido cítrico 3 mmol/l y 10 mmol/l de cada compuesto de sustrato.

Se analizaron los líquidos de reacción después de la reacción mediante cromatografía de líquidos de alto rendimiento (HPLC) en las siguientes condiciones.

15 Columna: CHIRALPAK AD-3 (4,6 x 250 mm, fabricada por Daicel Corporation)

Eluyente: hexano : etanol : ácido trifluoroacético = 95 : 5 : 0,1

Velocidad de flujo: 0,8 ml/min.

Temperatura: 30°C

Detección: UV 210 nm

20 Ejemplo 3

<Identificación de productos de reacción>

Cada líquido de reacción obtenido en el ejemplo 2 se derivatizó usando el método AccQ·Tag de Waters, y se sometió a análisis de separación usando un cromatógrafo de líquidos de alto rendimiento-espectrómetro de masas (CL-EM) en las siguientes condiciones para medir productos de reacción de aminoácidos.

25 Columna: XBridge C18 5 μm (2,1 x 150 mm, fabricada por Waters Corporation)

Eluyente A: acetato de amonio (10 mmol/l, pH 5)

Eluyente B: metanol (de 0 a 0,5 min. (0% → 1%), de 0,5 a 18 min. (1% → 5%), de 18 a 19 min. (5% → 9%), de 19 a 29,5 min. (9% → 17%), de 29,5 a 40 min. (17% → 60%), de 40 a 43 min. (60%))

Velocidad de flujo: 0,3 ml/min.

30 Temperatura: 30°C

Detección: espectrómetro de masas

35 Para los líquidos de reacción obtenidos usando los seis tipos de *E. coli*, se muestran aminoácidos de sustrato para los que se encontró la aparición de un nuevo pico en el análisis de CL-EM en la tabla 1. Dado que se encontró que los pesos moleculares de estos nuevos productos de reacción eran 16 veces mayores en comparación con los pesos moleculares de sus sustratos, se piensa que todos los nuevos productos de reacción son compuestos producidos por la adición de un átomo de oxígeno a sus sustratos.

Tabla 1

Plásmido	Aminoácido de sustrato (la cantidad de aumento en el peso molecular del producto de reacción)
pFoPA4H	Ácido L-pipecólico (16), L-prolina (16), L-leucina (16), D-norleucina (16)
pAoPA4H	Ácido L-pipecólico (16), L-prolina (16), L-leucina (16), D-norleucina (16)
pPcPA4H	Ácido L-pipecólico (16), L-leucina (16)
pGzPA4H	Ácido L-pipecólico (16), L-leucina (16)
pCgPA4H	Ácido L-pipecólico (16), L-leucina (16)

pEnPA4H	Ácido L-pipecólico (16), ácido D-pipecólico (16), L-prolina (16), D-prolina (16), L-leucina (16), L-valina (16), D-valina (16), L-norleucina (16), D-norleucina (16), ácido (S)-2-aminobutírico, L-norvalina (16)
---------	---

(1) El resultado del análisis de CL-EM del líquido de reacción que contenía *E. coli* transformada con pFoPA4H y L-prolina se muestra en la figura 1. Dado que el pico del producto de reacción mostró el mismo tiempo de retención que el de una sustancia patrón para *trans*-4-hidroxi-L-prolina, se asumió que el producto de reacción era *trans*-4-hidroxi-L-prolina. Se obtuvo el mismo resultado en los casos en los que se usó *E. coli* transformada con pAoPA4H o pEnPA4H.

(2) El resultado del análisis de CL-EM del líquido de reacción que contenía *E. coli* transformada con pFoPA4H y L-leucina se muestra en la figura 2. Dado que el pico del producto de reacción mostró el mismo tiempo de retención que el de una sustancia patrón para 4-hidroxi-L-leucina, se asumió que el producto de reacción era *trans*-4-hidroxi-L-leucina. Se obtuvo el mismo resultado en los casos en los que se usó *E. coli* transformada con pAoPA4H, pPcPA4H, pGzPA4H, pCgPA4H o pEnPA4H.

(3) Se recogió un producto de reacción obtenido haciendo reaccionar ácido L-pipecólico con *E. coli* transformada con pFoPA4H mediante cromatografía de líquidos de alto rendimiento (HPLC) en las siguientes condiciones.

Columna: TSKgel Amide80 (7,8 x 300 mm, fabricada por Tosoh Corporation)

Eluyente: acetato de amonio (10 mmol/l, pH 5) : acetonitrilo = 15 : 85

Velocidad de flujo: 2,3 ml/min.

Temperatura: 40°C

Detección: UV 210 nm

Se recogió el eluato que contenía el producto de reacción, y se secó a presión reducida usando un evaporador centrífugo. Se suspendió el residuo resultante en agua pesada, y se sometió a medición del espectro de resonancia magnética.

Basándose en los valores de desplazamiento químico obtenidos como resultado de ¹H-RMN (figura 6) y ¹³C-RMN (figura 7), la correlación obtenida como resultado de los análisis de HH-COSY y CH-HMQC, e información de una bibliografía (Molnar, T. *et al.*, 2008, Bioorg. Med. Chem. Lett., 18, 6290), se identificó el producto de reacción como ácido *trans*-4-hidroxi-L-pipecólico.

El resultado del análisis de CL-EM del líquido de reacción que contenía *E. coli* transformada con pFoPA4H y ácido L-pipecólico se muestra en la figura 3. La elución del pico del producto de reacción, ácido *trans*-4-hidroxi-L-pipecólico, se produjo a los 10,9 minutos. Se obtuvo el mismo resultado en los casos en los que se usó *E. coli* transformada con pAoPA4H, pPcPA4H, pGzPA4H, pCgPA4H o pEnPA4H.

(4) El resultado del análisis de CL-EM del líquido de reacción que contenía *E. coli* transformada con pEnPA4H y D-prolina se muestra en la figura 4. Dado que el pico del producto de reacción mostró el mismo tiempo de retención que el de una sustancia patrón para *cis*-4-hidroxi-D-prolina, se asumió que el producto de reacción era *cis*-4-hidroxi-D-prolina.

(5) El resultado del análisis de CL-EM del líquido de reacción que contenía *E. coli* transformada con pEnPA4H y ácido (S)-2-aminobutírico se muestra en la figura 5. Dado que el pico del producto de reacción mostró el mismo tiempo de retención que el de una sustancia patrón para L-homoserina, se asumió que el producto de reacción era L-homoserina.

Ejemplo 4

<Evaluación de la actividad de 4-hidroxilasa de ácido pipecólico para hidroxilación a hidroxiaminoácido>

Usando extractos libres de células obtenidos de las células obtenidas en el ejemplo 2, se llevó a cabo la reacción a 20°C a pH 6 durante 21 horas en tampón MES 50 mmol/l complementado con ácido 2-oxoglutarico 15 mmol/l, ácido ascórbico 5 mmol/l, clorhidrato de tris(2-carboxietil)fosfina 5 mmol/l, FeSO₄·7H₂O 0,5 mmol/l, ácido cítrico 3 mmol/l y 10 mmol/l de cada compuesto de sustrato. Se analizaron los líquidos de reacción después de la reacción mediante cromatografía de líquidos de alto rendimiento (HPLC) en las condiciones descritas en el ejemplo 2.

Para líquidos de reacción que usan dos tipos de extractos libres de células, se muestran hidroxiaminoácidos de sustrato para los que se encontró la aparición de un nuevo pico en el análisis de CL-EM en la tabla 2. Dado que se encontró que los pesos moleculares de estos nuevos productos de reacción eran 16 veces mayores en comparación con los pesos moleculares de sus sustratos, se cree que todos los nuevos productos de reacción son dihidroxiaminoácidos producidos por la adición de un átomo de oxígeno a hidroxiaminoácidos.

Tabla 2

Plásmido	Aminoácido de sustrato (la cantidad de aumento en el peso molecular del producto de reacción)
pFoPA4H	Ácido (5S)-hidroxi-L-pipecólico (16)
pAoPA4H	Ácido (5S)-hidroxi-L-pipecólico (16)
pPcPA4H	Ácido (5S)-hidroxi-L-pipecólico (16)
pGzPA4H	Ácido (5S)-hidroxi-L-pipecólico (16)
pCgPA4H	Ácido (5S)-hidroxi-L-pipecólico (16)
pEnPA4H	Ácido (5S)-hidroxi-L-pipecólico (16), (3R)-hidroxi-L-prolina (16)

Ejemplo 5 <Construcción de plásmidos que expresan el gen de 4-hidroxilasa de ácido pipecólico>

A partir de los resultados de análisis de secuencia de ADN recogido directamente de muestras de suelo, se descubrieron SEQ ID NO:15 y 17, que podían ser genes para hidroxilasas de aminoácido. Las secuencias de aminoácidos codificadas por las secuencias de ADN se muestran en SEQ ID NO:16 y 18.

A partir de cada secuencia de ADN, se amplificó un fragmento de ADN de aproximadamente 1 kpb mediante PCR. Cada fragmento obtenido se introdujo en pQE80L (fabricado por QIAGEN) según el método descrito en el ejemplo 1. Los plásmidos obtenidos se denominan a continuación en el presente documento pVsPA4H y pBsPA4H, respectivamente.

Ejemplo 6

<Evaluación de la actividad de 4-hidroxilasa de ácido pipecólico>

Usando los dos tipos de plásmidos pVsPA4H y pBsPA4H obtenidos en el ejemplo 5, se llevó a cabo la transformación de *E. coli* (*Escherichia coli*) Rosetta 2 (DE3) (fabricada por Merck Millipore Corporation) según un método común. Cada *E. coli* recombinante obtenida se cultivó con agitación a 30°C usando medio LB líquido complementado con ampicilina 50 mg/l, cloranfenicol 15 mg/l e IPTG 1 mmol/l durante 20 horas, y se recogieron las células bacterianas cultivadas. Usando extractos libres de células obtenidos de las células recogidas, se llevó a cabo la reacción a 20°C a pH 6 durante 21 horas en tampón MES 50 mmol/l complementado con ácido 2-oxoglutarico 15 mmol/l, ácido ascórbico 5 mmol/l, clorhidrato de tris(2-carboxietil)fosfina 5 mmol/l, FeSO₄·7H₂O 0,5 mmol/l, ácido cítrico 3 mmol/l y 10 mmol/l de compuesto de sustrato (ácido L-pipecólico).

Se analizaron los líquidos de reacción después de la reacción mediante la cromatografía de líquidos de alto rendimiento (HPLC) mencionada anteriormente. Como resultado, se encontró un compuesto que tenía el mismo peso molecular que el ácido trans-4-hidroxi-L-pipecólico en ambos líquidos de reacción.

Lista de secuencias

<110> API Corporation

<120> 4-hidroxilasa de ácido pipecólico y su uso para la síntesis de 4-hidroxiaminoácidos

<130> AP346-14516

<150> JP2014-017716

<151> 31-01-2014

<160> 18

<170> PatentIn versión 3.5

<210> 1

<211> 1014

<212> ADN

<213> *Fusarium oxysporum* c8D

<220>

<221> CDS

ES 2 773 889 T3

<222> (1)..(1014)

<400> 1

atg gcc gcc ctc aac gca gac acc ctc gac atg tcc ctc ttc ttc ggc	48
Met Ala Ala Leu Asn Ala Asp Thr Leu Asp Met Ser Leu Phe Phe Gly	
1 5 10 15	
acc ccc tcc cag aag caa gac ttc tgc gac tcc ctc ctg cgc ctg ctc	96
Thr Pro Ser Gln Lys Gln Asp Phe Cys Asp Ser Leu Leu Arg Leu Leu	
20 25 30	
aaa aag cgc ggg ggc gtc aag ctg atc aac cac ccc atc ccc tcc acc	144
Lys Lys Arg Gly Gly Val Lys Leu Ile Asn His Pro Ile Pro Ser Thr	
35 40 45	
tcc atc cac gag ctc ttc gca cag aca aag cgc ttc ttc aac ctc ccc	192
Ser Ile His Glu Leu Phe Ala Gln Thr Lys Arg Phe Phe Asn Leu Pro	
50 55 60	
ctc gag aca aaa atg ctc gcc aag cac cct ccc cag gca aac ccg aac	240
Leu Glu Thr Lys Met Leu Ala Lys His Pro Pro Gln Ala Asn Pro Asn	
65 70 75 80	
cgc ggg tac tcc ttc gtc ggc cag gaa aac gtc gcc aac atc agt ggt	288
Arg Gly Tyr Ser Phe Val Gly Gln Glu Asn Val Ala Asn Ile Ser Gly	
85 90 95	
tac gaa aag ggt ctt gga ccg cta aag act cgt gac atc aag gag act	336
Tyr Glu Lys Gly Leu Gly Pro Leu Lys Thr Arg Asp Ile Lys Glu Thr	
100 105 110	
gtg gac ttt gga tcg gcg aat gat gag ctt gtt gat aac ctt tgg gtt	384
Val Asp Phe Gly Ser Ala Asn Asp Glu Leu Val Asp Asn Leu Trp Val	
115 120 125	
ccg gag gag gag ttg ccc ggt ttt agg agc ttt atg gag ggg ttt tat	432
Pro Glu Glu Glu Leu Pro Gly Phe Arg Ser Phe Met Glu Gly Phe Tyr	
130 135 140	

gag ttg gct ttc aag acg gag atg cag ctt ctt gag gct ctt gct att	480
Glu Leu Ala Phe Lys Thr Glu Met Gln Leu Leu Glu Ala Leu Ala Ile	
145 150 155 160	
gct ctt ggt gtt tct ccc gat cac ttg aag tcg ctg cat aac cgt gct	528
Ala Leu Gly Val Ser Pro Asp His Leu Lys Ser Leu His Asn Arg Ala	
165 170 175	
gag aat gag ttt cga atc ttg cac tac ccc gcc att cct gcc tca gaa	576
Glu Asn Glu Phe Arg Ile Leu His Tyr Pro Ala Ile Pro Ala Ser Glu	
180 185 190	
ctc gca gac ggt aca gca acc cgc atc gca gag cac aca gat ttc gga	624
Leu Ala Asp Gly Thr Ala Thr Arg Ile Ala Glu His Thr Asp Phe Gly	
195 200 205	
acc atc acc atg ctt ttc caa gac tct gtc ggt ggc ctt cag gtc gag	672
Thr Ile Thr Met Leu Phe Gln Asp Ser Val Gly Gly Leu Gln Val Glu	
210 215 220	
gac caa gag aac ctc gga aca ttc aac aac gtc gaa tca gct tcc ccc	720
Asp Gln Glu Asn Leu Gly Thr Phe Asn Asn Val Glu Ser Ala Ser Pro	
225 230 235 240	
aca gac atc atc ctc aac atc ggt gat tct ctg cag agg ttg acg aat	768
Thr Asp Ile Ile Leu Asn Ile Gly Asp Ser Leu Gln Arg Leu Thr Asn	
245 250 255	
gac act ttc aag gct gcg tgt cat cgc gtt aca tac cct cct tct atc	816
Asp Thr Phe Lys Ala Ala Cys His Arg Val Thr Tyr Pro Pro Ser Ile	
260 265 270	
aaa gct ggt gat ggc gag cag gtc att cct gag aga tac tct atc gct	864
Lys Ala Gly Asp Gly Glu Gln Val Ile Pro Glu Arg Tyr Ser Ile Ala	
275 280 285	
tac ttt gct aag cct aac cga agc gct tcg ttg ttt ccg ctg aag gag	912
Tyr Phe Ala Lys Pro Asn Arg Ser Ala Ser Leu Phe Pro Leu Lys Glu	
290 295 300	
ttt att gag gag ggc gtt cct tgt aaa tat gag gat gta act gct tgg	960
Phe Ile Glu Glu Gly Val Pro Cys Lys Tyr Glu Asp Val Thr Ala Trp	
305 310 315 320	
gag tgg aat aac cgc cgt att gag aag ttg ttt tct gct gaa gct aag	1008
Glu Trp Asn Asn Arg Arg Ile Glu Lys Leu Phe Ser Ala Glu Ala Lys	
325 330 335	
gct taa	1014
Ala	

<210> 2

<211> 337

<212> PRT

5 <213> Fusarium oxysporum c8D

<400> 2

Met Ala Ala Leu Asn Ala Asp Thr Leu Asp Met Ser Leu Phe Phe Gly
1 5 10 15
Thr Pro Ser Gln Lys Gln Asp Phe Cys Asp Ser Leu Leu Arg Leu Leu
20 25 30
Lys Lys Arg Gly Gly Val Lys Leu Ile Asn His Pro Ile Pro Ser Thr
35 40 45
Ser Ile His Glu Leu Phe Ala Gln Thr Lys Arg Phe Phe Asn Leu Pro
50 55 60
Leu Glu Thr Lys Met Leu Ala Lys His Pro Pro Gln Ala Asn Pro Asn
65 70 75 80
Arg Gly Tyr Ser Phe Val Gly Gln Glu Asn Val Ala Asn Ile Ser Gly
85 90 95
Tyr Glu Lys Gly Leu Gly Pro Leu Lys Thr Arg Asp Ile Lys Glu Thr
100 105 110
Val Asp Phe Gly Ser Ala Asn Asp Glu Leu Val Asp Asn Leu Trp Val
115 120 125
Pro Glu Glu Glu Leu Pro Gly Phe Arg Ser Phe Met Glu Gly Phe Tyr
130 135 140
Glu Leu Ala Phe Lys Thr Glu Met Gln Leu Leu Glu Ala Leu Ala Ile
145 150 155 160
Ala Leu Gly Val Ser Pro Asp His Leu Lys Ser Leu His Asn Arg Ala
165 170 175
Glu Asn Glu Phe Arg Ile Leu His Tyr Pro Ala Ile Pro Ala Ser Glu
180 185 190
Leu Ala Asp Gly Thr Ala Thr Arg Ile Ala Glu His Thr Asp Phe Gly
195 200 205
Thr Ile Thr Met Leu Phe Gln Asp Ser Val Gly Gly Leu Gln Val Glu
210 215 220
Asp Gln Glu Asn Leu Gly Thr Phe Asn Asn Val Glu Ser Ala Ser Pro
225 230 235 240
Thr Asp Ile Ile Leu Asn Ile Gly Asp Ser Leu Gln Arg Leu Thr Asn
245 250 255
Asp Thr Phe Lys Ala Ala Cys His Arg Val Thr Tyr Pro Pro Ser Ile
260 265 270
Lys Ala Gly Asp Gly Glu Gln Val Ile Pro Glu Arg Tyr Ser Ile Ala
275 280 285
Tyr Phe Ala Lys Pro Asn Arg Ser Ala Ser Leu Phe Pro Leu Lys Glu
290 295 300
Phe Ile Glu Glu Gly Val Pro Cys Lys Tyr Glu Asp Val Thr Ala Trp
305 310 315 320
Glu Trp Asn Asn Arg Arg Ile Glu Lys Leu Phe Ser Ala Glu Ala Lys
325 330 335

Ala

<210> 3

<211> 987

5 <212> ADN

<213> Aspergillus oryzae RIB40

<220>

<221> CDS

<222> (1)..(987)

10 <400> 3

ES 2 773 889 T3

atg agc gaa gca ctg gac ctg agc ctg ctg aag ggt acg cct gaa caa	48
Met Ser Glu Ala Leu Asp Leu Ser Leu Leu Lys Gly Thr Pro Glu Gln	
1 5 10 15	
cgt gag cag atc agc gcc gag ctg ttg cac gcg ttg aaa acc cgt ggt	96
Arg Glu Gln Ile Ser Ala Glu Leu Leu His Ala Leu Lys Thr Arg Gly	
20 25 30	
ggc gtc aaa ctg aag aac cat ggt ctg ccg gat aag ctg gtg cac gag	144
Gly Val Lys Leu Lys Asn His Gly Leu Pro Asp Lys Leu Val His Glu	
35 40 45	

ES 2 773 889 T3

ctg ttc gat tgg act cgc aag ttc ttt gct ctg ccg cat gag gac aag Leu Phe Asp Trp Thr Arg Lys Phe Phe Ala Leu Pro His Glu Asp Lys 50 55 60	192
atg ctg gca aag cat ccg ccg cag gcg aat ccg aat cgt ggc tat tgt Met Leu Ala Lys His Pro Pro Gln Ala Asn Pro Asn Arg Gly Tyr Cys 65 70 75 80	240
tac gtt ggt cag gag tcg atc tct agc att tcc ggt tat gaa aaa ggc Tyr Val Gly Gln Glu Ser Ile Ser Ser Ile Ser Gly Tyr Glu Lys Gly 85 90 95	288
tta ccg caa ggt cgc ttt gtg cgc gac att aag gaa acg gtg gac ttc Leu Pro Gln Gly Arg Phe Val Arg Asp Ile Lys Glu Thr Val Asp Phe 100 105 110	336
ggc agc cca cgc gac gaa ctg gtc gat aac att tgg gtt ccg gag gaa Gly Ser Pro Arg Asp Glu Leu Val Asp Asn Ile Trp Val Pro Glu Glu 115 120 125	384
aaa ctg ccg ggt ttt cgt aag ttt atc gag gat ttc tac gaa acg tgc Lys Leu Pro Gly Phe Arg Lys Phe Ile Glu Asp Phe Tyr Glu Thr Cys 130 135 140	432
ttt aaa ctc gaa ctg gag atc ctg gcc gcc ctg gcg cgt gca ctg ggc Phe Lys Leu Glu Leu Glu Ile Leu Ala Ala Leu Ala Arg Ala Leu Gly 145 150 155 160	480
gtg gat gaa gat cac atg gtg tcc ctg cac aac aaa gcg gaa aac gag Val Asp Glu Asp His Met Val Ser Leu His Asn Lys Ala Glu Asn Glu 165 170 175	528
ttt cgt att ctg cac tat ccg gaa gtt ccg gca agc gag ctg gcg gat Phe Arg Ile Leu His Tyr Pro Glu Val Pro Ala Ser Glu Leu Ala Asp 180 185 190	576
ggt acc gct acc cgt atc gcg gag cat acc gat ttc ggt agc att acc Gly Thr Ala Thr Arg Ile Ala Glu His Thr Asp Phe Gly Ser Ile Thr 195 200 205	624
atg ctg ttt cag gac agc gtt ggt ggt ttg caa gtc gag gac caa caa Met Leu Phe Gln Asp Ser Val Gly Gly Leu Gln Val Glu Asp Gln Gln 210 215 220	672
aat ccg ggc gtg ttt cgt ggc atc gag agc gcg gac aaa acc gag atc Asn Pro Gly Val Phe Arg Gly Ile Glu Ser Ala Asp Lys Thr Glu Ile 225 230 235 240	720
att ctg aat att ggt gat agc atg cag cgt ctg acc aac gac acg ttc Ile Leu Asn Ile Gly Asp Ser Met Gln Arg Leu Thr Asn Asp Thr Phe 245 250 255	768
cgt gca gcg tgt cac cgt gtt acg tac ccg ccg agc gtc aaa gtt ggc Arg Ala Ala Cys His Arg Val Thr Tyr Pro Pro Ser Val Lys Val Gly 260 265 270	816
tct gaa gca gtc atc ccg gag cgc tat tcc atc gcc tac ttc gct aaa Ser Glu Ala Val Ile Pro Glu Arg Tyr Ser Ile Ala Tyr Phe Ala Lys 275 280 285	864
cct aac cgt agc gcg agc ttg ttt ccg ttc aaa gag ttc att acc cca Pro Asn Arg Ser Ala Ser Leu Phe Pro Phe Lys Glu Phe Ile Thr Pro 912	

290	295	300	
agc acc ccg tgc cgt tac gag gat atc aat gcg tgg gac ttc cag aat			960
Ser Thr Pro Cys Arg Tyr Glu Asp Ile Asn Ala Trp Asp Phe Gln Asn			
305	310	315	320
ctg cgc att agc cgc ctg ttc aag taa			987
Leu Arg Ile Ser Arg Leu Phe Lys			
	325		
<210> 4			
<211> 328			
<212> PRT			
5	<213> Aspergillus oryzae RIB40		
<400> 4			
Met Ser Glu Ala Leu Asp Leu Ser Leu Leu Lys Gly Thr Pro Glu Gln			
1	5	10	15
Arg Glu Gln Ile Ser Ala Glu Leu Leu His Ala Leu Lys Thr Arg Gly			
	20	25	30
Gly Val Lys Leu Lys Asn His Gly Leu Pro Asp Lys Leu Val His Glu			
	35	40	45
Leu Phe Asp Trp Thr Arg Lys Phe Phe Ala Leu Pro His Glu Asp Lys			
	50	55	60
Met Leu Ala Lys His Pro Pro Gln Ala Asn Pro Asn Arg Gly Tyr Cys			
65	70	75	80
Tyr Val Gly Gln Glu Ser Ile Ser Ser Ile Ser Gly Tyr Glu Lys Gly			
	85	90	95
Leu Pro Gln Gly Arg Phe Val Arg Asp Ile Lys Glu Thr Val Asp Phe			
	100	105	110
Gly Ser Pro Arg Asp Glu Leu Val Asp Asn Ile Trp Val Pro Glu Glu			
	115	120	125
Lys Leu Pro Gly Phe Arg Lys Phe Ile Glu Asp Phe Tyr Glu Thr Cys			
	130	135	140
Phe Lys Leu Glu Leu Glu Ile Leu Ala Ala Leu Ala Arg Ala Leu Gly			
145	150	155	160
Val Asp Glu Asp His Met Val Ser Leu His Asn Lys Ala Glu Asn Glu			
	165	170	175
Phe Arg Ile Leu His Tyr Pro Glu Val Pro Ala Ser Glu Leu Ala Asp			
	180	185	190
Gly Thr Ala Thr Arg Ile Ala Glu His Thr Asp Phe Gly Ser Ile Thr			
	195	200	205
Met Leu Phe Gln Asp Ser Val Gly Gly Leu Gln Val Glu Asp Gln Gln			
	210	215	220
Asn Pro Gly Val Phe Arg Gly Ile Glu Ser Ala Asp Lys Thr Glu Ile			
225	230	235	240
Ile Leu Asn Ile Gly Asp Ser Met Gln Arg Leu Thr Asn Asp Thr Phe			
	245	250	255
Arg Ala Ala Cys His Arg Val Thr Tyr Pro Pro Ser Val Lys Val Gly			
	260	265	270
Ser Glu Ala Val Ile Pro Glu Arg Tyr Ser Ile Ala Tyr Phe Ala Lys			
	275	280	285
Pro Asn Arg Ser Ala Ser Leu Phe Pro Phe Lys Glu Phe Ile Thr Pro			
	290	295	300
Ser Thr Pro Cys Arg Tyr Glu Asp Ile Asn Ala Trp Asp Phe Gln Asn			
305	310	315	320
Leu Arg Ile Ser Arg Leu Phe Lys			
	325		
<210> 5			

<211> 987

<212> ADN

<213> *Penicillium chrysogenum* Wisconsin 54-1255

<220>

5 <221> CDS

<222> (1)..(987)

<400> 5

ES 2 773 889 T3

atg agc caa gcc ctt gat ttg tct ctg ctg aaa ggt agc ccg gaa gaa	48
Met Ser Gln Ala Leu Asp Leu Ser Leu Leu Lys Gly Ser Pro Glu Glu	
1 5 10 15	
cg t gat gaa gtt tcc gca gcc ctg ctg gac acc ttg aaa acc cgt ggt	96
Arg Asp Glu Val Ser Ala Ala Leu Leu Asp Thr Leu Lys Thr Arg Gly	
20 25 30	
gtt gcg aaa ctg aag aac cac ggt ctg ccg gag gat ctg att gcc gag	144
Val Ala Lys Leu Lys Asn His Gly Leu Pro Glu Asp Leu Ile Ala Glu	
35 40 45	
atg ttc gat tac acg cgc cgt ttc ttc agc ttg tcc ctg gag gac aag	192
Met Phe Asp Tyr Thr Arg Arg Phe Phe Ser Leu Ser Leu Glu Asp Lys	
50 55 60	
atg acg gcg aag cac ccg cca gag gcg aac ccg aac cgc ggc tat agc	240
Met Thr Ala Lys His Pro Pro Glu Ala Asn Pro Asn Arg Gly Tyr Ser	
65 70 75 80	
tac gtt ggc cag gag agc gtg tct agc atc tca ggt tat gaa aag ggt	288
Tyr Val Gly Gln Glu Ser Val Ser Ser Ile Ser Gly Tyr Glu Lys Gly	
85 90 95	
tta ccg caa ggt aag act atc cgt gac att aaa gaa acc ctg gac atg	336
Leu Pro Gln Gly Lys Thr Ile Arg Asp Ile Lys Glu Thr Leu Asp Met	
100 105 110	
ggc agc ccg cac gac tcg ctg gtc gat aac att tgg gtg gca gaa gag	384
Gly Ser Pro His Asp Ser Leu Val Asp Asn Ile Trp Val Ala Glu Glu	
115 120 125	
aag ctg ccg ggt ttt cgc aag ttc atg gaa gat ttc tac gag agc tgt	432
Lys Leu Pro Gly Phe Arg Lys Phe Met Glu Asp Phe Tyr Glu Ser Cys	
130 135 140	
ttc aaa gtg gaa ctg gag att ctg gat gcc ctg gcg cag gcg ctg gaa	480
Phe Lys Val Glu Leu Glu Ile Leu Asp Ala Leu Ala Gln Ala Leu Glu	
145 150 155 160	
atc agc gct cct gat ctg cgt ctg ttg cac aat aaa gcg gag aat gag	528
Ile Ser Ala Pro Asp Leu Arg Leu Leu His Asn Lys Ala Glu Asn Glu	
165 170 175	
ttc cgc ctg ctg cat tac cct gct gtt ccg gct agc gca ctg gag gat	576
Phe Arg Leu Leu His Tyr Pro Ala Val Pro Ala Ser Ala Leu Glu Asp	
180 185 190	
ggc acc gca acg cgt atc gcg gag cac acg gac ttt ggc acc att acc	624
Gly Thr Ala Thr Arg Ile Ala Glu His Thr Asp Phe Gly Thr Ile Thr	
195 200 205	
atg ttg ttt cag gac agc acc ggc ggt ctg caa gtg gag gac caa acc	672
Met Leu Phe Gln Asp Ser Thr Gly Gly Leu Gln Val Glu Asp Gln Thr	

ES 2 773 889 T3

210	215	220	
aat ctg ggt acc ttt cat gac gtc gtt agc ggt ggc aaa agc gaa atc			720
Asn Leu Gly Thr Phe His Asp Val Val Ser Gly Gly Lys Ser Glu Ile			
225	230	235	240
att ctg aac atc ggc gat agc ctg cag cgt ctg acc aat gac acg ttc			768
Ile Leu Asn Ile Gly Asp Ser Leu Gln Arg Leu Thr Asn Asp Thr Phe			
	245	250	255
atg gcg gcg tgc cat cgt gtt acg tat ccg ccg acc gtc aaa gtg ggt			816
Met Ala Ala Cys His Arg Val Thr Tyr Pro Pro Thr Val Lys Val Gly			
	260	265	270
tct gac gtc gtg att ccg gag cgc tac tcc gtc gcc tat ttt gcg aag			864
Ser Asp Val Val Ile Pro Glu Arg Tyr Ser Val Ala Tyr Phe Ala Lys			
	275	280	285
ccg aat cgt atc gcg agc ctg ttt ccg ctg aaa aag ttt atc acc cca			912
Pro Asn Arg Ile Ala Ser Leu Phe Pro Leu Lys Lys Phe Ile Thr Pro			
	290	295	300
gca acg ccg tgc aag tat gaa gat att act gca tgg gac tac aac aat			960
Ala Thr Pro Cys Lys Tyr Glu Asp Ile Thr Ala Trp Asp Tyr Asn Asn			
	305	310	315
ctc cgt atc gcg aaa ctg ttc agc taa			987
Leu Arg Ile Ala Lys Leu Phe Ser			
	325		
<210> 6			
<211> 328			
<212> PRT			
5 <213> Penicillium chrysogenum Wisconsin 54-1255			
<400> 6			

```

Met Ser Gln Ala Leu Asp Leu Ser Leu Leu Lys Gly Ser Pro Glu Glu
1      5      10      15
Arg Asp Glu Val Ser Ala Ala Leu Leu Asp Thr Leu Lys Thr Arg Gly
20      25      30
Val Ala Lys Leu Lys Asn His Gly Leu Pro Glu Asp Leu Ile Ala Glu
35      40      45
Met Phe Asp Tyr Thr Arg Arg Phe Phe Ser Leu Ser Leu Glu Asp Lys
50      55      60
Met Thr Ala Lys His Pro Pro Glu Ala Asn Pro Asn Arg Gly Tyr Ser
65      70      75      80
Tyr Val Gly Gln Glu Ser Val Ser Ser Ile Ser Gly Tyr Glu Lys Gly
85      90      95
Leu Pro Gln Gly Lys Thr Ile Arg Asp Ile Lys Glu Thr Leu Asp Met
100     105     110
Gly Ser Pro His Asp Ser Leu Val Asp Asn Ile Trp Val Ala Glu Glu
115     120     125
Lys Leu Pro Gly Phe Arg Lys Phe Met Glu Asp Phe Tyr Glu Ser Cys
130     135     140
Phe Lys Val Glu Leu Glu Ile Leu Asp Ala Leu Ala Gln Ala Leu Glu
145     150     155     160
Ile Ser Ala Pro Asp Leu Arg Leu Leu His Asn Lys Ala Glu Asn Glu
165     170     175
Phe Arg Leu Leu His Tyr Pro Ala Val Pro Ala Ser Ala Leu Glu Asp
180     185     190
Gly Thr Ala Thr Arg Ile Ala Glu His Thr Asp Phe Gly Thr Ile Thr
195     200     205
Met Leu Phe Gln Asp Ser Thr Gly Gly Leu Gln Val Glu Asp Gln Thr
210     215     220
Asn Leu Gly Thr Phe His Asp Val Val Ser Gly Gly Lys Ser Glu Ile
225     230     235     240
Ile Leu Asn Ile Gly Asp Ser Leu Gln Arg Leu Thr Asn Asp Thr Phe
245     250     255
Met Ala Ala Cys His Arg Val Thr Tyr Pro Pro Thr Val Lys Val Gly
260     265     270
Ser Asp Val Val Ile Pro Glu Arg Tyr Ser Val Ala Tyr Phe Ala Lys
275     280     285
Pro Asn Arg Ile Ala Ser Leu Phe Pro Leu Lys Lys Phe Ile Thr Pro
290     295     300
Ala Thr Pro Cys Lys Tyr Glu Asp Ile Thr Ala Trp Asp Tyr Asn Asn
305     310     315     320
Leu Arg Ile Ala Lys Leu Phe Ser
325

```

<210> 7

<211> 1014

5 <212> ADN

<213> Gibberella zeae PH-1

<220>

<221> CDS

<222> (1)..(1014)

10 <400> 7

ES 2 773 889 T3

atg gct gcc ctg aac gtt gac acc ctg gac atg tcc ctt ttc ttc ggt	48
Met Ala Ala Leu Asn Val Asp Thr Leu Asp Met Ser Leu Phe Phe Gly	
1 5 10 15	
acc cct gat cag aag aaa gat ttc tgc gac tcg ctg ctg cgc ctg ctg	96
Thr Pro Asp Gln Lys Lys Asp Phe Cys Asp Ser Leu Leu Arg Leu Leu	
20 25 30	
aaa aag cgt ggc ggt gtg aag ctc gtt aac cac tct att ccg agc gag	144
Lys Lys Arg Gly Gly Val Lys Leu Val Asn His Ser Ile Pro Ser Glu	
35 40 45	
gat atc cat gag ctg ttt gcc caa acc aaa cgc ttc ttt gaa ctg ccg	192
Asp Ile His Glu Leu Phe Ala Gln Thr Lys Arg Phe Phe Glu Leu Pro	
50 55 60	
ctg gaa acg aaa atg att gcg aaa cat ccg cca cag gcg aat ccg aac	240
Leu Glu Thr Lys Met Ile Ala Lys His Pro Pro Gln Ala Asn Pro Asn	
65 70 75 80	
cgt ggt tac agc ttc gtg ggc cag gaa aac gtg gcg aat atc tcc ggt	288
Arg Gly Tyr Ser Phe Val Gly Gln Glu Asn Val Ala Asn Ile Ser Gly	
85 90 95	
tac gag aag ggt ctg ggc cca caa acc acg cgt gac att aaa gag act	336
Tyr Glu Lys Gly Leu Gly Pro Gln Thr Thr Arg Asp Ile Lys Glu Thr	
100 105 110	
gtt gat ttc ggt agc gcg acg gac gag ctg gtc gac aat atc tgg gtt	384
Val Asp Phe Gly Ser Ala Thr Asp Glu Leu Val Asp Asn Ile Trp Val	
115 120 125	
ccg gag gac aaa ctg ccg ggt ttt cgc aag ttc atg gaa ggt ttc tac	432
Pro Glu Asp Lys Leu Pro Gly Phe Arg Lys Phe Met Glu Gly Phe Tyr	

ES 2 773 889 T3

130	135	140	
gaa aag gcg ttt aag acc gaa atg cag ctg ttg gaa gct ttg gct atc Glu Lys Ala Phe Lys Thr Glu Met Gln Leu Leu Glu Ala Leu Ala Ile 145 150 155 160			480
gca ctg ggc gtt tca gca gac cac ctg aaa agc atc cac aat cgt gct Ala Leu Gly Val Ser Ala Asp His Leu Lys Ser Ile His Asn Arg Ala 165 170 175			528
gaa aac gag ttt cgt atc ttg cac tat ccg gca gtg ccg gcg tcc gag Glu Asn Glu Phe Arg Ile Leu His Tyr Pro Ala Val Pro Ala Ser Glu 180 185 190			576
ctg gcg gat ggc act gca acg cgc atc gca gag cac acc gat ttc ggt Leu Ala Asp Gly Thr Ala Thr Arg Ile Ala Glu His Thr Asp Phe Gly 195 200 205			624
acc att acg atg ctg ttc cag gac agc gtc ggt ggc tta cag gtc gaa Thr Ile Thr Met Leu Phe Gln Asp Ser Val Gly Gly Leu Gln Val Glu 210 215 220			672
gat caa gag aat ttg ggc cat ttc aac aac gtg gag agc gcg gca ccg Asp Gln Glu Asn Leu Gly His Phe Asn Asn Val Glu Ser Ala Ala Pro 225 230 235 240			720
acg gac att atc ctg aat att ggc gac agc ctg caa cgt ctg acc aat Thr Asp Ile Ile Leu Asn Ile Gly Asp Ser Leu Gln Arg Leu Thr Asn 245 250 255			768
gat acc ttt aaa gcc gcg tgt cat cgt gtc acc tac cct ccg agc att Asp Thr Phe Lys Ala Ala Cys His Arg Val Thr Tyr Pro Pro Ser Ile 260 265 270			816
aag gcg agc gac ggt gaa caa gtt att ccg gag cgc tat agc atc gcc Lys Ala Ser Asp Gly Glu Gln Val Ile Pro Glu Arg Tyr Ser Ile Ala 275 280 285			864
tat ttt gca aag ccg aac cgt tct gcc agc ctg ttc ccg ctg aaa gag Tyr Phe Ala Lys Pro Asn Arg Ser Ala Ser Leu Phe Pro Leu Lys Glu 290 295 300			912
ttt att gag gaa ggt atg ccg tgc aag tac gaa gat gtc acc gcg tgg Phe Ile Glu Glu Gly Met Pro Cys Lys Tyr Glu Asp Val Thr Ala Trp 305 310 315 320			960
gag tgg aat aac cgt cgt atc gag aag ctg ttt agc agc gat gcg aaa Glu Trp Asn Asn Arg Arg Ile Glu Lys Leu Phe Ser Ser Asp Ala Lys 325 330 335			1008
gcg taa Ala			1014
<210> 8			
<211> 337			
<212> PRT			
5 <213> Gibberella zeae PH-1			
<400> 8			

Met Ala Ala Leu Asn Val Asp Thr Leu Asp Met Ser Leu Phe Phe Gly
1 5 10 15
Thr Pro Asp Gln Lys Lys Asp Phe Cys Asp Ser Leu Leu Arg Leu Leu
20 25 30
Lys Lys Arg Gly Gly Val Lys Leu Val Asn His Ser Ile Pro Ser Glu
35 40 45
Asp Ile His Glu Leu Phe Ala Gln Thr Lys Arg Phe Phe Glu Leu Pro
50 55 60
Leu Glu Thr Lys Met Ile Ala Lys His Pro Pro Gln Ala Asn Pro Asn
65 70 75 80
Arg Gly Tyr Ser Phe Val Gly Gln Glu Asn Val Ala Asn Ile Ser Gly
85 90 95
Tyr Glu Lys Gly Leu Gly Pro Gln Thr Thr Arg Asp Ile Lys Glu Thr
100 105 110
Val Asp Phe Gly Ser Ala Thr Asp Glu Leu Val Asp Asn Ile Trp Val
115 120 125
Pro Glu Asp Lys Leu Pro Gly Phe Arg Lys Phe Met Glu Gly Phe Tyr
130 135 140
Glu Lys Ala Phe Lys Thr Glu Met Gln Leu Leu Glu Ala Leu Ala Ile
145 150 155 160
Ala Leu Gly Val Ser Ala Asp His Leu Lys Ser Ile His Asn Arg Ala
165 170 175
Glu Asn Glu Phe Arg Ile Leu His Tyr Pro Ala Val Pro Ala Ser Glu
180 185 190
Leu Ala Asp Gly Thr Ala Thr Arg Ile Ala Glu His Thr Asp Phe Gly
195 200 205
Thr Ile Thr Met Leu Phe Gln Asp Ser Val Gly Gly Leu Gln Val Glu
210 215 220
Asp Gln Glu Asn Leu Gly His Phe Asn Asn Val Glu Ser Ala Ala Pro
225 230 235 240
Thr Asp Ile Ile Leu Asn Ile Gly Asp Ser Leu Gln Arg Leu Thr Asn
245 250 255
Asp Thr Phe Lys Ala Ala Cys His Arg Val Thr Tyr Pro Pro Ser Ile
260 265 270
Lys Ala Ser Asp Gly Glu Gln Val Ile Pro Glu Arg Tyr Ser Ile Ala
275 280 285
Tyr Phe Ala Lys Pro Asn Arg Ser Ala Ser Leu Phe Pro Leu Lys Glu
290 295 300
Phe Ile Glu Glu Gly Met Pro Cys Lys Tyr Glu Asp Val Thr Ala Trp
305 310 315 320
Glu Trp Asn Asn Arg Arg Ile Glu Lys Leu Phe Ser Ser Asp Ala Lys
325 330 335

Ala

<210> 9

<211> 993

5 <212> ADN

<213> Colletotrichum gloeosporioides Nara gc5

<220>

<221> CDS

<222> (1)..(993)

10 <400> 9

ES 2 773 889 T3

atg	agc	gat	aaa	ggg	gca	acg	ttg	gac	atg	agc	ctg	ttt	ttc	ggg	acg	48
Met	Ser	Asp	Lys	Gly	Ala	Thr	Leu	Asp	Met	Ser	Leu	Phe	Phe	Gly	Thr	
1				5					10					15		
cca	gag	gaa	aaa	ggc	tat	ttc	tgt	acc	gaa	ctg	ttg	cgt	ctg	ctg	aag	96
Pro	Glu	Glu	Lys	Gly	Tyr	Phe	Cys	Thr	Glu	Leu	Leu	Arg	Leu	Leu	Lys	
			20					25					30			

ES 2 773 889 T3

ttg cgt ggt ggt gtt aag att cag aat cac tct atc ccg gac gaa gat	144
Leu Arg Gly Gly Val Lys Ile Gln Asn His Ser Ile Pro Asp Glu Asp	
35 40 45	
atc cac aag ctg ttc gac atg agc cgt aag ttt ttc gcc tta ccg ctg	192
Ile His Lys Leu Phe Asp Met Ser Arg Lys Phe Phe Ala Leu Pro Leu	
50 55 60	
gag act aaa atg gaa gcg aaa cat ccg ccg caa gct aat ccg aac cgc	240
Glu Thr Lys Met Glu Ala Lys His Pro Pro Gln Ala Asn Pro Asn Arg	
65 70 75 80	
ggc tac agc ttt atc ggc cag gag aat gtc gcc aac atc agc ggt tac	288
Gly Tyr Ser Phe Ile Gly Gln Glu Asn Val Ala Asn Ile Ser Gly Tyr	
85 90 95	
gag aaa ggc ctg ggc cct ggt aag acc cgt gat atc aaa gaa acc ctg	336
Glu Lys Gly Leu Gly Pro Gly Lys Thr Arg Asp Ile Lys Glu Thr Leu	
100 105 110	
gac atg ggt tcc gca cag gat gat ttg gtc gac aac ctg tgg att ccg	384
Asp Met Gly Ser Ala Gln Asp Asp Leu Val Asp Asn Leu Trp Ile Pro	
115 120 125	
gag gaa agc ctg ccg ggt ttt cgt ggc ttt atg gaa tcc ttc tac gaa	432
Glu Glu Ser Leu Pro Gly Phe Arg Gly Phe Met Glu Ser Phe Tyr Glu	
130 135 140	
aaa gca ttc aag acc gaa atg caa ctc ctg agc gcg ttg gca att gcg	480
Lys Ala Phe Lys Thr Glu Met Gln Leu Leu Ser Ala Leu Ala Ile Ala	
145 150 155 160	
ctg ggt gtt tcc gag gac cac atg aaa acg ctg cac aat cgc gcg gag	528
Leu Gly Val Ser Glu Asp His Met Lys Thr Leu His Asn Arg Ala Glu	
165 170 175	
aac gag ttc cgt ctg ctg cat tat ccg gct atc ccg gcg acc gag ctg	576
Asn Glu Phe Arg Leu Leu His Tyr Pro Ala Ile Pro Ala Thr Glu Leu	
180 185 190	
gca gat ggt acc gcg acg cgc att gcc gaa cac acc gat ttt ggt acg	624
Ala Asp Gly Thr Ala Thr Arg Ile Ala Glu His Thr Asp Phe Gly Thr	
195 200 205	
att acc atg ctg ttt caa gac agc gtc ggc ggt ctg caa gtg gag gac	672
Ile Thr Met Leu Phe Gln Asp Ser Val Gly Gly Leu Gln Val Glu Asp	
210 215 220	
caa acc cag ccg ggt gtg ttc cgt agc gtg gag agc gaa aag ccg acc	720
Gln Thr Gln Pro Gly Val Phe Arg Ser Val Glu Ser Glu Lys Pro Thr	
225 230 235 240	
gat atc att ctg aat atc ggc gac agc ctg cag cgc ctg acc aat gac	768
Asp Ile Ile Leu Asn Ile Gly Asp Ser Leu Gln Arg Leu Thr Asn Asp	
245 250 255	
acg ttc aga gcg gcg tgc cac cgt gtc acg tac ccg cca gct att aag	816
Thr Phe Arg Ala Ala Cys His Arg Val Thr Tyr Pro Pro Ala Ile Lys	
260 265 270	
gcg cgt gat aac gtg cag atc ccg gag cgc tat agc att gca tac ttt	864
Ala Arg Asp Asn Val Gln Ile Pro Glu Arg Tyr Ser Ile Ala Tyr Phe	
275 280 285	

ES 2 773 889 T3

gtt aaa cct aac cgt cat gcc tct ctg ctt ccg ctg aaa gag ttc atc 912
Val Lys Pro Asn Arg His Ala Ser Leu Leu Pro Leu Lys Glu Phe Ile
290 295 300

acc gac gcg acc ccg tgc cgt tat gaa gat gtt acg gcg tgg gag tgg 960
Thr Asp Ala Thr Pro Cys Arg Tyr Glu Asp Val Thr Ala Trp Glu Trp
305 310 315 320

aac aat cgc cgt att act aag ctg ttt ggc taa 993
Asn Asn Arg Arg Ile Thr Lys Leu Phe Gly
325 330

<210> 10

<211> 330

<212> PRT

5 <213> Colletotrichum gloeosporioides Nara gc5

<400> 10

Met	Ser	Asp	Lys	Gly	Ala	Thr	Leu	Asp	Met	Ser	Leu	Phe	Phe	Gly	Thr
1				5					10					15	
Pro	Glu	Glu	Lys	Gly	Tyr	Phe	Cys	Thr	Glu	Leu	Leu	Arg	Leu	Leu	Lys
			20					25					30		
Leu	Arg	Gly	Gly	Val	Lys	Ile	Gln	Asn	His	Ser	Ile	Pro	Asp	Glu	Asp
		35					40					45			
Ile	His	Lys	Leu	Phe	Asp	Met	Ser	Arg	Lys	Phe	Phe	Ala	Leu	Pro	Leu
	50					55					60				
Glu	Thr	Lys	Met	Glu	Ala	Lys	His	Pro	Pro	Gln	Ala	Asn	Pro	Asn	Arg
65					70					75					80
Gly	Tyr	Ser	Phe	Ile	Gly	Gln	Glu	Asn	Val	Ala	Asn	Ile	Ser	Gly	Tyr
				85				90						95	
Glu	Lys	Gly	Leu	Gly	Pro	Gly	Lys	Thr	Arg	Asp	Ile	Lys	Glu	Thr	Leu
			100					105					110		
Asp	Met	Gly	Ser	Ala	Gln	Asp	Asp	Leu	Val	Asp	Asn	Leu	Trp	Ile	Pro
		115				120					125				
Glu	Glu	Ser	Leu	Pro	Gly	Phe	Arg	Gly	Phe	Met	Glu	Ser	Phe	Tyr	Glu
	130					135					140				
Lys	Ala	Phe	Lys	Thr	Glu	Met	Gln	Leu	Leu	Ser	Ala	Leu	Ala	Ile	Ala
145					150					155					160
Leu	Gly	Val	Ser	Glu	Asp	His	Met	Lys	Thr	Leu	His	Asn	Arg	Ala	Glu
				165					170					175	
Asn	Glu	Phe	Arg	Leu	Leu	His	Tyr	Pro	Ala	Ile	Pro	Ala	Thr	Glu	Leu
			180					185					190		
Ala	Asp	Gly	Thr	Ala	Thr	Arg	Ile	Ala	Glu	His	Thr	Asp	Phe	Gly	Thr
		195					200					205			
Ile	Thr	Met	Leu	Phe	Gln	Asp	Ser	Val	Gly	Gly	Leu	Gln	Val	Glu	Asp
	210					215					220				
Gln	Thr	Gln	Pro	Gly	Val	Phe	Arg	Ser	Val	Glu	Ser	Glu	Lys	Pro	Thr
225				230						235					240
Asp	Ile	Ile	Leu	Asn	Ile	Gly	Asp	Ser	Leu	Gln	Arg	Leu	Thr	Asn	Asp
				245				250						255	
Thr	Phe	Arg	Ala	Ala	Cys	His	Arg	Val	Thr	Tyr	Pro	Pro	Ala	Ile	Lys
			260					265					270		
Ala	Arg	Asp	Asn	Val	Gln	Ile	Pro	Glu	Arg	Tyr	Ser	Ile	Ala	Tyr	Phe
		275					280					285			
Val	Lys	Pro	Asn	Arg	His	Ala	Ser	Leu	Leu	Pro	Leu	Lys	Glu	Phe	Ile
	290					295					300				
Thr	Asp	Ala	Thr	Pro	Cys	Arg	Tyr	Glu	Asp	Val	Thr	Ala	Trp	Glu	Trp
305					310					315					320
Asn	Asn	Arg	Arg	Ile	Thr	Lys	Leu	Phe	Gly						
				325					330						

<210> 11

<211> 1026

<212> ADN

5 <213> Emericella nidulans FGSC A4

<220>

<221> CDS

<222> (1)..(1026)

<400> 11

ES 2 773 889 T3

atg gat atc ccg acc ctg aat ttt cgt gac ttc acc agc ggt acc cag Met Asp Ile Pro Thr Leu Asn Phe Arg Asp Phe Thr Ser Gly Thr Gln 1 5 10 15	48
agc caa cgt gac acg ttt tgt agc aac ctg tac acc tcc ctg agc acg Ser Gln Arg Asp Thr Phe Cys Ser Asn Leu Tyr Thr Ser Leu Ser Thr 20 25 30	96
ctg ggc ttc gtg aaa atc aag aac cac acg att ccg gat gag att tta Leu Gly Phe Val Lys Ile Lys Asn His Thr Ile Pro Asp Glu Ile Leu 35 40 45	144
gac cag gtg ttt gat tgg tct aaa cac ttc ttt gcg ctg ccg ctg gag Asp Gln Val Phe Asp Trp Ser Lys His Phe Phe Ala Leu Pro Leu Glu 50 55 60	192
tcc aaa act ctg gca gcg cat ccg gtc caa gca aac cca cat cgt ggc Ser Lys Thr Leu Ala Ala His Pro Val Gln Ala Asn Pro His Arg Gly 65 70 75 80	240
tgg tcc tgt gtc ggc cag gag aag ctg agc gtt att cgc cag ggt aag Trp Ser Cys Val Gly Gln Glu Lys Leu Ser Val Ile Arg Gln Gly Lys 85 90 95	288
gca gtg ttc gat ctg aaa gaa agc ttc gat ctg ggt ccg gaa aat gac Ala Val Phe Asp Leu Lys Glu Ser Phe Asp Leu Gly Pro Glu Asn Asp 100 105 110	336
ccg ctg tat cct aat atc ttc ccg gac gat tcg gtc atc ccg ggt ttt Pro Leu Tyr Pro Asn Ile Phe Pro Asp Asp Ser Val Ile Pro Gly Phe 115 120 125	384
cgt ccg ttc atg gag agc ttt tac gcc caa tgc caa gcc ctg cac ttg Arg Pro Phe Met Glu Ser Phe Tyr Ala Gln Cys Gln Ala Leu His Leu 130 135 140	432
act ctt ctg tcc gcc att gca ctg agc ctg aat cag gat gca agc ttc Thr Leu Leu Ser Ala Ile Ala Leu Ser Leu Asn Gln Asp Ala Ser Phe 145 150 155 160	480
ctg agc gac cgt tgc ggt acg aac agc tca gag ctg cgc ctg aac cac Leu Ser Asp Arg Cys Gly Thr Asn Ser Ser Glu Leu Arg Leu Asn His 165 170 175	528
tat cct gcg acc aag att agc gat ttg cag tct ggc aaa aag atg cgt Tyr Pro Ala Thr Lys Ile Ser Asp Leu Gln Ser Gly Lys Lys Met Arg 180 185 190	576
att agc tct cat acg gac ttt ggc acg att acc ctg ttg tgg cag gac Ile Ser Ser His Thr Asp Phe Gly Thr Ile Thr Leu Leu Trp Gln Asp 195 200 205	624

ES 2 773 889 T3

ggt gtt ggt ggt ctg gaa gtg gag gat caa aat cac gag ggt gtc tac	672
Gly Val Gly Gly Leu Glu Val Glu Asp Gln Asn His Glu Gly Val Tyr	
210 215 220	
ttg ccg gtc ggc ccg acc agc acc gaa tcg agc agc ggt gcg gac acc	720
Leu Pro Val Gly Pro Thr Ser Thr Glu Ser Ser Ser Gly Ala Asp Thr	
225 230 235 240	
gac tgt gac agc ggt cgc gag atg atc gtg aac gtt ggc gac tgc ctg	768
Asp Cys Asp Ser Gly Arg Glu Met Ile Val Asn Val Gly Asp Cys Leu	
245 250 255	
cag cgt tgg acc aac gat cgt ctg cgt agc gcg aac cac cgc gtt acc	816
Gln Arg Trp Thr Asn Asp Arg Leu Arg Ser Ala Asn His Arg Val Thr	
260 265 270	
ctg ccg gct gaa atg aaa gac aag agc cgt ccg gag atc agc aat gac	864
Leu Pro Ala Glu Met Lys Asp Lys Ser Arg Pro Glu Ile Ser Asn Asp	
275 280 285	
ctc gtt ccg gat cgc tat agc gtt gcg tac ttc ggt aaa cca gat cgt	912
Leu Val Pro Asp Arg Tyr Ser Val Ala Tyr Phe Gly Lys Pro Asp Arg	
290 295 300	
ggt gcg ctg gtg gcc gct atc ccg gaa ctg gtt gag gaa ggc gaa gaa	960
Gly Ala Leu Val Ala Ala Ile Pro Glu Leu Val Glu Glu Gly Glu Glu	
305 310 315 320	
gtg cgt tac aag ggt ggc atg acc gcg tgg gag tac aat cag agc cgc	1008
Val Arg Tyr Lys Gly Gly Met Thr Ala Trp Glu Tyr Asn Gln Ser Arg	
325 330 335	
ctg ttg caa acg tat taa	1026
Leu Leu Gln Thr Tyr	
340	

<210> 12

<211> 341

<212> PRT

5 <213> Emericella nidulans FGSC A4

<400> 12

Met Asp Ile Pro Thr Leu Asn Phe Arg Asp Phe Thr Ser Gly Thr Gln
1 5 10 15
Ser Gln Arg Asp Thr Phe Cys Ser Asn Leu Tyr Thr Ser Leu Ser Thr
20 25 30
Leu Gly Phe Val Lys Ile Lys Asn His Thr Ile Pro Asp Glu Ile Leu
35 40 45
Asp Gln Val Phe Asp Trp Ser Lys His Phe Phe Ala Leu Pro Leu Glu
50 55 60
Ser Lys Thr Leu Ala Ala His Pro Val Gln Ala Asn Pro His Arg Gly
65 70 75 80
Trp Ser Cys Val Gly Gln Glu Lys Leu Ser Val Ile Arg Gln Gly Lys
85 90 95
Ala Val Phe Asp Leu Lys Glu Ser Phe Asp Leu Gly Pro Glu Asn Asp
100 105 110
Pro Leu Tyr Pro Asn Ile Phe Pro Asp Asp Ser Val Ile Pro Gly Phe
115 120 125
Arg Pro Phe Met Glu Ser Phe Tyr Ala Gln Cys Gln Ala Leu His Leu
130 135 140
Thr Leu Leu Ser Ala Ile Ala Leu Ser Leu Asn Gln Asp Ala Ser Phe
145 150 155 160
Leu Ser Asp Arg Cys Gly Thr Asn Ser Ser Glu Leu Arg Leu Asn His
165 170 175
Tyr Pro Ala Thr Lys Ile Ser Asp Leu Gln Ser Gly Lys Lys Met Arg
180 185 190
Ile Ser Ser His Thr Asp Phe Gly Thr Ile Thr Leu Leu Trp Gln Asp
195 200 205
Gly Val Gly Gly Leu Glu Val Glu Asp Gln Asn His Glu Gly Val Tyr
210 215 220
Leu Pro Val Gly Pro Thr Ser Thr Glu Ser Ser Ser Gly Ala Asp Thr
225 230 235 240
Asp Cys Asp Ser Gly Arg Glu Met Ile Val Asn Val Gly Asp Cys Leu
245 250 255
Gln Arg Trp Thr Asn Asp Arg Leu Arg Ser Ala Asn His Arg Val Thr
260 265 270
Leu Pro Ala Glu Met Lys Asp Lys Ser Arg Pro Glu Ile Ser Asn Asp
275 280 285
Leu Val Pro Asp Arg Tyr Ser Val Ala Tyr Phe Gly Lys Pro Asp Arg
290 295 300
Gly Ala Leu Val Ala Ala Ile Pro Glu Leu Val Glu Glu Gly Glu Glu
305 310 315 320
Val Arg Tyr Lys Gly Gly Met Thr Ala Trp Glu Tyr Asn Gln Ser Arg
325 330 335
Leu Leu Gln Thr Tyr
340

<210> 13

<211> 28

5 <212> ADN

<213> Secuencia artificial

<220>

<223> cebador de PCR

<400> 13

10 aaggatccat ggccgccctc aacgcaga

28

<210> 14

<211> 28

<212> ADN
 <213> Secuencia artificial
 <220>
 <223> cebador de PCR
 5 <400> 14
 ccaagctttt aagccttagc ttcagcag 28
 <210> 15
 <211> 756
 <212> ADN
 10 <213> Desconocido
 <220>
 <223> ambiental
 <220>
 <221> CDS
 15 <222> (1)..(753)
 <400> 15

```

atg acc acc gcc gcc ttc acg cct ccc ttc atc gat ccc gcc gaa ctg      48
Met Thr Thr Ala Ala Phe Thr Pro Pro Phe Ile Asp Pro Ala Glu Leu
1      5      10      15
ccc gcc gcg ctg cgc gag cag ggc tac gcg gtg ctg agc ccc caa ggc      96
Pro Ala Ala Leu Arg Glu Gln Gly Tyr Ala Val Leu Ser Pro Gln Gly
20      25      30
gtc agc gac tgg ctc ggc gca ccg ctt gcg caa ctc gag gca ctg cac      144
Val Ser Asp Trp Leu Gly Ala Pro Leu Ala Gln Leu Glu Ala Leu His
35      40      45
gcg gac tgg aac aac ctt cca ccc gac gac tac ctg aag gac ggc ggc      192
Ala Asp Trp Asn Asn Leu Pro Pro Asp Asp Tyr Leu Lys Asp Gly Gly
50      55      60
cgc tac cgc acg cgc cgc cat gca tgc ttc acc gtc gag gcc ggc gac      240
Arg Tyr Arg Thr Arg Arg His Ala Cys Phe Thr Val Glu Ala Gly Asp
65      70      75      80
aac gcc acg ctc gaa cag gtg ccg cat cgc gcg cac tgg cag ccg gtc      288
Asn Ala Thr Leu Glu Gln Val Pro His Arg Ala His Trp Gln Pro Val
85      90      95
gaa tac aac gcg ctg cac ggg ggc atg cag cgc tgg ttt gcg ccg atg      336
Glu Tyr Asn Ala Leu His Gly Gly Met Gln Arg Trp Phe Ala Pro Met
100      105      110
ctg gcc gag acc atc gtg cag cca gtg tgg caa cgc ctg ctg cag cgc      384
Leu Ala Glu Thr Ile Val Gln Pro Val Trp Gln Arg Leu Leu Gln Arg
115      120      125
ctc ggt acg gcc gcg agc gac atg cgg ggt acg cca cag aaa tgg ttt      432
Leu Gly Thr Ala Ala Ser Asp Met Arg Gly Thr Pro Gln Lys Trp Phe
130      135      140
gtc gag gcg cat caa ttc cgc atc gac acc gca ggc ggc atc ggc cgt      480
Val Glu Ala His Gln Phe Arg Ile Asp Thr Ala Gly Gly Ile Gly Arg
145      150      155      160
ccg acc ccc gag ggc gcg cac cgc gac ggc gtc gac ctg gtg gcc gtc      528
Pro Thr Pro Glu Gly Ala His Arg Asp Gly Val Asp Leu Val Ala Val
165      170      175
gcg ttg gtc ggt cgt caa ggc atc aag ggc ggc gag acg cgc gtg ttc      576
Ala Leu Val Gly Arg Gln Gly Ile Lys Gly Gly Glu Thr Arg Val Phe
180      185      190
gag gcc aac ggc cgt cgc ggc gag cga ttc acc atg acc gag ccc tgg      624
Glu Ala Asn Gly Arg Arg Gly Glu Arg Phe Thr Met Thr Glu Pro Trp
195      200      205
acg ctg ctc ctg ctc gac gac gca cgc gtc att cac gaa tcg acg ccg      672
Thr Leu Leu Leu Leu Asp Asp Ala Arg Val Ile His Glu Ser Thr Pro
210      215      220
atc cag cca ctc gaa gag ggc ggt gcg ggt tgg cgc gat acg ctc gtg      720
Ile Gln Pro Leu Glu Glu Gly Gly Ala Gly Trp Arg Asp Thr Leu Val
225      230      235      240
atc acc tgc cga tcg caa ggt ttc cag ggc gat taa      756
Ile Thr Cys Arg Ser Gln Gly Phe Gln Gly Asp
245      250

```

<210> 16

<211> 251

<212> PRT

5 <213> Desconocido

<220>

<223> ambiental

<400> 16

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

<210> 17

<211> 840

5 <212> ADN

<213> Desconocido

<220>

<223> ambiental

<220>

10 <221> CDS

<222> (1)..(837)

<400> 17

atg gaa gtg cca aat tcc agc ttc aat acg aat tcg ccg tta cgg gag	48
Met Glu Val Pro Asn Ser Ser Phe Asn Thr Asn Ser Pro Leu Arg Glu	
1 5 10 15	
tcg atc tct tcg att gta atg aaa ggg tat gcg att cta aaa att cca	96
Ser Ile Ser Ser Ile Val Met Lys Gly Tyr Ala Ile Leu Lys Ile Pro	
20 25 30	
gtc cat gcg caa aca gcc atg tgg aag gtg ttg gac aac atc gac tcg	144
Val His Ala Gln Thr Ala Met Trp Lys Val Leu Asp Asn Ile Asp Ser	
35 40 45	
gtt tcg cct gat ctt cgg cga gac ttc agt ttt ccg gag gtt acc gat	192
Val Ser Pro Asp Leu Arg Asp Phe Ser Phe Pro Glu Val Thr Asp	
50 55 60	
gga ttt tta tct gtc ggc ggt gag tac gcc aag tac acc agc agc atc	240
Gly Phe Leu Ser Val Gly Gly Glu Tyr Ala Lys Tyr Thr Ser Ser Ile	
65 70 75 80	
gat ttg tgc gat cgc ttc tgc ttc tgg cat aag aat cgt gaa atg cat	288
Asp Leu Cys Asp Arg Phe Cys Phe Trp His Lys Asn Arg Glu Met His	
85 90 95	
caa ggg aaa gaa ttc gca aac gac gaa gtc tat agg acc ata aag acc	336
Gln Gly Lys Glu Phe Ala Asn Asp Glu Val Tyr Arg Thr Ile Lys Thr	
100 105 110	
tgc gag atg gaa atg cac acc atg gcg cag caa ttg att tcg gag ttg	384
Cys Glu Met Glu Met His Thr Ala Gln Gln Leu Ile Ser Glu Leu	
115 120 125	
tgg gac ttt ttc aat agc ccc gat ccg gtg aat att cga gac agc tcg	432
Trp Asp Phe Phe Asn Ser Pro Asp Pro Val Asn Ile Arg Asp Ser Ser	
130 135 140	
tat ctg cag ttg tgc atg tac gca agt cag tac cac gcg aat gat cgc	480
Tyr Leu Gln Leu Cys Met Tyr Ala Ser Gln Tyr His Ala Asn Asp Arg	
145 150 155 160	
ggc tat ctg cag gac cgc cat gaa gat ggg cat ttg atc acg ctg atc	528
Gly Tyr Leu Gln Asp Arg His Glu Asp Gly His Leu Ile Thr Leu Ile	
165 170 175	
aaa ccg acg cgc gat ggc ctc gtc att ttc cct ggc ggc ccg gaa gac	576
Lys Pro Thr Arg Asp Gly Leu Val Ile Phe Pro Gly Gly Pro Glu Asp	
180 185 190	
tgc gaa gtt cca gta ttt ctg cgt gat gat gaa ttg ctc gtt att gca	624
Cys Glu Val Pro Val Phe Leu Arg Asp Asp Glu Leu Leu Val Ile Ala	
195 200 205	
ggg tca ctg ctg acg gca atg tcg gat gga aga att ccc ccg atg tat	672
Gly Ser Leu Leu Thr Ala Met Ser Asp Gly Arg Ile Pro Pro Met Tyr	
210 215 220	
cac gcc gtt aaa aat cca ttt gtg cag atg gaa aga aaa tcc gtt gtg	720
His Ala Val Lys Asn Pro Phe Val Gln Met Glu Arg Lys Ser Val Val	
225 230 235 240	
tat ttt gcc att ccg gat ctt tcg cag acc tat acg aca ttg ctc gcc	768
Tyr Phe Ala Ile Pro Asp Leu Ser Gln Thr Tyr Thr Thr Leu Leu Ala	
245 250 255	
aaa aat ccg atc aat gtc gga gaa ttg gcg gat gaa agt cac cgg gcc	816
Lys Asn Pro Ile Asn Val Gly Glu Leu Ala Asp Glu Ser His Arg Ala	
260 265 270	
ttc ggc aat acg gcc ttg atc taa	840
Phe Gly Asn Thr Ala Leu Ile	
275	

<210> 18

<211> 279

5 <212> PRT

<213> Desconocido

<220>

<223> ambiental

<400> 18

```

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

```

REIVINDICACIONES

1. Uso de la proteína 4-hidroxilasa de ácido pipercolico que tiene una actividad para reaccionar con ácido L-pipercolico en presencia de ácido 2-oxoglutarico e iones de hierro (II), para producir ácido trans-4-hidroxi-L-pipercolico, en el que la proteína 4-hidroxilasa de ácido pipercolico se selecciona del grupo que consiste en los siguientes (A), (B) y (C):

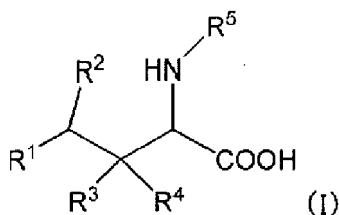
(A) un polipéptido que comprende la secuencia de aminoácidos representada por SEQ ID NO:2, 4, 6, 8, 10, 12, 16 ó 18;

(B) un polipéptido que comprende la secuencia de aminoácidos representada por SEQ ID NO:2, 4, 6, 8, 10, 12, 16 ó 18 excepto que se delecionan, sustituyen y/o añaden de uno a 100 aminoácidos, y que tiene actividad de 4-hidroxilasa de ácido pipercolico; y

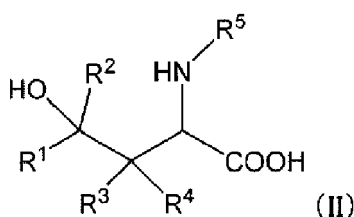
(C) un polipéptido que comprende una secuencia de aminoácidos que no es menos del 80% idéntica a la secuencia de aminoácidos representada por SEQ ID NO:2, 4, 6, 8, 10, 12, 16 ó 18, y que tiene actividad de 4-hidroxilasa de ácido pipercolico.

2. Uso de la proteína 4-hidroxilasa de ácido pipercolico según la reivindicación 1, proteína que se produce como proteína recombinante por un huésped que no tiene capacidad de glicosilación.

3. Método para producir 4-hidroxi-aminoácido, comprendiendo dicho método hacer reaccionar dicha proteína 4-hidroxilasa de ácido pipercolico como se define en la reivindicación 1 ó 2, una(s) célula(s) que comprende(n) dicha proteína, un producto tratado de dicha(s) célula(s) y/o un líquido de cultivo obtenido cultivando dicha(s) célula(s), comprendiendo dichos producto tratado y líquido de cultivo dicha proteína, con α -aminoácido para producir 4-hidroxi-aminoácido, en el que dicho α -aminoácido se representa mediante la siguiente fórmula general (I):



en la que cada uno de R¹, R² y R⁵ representa un átomo de hidrógeno o alquilo C₁-C₃; cada uno de R³ y R⁴ representa un átomo de hidrógeno, alquilo C₁-C₃ o hidroxilo; y alternativamente R² puede unirse a R⁵ o al átomo de nitrógeno del grupo amino para formar una estructura de anillo, y dicho 4-hidroxi-aminoácido se representa mediante la siguiente fórmula general (II):



en la que cada uno de R¹, R² y R⁵ representa un átomo de hidrógeno o alquilo C₁-C₃; cada uno de R³ y R⁴ representa un átomo de hidrógeno, alquilo C₁-C₃ o hidroxilo; y alternativamente R² puede unirse a R⁵ o al átomo de nitrógeno del grupo amino para formar una estructura de anillo.

4. Método para producir 4-hidroxi-aminoácido según la reivindicación 3, en el que dicho α -aminoácido es ácido pipercolico, y dicho 4-hidroxi-aminoácido es ácido 4-hidroxi-L-pipercolico.
5. Método para producir 4-hidroxi-aminoácido según la reivindicación 3 ó 4, en el que dicha célula que comprende 4-hidroxilasa de ácido pipercolico es una célula transformada con un ADN que codifica para proteína 4-hidroxilasa de ácido pipercolico.
6. Método para producir 4-hidroxi-aminoácido según la reivindicación 5, en el que dicho ADN que codifica para proteína 4-hidroxilasa de ácido pipercolico se selecciona del grupo que consiste en los siguientes (D), (E) y (F):

(D) un ADN que comprende la secuencia de nucleótidos representada por SEQ ID NO:1, 3, 5, 7, 9, 11, 15 ó 17;

- (E) un ADN que comprende la secuencia de nucleótidos representada por SEQ ID NO:1, 3, 5, 7, 9, 11, 15 ó 17 excepto que se sustituyen, delecionan y/o añaden de uno a 300 nucleótidos, y que codifica para un polipéptido que tiene actividad de 4-hidroxilasa de ácido pipecólico; y
- 5 (F) un ADN que comprende una secuencia de nucleótidos que se hibrida con la cadena complementaria de la secuencia de nucleótidos representada por SEQ ID NO:1, 3, 5, 7, 9, 11, 15 ó 17 en condiciones rigurosas, donde el lavado se lleva a cabo en presencia de 0,1x SSC y SDS al 0,1% a 60°C, y que codifica para un polipéptido que tiene actividad de 4-hidroxilasa de ácido pipecólico.
7. Microorganismo transformado con un vector de expresión que comprende un ADN que codifica para y que expresa una proteína 4-hidroxilasa de ácido pipecólico seleccionada del grupo que consiste en los
- 10 siguientes (D), (E) y (F):
- (D) un ADN que comprende la secuencia de nucleótidos representada por SEQ ID NO:1, 3, 5, 7, 9, 11, 15 ó 17;
- (E) un ADN que comprende la secuencia de nucleótidos representada por SEQ ID NO:1, 3, 5, 7, 9, 11, 15 ó 17 excepto que se sustituyen, delecionan y/o añaden de uno a 300 nucleótidos, y que codifica para un
- 15 polipéptido que tiene actividad de 4-hidroxilasa de ácido pipecólico; y
- (F) un ADN que comprende una secuencia de nucleótidos que se hibrida con la cadena complementaria de la secuencia de nucleótidos representada por SEQ ID NO:1, 3, 5, 7, 9, 11, 15 ó 17 en condiciones rigurosas, donde el lavado se lleva a cabo en presencia de 0,1x SSC y SDS al 0,1% a 60°C, y que codifica para un polipéptido que tiene actividad de 4-hidroxilasa de ácido pipecólico.
- 20 8. Microorganismo según la reivindicación 7, que se selecciona del grupo que consiste en los géneros *Escherichia*, *Bacillus*, *Pseudomonas* y *Corynebacterium*.
9. Proteína 4-hidroxilasa de ácido pipecólico que tiene una actividad para reaccionar con ácido L-pipecólico en presencia de ácido 2-oxoglutarico e iones de hierro (II), para producir ácido trans-4-hidroxi-L-pipecólico, en la que la proteína 4-hidroxilasa de ácido pipecólico se selecciona del grupo que consiste en los siguientes
- 25 (A), (B) y (C):
- (A) un polipéptido que comprende la secuencia de aminoácidos representada por SEQ ID NO: 18;
- (B) un polipéptido que comprende la secuencia de aminoácidos representada por SEQ ID NO: 18 excepto que se delecionan, sustituyen y/o añaden de uno a 100 aminoácidos, y que tiene actividad de 4-hidroxilasa de ácido pipecólico; y
- 30 (C) un polipéptido que comprende una secuencia de aminoácidos que no es menos del 80% idéntica a la secuencia de aminoácidos representada por SEQ ID NO: 18, y que tiene actividad de 4-hidroxilasa de ácido pipecólico.

Fig. 1

segmento n.º 1

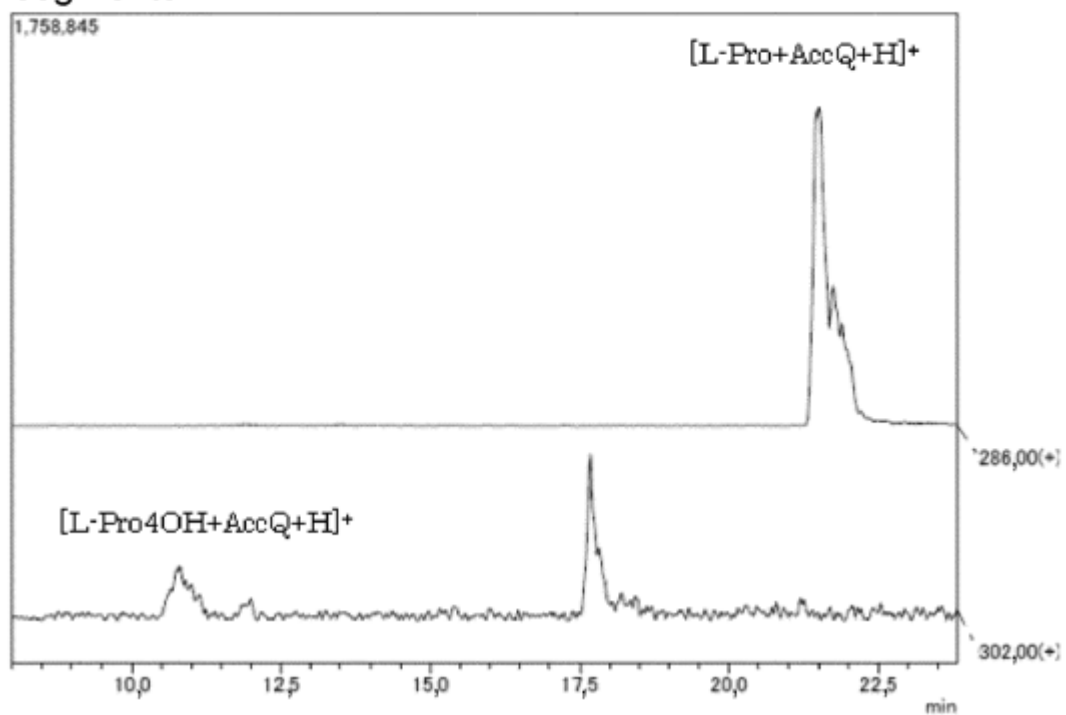


Fig. 2

segmento n.º 1

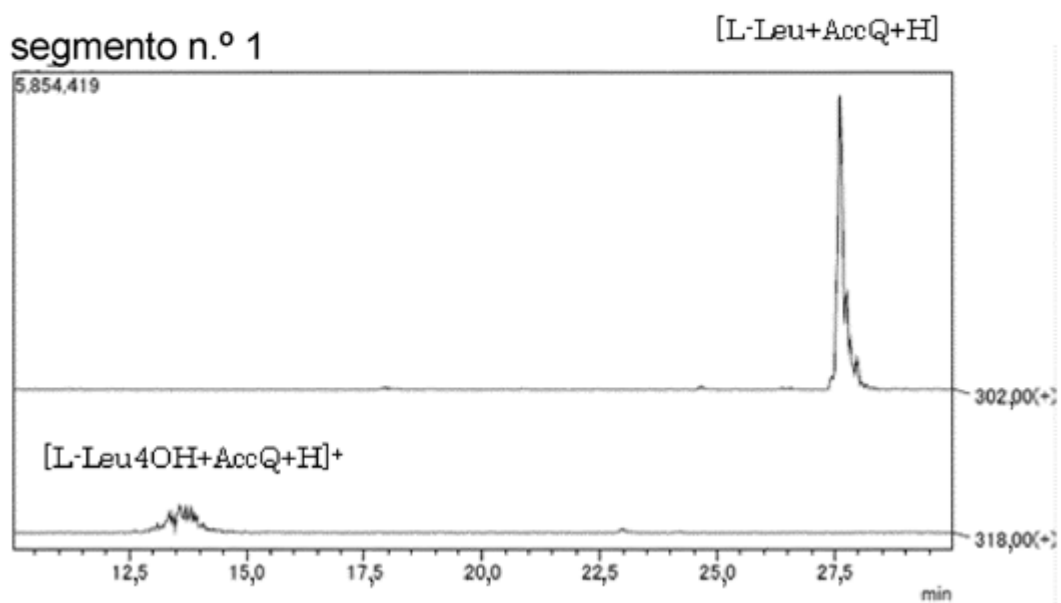


Fig. 3

segmento n.º 1

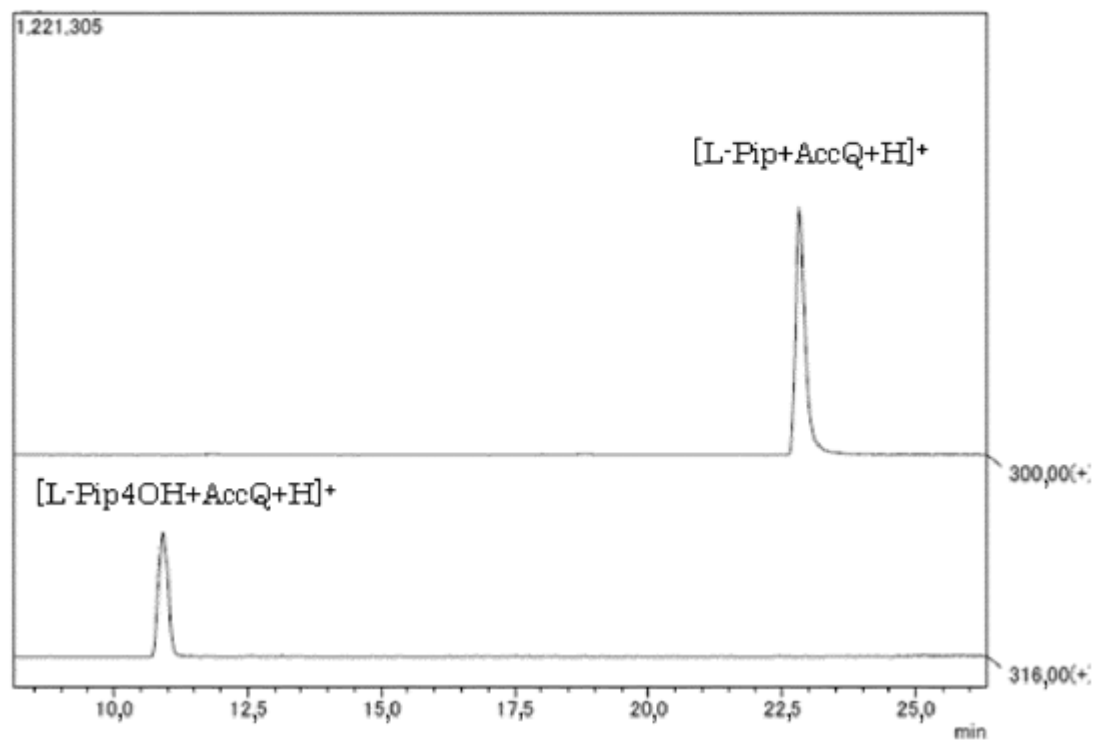


Fig. 4

segmento n.º 1

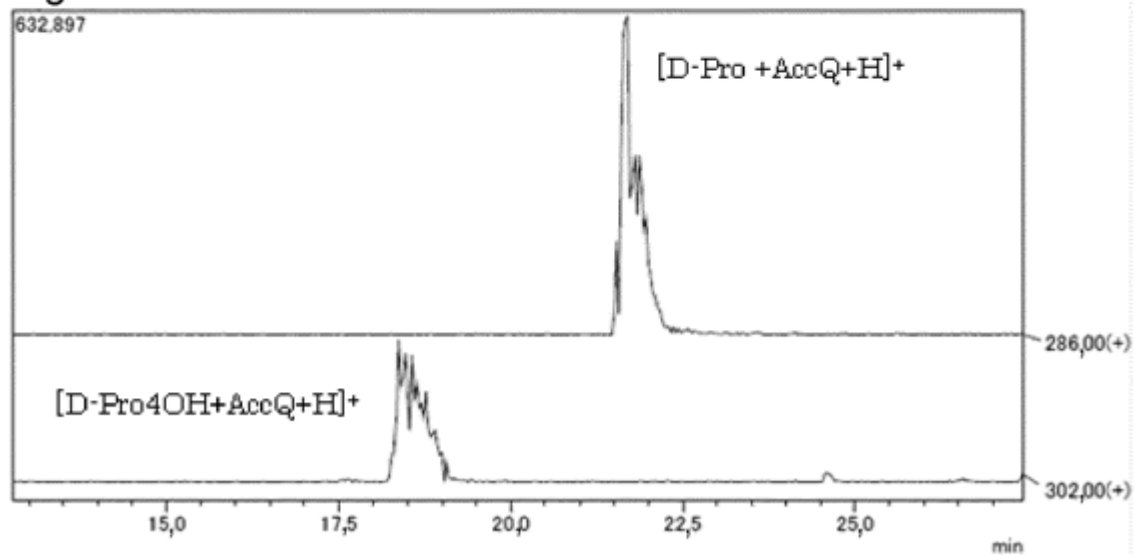


Fig. 5

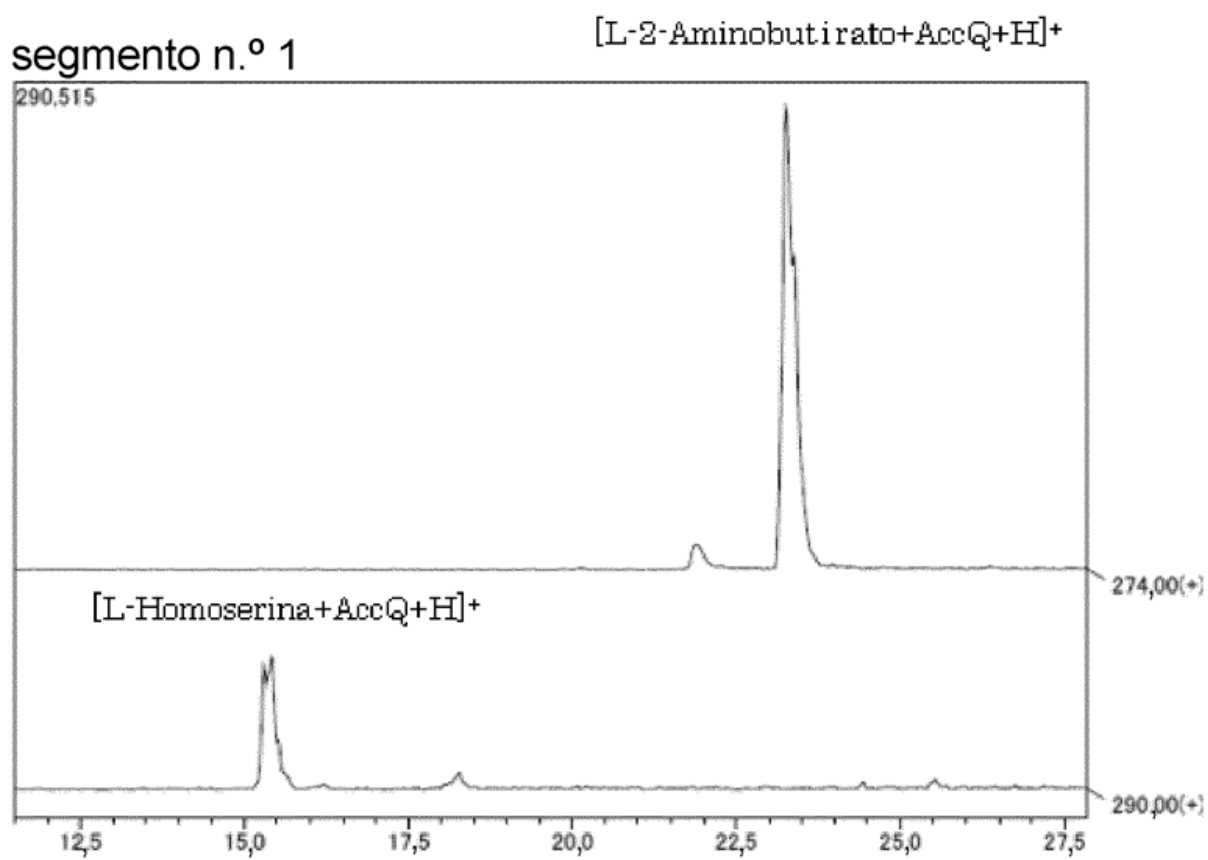


Fig. 6

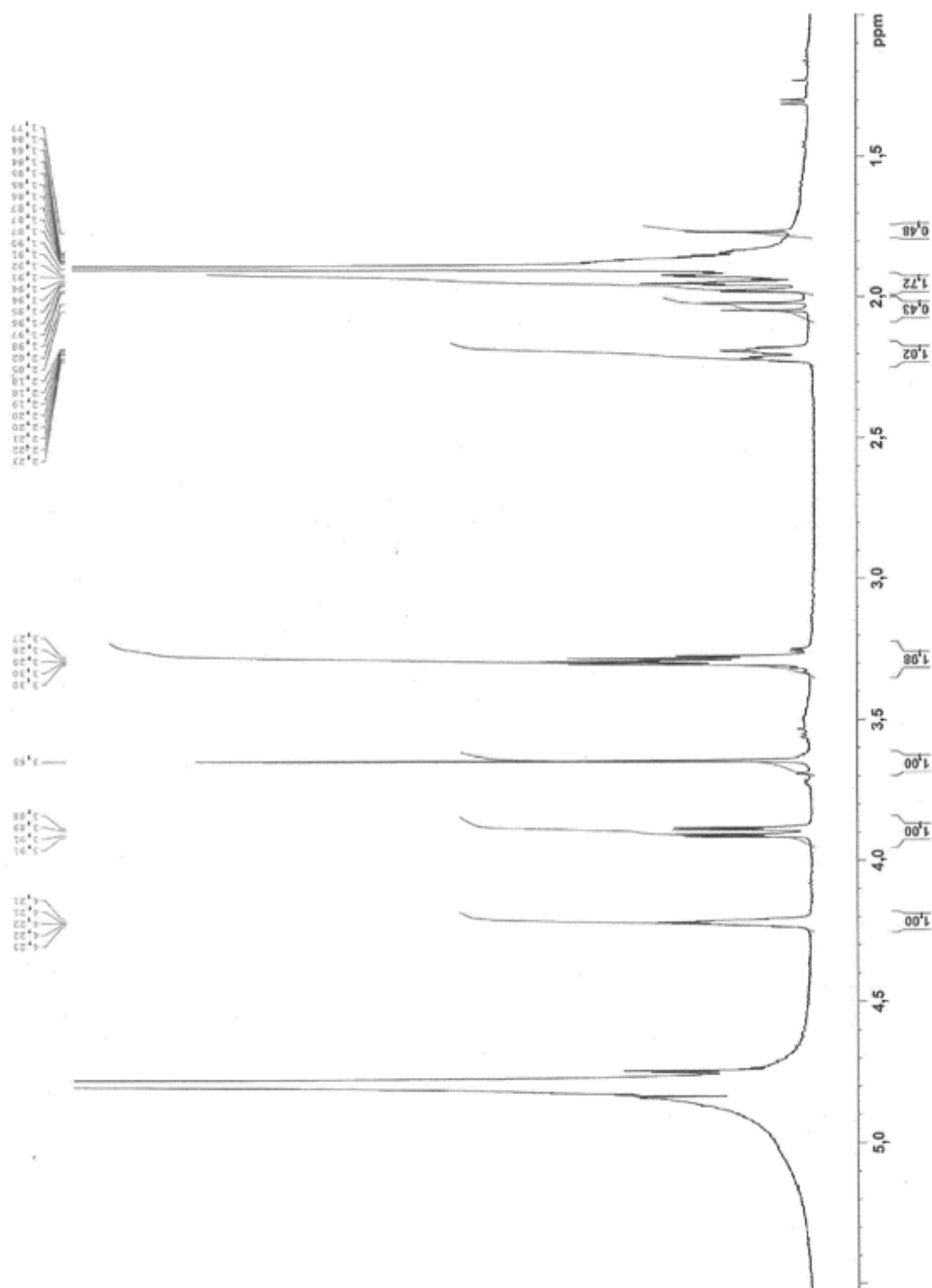


Fig. 7

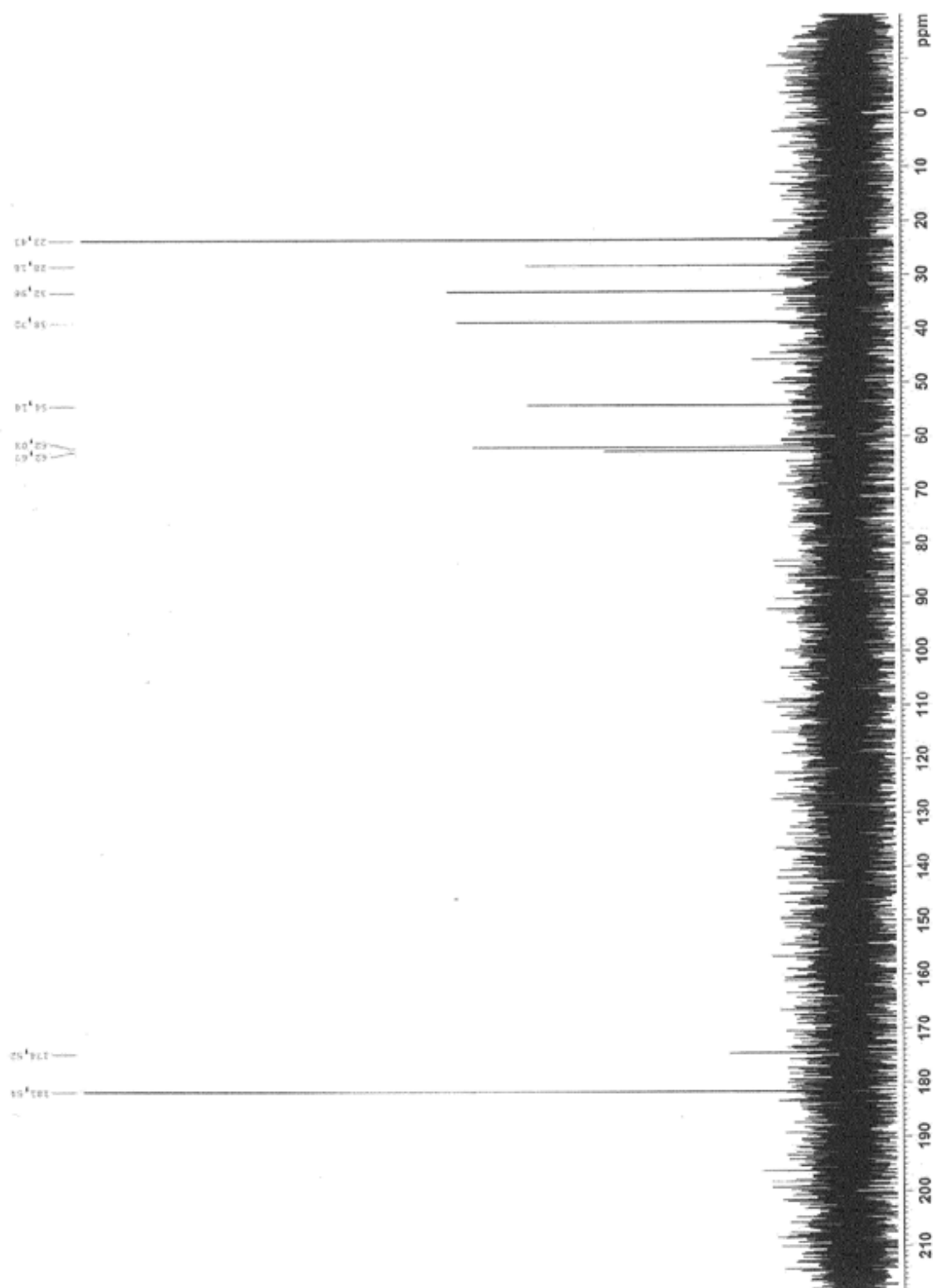
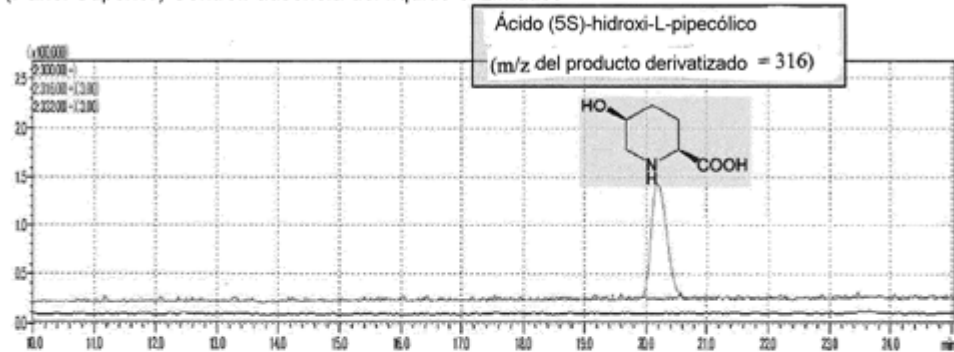


Fig. 8

(Panel Superior) Control: ausencia del líquido enzimático



(Panel inferior) Adición de un extracto libre de células obtenido de *E. coli* transformada con pEnPA4H

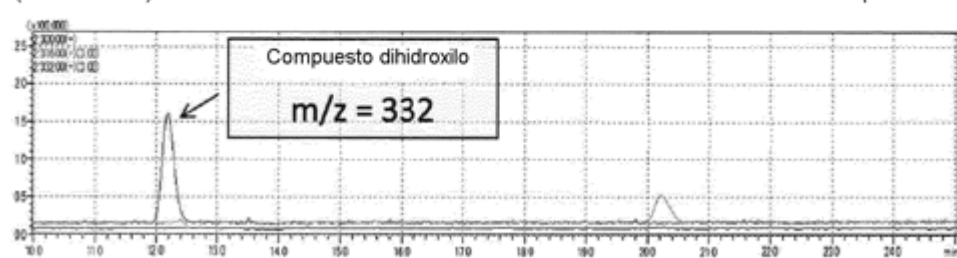
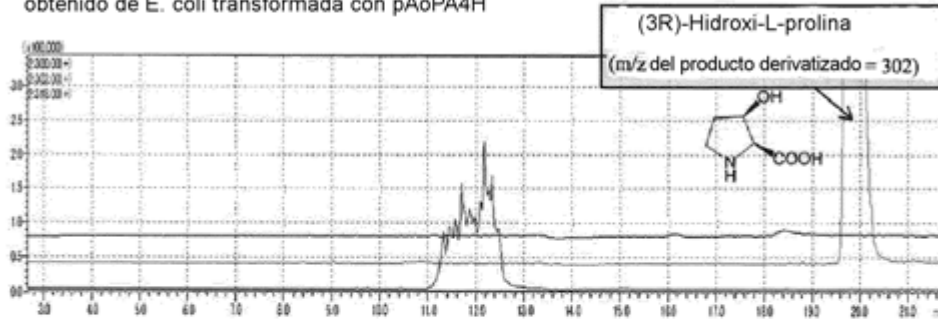


Fig. 9

(Panel superior) Control: adición de un extracto libre de células obtenido de *E. coli* transformada con pAoPA4H



(Panel inferior) Adición de un extracto libre de células obtenido de *E. coli* transformada con pEnPA4H

