

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 703 770**

51 Int. Cl.:

C12P 17/04 (2006.01)

C12N 9/90 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **02.06.2010 PCT/EP2010/057696**

87 Fecha y número de publicación internacional: **09.12.2010 WO10139719**

96 Fecha de presentación y número de la solicitud europea: **02.06.2010 E 10724460 (0)**

97 Fecha y número de publicación de la concesión europea: **10.10.2018 EP 2438182**

54 Título: **Preparación biocatalítica de ambroxán**

30 Prioridad:

05.06.2009 EP 09162104

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

12.03.2019

73 Titular/es:

**BASF SE (100.0%)
Carl-Bosch-Strasse 38
67056 Ludwigshafen am Rhein, DE**

72 Inventor/es:

**BREUER, MICHAEL;
HÖRSTER, ANDREA y
HAUER, BERNHARD**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 703 770 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).

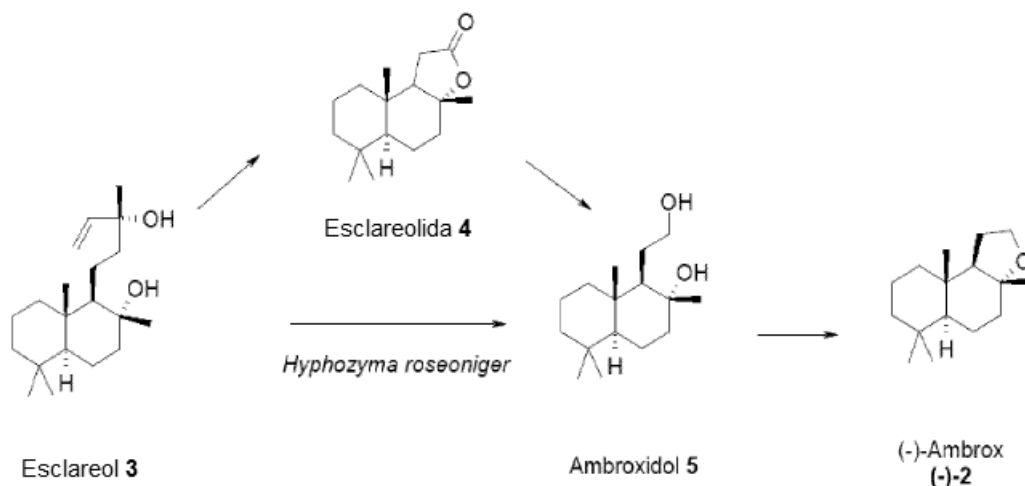
DESCRIPCIÓN

Preparación biocatalítica de ambroxán

La invención se refiere a un procedimiento para la preparación biocatalítica de ambroxán.

Los compuestos con la estructura principal del dodecahidronafto[2,1-*b*]furano son de gran interés económico como productos químicos aromáticos. En este caso se menciona el Compuesto 2, el (3a*R*,5a*S*,9a*S*,9b*R*)-3a,6,6,9a-tetrametildodecahidronafto-[2,1-*b*]-furano) conocido como el estereoisómero levógiro ((-)-2) del ambroxán.

Originalmente obtenido a partir del ámbar del cachalote, actualmente existen principalmente dos rutas que permiten el acceso a ambroxán. Esclareol (3), un ingrediente de la esclárea (*Salvia sclarea*) sirve a este respecto con frecuencia como material de partida adecuado para material semisintético, debido a que contiene ya información óptica para el compuesto ((-)-2). A este respecto, la degradación oxidativa puede llevarse a cabo con ácido crómico, permanganato, H₂O₂ u ozono [Stoll et al.; *Helv Chim Acta* (1950), 33: 1251]. La esclareolida obtenida (4) se convierte mediante una reacción posterior (por ejemplo con LiAlH₄ o NaBH₄) en el ambrox-1,4-diol (5) [Mookherjee et al.; *Perfumer and Flavourist* (1990), 15: 27]. La preparación de compuesto (4) a partir de esclareol (3) puede tener lugar también mediante una biotransformación con *Hyphozyma roseoniger* [documento EP 204009].



Ambrox-1,4-diol (5) puede ciclarse por último con una serie de procedimientos químicos para dar el compuesto ((-)-2). Para el racemato de ambroxán rac-2 se trató accesos, entre otros, a través de ácido homofarnesílico [documento US 513.270; Lucius et al.; *Chem Ber* (1960), 93: 2663] y 4-(2,6,6-trimetil-ciclohex-1-enil)-butan-2-ona [Büchi et al.; *Helv Chim Acta* (1989), 72: 996]. El volumen de mercado de ambroxán en 2002 era en aquel momento en promedio de 20 toneladas al año. Para ello es necesaria una base de partida de aproximadamente 33 toneladas de esclareol al año. Para la producción de una tonelada de ambroxán se necesitan 207 toneladas de distintas sustancias individuales, que a su vez requieren la acumulación de 206 toneladas de desecho. Las sustancias generadas presentan diferentes, pero en conjunto relativamente fuertes efectos sobre el medio ambiente y la salud [Deutsche Bundesstiftung Umwelt]. Por lo tanto, esta síntesis es de muy alto consumo energético y requiere el uso de productos químicos contaminantes.

La síntesis biocatalítica de compuesto ((-)-2) se describe en la bibliografía [Neumann et al.; *Biol Chem Hoppe Seyler* (1986), 367: 723]. A este respecto, la molécula se obtiene a partir de homofarnesol (Compuesto (1), (3*Z*,7*E*)-4,8,12-trimetiltrideca-3,7,11-trien-1-ol). Como catalizador se usó la escualeno-hopeno ciclasa (SHC) de *Alicyclobacillus acidocaldarius* (anteriormente *Bacillus acidocaldarius*). La enzima cataliza naturalmente la ciclación de escualeno para dar hopano. Como reacción secundaria, esta SHC puede convertir obviamente también el Compuesto (1) para dar ambroxán ((-)-2). El biocatalizador puede prepararse de manera recombinante [Ochs D. et al.; *J Bacteriol* (1992), 174: 298]. La reacción de ciclación de homofarnesol para dar ambroxán asciende según Neumann et al. no obstante únicamente al 1,2 % (calculado a partir de los picos de CG) y la actividad específica con respecto a la ciclación de homofarnesol se indica 0,02 mU/mg de proteína.

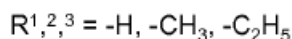
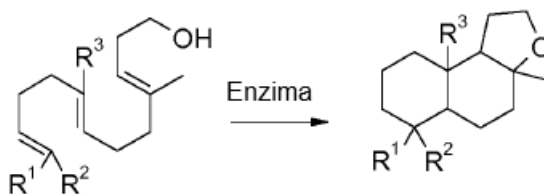
La solicitud de patente JP2009060799 describe la síntesis biocatalítica de (-)-ambroxán a partir de homofarnesol por medio de una SHC, teniendo lugar la reacción en un disolvente a un pH de 5,2-6,6. Los ejemplos experimentales usan la SHC de *A. acidocaldarius*.

Por lo tanto era objetivo de la presente invención proporcionar un nuevo procedimiento para la preparación de ambroxán así como derivados del mismo, que ofrezca ventajas con respecto a las síntesis técnicamente costosas del estado de la técnica. Mediante el uso de procedimientos biotécnicos se evitará la generación de sustancias

contaminantes y se reducirá drásticamente el consumo de energía. Otro objetivo era reducir los gastos generados adicionalmente por el uso de materias primas fácilmente accesibles así como una disminución del número de reacciones químicas (o etapas). Asimismo se conseguirá una productividad elevada.

5 Este objetivo se consigue mediante un procedimiento para la preparación de derivados de ambroxán, preferentemente ambroxán, de fórmula general (2) caracterizado porque se hacen reaccionar derivados de homofarnesol de fórmula general (1) de manera biocatalítica para dar los derivados de ambroxán correspondientes por medio de un polipéptido con la actividad de una homofarnesol-ambroxán ciclasa como enzima, en el que la actividad del polipéptido es la reacción de homofarnesol para dar ambroxán con sustrato principal homofarnesol, caracterizado porque la enzima es un polipéptido, que se codifica por una molécula de ácido nucleico que comprende al menos una molécula de ácido nucleico seleccionada del grupo que consiste en:

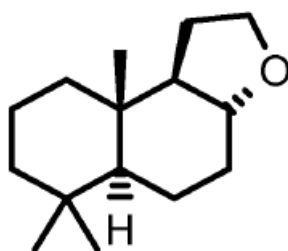
- 10 a) molécula de ácido nucleico que codifica un polipéptido que comprende la SEQ ID NO: 2, 5 o 10;
 b) molécula de ácido nucleico que comprende al menos un polinucleótido de la secuencia mostrada en SEQ ID NO: 1;
 15 c) molécula de ácido nucleico que codifica un polipéptido cuya secuencia presenta una identidad de al menos el 60 % con respecto a las secuencias SEQ ID NO: 2, 5 o 10.



20 Derivados son en particular estereoisómeros, preferentemente enantiómeros pero también diastereómeros del compuesto (2). En una realización de la presente invención, derivados de homofarnesol o ambroxán son Compuestos sustituidos (1) y (2), siendo los sustituyentes inertes con respecto a la reacción biocatalizada. Con ello se expresan en particular compuestos de las fórmulas estructurales representadas a continuación.

En el sentido de la invención, los términos "Compuesto (1)", "homofarnesol", "4,8,12-trimetiltrideca-3,7,11-trien-1-ol)" y derivados del homofarnesol o los términos "Compuesto (2)", "ambroxán", "dodecahidronafto[2,1-b]furano" y derivados del ambroxán tienen el mismo significado y pueden intercambiarse y sustituirse uno por otro, siempre que no se defina expresamente nada al contrario.

25 En una variante preferida de la invención, se forma el ambroxán levógiro de fórmula ((-)-2):



(-)-Ambroxán 2

Los polipéptidos con la actividad de una homofarnesol-ambroxán ciclasa representan una nueva clase de enzimas.

30 El término "actividad" describe la capacidad de una enzima para convertir un sustrato en un producto. La actividad puede determinarse en una denominada prueba de actividad a través del aumento del producto, la disminución del sustrato (o educto) o a través de una combinación de estos parámetros en función del tiempo.

Las enzimas usadas de acuerdo con la invención tienen una actividad para la reacción de homofarnesol para dar

ambroxán.

En el sentido de la invención, el sustrato principal es el compuesto químico que en comparación con todos los demás compuestos, que pueden convertirse por la enzima, presenta en porcentaje en moles la mayor proporción como componente de reacción de la homofarnesol-ambroxán ciclasa.

- 5 En el sentido de la invención, el sustrato principal es homofarnesol y por lo tanto la actividad principal y con ello la reacción principal de una homofarnesol-ambroxán ciclasa la reacción con el sustrato principal homofarnesol.

La actividad de la homofarnesol-ambroxán ciclasa se define en una variante de la invención a través del rendimiento en porcentaje en moles. Preferentemente, con la reacción de homofarnesol para dar ambroxán en presencia de una enzima de la nueva clase de homofarnesol-ambroxán ciclasas, se obtiene un rendimiento de ambroxán en porcentaje en moles con respecto a los moles empleados de homofarnesol de 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100; de manera especialmente preferente, el rendimiento se encuentra entre 5 y 100, 10 y 100, 20 y 100, 25 y 100, 30 y 100, 35 y 100, en particular entre 40 y 100, 45 y 100, 50 y 100, 60 y 100, 70 y 100 porcentaje en moles.

La actividad de la homofarnesol-ambroxán ciclasa, en otra variante de la invención se define a través de la conversión (cantidad de producto/(cantidad de producto + cantidad de educto residual) * 100) en porcentaje en moles. Preferentemente, con la reacción de homofarnesol para dar ambroxán en presencia de una enzima de la nueva clase de homofarnesol-ambroxán ciclasas, se obtiene una conversión de ambroxán en porcentaje en moles de 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100; de manera especialmente preferente, la conversión se encuentra entre 5 y 100, 10 y 100, 20 y 100, 25 y 100, 30 y 100, 35 y 100, en particular entre 40 y 100, 45 y 100, 50 y 100, 60 y 100, 70 y 100.

25 En una realización preferida de la invención, el rendimiento y/o la conversión se determina a lo largo de un periodo de tiempo definido de por ejemplo 4, 6, 8, 10, 12, 16, 20, 24, 36 o 48 horas en el que la reacción de homofarnesol para dar ambroxán tiene lugar por medio de las ciclasas usadas de acuerdo con la invención. En otra variante, la reacción se lleva a cabo en condiciones definidas de manera precisa, de por ejemplo 25, 30, 40, 50 o 60 °C. En particular, tiene lugar la determinación del rendimiento y/o de la conversión llevándose a cabo la reacción para la reacción de homofarnesol para dar ambroxán por medio de las ciclasas usadas de acuerdo con la invención a 30 °C durante 16 horas.

30 En una realización de la invención, para la determinación del rendimiento y/o de la conversión se hace reaccionar una solución 10 mM de homofarnesol (tamponada con citrato) con una solución de ciclasa, encontrándose la enzima como extracto de proteína de membrana de células que expresan homofarnesol-ambroxán ciclasas (aisladas de acuerdo con [Ochs D. et al.; J Bacteriol (1992), 174: 298]) en una concentración de contenido de proteína del 0,08 por ciento en peso.

35 En otra realización de la presente invención, una homofarnesol-ambroxán ciclasa se caracteriza porque con la reacción de homofarnesol para dar ambroxán en las mismas condiciones muestra una conversión y/o rendimiento de 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 500, 1000 veces o superior en comparación con la escualeno-hopeno ciclasa (SHC) de *Alicyclobacillus acidocaldarius* (anteriormente *Bacillus acidocaldarius*). El término condición se refiere en este caso a condiciones de reacción tales como concentración de sustrato, concentración de enzima, periodo de tiempo de reacción y/o temperatura.

45 En una variante de la presente invención, una homofarnesol-ambroxán ciclasa se caracteriza por cualquiera o varias, en cualquier combinación aleatoria, de las definiciones mencionadas anteriormente.

Asimismo, se divulga un procedimiento para la preparación de ambroxán, en el que

- a) se pone en contacto homofarnesol con la homofarnesol-ambroxán ciclasa y/o se incuba, y
50 b) se aísla ambroxán.

En una realización de la invención, se pone en contacto homofarnesol con la homofarnesol-ambroxán ciclasa en un medio y/o se incuba de modo que tiene lugar una reacción de homofarnesol para dar ambroxán en presencia de la ciclasa. Preferentemente, en el caso del medio se trata de un medio de reacción acuoso. En el caso de los medios de reacción acuosos se trata preferentemente de soluciones tamponadas, que por regla general presentan un valor de pH de preferentemente de 5 a 8. Como tampón puede usarse un tampón de citrato, fosfato, TRIS (tris(hidroximetil)-aminometano), MES (ácido 2-(N-morfolino)etanosulfónico). Asimismo el medio de reacción puede contener también aditivos adicionales, tales como por ejemplo detergentes (por ejemplo taurodesoxicolato).

El sustrato (1) se emplea preferentemente en una concentración de 5 - 100 mM, de manera especialmente

preferente de 15 - 25 mM en la reacción enzimática y puede seguirse de manera continua o discontinua.

5 La ciclación enzimática tiene lugar por regla general a una temperatura de reacción por debajo de la temperatura de desactivación de la ciclasa empleada y por encima de -10 °C. De manera especialmente preferente se encuentra en el intervalo de 0 a 100 °C, en particular de 15 a 60 °C y en especial de 20 a 40 °C, por ejemplo a aproximadamente 30 °C.

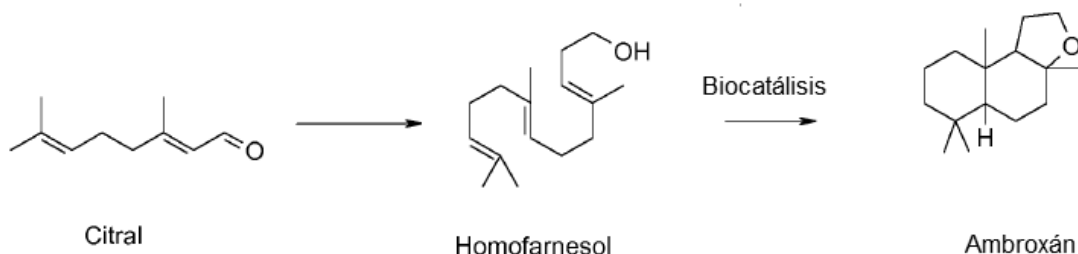
El producto de reacción ambroxán puede extraerse con disolventes orgánicos, seleccionados del grupo de los mencionados anteriormente, y dado el caso destilarse para la purificación.

10 Junto a estos sistemas acuosos de una sola fase, en otra variante de la invención se emplean también sistemas de dos fases. A este respecto se usan líquidos iónicos como segunda fase, pero preferentemente medios de reacción orgánicos, no miscibles con agua, como segunda fase. De esta manera los productos de reacción se acumulan en la fase orgánica. Después de la reacción, ambroxán en la fase orgánica puede separarse fácilmente de la fase acuosa, que contiene el biocatalizador.

15 Por medios de reacción no acuosos se entienden medios de reacción que contienen menos del 1 % en peso, preferentemente menos del 0,5 % en peso de agua, con respecto al peso total del medio de reacción líquido. En particular, la reacción puede llevarse a cabo en un disolvente orgánico.

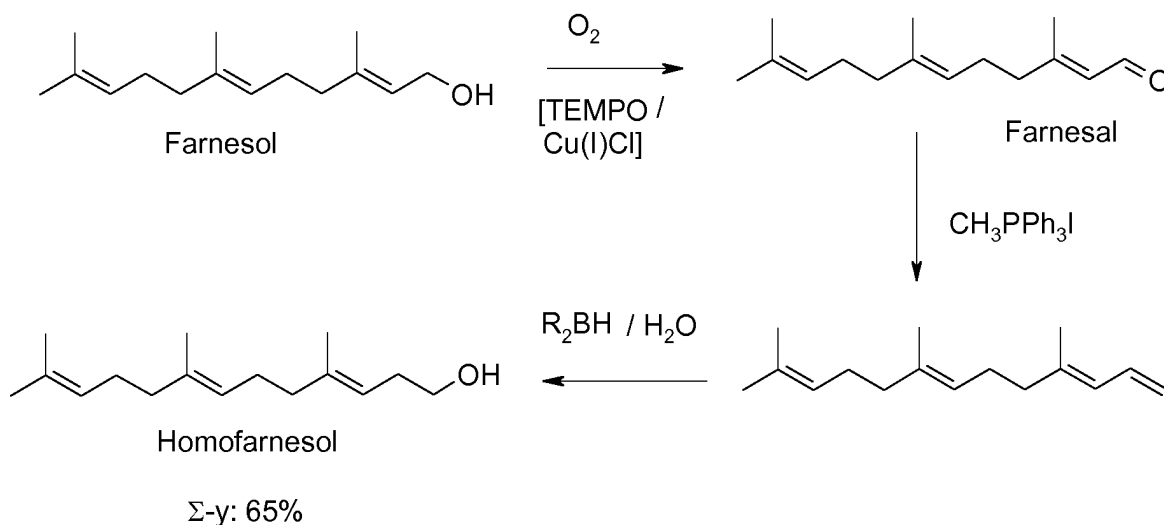
20 Disolventes orgánicos adecuados son por ejemplo hidrocarburos alifáticos, preferentemente con 5 a 8 átomos de carbono, tales como pentano, ciclopentano, hexano, ciclohexano, heptano, octano o ciclooctano, hidrocarburos alifáticos halogenados, preferentemente con uno o dos átomos de carbono, tales como diclorometano, cloroformo, tetraclorocarbono, dicloroetano o tetracloroetano, hidrocarburos aromáticos, tales como benceno, tolueno, los xilenos, clorobenceno o diclorobenceno, éteres alifáticos acíclicos y cíclicos o alcoholes, preferentemente con 4 a 8 átomos de carbono, tales como etanol, isopropanol, dietil éter, metil-terc-butil éter, etil-terc-butil éter, dipropil éter, diisopropil éter, dibutil éter, tetrahidrofurano o ésteres tales como acetato de etilo o acetato de n-butilo o cetonas tales como metilisobutil cetona o dioxano o mezclas de los mismos. De manera especialmente preferente se usan los mencionados anteriormente heptano, metil-terc-butil éter, diisopropil éter, tetrahidrofurano, acetato de etilo.

25 En una realización de la presente invención, como material de partida para la síntesis de ambroxán se emplea citral. Esto es una ventaja particular del procedimiento de acuerdo con la invención, dado que citral se encuentra disponible de manera económica en grandes cantidades. Citral se hace reaccionar en una síntesis química clásica para dar homofarnesol.



30 En una realización de la presente invención, la reacción de citral para dar homofarnesol tiene lugar en las siguientes etapas (por ejemplo: JOC, (1992), 57, 2794):

El procedimiento de acuerdo con la invención ofrece una ventaja adicional porque la reacción total de homofarnesol para dar ambroxán tiene lugar en sistemas acuosos de una sola fase pero también en sistemas de dos fases.



En el caso de los sistemas de dos fases se emplean los mencionados anteriormente. Preferentemente se usan los disolventes orgánicos, no miscibles con agua, mencionados anteriormente, como segunda fase. De esta manera los productos de reacción se acumulan en la fase orgánica. Después de la reacción, ambroxán en la fase orgánica puede separarse fácilmente de la fase acuosa, que contiene el biocatalizador.

También en la realización de la invención con citral como material de partida, se pone en contacto homofarnesol con la homofarnesol-ambroxán ciclasa en un medio y/o se incubaba de modo que tiene lugar una reacción de homofarnesol para dar ambroxán en presencia de la ciclasa. Preferentemente, en el caso del medio se trata de un medio de reacción acuoso. En el caso de los medios de reacción acuosos se trata preferentemente de soluciones tamponadas, que por regla general presentan un valor de pH de preferentemente de 5 a 8. Como tampón puede usarse un tampón de citrato, fosfato, TRIS (tris(hidroximetil)-aminometano), MES (ácido 2-(N-morfolino)etanosulfónico). Asimismo el medio de reacción puede contener también aditivos adicionales, tales como por ejemplo detergentes (taurodesoxicolato, u otros).

En una realización, el producto de reacción ambroxán se extraerá con disolventes orgánicos seleccionados del grupo de los mencionados anteriormente y dado el caso se destilarán para la purificación.

Asimismo, se divulga un procedimiento para la preparación biocatalítica de ambroxán, caracterizado porque la enzima es un polipéptido, que se codifica por una molécula de ácido nucleico que comprende al menos una molécula de ácido nucleico seleccionada del grupo que consiste en:

a) molécula de ácido nucleico que codifica un polipéptido que comprende la SEQ ID NO: 2, 5, 6, 7, 8, 9, 10 u 11;

b) molécula de ácido nucleico que comprende al menos un polinucleótido de la secuencia mostrada en SEQ ID NO: 1;

c) molécula de ácido nucleico que codifica un polipéptido cuya secuencia presenta una identidad de al menos el 46 % con respecto a las secuencias SEQ ID NO: 2, 5, 6, 7, 8, 9, 10 u 11;

d) molécula de ácido nucleico según (a) a (c), que para un polipéptido funcionalmente equivalente o un fragmento de la secuencia de acuerdo con la SEQ ID NO: 2, 5, 6, 7, 8, 9, 10 u 11;

e) molécula de ácido nucleico, que codifica para un polipéptido funcionalmente equivalente con la actividad de una homofarnesol-ambroxán ciclasa, que se obtiene amplificándose una molécula de ácido nucleico de un banco de ADNc o de un ADN genómico por medio de los cebadores de acuerdo con la n.º 3 y 4, o sintetizándose químicamente la molécula de ácido nucleico mediante síntesis *de novo*;

f) molécula de ácido nucleico, que codifica para un polipéptido funcionalmente equivalente con la actividad de una homofarnesol-ambroxán ciclasa, que en condiciones rigurosas híbrida con una molécula de ácido nucleico de acuerdo con (a) a (c);

g) molécula de ácido nucleico que codifica para un polipéptido funcionalmente equivalente con la actividad de una homofarnesol-ambroxán ciclasa, que puede aislarse a partir de un banco de ADN con el uso de una molécula de ácido nucleico de acuerdo con (a) a (c) o sus fragmentos parciales de al menos 15 nt, preferentemente 20 nt, 30 nt, 50 nt, 100 nt, 200 nt o 500 nt como sonda en condiciones de hibridación rigurosas; y

h) molécula de ácido nucleico, que codifica para un polipéptido funcionalmente equivalente con la actividad de

una homofarnesol-ambroxán ciclasa, presentando la secuencia del polipéptido una identidad de al menos el 46 % con respecto a las secuencias SEQ ID NO: 2, 5, 6, 7, 8, 9, 10 u 11;

5 i) molécula de ácido nucleico, que codifica para un polipéptido funcionalmente equivalente con la actividad de una homofarnesol-ambroxán ciclasa, donde el polipéptido se codifica por una molécula de ácido nucleico seleccionada del grupo de los descritos en a) a h) y se aisló o puede aislarse por medio de un anticuerpo monoclonal,

j) molécula de ácido nucleico, que codifica para un polipéptido funcionalmente equivalente con la actividad de una homofarnesol-ambroxán ciclasa, presentando el polipéptido un sitio de unión análogo o similar, tal como un polipéptido codificado por una molécula de ácido nucleico seleccionada del grupo de los descritos en a) a h).

10 Un sitio de unión análogo o similar en el sentido anterior es un motivo o dominio conservado de la secuencia de aminoácidos con una homología del 80 %, de manera especialmente preferente del 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, en particular 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % o 100 %, que garantiza la unión del mismo sustrato, en particular homofarnesol.

15 Preferentemente, la molécula de ácido nucleico c) presenta una identidad con la SEQ ID NO: 1 de al menos el 46 %, 47 %, 48 %, 49 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, de manera especialmente preferente del 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, en particular 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 %.

Asimismo, un polipéptido funcionalmente equivalente presenta una identidad con la SEQ ID NO: 2 de al menos el 46 %, 47 %, 48 %, 49 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, de manera especialmente preferente del 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, en particular 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 %.

20 En lugar del término "identidad" puede usarse con el mismo significado también el término "homólogo" u "homología".

La identidad entre dos secuencias de ácido nucleico o secuencias de polipéptido se calcula mediante comparación con ayuda del programa BESTFIT que se basa en el algoritmo de Smith, T.F. y Waterman, M.S (Adv. Appl. Math. 2: 482-489 (1981)).

25 Preferentemente, la identidad entre dos secuencias de ácido nucleico o secuencias de polipéptido se define por la identidad de la secuencia de ácido nucleico/secuencia de polipéptido a lo largo de la longitud de secuencia total en cada caso, tal como se calcula mediante comparación con ayuda del programa GAP que se basa en el algoritmo de Needleman, S.B. y Wunsch, C.D. (J. Mol. Biol. 48: 443-453)).

30 Preferentemente, las comparaciones de identidad se llevan a cabo con el ajuste de los siguientes parámetros para aminoácidos:

penalización de creación de huecos: 8; penalización de extensión de huecos: 2
y los siguientes parámetros para ácidos nucleicos:

penalización de creación de huecos: 50; penalización de extensión de huecos: 3.

35 En una realización, la identidad entre dos secuencias de ácido nucleico o secuencias de polipéptido se propone mediante comparación con ayuda del programa BLASTP 2.2.20+ con ajustes convencionales tal como en NCBI Blast (Referencia: Altschul et al., (1997), Nucleic Acids Res. 25:3389-3402. Reference for compositional score matrix adjustment: Altschul et al., (2005), FEBS J. 272:5101-5109.)

Asimismo se divulgan otros homólogos o equivalentes funcionales de la SEQ ID NO: 1, que hibridan con esta secuencia de ácido nucleico en condiciones rigurosas.

40 Los "equivalentes funcionales" describen en principio en este caso secuencias de ácido nucleico que en condiciones estándar hibridan con una secuencia de ácido nucleico o partes de una secuencia de ácido nucleico y son capaces de provocar la expresión de una proteína con las mismas propiedades que las de la homofarnesol-ambroxán ciclasa en una célula o un organismo.

45 Para la hibridación se usan ventajosamente oligonucleótidos cortos con una longitud de aproximadamente 10-50 pb, preferentemente 15-40 pb por ejemplo de las regiones conservadas u otras, que pueden determinarse a través de comparaciones con otros genes relacionados de manera conocida por el experto en la materia. Pueden usarse también fragmentos más largos de los ácidos nucleicos de acuerdo con la invención con una longitud de 100-500 pb o las secuencias completas para la hibridación. En función del ácido nucleico/oligonucleótido usado, de la longitud del fragmento o de la secuencia completa o en función de qué tipo de ácido nucleico, es decir, ADN o ARN, se usan para la hibridación, varían las condiciones estándar. De este modo, por ejemplo las temperaturas de fusión para ADN:hibridos de ADN es aproximadamente 10 °C menor que la de ADN:hibridos de ARN de igual longitud.

50 Por condiciones de hibridación estándar se entienden por ejemplo en función del ácido nucleico temperaturas entre 42 y 58 °C en una solución tampón acuosa con una concentración entre 0,1 y 5 x SSC (1 X SSC = NaCl 0,15 M, citrato de sodio 15 mM, pH 7,2) o adicionalmente en presencia del 50 % de formamida tal como por ejemplo 42 °C en 5 x SSC, 50 % de formamida. Ventajosamente, las condiciones de hibridación para ADN:hibridos de ADN son 0,1

x SSC y temperaturas entre aproximadamente 20 °C y 65 °C, preferentemente entre aproximadamente 30 °C y 45 °C. Para ADN:híbridos de ARN, las condiciones de hibridación son ventajosamente 0,1 x SSC y temperaturas entre aproximadamente 30 °C y 65 °C, preferentemente entre aproximadamente 45 °C y 55 °C. Estas temperaturas indicadas para la hibridación son valores de temperatura de fusión calculados a modo de ejemplo para un ácido nucleico con una longitud de aproximadamente 100 nucleótidos y un contenido de G + C del 50 % en ausencia de formamida. Las condiciones experimentales para la hibridación de ADN se describen en libros de texto especializados de Genética, tales como por ejemplo Sambrook et al., "Molecular Cloning", Cold Spring Harbor Laboratory, 1989, y pueden calcularse según fórmulas conocidas por el experto por ejemplo en función de la longitud de los ácidos nucleicos, del tipo de híbridos o del contenido de G + C. El experto puede extraer informaciones adicionales para la hibridación de los siguientes libros de texto: Ausubel et al. (eds), 1985, "Current Protocols in Molecular Biology", John Wiley & Sons, Nueva York; Harnes and Higgins (eds), 1985, "Nucleic Acids Hybridization: A Practical Approach", IRL Press at Oxford University Press, Oxford; Brown (ed), 1991, Essential Molecular Biology: A Practical Approach, IRL Press at Oxford University Press, Oxford.

Por un equivalente funcional se entienden además también secuencias de ácido nucleico que con una secuencia de ácido nucleico determinada ("secuencia de ácido nucleico original) son homólogas o idénticas hasta un porcentaje definido y que presentan la misma actividad que las secuencias de ácido nucleico originales, asimismo en particular también mutaciones naturales o artificiales de estas secuencias de ácido nucleico.

"Equivalentes funcionales" o análogos de las enzimas divulgadas en concreto son en el contexto de la presente divulgación polipéptidos distintos de los mismos, que tienen además la actividad biológica deseada, tal como por ejemplo actividad principal, especificidad de sustrato. De este modo, por ejemplo por "equivalentes funcionales" se entienden enzimas que catalizan la reacción modelo y que presentan al menos el 20 %, 30 %, 50 %, 60 %, 65 %, 70 %, 75 %, 80 %, de manera especialmente preferente del 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, en particular 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de la actividad de una enzima, que comprende una de las secuencias de aminoácidos expuestas en las SEQ ID NO: 2, 5, 6, 7, 8, 9, 10 u 11.

Por "equivalentes funcionales" se entienden en particular también mutantes, que en al menos una posición de secuencia de las secuencias de aminoácidos mencionadas anteriormente presentan un aminoácido distinto del aminoácido mencionado en concreto pero, a pesar de ello, tienen una de las actividades biológicas mencionadas anteriormente. Los "equivalentes funcionales" abarcan por lo tanto los mutantes que pueden obtenerse mediante una o varias adiciones, sustituciones, deleciones y/o inversiones de aminoácido, pudiendo aparecer las modificaciones mencionadas en cualquier posición de secuencia, siempre que lleven a un mutante con el perfil de propiedades divulgado en este caso. Equivalencia funcional se da en particular también entonces cuando los patrones de reactividad entre mutante y polipéptido no modificado coinciden cualitativamente, es decir, por ejemplo se hacen reaccionar sustratos iguales con diferente velocidad.

Ejemplos de sustituciones de aminoácido adecuadas se desprenden de la siguiente Tabla:

Reto original	Ejemplos de sustitución
Ala	Ser
Arg	Lys
Asn	Gln; His
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Pro
His	Asn; Gln
Ile	Leu; Val
Leu	Ile; Val
Lys	Arg; Gln; Glu
Met	Leu; Ile
Phe	Met; Leu; Tyr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp; Phe
Val	Ile; Leu

"Equivalentes funcionales" en el sentido anterior son también "precursores" de los polipéptidos descritos así como "derivados funcionales".

"Precursores" son a este respecto etapas previas naturales o sintéticas de los polipéptidos con o sin la actividad biológica deseada.

"Derivados funcionales" de los polipéptidos puede prepararse asimismo en grupos laterales de aminoácido

funcionales o en sus extremos N o C terminales con ayuda de técnicas conocidas. Los derivados de este tipo comprenden por ejemplo ésteres alifáticos de grupos ácido carboxílico, amidas de grupos ácido carboxílico, que pueden obtenerse mediante reacción con amoníaco o con una amina primaria o secundaria; derivados de N-acilo de grupos amino libres, preparados mediante reacción con grupos acilo; o derivados de O-acilo de grupos hidroxilo libres, preparados mediante reacción con grupos acilo.

En el caso de una posible glicosilación de proteína, los "equivalentes funcionales" abarcan proteínas del tipo designado anteriormente en forma desglucosilada o glicosilada así como formas diferentes que pueden obtenerse mediante modificación del patrón de glicosilación.

Los "equivalentes funcionales" abarcan naturalmente también polipéptidos a los que puede accederse a partir de otros organismos, así como variantes que existen naturalmente. Por ejemplo, mediante comparación de secuencias puede establecerse zonas de regiones de secuencia homólogas y siguiendo las especificaciones de la invención pueden averiguarse enzimas equivalentes. Los "equivalentes funcionales" abarcan asimismo fragmentos, preferentemente dominios o motivos de secuencia individuales, de los polipéptidos de acuerdo con la invención, que presentan por ejemplo la función biológica deseada.

Los "equivalentes funcionales" son además proteínas de fusión que presentan una de las secuencias de polipéptido mencionadas anteriormente o equivalentes funcionales derivados de las mismas y al menos una secuencia heteróloga adicional, funcionalmente distinta de la misma, en enlace funcional N o C terminal (es decir, sin perjuicio funcional esencial mutuo de las partes de proteína de fusión). Ejemplos no limitativos de secuencias heterólogas de este tipo son por ejemplo péptidos señal o enzimas.

Homólogos de las proteínas divulgadas en este caso pueden identificarse mediante selección de bancos combinatorios de mutantes, tales como por ejemplo mutantes de acortamiento. Por ejemplo puede generarse un banco variado de variantes de proteína mediante mutagénesis combinatoria en el plano de ácido nucleico, tal como por ejemplo mediante ligación enzimática de una mezcla de oligonucleótidos sintéticos. Existe una pluralidad de procedimientos que pueden usarse para la preparación de bancos de homólogos potenciales a partir de una secuencia de oligonucleótidos degenerada. La síntesis química de una secuencia génica generada puede llevarse a cabo en un sintetizador automático de ADN, y el gen sintético puede ligarse entonces en un vector de expresión adecuado. El uso de un conjunto de genes degenerado permite la provisión de todas las secuencias en una mezcla, que codifican el conjunto deseado de secuencias de proteína potenciales. Procedimientos para la síntesis de oligonucleótidos degenerados son conocidos por el experto en la materia (por ejemplo Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura et al. (1984) *Annu. Rev. Biochem.* 53:323; Itakura et al., (1984) *Science* 198:1056; Ike et al. (1983) *Nucleic Acids Res.* 11:477).

En el estado de la técnica son conocidas varias técnicas para la selección de productos génicos de bancos combinatorios, que se han producido mediante mutaciones puntuales o acortamiento, y para la selección de bancos de ADNc sobre productos génicos con una propiedad seleccionada. Estas técnicas pueden adaptarse a la selección rápida de los bancos de genes, que se han generado mediante mutagénesis combinatoria de homólogos de acuerdo con la invención. Las técnicas usadas con la mayor frecuencia para la selección de grandes bancos de genes, que están sujetas a un análisis con alto rendimiento, comprenden la clonación del banco de genes en vectores de expresión replicables, transformación de células adecuadas con el banco de vectores resultante y expresión de los genes recombinantes en condiciones en las que la detección de la actividad deseada facilita el aislamiento del vector que codifica el gen, cuyo producto se detectó. La Recursive-Ensemble-Mutagenese (REM) (mutagénesis de conjunto recursiva), una técnica que amplía la frecuencia de mutantes funcionales en los bancos, puede usarse en combinación con las pruebas de selección para identificar homólogos (Arkin y Yourvan (1992) *PNAS* 89:7811-7815; Delgrave et al. (1993) *Protein Engineering* 6(3):327-331).

Se divulgan además secuencias de ácido nucleico (secuencias de ADN y ARN mono- y bicatenarios, tales como por ejemplo ADNc y ARNm), que codifican para una enzima con actividad ciclasa de acuerdo con la invención. Se prefieren secuencias de ácido nucleico que codifican por ejemplo para secuencias de aminoácidos de acuerdo con las SEQ ID NO: 2, 5, 6, 7, 8, 9, 10 u 11 secuencias parciales características de las mismas.

Todas las secuencias de ácido nucleico mencionadas en el presente documento pueden prepararse de manera en sí conocida mediante síntesis química a partir de los elementos constructivos de nucleótido, tales como por ejemplo mediante condensación de fragmentos de elementos constructivos de ácido nucleico complementarios, solapantes, individuales, de la doble hélice. La síntesis química de oligonucleótidos puede tener lugar por ejemplo, de manera conocida, según el procedimiento de fosfoamidita (Voet, Voet, 2ª edición, Wiley Press New York, páginas 896-897). La adición de oligonucleótidos sintéticos y la compleción de huecos con ayuda del fragmento de la ADN polimerasa y reacciones de ligación así como procedimientos de clonación generales se describen en Sambrook et al. (1989), *Molecular Cloning: A laboratory manual*, Cold Spring Harbor Laboratory Press.

Otras configuraciones para llevar a cabo el procedimiento biocatalítico de acuerdo con la invención para la preparación de ambroxán:

El procedimiento de acuerdo con la invención se caracteriza porque la enzima se encuentra en una forma seleccionada del grupo que consiste en:

a) polipéptido libre, dado el caso purificado o parcialmente purificado con la actividad de una homofarnesol-ambroxán ciclasa;

b) polipéptido inmovilizado con la actividad de una homofarnesol-ambroxán ciclasa;

c) polipéptido aislado a partir de células de acuerdo con a) o b);

5 d) célula completa, dado el caso células quiescentes o disgregadas, que contiene al menos un polipéptido con la actividad de una homofarnesol-ambroxán ciclasa;

e) lisado celular u homogeneizado celular de las células descritas en d).

10 En una realización de la invención, las células son microorganismos, preferentemente microorganismos transgénicos que expresan al menos una molécula heteróloga de ácido nucleico que codifica para un polipéptido con la actividad de una homofarnesol-ambroxán ciclasa.

Además se divulga por lo tanto también un constructo génico o un vector que comprende una molécula de ácido nucleico que codifica para un polipéptido con la actividad de una homofarnesol-ambroxán ciclasa, preferentemente una molécula de ácido nucleico que comprende al menos una molécula de ácido nucleico seleccionada del grupo que consiste en:

15 a) molécula de ácido nucleico que codifica un polipéptido que comprende la SEQ ID NO: 2, 5, 6, 7, 8, 9, 10 u 11;

b) molécula de ácido nucleico que comprende al menos un polinucleótido de la secuencia mostrada en SEQ ID NO: 1;

c) molécula de ácido nucleico que codifica un polipéptido cuya secuencia presenta una identidad de al menos el 46 % con respecto a las secuencias SEQ ID NO: 2, 5, 6, 7, 8, 9, 10 u 11;

20 d) molécula de ácido nucleico según (a) a (c), que para un fragmento de las secuencias de acuerdo con las SEQ ID NO: 2, 5, 6, 7, 8, 9, 10 u 11;

e) molécula de ácido nucleico, que codifica para un polipéptido con la actividad de una homofarnesol-ambroxán ciclasa, que se obtiene amplificándose una molécula de ácido nucleico de un banco de datos de ADNc o de un banco de datos de genoma por medio de los cebadores de acuerdo con las secuencias n.º 3 y 4;

25 f) molécula de ácido nucleico, que codifica para un polipéptido con la actividad de una homofarnesol-ambroxán ciclasa, que en condiciones rigurosas hibrida con una molécula de ácido nucleico de acuerdo con (a) a (c);

30 g) molécula de ácido nucleico que codifica para un polipéptido con la actividad de una homofarnesol-ambroxán ciclasa, que puede aislarse a partir de un banco de ADN con el uso de una molécula de ácido nucleico de acuerdo con (a) a (c) o sus fragmentos parciales de al menos 15 nt, preferentemente 20 nt, 30 nt, 50 nt, 100 nt, 200 nt o 500 nt como sonda en condiciones de hibridación rigurosas; y

h) molécula de ácido nucleico, que codifica para un polipéptido con la actividad de una homofarnesol-ambroxán ciclasa, presentando la secuencia del polipéptido una identidad de al menos el 46 % con respecto a las secuencias SEQ ID NO: 2, 5, 6, 7, 8, 9, 10 u 11;

35 i) molécula de ácido nucleico, que codifica para un polipéptido funcionalmente equivalente con la actividad de una homofarnesol-ambroxán ciclasa, donde el polipéptido se codifica por una molécula de ácido nucleico seleccionada del grupo de los descritos en a) a h) y se aisló o puede aislarse por medio de un anticuerpo monoclonal,

40 j) molécula de ácido nucleico, que codifica para un polipéptido funcionalmente equivalente con la actividad de una homofarnesol-ambroxán ciclasa, presentando el polipéptido un sitio de unión análogo o similar, tal como un polipéptido codificado por una molécula de ácido nucleico seleccionada del grupo de los descritos en a) a h).

Asimismo, en este caso se divulgan asimismo las células huésped que contienen un constructo génico o un vector tal como se describe anteriormente. Para ello se introducen las secuencias de ácido nucleico usadas de manera ventajosa en un constructo génico transgénico, que pueden garantizar una expresión transgénica de una homofarnesol-ambroxán ciclasa, en un organismo, preferentemente microorganismo.

45 En los constructos génicos, a este respecto, una molécula de ácido nucleico que codifica para una homofarnesol-ambroxán-ciclasa se encuentra preferentemente en enlace funcional con al menos un elemento control genético (por ejemplo un promotor y/o terminador), que garantiza una expresión en un organismo, preferentemente microorganismo.

50 Por un enlace funcional se entiende por ejemplo la disposición secuencial de un promotor con la secuencia de ácido nucleico que va a expresarse que codifica para una homofarnesol-ambroxán ciclasa (por ejemplo la secuencia de

acuerdo con la SEQ ID NO: 1) y dado el caso otros elementos reguladores tales como por ejemplo un terminador, de tal manera que cada uno de los elementos reguladores pueda cumplir su función en la expresión transgénica de la secuencia de ácido nucleico. Para ello no es realmente necesario un enlace directo en el sentido químico. La preparación de un enlace funcional como también la preparación del constructo genómico puede realizarse por medio de técnicas de recombinación y clonación usuales, tal como se describen por ejemplo en Maniatis T, Fritsch EF y Sambrook J (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor (NY), en Silhavy TJ, Berman ML y Enquist LW (1984) *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor (NY), en Ausubel FM et al. (1987). Entre ambas secuencias pueden situarse en cambio también otras secuencias que tienen por ejemplo la función de un ligador con determinados sitios de corte de enzima de restricción o de un péptido señal. También la inserción de secuencias puede llevar a la expresión de proteínas de fusión. Preferentemente, el constructo génico, que se compone de un enlace de promotor y secuencia de ácido nucleico que va a expresarse, puede encontrarse integrado en un vector e insertarse mediante por ejemplo transformación en un microorganismo.

Las secuencias de ácido nucleico contenidas en los constructos génicos o vectores pueden estar enlazadas funcionalmente con otras secuencias de control genéticas junto a un promotor. El término de las secuencias control genéticas ha de entenderse en un sentido amplio y quiere decir todas aquellas secuencias que tienen una influencia sobre la conclusión o la función del constructo genómico. Las secuencias control genéticas modifican por ejemplo la transcripción y traducción en organismos procariontes o eucariotas. Preferentemente, los constructos génicos comprenden una secuencia control, así como dado el caso otros elementos reguladores habituales, en concreto en cada caso enlazados funcionalmente con la secuencia de ácido nucleico que va a expresarse de manera transgénica.

Como secuencias control se entienden aquellas que posibilitan una recombinación homóloga o inserción en el genoma de un organismo huésped o permiten la eliminación del genoma. En la recombinación homóloga puede intercambiarse de manera dirigida por ejemplo la secuencia codificante de un gen endógeno determinado por la secuencia que codifica para un ARNbc.

Un constructo génico y los vectores derivados del mismo pueden contener otros elementos funcionales. El término elemento funcional ha de entenderse en un sentido amplio y quiere decir todos aquellos elementos que tienen una influencia sobre la preparación, proliferación o función de los casetes de expresión de acuerdo con la invención, vectores u organismos transgénicos. A modo de ejemplo pero de manera no limitativa pueden mencionarse:

a) Marcadores de selección, que confieren una resistencia contra antibióticos o biocidas tales como por ejemplo kanamicina, G 418, bleomicina, higromicina, etc.

b) Genes indicadores, que codifican para proteínas fácilmente cuantificables y a través de color propio o actividad enzimática garantizan una evaluación de la eficiencia de transformación o del sitio o instante de expresión. Se prefieren muy especialmente a este respecto proteínas indicadoras (Schenborn E, Groskreutz D. *Mol Biotechnol.* 1999; 13(1):29-44) tal como la "green fluorescence protein" (GFP) (proteína fluorescente verde) (Sheen et al. (1995) *Plant Journal* 8(5):777-784), la cloranfenicoltransferasa, una luciferasa (Ow et al. (1986) *Science* 234:856-859), el gen de eucorina (Prasher et al. (1985) *Biochem Biophys Res Commun* 126(3):1259-1268), la beta-galactosidasa, se prefiere muy especialmente la β -glucuronidasa (Jefferson et al. (1987) *EMBO J* 6:3901-3907).

c) Orígenes de la replicación, que garantizan una propagación de los casetes de expresión de acuerdo con la invención o vectores en por ejemplo *E.coli*. A modo de ejemplo se mencionan ORI (origin of DNA replication) (origen de replicación de ADN), el pBR322 ori o el P15A ori (Sambrook et al.: *Molecular Cloning. A Laboratory Manual*, 2ª ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Para la selección de células recombinadas de manera homóloga o también transformadas de manera satisfactoria es necesario por regla general introducir adicionalmente un marcador seleccionable que confiera a las células recombinadas de manera satisfactoria una resistencia contra un biocida o un antibiótico. El marcador de selección permite la selección de las células transformadas con respecto a las no transformadas (McCormick et al. (1986) *Plant Cell Reports* 5:81-84).

La introducción del constructo génico en un organismo puede realizarse de manera ventajosa con el uso de vectores, en los que están contenidos los constructos génicos. Otro objeto de la invención se refiere por lo tanto a dichos vectores transgénicos, que comprenden un constructo génico transgénico para una homofarnesol-ambroxán ciclasa.

Los vectores pueden ser a modo de ejemplo plásmidos, cósmidos, fagos, virus. El constructo génico puede introducirse en el vector (preferentemente un vector de plásmido) a través de un sitio de corte de restricción adecuado. El vector generado se introduce en primer lugar en *E.coli*. Se seleccionan *E.coli* correctamente transformadas, se cultivan y se obtiene el vector recombinante con procedimientos conocidos por el experto en la materia. Análisis de restricción y secuenciación pueden servir para comprobar la etapa de clonación. Se prefieren aquellos vectores que permiten una integración estable del casete de expresión en el genoma huésped.

La preparación de un organismo transgénico correspondiente tiene lugar por ejemplo por medio de transformación o transfección por medio de las proteínas o ácidos nucleicos correspondientes. La preparación de un organismo transformado (o de una célula transformada) requiere que el ADN correspondiente (por ejemplo el vector de expresión), ARN o proteína se introduzca en la célula huésped correspondiente. Para este proceso, que se denomina transformación (o transducción o transfección), se encuentra disponible una pluralidad de procedimientos (Keown et al. (1990) *Methods in Enzymology* 185:527-537). De este modo, el ADN o ARN puede introducirse a modo de ejemplo directamente mediante microinyección o mediante bombardeo con micropartículas recubiertas con ADN. También la célula puede permeabilizarse químicamente, por ejemplo con polietilenglicol, de modo que el ADN puede llegar mediante difusión a la célula. El ADN puede realizarse también mediante fusión de protoplastos con otras unidades que contienen ADN tales como minicélulas, células, lisosomas o liposomas. La electroporación es un procedimiento adecuado adicional para la introducción de ADN, en el que las células se permeabilizan de manera reversible mediante un impulso eléctrico.

Células transformadas de manera estable, es decir, aquellas que contienen el ADN introducido integrado en el ADN de la célula huésped, pueden seleccionarse con respecto a las no transformadas, cuando un marcador seleccionable es un constituyente del ADN introducido. Como marcador puede funcionar a modo de ejemplo cualquier gen que es capaz de conferir una resistencia contra antibióticos (tales como kanamicina, G 418, bleomicina, higromicina, etc.) (véase anteriormente). Las células transformadas, que expresan un gen marcador de este tipo, son capaces de sobrevivir en presencia de concentraciones de un antibiótico correspondiente, que destruyen un tipo natural no transformado. 89:525-533 usa.

"Transgén" o "recombinante" significa con respecto, por ejemplo, a una secuencia de ácido nucleico, un constructo génico o un vector que contiene dicha secuencia de ácido nucleico o un organismo transformado con dicha secuencia de ácido nucleico, constructo génico o vector, todas aquellas construcciones realizadas mediante procedimientos de ingeniería genética, en los que o bien

- a) la secuencia de ácido nucleico que codifica para una homofarnesol-ambroxán ciclasa, o
- b) una secuencia control genética enlazada funcionalmente con dicha secuencia de ácido nucleico en a), por ejemplo un promotor funcional, o
- c) (a) y (b)

no se encuentran en su entorno genético natural o se modificaron mediante procedimientos de ingeniería genética, en los que la modificación a modo de ejemplo puede ser sustituciones, adiciones, deleciones, inversión o inserciones de uno o varios restos de nucleótido. Un entorno genético natural significa el locus cromosómico natural en el organismo de origen o la presencia en una biblioteca genómica. En el caso de una biblioteca genómica se mantiene preferentemente al menos en parte el entorno genético natural de la secuencia de ácido nucleico.

Organismos huésped o de partida preferidos como organismos transgénicos son principalmente microorganismos de acuerdo con la definición mencionada anteriormente. Están incluidos en el contexto de la invención en particular microorganismos, preferentemente aquellas células transgénicas o recombinantes, que se seleccionan de bacterias, cianobacterias, hongos y levaduras. Preferentemente la célula se selecciona de bacterias de los géneros *Escherichia*, *Corynebacterium*, *Ralstonia*, *Clostridium*, *Pseudomonas*, *Bacillus*, *Zymomonas*, *Rhodobacter*, *Streptomyces*, *Burkholderia*, *Lactobacillus* y *Lactococcus*. De manera especialmente preferente, la célula se selecciona de bacterias de las especies *Escherichia coli*, *Pseudomonas putida*, *Burkholderia glumae*, *Streptomyces lividans*, *Streptomyces coelicolor* y *Zymomonas mobilis*.

Como organismo donante, es decir, organismo a partir del cual se aísla una homofarnesol-ambroxán ciclasa, se mencionan *Methylococcus capsalatus*, *Rhodopseudomonas palustris*, *Bradyrhizobium japonicum*, *Frankia spec.*, *Streptomyces coelicolor*, *Rhodopseudomonas palustris*, *Rhodopseudomonas palent*, *Frankia alni*, *Bacillus anthracis*, *Burkholderia ambifaria*, en particular *Zymomonas mobilis* y *Bradyrhizobium japonicum*.

Los organismos transgénicos, en particular *Streptomyces coelicolor* o *Zymomonas mobilis*, que presentan homofarnesol-ambroxán ciclasas endógenas, muestran en una variante de la invención una sobreexpresión de las homofarnesol-ambroxán ciclasas. Sobreexpresión significa cualquier forma de expresión de las homofarnesol-ambroxán ciclasas que puede encontrarse además de la expresión original del tipo natural.

Asimismo se divulga la preparación de *E. coli* transgénicas, que expresan el gen que codifica para las homofarnesol-ambroxán ciclasas, preferentemente SEQ ID NO: 1. Para ello se amplifica con cebadores, por ejemplo Zm-SHC_fw y Zm-SHC_rev (SEQ ID NO: 3 o 3) el gen de la ciclasa, preferentemente de *Zymomonas mobilis* (SEQ ID NO: 1). Los cebadores se mezclan de manera equimolar. Las PCR con ADN genómico de un organismo donante, preferentemente de *Z. mobilis* (LU8910 = ATCC31821) se llevan a cabo según instrucciones del fabricante con Pwo-Polymerase (Roche Applied Science) y el siguiente programa de gradiente de temperatura: 95 °C durante 3 min; 30 ciclos con 95 °C durante 30 s, 50 °C durante 30 s, y 72 °C durante 3 min; 72 °C durante 10 min.; 4 °C hasta su uso. El producto de PCR se aísla mediante electroforesis en gel de agarosa (Gel E al 1,2 %, Invitrogen) y cromatografía en columna (Kit GFX, Amersham Pharmacia) y a continuación se secuencian (cebador de secuenciación: Zm-SHC_fw y Zm-SHC_rev). El producto de PCR se digirió con las endonucleasas de restricción, preferentemente *NdeI* y *BamHI*, y se liga en un vector cortado de manera correspondiente, preferentemente vector pDHE1650 [documento WO

ES 2 703 770 T3

200132890 A1]. El plásmido así obtenido se transforma en *E. coli*, preferentemente en la cepa TG10 pAgro4 de *E. coli* [Takeshita S. et al.; Gene (1987), 61: 63] pHSG575.

5 Inoculadas a partir de un cultivo previo correspondiente, se cultivan las *E. coli*, preferentemente se fijan en LBamp/Spec/Cm (ampicilina 100 µg/l; espectinomicina 100 µg/l; cloranfenicol 20 µg/l), IPTG 0,1 mM, ramnosa 0,5 g/l en 16 h a 37 °C, y a continuación se centrifugan a 5000*g/10 min y dado el caso se almacenan a 4 °C.

En otra realización, se aísla la ciclasa a partir del organismo donante o a partir del organismo huésped transgénico y dado el caso se purifica.

10 A partir de las células se prepara un extracto de proteína suspendiéndose el sedimento celular en tampón de disgregación (Tris/HCl 0,2 M, EDTA 0,5 M, pH 8,0), 375 U de benzonasa (por ejemplo Novagen, 25 U/µl), 40 µl de PMSF (100 mM, disuelto en *i*-PropOH), sacarosa 5,3 g/100 ml y lisozima aproximadamente 3,3 mg/100 ml. La preparación se mezcla y se incuba durante 30 min sobre hielo. A continuación se congela dado el caso la mezcla a -20 °C. Después de la descongelación de la preparación se rellena con agua destilada y se incuba de nuevo durante 30 min sobre hielo. A continuación se disgregan las células 3 veces durante 3 min con ultrasonidos.

15 Después de la disgregación se separaron por centrifugación los detritos celulares en 60 min a 4 °C y 26900 *g. El sobrenadante se retira y el sedimento se resuspende en tampón de solubilización (Tris/HCl 50 mM, MgCl₂x6H₂O 10 mM, Triton X-100 al 1 %, pH 8,0) y se homogeneiza durante aproximadamente 5 min, por ejemplo con un Potter. A continuación se mantiene la suspensión durante 30 min sobre hielo. El extracto homogeneizado se centrifuga de nuevo durante 1 h a 4 °C y 26900 *g y se retira el sedimento. El extracto puede emplearse para pruebas enzimáticas y puede almacenarse durante varias semanas sin pérdida de actividad a -20 °C. El contenido de proteína se encuentra en el intervalo de 1 mg /ml

20 Las ciclasas empleadas de acuerdo con la invención pueden usarse como enzima libre o inmovilizada.

25 Las ciclasas usadas de acuerdo con la invención pueden emplearse de manera libre o inmovilizada. Por una enzima inmovilizada se entiende una enzima que está fijada a un soporte inerte. Materiales de soporte adecuados así como las enzimas inmovilizadas sobre los mismos son conocidos por el documento EP-A-1149849, el documento EP-A-1 069 183 y el documento DE-OS 100193773 así como por las citas bibliográficas citadas en los mismos. A este respecto se hace referencia a la divulgación de estos documentos en su totalidad. A los materiales de soporte adecuados pertenecen por ejemplo arcillas, minerales de arcilla, tales como caolinita, tierras de diatomeas, perlita, dióxido de silicio, óxido de aluminio, carbonato de sodio, carbonato de calcio, polvos de celulosa, materiales de intercambio aniónico, polímeros sintéticos, tales como poliestireno, resinas acrílicas, resinas de fenol-formaldehído, poliuretanos y poliolefinas, tales como polietileno y polipropileno. Los materiales de soporte se emplean para la preparación de las enzimas soportadas habitualmente en una forma particulada, finamente dividida, prefiriéndose formas porosas. El tamaño de partícula del material de soporte asciende habitualmente a no más de 5 mm, en particular no más de 2 mm (curva granulométrica). De manera análoga, en el caso del uso de la ciclasa como catalizador de célula completa, puede seleccionarse una forma libre o inmovilizada. Materiales de soporte son por ejemplo alginato de Ca, y carragenano. Enzimas como también células pueden reticularse también directamente con glutaraldehído (reticulación para dar CLEA). Procedimientos de inmovilización correspondientes y adicionales se describen por ejemplo en J. Lalonde y A. Margolin "Immobilization of Enzymes" en K. Drauz y H. Waldmann, Enzyme Catalysis in Organic Synthesis 2002, Vol.III, 991-1032, Wiley-VCH, Weinheim.

35 Pueden usarse además células completas, dado el caso células quiescentes o disgregadas, lisado celular o homogeneizado celular.

40 Para el procedimiento de acuerdo con la invención pueden usarse células en crecimiento, que contienen ácidos nucleicos que codifican para la ciclasa, constructos de ácido nucleico o vectores. También pueden usarse células quiescentes o disgregadas. Por células disgregadas se entienden por ejemplo células que se han hecho permeables a través de un tratamiento con por ejemplo disolventes, o células que se han roto a través de un tratamiento enzimático, a través de un tratamiento mecánico (por ejemplo prensa francesa o ultrasonidos) o a través de otro procedimiento. Los extractos brutos así obtenidos son adecuados de manera ventajosa para el procedimiento de acuerdo con la invención. También pueden usarse para el procedimiento enzimas purificadas o no purificadas. Asimismo son adecuados microorganismos inmovilizados o enzimas que pueden emplearse ventajosamente en la reacción.

45 El procedimiento de acuerdo con la invención puede hacerse funcionar de manera discontinua, semi-discontinua o continua.

La reacción puede llevarse a cabo tanto en un modo de proceder discontinuo como también en alimentación discontinua.

El producto se purifica a continuación mediante extracción y/o destilación.

55 Otro objeto de la presente invención es el uso de un polipéptido con la actividad de una homofarnesol-ambroxán ciclasa para la reacción biocatalítica de homofarnesol para dar ambroxán, caracterizado porque el polipéptido se

codifica por una molécula de ácido nucleico que comprende al menos una molécula de ácido nucleico seleccionada del grupo que consiste en:

- a) molécula de ácido nucleico que codifica un polipéptido que comprende la SEQ ID NO: 2, 5 o 10;
- 5 b) molécula de ácido nucleico que comprende al menos un polinucleótido de la secuencia mostrada en SEQ ID NO: 1;
- c) molécula de ácido nucleico que codifica un polipéptido cuya secuencia presenta una identidad de al menos el 60 % con respecto a las secuencias SEQ ID NO: 2, 5 o 10;

Los siguientes Ejemplos ilustrarán la invención, pero sin limitarla. En este sentido se hace referencia a las Figuras adjuntas, a este respecto muestra:

10 Parte experimental

Ejemplos

Ejemplo 1 Clonación de Zm-SHC y expresión en *E. coli*

Con los oligonucleótidos Zm-SHC_fw y Zm-SHC_rev puede amplificarse el gen de la ciclasa a partir de *Zymomonas mobilis*.

15 Cebador:

N.º de cebador	Secuencia (5'->3')	Posición
Zm-SHC_fw	gcgctgtttcatatgggtattgaca	Cebador N-terminal
Zm-SHC_rev	gcgcttaccctggatcctcgaaaat	Cebador C-terminal

Se mezclaron de manera equimolar respectivamente 100 ng de los cebadores Zm-SHC_fw y Zm-SHC_rev. La PCR con ADN genómico de *Z. mobilis* (ATCC31821) se llevó a cabo según las instrucciones del fabricante con Pwo-Polymerase (Roche Applied Science) y el siguiente programa de gradiente de temperatura: 95 °C durante 3 min; 30 ciclos con 95 °C durante 30 s, 50 °C durante 30 s, y 72 °C durante 3 min; 72 °C durante 10 min.; 4 °C hasta su uso. El producto de PCR (~2,2 kB) se aisló mediante electroforesis en gel de agarosa (Gel E al 1,2 %, Invitrogen) y cromatografía en columna (Kit GFX, Amersham Pharmacia) y a continuación se secuenció (cebador de secuenciación: Zm-SHC_fw y Zm-SHC_rev). La secuencia obtenida corresponde a la secuencia publicada.

El producto de PCR se digirió con las endonucleasas de restricción *NdeI* y *BamHI* y se ligó en vector pDHE19.2 digerido de manera correspondiente. La secuenciación de los plásmidos resultantes dio como resultado la secuencia de ácido nucleico representada en la Seq-ID1.

El plásmido pDHE-Zm-SHC-1 se transformó en la cepa TG10 pAgro4 pHSG575 de *E. coli* [Takeshita et al., Gene 1987, 61:63-74; Tomoyasu et al., Mol Microbiol 2001, 40:397-413]. Las *E. coli* recombinantes reciben la denominación *E. coli* LU15568.

Ejemplo 2a: Provisión de homofarnesol ciclasa recombinante a partir de *Z. mobilis* (SEQ ID NO: 2)

30 Inoculada a partir de 2 ml de un cultivo previo correspondiente, *E. coli* LU 15568 se fijó en 20 ml de LB-Amp/Spec/Cm (ampicilina 100 µg/l; espectinomicina 100 µg/l; cloranfenicol 20 µg/l), IPTG 0,1 mM, ramnosa 0,5 g/l en matraz Erlenmeyer 100 ml (Schikanen) durante 16 h a 37 °C, se centrifugó a 5000*g/10 min y se almacenó a 4 °C. Se preparó extracto de proteína suspendiéndose el sedimento celular en 15 ml de tampón de disgregación (Tris/HCl 0,2 M, EDTA 0,5 M, pH 8,0), 375 U de benzonasa (por ejemplo Novagen, 25 U/µl), 40 µl de PMSF (100 mM, disuelto en *i*-PropOH), 0,8 g de sacarosa y aproximadamente 0,5 mg de lisozima. La preparación se mezcló y se incubó durante 30 min sobre hielo. A continuación se congeló la mezcla a -20 °C.

Después de la descongelación de la preparación se rellenó con agua destilada hasta aproximadamente 40 ml y se incubó de nuevo durante 30 min sobre hielo.

40 A continuación se disgregaron las células 3 veces durante 3 min con ultrasonidos (HTU-Soni 130, empresa G. Heinemann, Schwäbisch-Hall, amplitud 80 %, 15" de pulso / 15" de pausa).

Después de la disgregación se separaron por centrifugación los detritos celulares en 60 min a 4 °C y 26900 *g. El sobrenadante se retiró y el sedimento se resuspendió en 100 ml de tampón de solubilización (Tris/HCl 50 mM, MgCl₂x6H₂O 10 mM, Triton X-100 al 1 %, pH 8,0) y se trató en Potter durante aproximadamente 5 min. A continuación se mantuvo la suspensión durante 30 min sobre hielo.

45 El extracto homogeneizado se centrifugó de nuevo durante 1 h a 4 °C y 26900 *g y se retiró el sedimento. El extracto

se empleó para las pruebas enzimáticas y puede almacenarse durante varias semanas sin pérdida de actividad a -20 °C. El contenido de proteína se encuentra en el intervalo de 1 mg /ml.

Ejemplo 2b: Provisión de homofarnesol ciclasa recombinante a partir de *Bradyrhizobium japonicum* (SEQ ID NO: 5)

Se preparó homofarnesol ciclasa a partir de *Bradyrhizobium japonicum* tal como en el Ejemplo 2a.

5 Ejemplo 3a: Determinación de actividad de la homofarnesol ciclasa recombinante (SEQ ID NO: 2) a partir de *E. coli* LU15568

10 Con la preparación de proteína descrita en el Ejemplo 2a se incubó homofarnesol (1, (3Z,7E)-4,8,12-trimetiltrideca-3,7,11-trien-1-ol). En detalle se mezclaron entre sí 4 ml de preparación de proteína, 0,5 ml de tampón citrato de Na (citrato de sodio 1 M, pH 4,9), 0,5 ml de solución de homofarnesol (100 mM en tampón citrato de Na 0,1 M, pH 6,5 con 2 % (p/p) de taurodesoxicolato) y se incubó con agitación a 30 °C. Una reacción control se incubó en la misma composición pero a 60 °C.

Después de una incubación de 16 horas se extrajeron las preparaciones con 10 ml de hexano/ *n*-propanol 3:2 y se concentró la fase orgánica hasta sequedad.

El residuo se recogió en 200 µl diclorometano y se empleó para el análisis de CG o CG/EM.

15 Con la analítica representada en el Ejemplo 4 pudo establecerse una conversión del 41 % (»20,5 µmol de 2).

En comparación, en reacciones de homofarnesol con la escualeno-hopeno ciclasa conocida de *Alicyclobacillus acidocaldarius* se consiguió una conversión del 1,2 %.

Ejemplo 3b: Determinación de actividad de la homofarnesol ciclasa recombinante (SEQ ID NO: 5) a partir de *E. coli*

20 La actividad enzimática del biocatalizador preparados de manera recombinante en *E. coli* a partir de *Bradyrhizobium japonicum* corresponde a la de Zm-SHC. Después de una incubación de 3 horas de una solución 20 mM de homofarnesol con homogeneizado celular (cantidad de proteína: 31 mg) a 37 °C se hizo reaccionar un 26,4 % de homofarnesol para dar ambroxán. En las mismas condiciones se consiguió con Zm-SHC recombinante una conversión del 22,4 %.

Ejemplo 4 Analítica

25 *Conversión/Cuantificación*

La conversión de homofarnesol (1) en ambroxán (2) puede determinarse con el siguiente sistema de CG:
Columna: Optima 1 de 10 m

Perfil de temperatura:

0':	100 °C	
		15 °C/min
14,7':	320 °C	
34,7':	320 °C	

Temperatura de inyector: 320 °C

TA: homofarnesol aproximadamente 12,5'; ambroxán: aproximadamente 11,4'

30 Con material auténtico se creó una serie de calibración, por medio de la que se determinó la concentración de muestras desconocidas.

Identificación

35 La identificación tuvo lugar por medio de CG capilar/EM de los iones positivos tras ionización por choque de electrones (EI) e ionización química (CI) con amoníaco. Aparatos: CG (HP 6890) acoplado con dos MSD (HP 5973) para ionización EI e CI.

Condiciones de CG:

Columna de separación:	DB-1
Longitud:	30 m
Diámetro interno:	250 µm
Grosor de película:	0,25 µm
Gas portador:	He
Flujo:	1,2 ml/min

ES 2 703 770 T3

Relación de separación: 1:60
 Temperatura del horno: 100 °C, 5K/min a 200 °C, 5 min isoterma, 30 K/min a 300 °C, 50 min isoterma
 Temperatura de inyector: 250 °C
 Temperatura de línea de transferencia: 300 °C
 Cantidad de dosificación: 0,2 µl
 Preparación de muestras: Inyección directa

CONDICIONES DE EM

EI: Zona de exploración: 25- 750 uma Energía de ionización: 70 eV	CI: Zona de exploración: 75-800 uma
---	--

LISTADO DE SECUENCIAS

	<code><110> BASF SE</code>		
	<code><120> Preparación biocatalítica de ambroxán</code>		
5	<code><130> PF 62207 PCT</code>		
	<code><160> 11</code>		
	<code><170> PatentIn versión 3.5</code>		
	<code><210> 1</code>		
10	<code><211> 6525</code>		
	<code><212> ADN</code>		
	<code><213> Zymomonas mobilis</code>		
	<code><220></code>		
	<code><221> CDS</code>		
	<code><222> (154)..(2328)</code>		
15	<code><400> 1</code>		
	<code>cgatcaccac aattcagcaa attgtgaaca tcatcacggt catctttccc tggttgcaa</code>		60
	<code>tggccattt tctgtcagt aacgagaagg tcgogaattc aggcgctttt tagactggtc</code>		120
	<code>gtaatgaaca attcttaaga aggagatata cat atg ggt att gac aga atg aat</code>		174
		<code>Met Gly Ile Asp Arg Met Asn</code> <code>1 5</code>	
	<code>agc tta agt cgc ttg tta atg aag aag att ttc ggg gct gaa aaa acc</code>		222
	<code>Ser Leu Ser Arg Leu Leu Met Lys Lys Ile Phe Gly Ala Glu Lys Thr</code>	<code>10 15 20</code>	
	<code>tcg tat aaa cgg gct tcc gat acc ata atc gga acg gat acc ctg aaa</code>		270
	<code>Ser Tyr Lys Pro Ala Ser Asp Thr Ile Ile Gly Thr Asp Thr Leu Lys</code>	<code>25 30 35</code>	
	<code>aga cgg aac cgg cgg cct gaa cgg acg gca aaa gtc gac aaa acg ata</code>		318
	<code>Arg Pro Asn Arg Arg Pro Glu Pro Thr Ala Lys Val Asp Lys Thr Ile</code>	<code>40 45 50 55</code>	
	<code>ttc aag act atg ggg aat agt ctg aat aat acc ctt gtt tca gcc tgt</code>		366
	<code>Phe Lys Thr Met Gly Asn Ser Leu Asn Asn Thr Leu Val Ser Ala Cys</code>	<code>60 65 70</code>	
	<code>gac tgg ttg atc gga caa caa aag ccc gat ggt cat tgg gtc ggt gcc</code>		414
	<code>Asp Trp Leu Ile Gly Gln Gln Lys Pro Asp Gly His Trp Val Gly Ala</code>	<code>75 80 85</code>	
	<code>gtg gaa tcc aat gct tcg atg gaa gca gaa tgg tgt ctg gcc ttg tgg</code>		462
	<code>Val Glu Ser Asn Ala Ser Met Glu Ala Glu Trp Cys Leu Ala Leu Trp</code>	<code>90 95 100</code>	
	<code>ttt ttg ggt ctg gaa gat cat cgg ctt cgt cca aga ttg ggc aat gct</code>		510
	<code>Phe Leu Gly Leu Glu Asp His Pro Leu Arg Pro Arg Leu Gly Asn Ala</code>	<code>105 110 115</code>	
	<code>ctt ttg gaa atg cag cgg gaa gat ggc tct tgg gga gtc tat ttc ggc</code>		558
	<code>Leu Leu Glu Met Gln Arg Glu Asp Gly Ser Trp Gly Val Tyr Phe Gly</code>	<code>120 125 130 135</code>	

ES 2 703 770 T3

gct gga aat ggc gat atc aat gcc acg gtt gaa gcc tat gcg gcc ttg Ala Gly Asn Gly Asp Ile Asn Ala Thr Val Glu Ala Tyr Ala Ala Leu 140 145 150	606
cgg tct ttg ggg tat tct gcc gat aat cct gtt ttg aaa aaa gcg gca Arg Ser Leu Gly Tyr Ser Ala Asp Asn Pro Val Leu Lys Lys Ala Ala 155 160 165	654
gca tgg att gct gaa aaa ggc gga tta aaa aat atc cgt gtc ttt acc Ala Trp Ile Ala Glu Lys Gly Gly Leu Lys Asn Ile Arg Val Phe Thr 170 175 180	702
cgt tat tgg ctg gcg ttg atc ggg gaa tgg cct tgg gaa aag acc cct Arg Tyr Trp Leu Ala Leu Ile Gly Glu Trp Pro Trp Glu Lys Thr Pro 185 190 195	750
aac ctt ccc cct gaa att atc tgg ttc cct gat aat ttt gtc ttt tcg Asn Leu Pro Pro Glu Ile Ile Trp Phe Pro Asp Asn Phe Val Phe Ser 200 205 210 215	798
att tat aat ttt gcc caa tgg gcg cgg gca acc atg gtg ccg att gct Ile Tyr Asn Phe Ala Gln Trp Ala Arg Ala Thr Met Val Pro Ile Ala 220 225 230	846
att ctg tcc gcg aga cga cca agc cgc ccg ctg cgc cct caa gac cga Ile Leu Ser Ala Arg Arg Pro Ser Arg Pro Leu Arg Pro Gln Asp Arg 235 240 245	894
ttg gat gaa ctg ttt cca gaa ggc cgc gct cgc ttt gat tat gaa ttg Leu Asp Glu Leu Phe Pro Glu Gly Arg Ala Arg Phe Asp Tyr Glu Leu 250 255 260	942
ccg aaa aaa gaa ggc atc gat ctt tgg tcg caa ttt ttc cga acc act Pro Lys Lys Glu Gly Ile Asp Leu Trp Ser Gln Phe Phe Arg Thr Thr 265 270 275	990
gac cgt gga tta cat tgg gtt cag tcc aat ctg tta aag cgc aat agc Asp Arg Gly Leu His Trp Val Gln Ser Asn Leu Leu Lys Arg Asn Ser 280 285 290 295	1038
ttg cgt gaa gcc gct atc cgt cat gtt ttg gaa tgg att atc cgg cat Leu Arg Glu Ala Ala Ile Arg His Val Leu Glu Trp Ile Ile Arg His 300 305 310	1086
cag gat gcc gat ggc ggt tgg ggt gga att cag cca cct tgg gtc tat Gln Asp Ala Asp Gly Gly Trp Gly Gly Ile Gln Pro Pro Trp Val Tyr 315 320 325	1134
ggt ttg atg gcg tta cat ggt gaa ggc tat cag ctt tat cat ccg gtg Gly Leu Met Ala Leu His Gly Glu Gly Tyr Gln Leu Tyr His Pro Val 330 335 340	1182
atg gcc aag gct ttg tcg gct ttg gat gat ccc ggt tgg cga cat gac Met Ala Lys Ala Leu Ser Ala Leu Asp Asp Pro Gly Trp Arg His Asp 345 350 355	1230
aga ggc gag tct tct tgg ata cag gcc acc aat agt ccg gta tgg gat Arg Gly Glu Ser Ser Trp Ile Gln Ala Thr Asn Ser Pro Val Trp Asp 360 365 370 375	1278
aca atg ttg gcc ttg atg gcg tta aaa gac gcc aag gcc gag gat cgt Thr Met Leu Ala Leu Met Ala Leu Lys Asp Ala Lys Ala Glu Asp Arg 380 385 390	1326

ES 2 703 770 T3

ttt acg ccg gaa atg gat aag gcc gcc gat tgg ctt ttg gct cga cag	1374
Phe Thr Pro Glu Met Asp Lys Ala Ala Asp Trp Leu Leu Ala Arg Gln	
395 400 405	
gtc aaa gtc aaa ggc gat tgg tca atc aaa ctg ccc gat gtt gaa ccc	1422
Val Lys Val Lys Gly Asp Trp Ser Ile Lys Leu Pro Asp Val Glu Pro	
410 415 420	
ggt gga tgg gca ttt gaa tat gcc aat gat cgc tat ccc gat acc gat	1470
Gly Gly Trp Ala Phe Glu Tyr Ala Asn Asp Arg Tyr Pro Asp Thr Asp	
425 430 435	
gat acc gcc gtc gct ttg atc gcc ctt tcc tct tat cgt gat aag gag	1518
Asp Thr Ala Val Ala Leu Ile Ala Leu Ser Ser Tyr Arg Asp Lys Glu	
440 445 450 455	
gag tgg caa aag aaa ggc gtt gag gac gcc att acc cgt ggg gtt aat	1566
Glu Trp Gln Lys Lys Gly Val Glu Asp Ala Ile Thr Arg Gly Val Asn	
460 465 470	
tgg ttg atc gcc atg caa agc gaa tgt ggc ggt tgg gga gcc ttt gat	1614
Trp Leu Ile Ala Met Gln Ser Glu Cys Gly Gly Trp Gly Ala Phe Asp	
475 480 485	
aag gat aat aac aga agt atc ctt tcc aaa att cct ttt tgt gat ttc	1662
Lys Asp Asn Asn Arg Ser Ile Leu Ser Lys Ile Pro Phe Cys Asp Phe	
490 495 500	
gga gaa tct att gat ccg cct tca gtc gat gta acg gcg cat gtt tta	1710
Gly Glu Ser Ile Asp Pro Pro Ser Val Asp Val Thr Ala His Val Leu	
505 510 515	
gag gcc ttt ggc acc ttg gga ctg tcc cgc gat atg ccg gtc atc caa	1758
Glu Ala Phe Gly Thr Leu Gly Leu Ser Arg Asp Met Pro Val Ile Gln	
520 525 530 535	
aaa gcg atc gac tat gtc cgt tcc gaa cag gaa gcc gaa ggc gcg tgg	1806
Lys Ala Ile Asp Tyr Val Arg Ser Glu Gln Glu Ala Glu Gly Ala Trp	
540 545 550	
ttt ggt cgt tgg ggc gtt aat tat atc tat ggc acc ggt gcg gtt ctg	1854
Phe Gly Arg Trp Gly Val Asn Tyr Ile Tyr Gly Thr Gly Ala Val Leu	
555 560 565	
cct gct ttg gcg gcg atc ggt gaa gat atg acc cag cct tac atc acc	1902
Pro Ala Leu Ala Ala Ile Gly Glu Asp Met Thr Gln Pro Tyr Ile Thr	
570 575 580	
aag gct tgc gat tgg ctg gtc gca cat cag cag gaa gac ggc ggt tgg	1950
Lys Ala Cys Asp Trp Leu Val Ala His Gln Gln Glu Asp Gly Gly Trp	
585 590 595	
ggc gaa agc tgc tct tcc tat atg gag att gat tcc att ggg aag ggc	1998
Gly Glu Ser Cys Ser Ser Tyr Met Glu Ile Asp Ser Ile Gly Lys Gly	
600 605 610 615	
cca acc acg ccg tcc cag act gct tgg gct ttg atg ggg ttg atc gcg	2046
Pro Thr Thr Pro Ser Gln Thr Ala Trp Ala Leu Met Gly Leu Ile Ala	
620 625 630	
gcc aat cgt ccc gaa gat tat gaa gcc att gcc aag gga tgc cat tat	2094
Ala Asn Arg Pro Glu Asp Tyr Glu Ala Ile Ala Lys Gly Cys His Tyr	

ES 2 703 770 T3

ctg att gat cgc caa gag cag gat ggt agc tgg aaa gaa gaa gaa ttc	2142
Leu Ile Asp Arg Gln Glu Gln Asp Gly Ser Trp Lys Glu Glu Glu Phe	
650 655 660	
acc ggc acc gga ttc ccc ggt tat ggc gtg ggt cag acg atc aag ttg	2190
Thr Gly Thr Gly Phe Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu	
665 670 675	
gat gat ccg gct tta tcg aaa cga ttg ctt caa ggc gct gaa ctg tca	2238
Asp Asp Pro Ala Leu Ser Lys Arg Leu Leu Gln Gly Ala Glu Leu Ser	
680 685 690 695	
cgg gcg ttt atg ctg cgt tat gat ttt tat cgg caa ttc ttc ccg att	2286
Arg Ala Phe Met Leu Arg Tyr Asp Phe Tyr Arg Gln Phe Phe Pro Ile	
700 705 710	
atg gcg tta agt cgg gca gag aga ctg att gat ttg aat aat	2328
Met Ala Leu Ser Arg Ala Glu Arg Leu Ile Asp Leu Asn Asn	
715 720 725	
tgatagtatt ggggaggagg agtcttttta aaagagacta ctccgtccta ttttcgagga	2388
tccgtcgacc tgcagccaag cttggctggt ttggcggatg agagaagatt ttcagcctga	2448
tacagattaa atcagaacgc agaagcggtc tgataaaaca gaatttgcct gccggcagta	2508
gcgcggtggt cccacctgac cccatgccga actcagaagt gaaacgccgt agcggcagatg	2568
gtagtgtggg gtctcccat gcgagagtag ggaactgcca ggcatcaaat aaaacgaaag	2628
gctcagtcga aagactgggc ctttcgtttt atctgttggt tgtcggtgaa cgctctcctg	2688
agtaggacaa atccgccggg agcggatttg aacggtgcga agcaacggcc cggagggtgg	2748
cgggcaggac gcccgccata aactgccagg catcaaatta agcagaaggc catcctgacg	2808
gatggccttt ttgcgtttct acaaactctt ttgtttattt ttctaatac attcaaatat	2868
gtatccgctc atgagacaat aaccctgata aatgcttcaa taatattgaa aaaggaagag	2928
tatgagtatt caacatttcc gtgtcgccct tattcccttt tttgcggcat tttgccttcc	2988
tgtttttgct caccagaaa cgctggtgaa agtaaaagat gctgaagatc agttgggtgc	3048
acgagtgggt tacatcgaac tggatctcaa cagcggtaag atccttgaga gttttcgccc	3108
cgaagaacgt tttccaatga tgagcacttt taaagttctg ctatgtggcg cgttattatc	3168
ccgtgttgac gccgggcaag agcaactcgg tcgccgcata cactattctc agaatgactt	3228
ggttgagtac tcaccagtca cagaaaagca tcttacggat ggcatgacag taagagaatt	3288
atgcagtgct gccataacca tgagtgataa cactgcggcc aacttacttc tgacaacgat	3348
cggaggaccg aaggagctaa ccgctttttt gcacaacatg ggggatcatg taactcgcct	3408
tgatcgttgg gaaccggagc tgaatgaagc cataccaaac gacgagcgtg acaccacgat	3468
gcctgtagca atggcaacaa cgttgcgcaa actattaact ggcgaactac ttactctagc	3528
ttcccggcaa caattaatag actggatgga ggcggataaa gttgcaggac cacttctgcg	3588

ES 2 703 770 T3

ctcggccctt ccggctggct ggtttattgc tgataaatct ggagccggtg agcgtgggtc 3648
tcgcggtatc attgcagcac tggggccaga tggtaaagccc tcccgtatcg tagttatcta 3708
cacgacgggg agtcaggcaa ctatggatga acgaaataga cagatcgctg agataggtgc 3768
ctcactgatt aagcattggt aactgtcaga ccaagtttac tcatatatac tttagattga 3828
tttaaaactt catttttaat ttaaaaggat ctaggtgaag atcctttttg ataatctcat 3888
gacccaaaatc ccttaacgtg agttttcgtt cactgagcg tcagaccccg tagaaaagat 3948
caaaggatct tcttgagatc cttttttct ggcgtaatc tgctgcttgc aaacaaaaaa 4008
accaccgcta ccagcggtg tttgtttgcc ggatcaagag ctaccaactc tttttccgaa 4068
ggtaactggc ttcagcagag cgcagatacc aaatactgtc cttctagtgt agccgtagtt 4128
aggccaccac ttcaagaact ctgtagcacc gcctacatac ctgctctgc taatcctggt 4188
accagtggct gctgccagtg gcgataagtc gtgtcttacc gggttggact caagacgata 4248
gttaccggat aaggcgcagc ggtcgggctg aacggggggt tcgtgcacac agcccagctt 4308
ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag aaagcggcac 4368
gcttcccgaa gggagaaaagg cggacaggta tccggtaagc ggcagggctg gaacaggaga 4428
gcgcacgagg gagcttccag ggggaaacgc ctggtatctt tatagtctg tcgggtttcg 4488
ccacctctga cttgagcgtc gatttttgtg atgctcgtca gggggcgga gcctatggaa 4548
aaacgccagc aacgcggcct ttttacggtt cctggccttt tgctggcctt ttgctcacat 4608
gttctttcct gcgttatccc ctgattctgt ggataaccgt attacgcct ttgagtgagc 4668
tgataccgct cggccgagcc gaacgaccga gcgcagcagag tcagtgagcg aggaagcggga 4728
agagcgcctg atgcggtatt ttctccttac gcctctgtgc ggtatttcac accgcatata 4788
tggtgcactc tcagtacaat ctgctctgat gccgcatagt taagccagta tacactccgc 4848
tatcgctacg tgactgggtc atggctgcgc cccgacaccc gccaacaccc gctgacgcgc 4908
cctgacgggc ttgtctgctc ccggcatccg cttacagaca agctgtgacc gtctccggga 4968
gctgcatgtg tcagaggttt tcaccgtcat caccgaaacg cgcgaggcag ctgcggtaaa 5028
gctcatcagc gtggtcgtga agcgattcac agatgtctgc ctgttcatcc gcgtccagct 5088
cgttgagttt ctccagaagc gttaatgtct ggcttctgat aaagcgggcc atgttaaggg 5148
cggttttttc ctgtttggtc actgatgcct ccgtgtaagg gggatttctg ttcattggggg 5208
taatgatacc gatgaaacga gagaggatgc tcacgatacg ggttactgat gatgaacatg 5268
cccggttact ggaacgttgt gagggtaaac aactggcgggt atggatgcgg cgggaccaga 5328
gaaaaatcac tcagggtcaa tgccagcgtc tcgttaatac agatgtaggt gttccacagg 5388
gtagccagca gcctcctgcg atgcagatcc ggaacataat ggtgcagggc gctgacttcc 5448

ES 2 703 770 T3

gcgtttccag actttacgaa acacggaaac cgaagacat tcatgttggt gctcaggtcg 5508
 cagacgtttt gcagcagcag tcgcttcacg ttcgctcgcg tatcggtgat tcattctgct 5568
 aaccagtaag gcaacccccgc cagcctagcc gggtcctcaa cgacaggagc acgatcatgc 5628
 gcacccgtgg ccaggaccca acgctgcccg agatgcgccg cgtgcggctg ctggagatgg 5688
 cggacgcgat ggatatgttc tgccaagggt tggtttgccg attcacagtt ctccgcaaga 5748
 attgattggc tccaattctt ggagtgggtga atccgtagc gaggtgccgc cggcttccat 5808
 tcaggtcgag gtggcccggc tccatgcacc gcgacgcaac gcggggaggc agacaagga 5868
 tagggcggcg cctacaatcc atgccaaacc gttccatgtg ctgccgagg cggcataaat 5928
 cgccgtgacg atcagcggtc caatgatcga agttaggctg gtaagagccg cgagcgatcc 5988
 ttgaagctgt ccctgatggt cgtcatctac ctgcctggac agcatggcct gcaacgcggg 6048
 catcccgatg ccgccggaag cgagaagaat cataatgggg aaggccatcc agcctcgcgt 6108
 cgcaacgcc agcaagacgt agcccagcgc gtcggccgcc atgccggcga taatggcctg 6168
 cttctcggcg aaacgtttgg tggcgggacc agtgacgaag gcttgagcga gggcgtgcaa 6228
 gattccgaat accgcaagcg acaggccgat catcgctcgcg ctccagcgaag agcggtcctc 6288
 gccgaaaatg acccagagcg ctgccggcac ctgtcctacg agttgcatga taaagaagac 6348
 agtcataagt gcggcgacga tagtcatgcc ccgcgcccac cggaaggagc tgactgggtt 6408
 gaaggctctc aagggcatcg gtcgacgctc tcccttatgc gactcctgca ttaggaagca 6468
 gccagtagt aggttgaggc cgttgagcac cgccgccgca aggaatggtg catgcat 6525

<210> 2

<211> 725

<212> PRT

5 <213> Zymomonas mobilis

<400> 2

Met Gly Ile Asp Arg Met Asn Ser Leu Ser Arg Leu Leu Met Lys Lys
 1 5 10 15

Ile Phe Gly Ala Glu Lys Thr Ser Tyr Lys Pro Ala Ser Asp Thr Ile
 20 25 30

Ile Gly Thr Asp Thr Leu Lys Arg Pro Asn Arg Arg Pro Glu Pro Thr
 35 40 45

Ala Lys Val Asp Lys Thr Ile Phe Lys Thr Met Gly Asn Ser Leu Asn
 50 55 60

Asn Thr Leu Val Ser Ala Cys Asp Trp Leu Ile Gly Gln Gln Lys Pro
 65 70 75 80

ES 2 703 770 T3

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

ES 2 703 770 T3

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

ES 2 703 770 T3

		580						585						590			
	Gln	Gln	Glu	Asp	Gly	Gly	Trp	Gly	Glu	Ser	Cys	Ser	Ser	Tyr	Met	Glu	
			595					600					605				
	Ile	Asp	Ser	Ile	Gly	Lys	Gly	Pro	Thr	Thr	Pro	Ser	Gln	Thr	Ala	Trp	
		610					615					620					
	Ala	Leu	Met	Gly	Leu	Ile	Ala	Ala	Asn	Arg	Pro	Glu	Asp	Tyr	Glu	Ala	
		625				630					635					640	
	Ile	Ala	Lys	Gly	Cys	His	Tyr	Leu	Ile	Asp	Arg	Gln	Glu	Gln	Asp	Gly	
					645					650					655		
	Ser	Trp	Lys	Glu	Glu	Glu	Phe	Thr	Gly	Thr	Gly	Phe	Pro	Gly	Tyr	Gly	
				660					665					670			
	Val	Gly	Gln	Thr	Ile	Lys	Leu	Asp	Asp	Pro	Ala	Leu	Ser	Lys	Arg	Leu	
			675					680					685				
	Leu	Gln	Gly	Ala	Glu	Leu	Ser	Arg	Ala	Phe	Met	Leu	Arg	Tyr	Asp	Phe	
		690					695					700					
	Tyr	Arg	Gln	Phe	Phe	Pro	Ile	Met	Ala	Leu	Ser	Arg	Ala	Glu	Arg	Leu	
		705				710					715					720	
	Ile	Asp	Leu	Asn	Asn												
					725												

5 <210> 3
 <211> 25
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Cebador
 <400> 3
 gcgctgtttc atatgggtat tgaca 25

10 <210> 4
 <211> 25
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> Cebador
 <400> 4
 gcgcttacc tggatcctcg aaaat 25

20 <210> 5
 <211> 684
 <212> PRT
 <213> Bradyrhizobium japonicum

ES 2 703 770 T3

<400> 5

Met Thr Val Thr Ser Ser Ala Ser Ala Arg Ala Thr Arg Asp Pro Gly
 1 5 10 15

Asn Tyr Gln Thr Ala Leu Gln Ser Thr Val Arg Ala Ala Ala Asp Trp
 20 25 30

Leu Ile Ala Asn Gln Lys Pro Asp Gly His Trp Val Gly Arg Ala Glu
 35 40 45

Ser Asn Ala Cys Met Glu Ala Gln Trp Cys Leu Ala Leu Trp Phe Met
 50 55 60

Gly Leu Glu Asp His Pro Leu Arg Lys Arg Leu Gly Gln Ser Leu Leu
 65 70 75 80

Asp Ser Gln Arg Pro Asp Gly Ala Trp Gln Val Tyr Phe Gly Ala Pro
 85 90 95

Asn Gly Asp Ile Asn Ala Thr Val Glu Ala Tyr Ala Ala Leu Arg Ser
 100 105 110

Leu Gly Phe Arg Asp Asp Glu Pro Ala Val Arg Arg Ala Arg Glu Trp
 115 120 125

Ile Glu Ala Lys Gly Gly Leu Arg Asn Ile Arg Val Phe Thr Arg Tyr
 130 135 140

Trp Leu Ala Leu Ile Gly Glu Trp Pro Trp Glu Lys Thr Pro Asn Ile
 145 150 155 160

Pro Pro Glu Val Ile Trp Phe Pro Leu Trp Phe Pro Phe Ser Ile Tyr
 165 170 175

Asn Phe Ala Gln Trp Ala Arg Ala Thr Leu Met Pro Ile Ala Val Leu
 180 185 190

Ser Ala Arg Arg Pro Ser Arg Pro Leu Pro Pro Glu Asn Arg Leu Asp
 195 200 205

Ala Leu Phe Pro His Gly Arg Lys Ala Phe Asp Tyr Glu Leu Pro Val
 210 215 220

ES 2 703 770 T3

Lys Ala Gly Ala Gly Gly Trp Asp Arg Phe Phe Arg Gly Ala Asp Lys
 225 230 235 240

Val Leu His Lys Leu Gln Asn Leu Gly Asn Arg Leu Asn Leu Gly Leu
 245 250 255

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

Gln Asp Phe Asp Gly Ala Trp Gly Gly Ile Gln Pro Pro Trp Ile Tyr
 275 280 285

Gly Leu Met Ala Leu Tyr Ala Glu Gly Tyr Pro Leu Asn His Pro Val
 290 295 300

Leu Ala Lys Gly Leu Asp Ala Leu Asn Asp Pro Gly Trp Arg Val Asp
 305 310 315 320

Val Gly Asp Ala Thr Tyr Ile Gln Ala Thr Asn Ser Pro Val Trp Asp
 325 330 335

Thr Ile Leu Thr Leu Leu Ala Phe Asp Asp Ala Gly Val Leu Gly Asp
 340 345 350

Tyr Pro Glu Ala Val Asp Lys Ala Val Asp Trp Val Leu Gln Arg Gln
 355 360 365

Val Arg Val Pro Gly Asp Trp Ser Met Lys Leu Pro His Val Lys Pro
 370 375 380

Gly Gly Trp Ala Phe Glu Tyr Ala Asn Asn Tyr Tyr Pro Asp Thr Asp
 385 390 395 400

Asp Thr Ala Val Ala Leu Ile Ala Leu Ala Pro Leu Arg His Asp Pro
 405 410 415

Lys Trp Lys Ala Lys Gly Ile Asp Glu Ala Ile Gln Leu Gly Val Asp
 420 425 430

Trp Leu Ile Gly Met Gln Ser Gln Gly Gly Gly Trp Gly Ala Phe Asp
 435 440 445

Lys Asp Asn Asn Gln Lys Ile Leu Thr Lys Ile Pro Phe Cys Asp Tyr
 450 455 460

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

ES 2 703 770 T3

Glu Ala Phe Gly Lys Leu Gly Ile Ser Arg Asn His Pro Ser Met Val
 485 490 495

Gln Ala Leu Asp Tyr Ile Arg Arg Glu Gln Glu Pro Ser Gly Pro Trp
 500 505 510

Phe Gly Arg Trp Gly Val Asn Tyr Val Tyr Gly Thr Gly Ala Val Leu
 515 520 525

Pro Ala Leu Ala Ala Ile Gly Glu Asp Met Thr Gln Pro Tyr Ile Gly
 530 535 540

Arg Ala Cys Asp Trp Leu Val Ala His Gln Gln Ala Asp Gly Gly Trp
 545 550 555 560

Gly Glu Ser Cys Ala Ser Tyr Met Asp Val Ser Ala Val Gly Arg Gly
 565 570 575

Thr Thr Thr Ala Ser Gln Thr Ala Trp Ala Leu Met Ala Leu Leu Ala
 580 585 590

Ala Asn Arg Pro Gln Asp Lys Asp Ala Ile Glu Arg Gly Cys Met Trp
 595 600 605

Leu Val Glu Arg Gln Ser Ala Gly Thr Trp Asp Glu Pro Glu Phe Thr
 610 615 620

Gly Thr Gly Phe Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu Asn
 625 630 635 640

Asp Pro Ala Leu Ser Gln Arg Leu Met Gln Gly Pro Glu Leu Ser Arg
 645 650 655

Ala Phe Met Leu Arg Tyr Gly Met Tyr Arg His Tyr Phe Pro Leu Met
 660 665 670

Ala Leu Gly Arg Ala Leu Arg Pro Gln Ser His Ser
 675 680

- <210> 6
- <211> 657
- <212> PRT
- <213> Burkholderia ambifaria

5

<400> 6

Met Asn Asp Leu Thr Glu Met Ala Thr Leu Ser Ala Gly Thr Val Pro
 1 5 10 15

ES 2 703 770 T3

Ala Gly Leu Asp Ala Ala Val Ala Ser Ala Thr Asp Ala Leu Leu Ala
 20 25 30

Ala Gln Asn Ala Asp Gly His Trp Val Tyr Glu Leu Glu Ala Asp Ser
 35 40 45

Thr Ile Pro Ala Glu Tyr Val Leu Leu Val His Tyr Leu Gly Glu Thr
 50 55 60

Pro Asn Leu Glu Leu Glu Gln Lys Ile Gly Arg Tyr Leu Arg Arg Val
 65 70 75 80

Gln Gln Ala Asp Gly Gly Trp Pro Leu Phe Thr Asp Gly Ala Pro Asn
 85 90 95

Ile Ser Ala Ser Val Lys Ala Tyr Phe Ala Leu Lys Val Ile Gly Asp
 100 105 110

Asp Glu Asn Ala Glu His Met Gln Arg Ala Arg Arg Ala Ile Gln Ala
 115 120 125

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

Leu Tyr Gly Ala Ile Pro Trp Arg Ala Val Pro Met Met Pro Val Glu
 145 150 155 160

Ile Met Leu Leu Pro Gln Trp Phe Pro Phe His Leu Ser Lys Val Ser
 165 170 175

Tyr Trp Ala Arg Thr Val Ile Val Pro Leu Leu Val Leu Asn Ala Lys
 180 185 190

Arg Pro Ile Ala Lys Asn Pro Arg Gly Val Arg Ile Asp Glu Leu Phe
 195 200 205

Val Asp Pro Pro Val Asn Ala Gly Leu Leu Pro Arg Gln Gly His Gln
 210 215 220

Ser Pro Gly Trp Phe Ala Phe Phe Arg Val Val Asp His Ala Leu Arg
 225 230 235 240

Ala Ala Asp Gly Leu Phe Pro Asn Tyr Thr Arg Glu Arg Ala Ile Arg
 245 250 255

Gln Ala Val Ser Phe Val Asp Glu Arg Leu Asn Gly Glu Asp Gly Leu
 260 265 270

ES 2 703 770 T3

Gly Ala Ile Tyr Pro Ala Met Ala Asn Ala Val Met Met Tyr Asp Val
 275 280 285

Leu Gly Tyr Ala Glu Asp His Pro Asn Arg Ala Ile Ala Arg Lys Ser
 290 300

Ile Glu Lys Leu Leu Val Val Gln Glu Asp Glu Ala Tyr Cys Gln Pro
 305 310 315 320

Cys Leu Ser Pro Val Trp Asp Thr Ser Leu Ala Ala His Ala Leu Leu
 325 330 335

Glu Thr Gly Asp Ala Arg Ala Glu Glu Ala Val Ile Arg Gly Leu Glu
 340 345 350

Trp Leu Arg Pro Leu Gln Ile Leu Asp Val Arg Gly Asp Trp Ile Ser
 355 360 365

Arg Arg Pro His Val Arg Pro Gly Gly Trp Ala Phe Gln Tyr Ala Asn
 370 375 380

Pro His Tyr Pro Asp Val Asp Asp Thr Ala Val Val Ala Val Ala Met
 385 390 395 400

Asp Arg Val Gln Lys Leu Lys His Asn Asp Ala Phe Arg Asp Ser Ile
 405 410 415

Ala Arg Ala Arg Glu Trp Val Val Gly Met Gln Ser Ser Asp Gly Gly
 420 425 430

Trp Gly Ala Phe Glu Pro Glu Asn Thr Gln Tyr Tyr Leu Asn Asn Ile
 435 440 445

Pro Phe Ser Asp His Gly Ala Leu Leu Asp Pro Pro Thr Ala Asp Val
 450 455 460

Ser Gly Arg Cys Leu Ser Met Leu Ala Gln Leu Gly Glu Thr Pro Leu
 465 470 475 480

Asn Ser Glu Pro Ala Arg Arg Ala Leu Asp Tyr Met Leu Lys Glu Gln
 485 490 495

Glu Pro Asp Gly Ser Trp Tyr Gly Arg Trp Gly Met Asn Tyr Val Tyr
 500 505 510

Gly Thr Trp Thr Ala Leu Cys Ala Leu Asn Ala Ala Gly Leu Thr Pro

ES 2 703 770 T3

515 520 525

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

Asn Lys Asp Gly Gly Trp Gly Glu Asp Gly Asp Ser Tyr Lys Leu Asn
545 550 555 560

Tyr Arg Gly Phe Glu Gln Ala Pro Ser Thr Ala Ser Gln Thr Ala Trp
565 570 575

Ala Leu Leu Gly Leu Met Ala Ala Gly Glu Val Asn Asn Pro Ala Val
580 585 590

Ala Arg Gly Val Glu Tyr Leu Ile Ala Glu Gln Lys Glu His Gly Leu
595 600 605

Trp Asp Glu Thr Arg Phe Thr Ala Thr Gly Phe Pro Arg Val Phe Tyr
610 615 620

Leu Arg Tyr His Gly Tyr Arg Lys Phe Phe Pro Leu Trp Ala Leu Ala
625 630 635 640

Arg Tyr Arg Asn Leu Lys Arg Asn Asn Ala Thr Arg Val Thr Phe Gly
645 650 655

Leu

<210> 7
<211> 682
<212> PRT
<213> Burkholderia ambifaria

5

<400> 7

Met Ile Arg Arg Met Asn Lys Ser Gly Pro Ser Pro Trp Ser Ala Leu
1 5 10 15

Asp Ala Ala Ile Ala Arg Gly Arg Asp Ala Leu Met Arg Leu Gln Gln
20 25 30

Pro Asp Gly Ser Trp Cys Phe Glu Leu Glu Ser Asp Ala Thr Ile Thr
35 40 45

Ala Glu Tyr Ile Leu Met Met His Phe Met Asp Lys Ile Asp Asp Ala
50 55 60

Arg Gln Glu Lys Met Ala Arg Tyr Leu Arg Ala Ile Gln Arg Leu Asp

ES 2 703 770 T3

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

ES 2 703 770 T3

Val Trp Asp Thr Ala Trp Ser Thr Met Ala Leu Glu Gln Ala Arg Gly
 325 330 335

Val Ala Val Pro Glu Ala Gly Ala Pro Ala Ser Ala Leu Asp Glu Leu
 340 345 350

Asp Ala Arg Ile Ala Arg Ala Tyr Asp Trp Leu Ala Glu Arg Gln Val
 355 360 365

Asn Asp Leu Arg Gly Asp Trp Ile Glu Asn Ala Pro Ala Asp Thr Gln
 370 375 380

Pro Gly Gly Trp Ala Phe Gln Tyr Ala Asn Pro Tyr Tyr Pro Asp Ile
 385 390 395 400

Asp Asp Ser Ala Val Val Thr Ala Met Leu Asp Arg Arg Gly Arg Thr
 405 410 415

His Arg Asn Ala Asp Gly Ser His Pro Tyr Ala Ala Arg Val Ala Arg
 420 425 430

Ala Leu Asp Trp Met Arg Gly Leu Gln Ser Arg Asn Gly Gly Phe Ala
 435 440 445

Ala Phe Asp Ala Asp Cys Asp Arg Leu Tyr Leu Asn Ala Ile Pro Phe
 450 455 460

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

Arg Val Leu Leu Cys Phe Gly Val Thr Lys Arg Ala Asp Asp Arg Ala
 485 490 495

Ser Leu Ala Arg Ala Ile Asp Tyr Val Lys Arg Thr Gln Gln Pro Asp
 500 505 510

Gly Ser Trp Trp Gly Arg Trp Gly Thr Asn Tyr Leu Tyr Gly Thr Trp
 515 520 525

Ser Val Leu Ala Gly Leu Ala Leu Ala Gly Glu Asp Pro Ser Gln Pro
 530 535 540

Tyr Ile Ala Arg Ala Leu Ala Trp Leu Arg Ala Arg Gln His Ala Asp
 545 550 555 560

Gly Gly Trp Gly Glu Thr Asn Asp Ser Tyr Ile Asp Pro Ala Leu Ala
 565 570 575

ES 2 703 770 T3

Gly Thr Asn Ala Gly Glu Ser Thr Ser Asn Cys Thr Ala Trp Ala Leu
 580 585 590

Leu Ala Gln Met Ala Phe Gly Asp Gly Glu Ser Glu Ser Val Arg Arg
 595 600 605

Gly Ile Ala Tyr Leu Gln Ser Val Gln Gln Asp Asp Gly Phe Trp Trp
 610 615 620

His Arg Ser His Asn Ala Pro Gly Phe Pro Arg Ile Phe Tyr Leu Lys
 625 630 635 640

Tyr His Gly Tyr Thr Ala Tyr Phe Pro Leu Trp Ala Leu Ala Arg Tyr
 645 650 655

Arg Arg Leu Ala Gly Gly Val Ser Ala Ala Gly Ala His Ala Val Pro
 660 665 670

Ala Ser Thr Gly Ala Asp Ala Ala Leu Ala
 675 680

<210> 8
 <211> 617
 <212> PRT
 <213> Bacillus anthracis

5

<400> 8

Met Leu Leu Tyr Glu Lys Ala His Glu Glu Ile Val Arg Arg Ala Thr
 1 5 10 15

Ala Leu Gln Thr Met Gln Trp Gln Asp Gly Thr Trp Arg Phe Cys Phe
 20 25 30

Glu Gly Ala Pro Leu Thr Asp Cys His Met Ile Phe Leu Leu Lys Leu
 35 40 45

Leu Gly Arg Asp Lys Glu Ile Glu Pro Phe Val Glu Arg Val Ala Ser
 50 55 60

Leu Gln Thr Asn Glu Gly Thr Trp Lys Leu His Glu Asp Glu Val Gly
 65 70 75 80

Gly Asn Leu Ser Ala Thr Ile Gln Ser Tyr Ala Ala Leu Leu Ala Ser
 85 90 95

Lys Lys Tyr Thr Lys Glu Asp Ala Asn Met Lys Arg Ala Glu Asn Phe
 100 105 110

ES 2 703 770 T3

Ile Gln Glu Arg Gly Gly Val Ala Arg Ala His Phe Met Thr Lys Phe
115 120 125

Leu Leu Ala Ile His Gly Glu Tyr Glu Tyr Pro Ser Leu Phe His Leu
130 135 140

Pro Thr Pro Ile Met Phe Leu Gln Asn Asp Ser Pro Phe Ser Ile Phe
145 150 155 160

Glu Leu Ser Ser Ser Ala Arg Ile His Leu Ile Pro Met Met Leu Cys
165 170 175

Leu Asn Lys Arg Phe Arg Val Gly Lys Lys Leu Leu Pro Asn Leu Asn
180 185 190

His Ile Ala Gly Gly Gly Gly Glu Trp Phe Arg Glu Asp Arg Ser Pro
195 200 205

Val Phe Gln Thr Leu Leu Ser Asp Val Lys Gln Ile Ile Ser Tyr Pro
210 215 220

Leu Ser Leu His His Lys Gly Tyr Glu Glu Ile Glu Arg Phe Met Lys
225 230 235 240

Glu Arg Ile Asp Glu Asn Gly Thr Leu Tyr Ser Tyr Ala Thr Ala Ser
245 250 255

Phe Tyr Met Ile Tyr Ala Leu Leu Ala Leu Gly His Ser Leu Gln Ser
260 265 270

Ser Met Ile Gln Lys Ala Ile Ala Gly Ile Thr Ser Tyr Ile Trp Lys
275 280 285

Met Glu Arg Gly Asn His Leu Gln Asn Ser Pro Ser Thr Val Trp Asp
290 295 300

Thr Ala Leu Leu Ser Tyr Ala Leu Gln Glu Ala Gln Val Ser Lys Asp
305 310 315 320

Asn Lys Met Ile Gln Asn Ala Thr Ala Tyr Leu Leu Lys Lys Gln His
325 330 335

Thr Lys Lys Ala Asp Trp Ser Val His Ala Pro Ala Leu Thr Pro Gly
340 345 350

Gly Trp Gly Phe Ser Asp Val Asn Thr Thr Ile Pro Asp Ile Asp Asp
355 360 365

ES 2 703 770 T3

Thr Thr Ala Val Leu Arg Ala Leu Ala Arg Ser Arg Gly Asn Lys Asn
 370 375 380

Ile Asp Asn Ala Trp Lys Lys Gly Gly Asn Trp Ile Lys Gly Leu Gln
 385 390 395 400

Asn Asn Asp Gly Gly Trp Gly Ala Phe Glu Lys Gly Val Thr Ser Lys
 405 410 415

Leu Leu Ala Lys Leu Pro Ile Glu Asn Ala Ser Asp Met Ile Thr Asp
 420 425 430

Pro Ser Thr Pro Asp Ile Thr Gly Arg Val Leu Glu Phe Phe Gly Thr
 435 440 445

Tyr Ala Gln Asn Glu Leu Pro Glu Lys Gln Ile Gln Arg Ala Ile Asn
 450 455 460

Trp Leu Met Asn Val Gln Glu Glu Asn Gly Ser Trp Tyr Gly Lys Trp
 465 470 475 480

Gly Ile Cys Tyr Leu Tyr Gly Thr Trp Ala Val Met Thr Gly Leu Arg
 485 490 495

Ser Leu Gly Ile Pro Ser Ser Asn Pro Ser Leu Thr Arg Ala Ala Ser
 500 505 510

Trp Leu Glu His Ile Gln His Glu Asp Gly Gly Trp Gly Glu Ser Cys
 515 520 525

His Ser Ser Val Glu Lys Arg Phe Val Thr Leu Pro Phe Ser Thr Pro
 530 535 540

Ser Gln Thr Ala Trp Ala Leu Asp Ala Leu Ile Ser Tyr Tyr Asp Thr
 545 550 555 560

Glu Thr Pro Ala Ile Arg Lys Gly Val Ser Tyr Leu Leu Ser Asn Pro
 565 570 575

Tyr Val Asn Glu Arg Tyr Pro Thr Gly Thr Gly Leu Pro Gly Ala Phe
 580 585 590

Tyr Ile Arg Tyr His Ser Tyr Ala His Ile Tyr Pro Leu Leu Thr Leu
 595 600 605

Ala His Tyr Ile Lys Lys Tyr Arg Lys

610

615

ES 2 703 770 T3

<210> 9
 <211> 720
 <212> PRT
 <213> Frankia alni

5 <400> 9

Met Pro Ala Gly Val Gly Val Leu Val Trp Leu Asp Gln Arg Leu Arg
 1 5 10 15

Ala Met Gly Arg Pro Asp Leu Val Thr Thr Thr Gly Gly Ala Glu Ile
 20 25 30

Pro Phe Val Leu Val Ala Ala Thr Ala Ser Thr Val Gly Val Ala Leu
 35 40 45

Ala Leu Arg Arg Pro Arg His Pro Val Gly Trp Leu Phe Leu Ala Leu
 50 55 60

Gly Gly Val Leu Leu Leu Ser Gly Gly Thr Gln Gly Tyr Ala Ala Tyr
 65 70 75 80

Gly Ala Val Ala Arg Pro Gly Arg Leu Pro Ala Ala Asp Leu Val Ala
 85 90 95

Ile Tyr Ala Asp Ala Gly Phe Ile Pro Trp Leu Val Leu Val Ala Leu
 100 105 110

Ile Leu His Leu Thr Pro Thr Gly Arg Pro Leu Ser Ala Arg Trp Gly
 115 120 125

Arg Ile Ala Leu Ala Thr Ala Val Ala Gly Gly Leu Trp Leu Leu Val
 130 135 140

Gly Leu Val Thr Thr Glu Thr Met Gln Pro Pro Phe Gln Ser Val Thr
 145 150 155 160

Asn Pro Leu Leu Ile Gly Gly Pro Leu Gly Pro Leu Leu Val Ala Arg
 165 170 175

Arg Val Leu Gly Leu Ala Thr Gly Ala Gly Val Val Leu Ala Ala Val
 180 185 190

Ser Leu Ile Val Arg Phe Arg Arg Ser Val Asp Val Glu Arg Arg Gln
 195 200 205

Leu Leu Trp Val Ala Val Ala Ala Val Pro Leu Pro Val Leu Met Ala

ES 2 703 770 T3

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

ES 2 703 770 T3

Ala Glu Arg Arg Thr Ile Glu Arg Asn Leu His Asp Gly Ala Gln Gln
465 470 475 480

Arg Leu Leu Ala Leu Ala Leu Gln Leu Arg Ala Val Gln Leu Gly Gly
485 490 495

Asp Glu Ala Ser Leu Arg Gln Ala Ile Ser Thr Gly Ile Asp Gln Leu
500 505 510

Gln Ala Ala Val Val Glu Leu Arg Glu Leu Ala Asn Gly Leu His Pro
515 520 525

Ala Val Leu Ala Asp Gly Gly Leu Ala Ala Ala Leu Asp Asp Val Ala
530 535 540

Ala Arg Thr Pro Val Pro Ile Lys Ile Ser Ala Pro Asp Arg Arg Tyr
545 550 555 560

Pro Pro Asp Leu Glu Ala Ala Ala Trp Phe Ile Ala Cys Glu Ala Met
565 570 575

Ala Asn Ala Val Lys His Ala His Pro Thr Thr Ile Ala Val Asp Val
580 585 590

Ser Ala Pro Asp Gly Gln Leu Ile Val Glu Val Arg Asp Asp Gly Ile
595 600 605

Gly Gly Ala Gln Pro Ser Gly Pro Gly Leu Arg Gly Ile Ala Asp Arg
610 615 620

Ala Glu Ala Phe Gly Gly Ser Leu Thr Val His Thr Asp Pro Gly Thr
625 630 635 640

Gly Thr Thr Ile Arg Ala Leu Leu His Arg Arg Ser Pro Leu Ser Ser
645 650 655

Gly Arg Arg Ser Val Met Ile Glu Gly Cys Val Asp Val Val Ala Val
660 665 670

Arg Arg Phe Arg Cys Arg Ser Ser Arg Gly Ser Gly Ser Arg Arg Arg
675 680 685

Arg Ser Ser Trp Arg Cys Gly Gly Ile Cys Gly Ser Arg Cys Arg Thr
690 695 700

Gly Met Ser Arg Ser Cys Ser Arg Asn Ala Ala Ser Lys Leu Ile Thr
705 710 715 720

ES 2 703 770 T3

<211> 685
 <212> PRT
 <213> Rhodopseudomonas palent
 <400> 10

```

Met Asp Ser Ile Leu Ala Pro Arg Ala Asp Ala Pro Arg Asn Ile Asp
 1          5          10          15

Gly Ala Leu Arg Glu Ser Val Gln Gln Ala Ala Asp Trp Leu Val Ala
 20          25          30

Asn Gln Lys Pro Asp Gly His Trp Val Gly Arg Ala Glu Thr Asn Ala
 35          40          45

Thr Met Glu Ala Gln Trp Cys Leu Ala Leu Trp Phe Leu Gly Leu Glu
 50          55          60

Asp His Pro Leu Arg Val Arg Leu Gly Arg Ala Leu Leu Asp Thr Gln
 65          70          75          80

Arg Pro Asp Gly Ala Trp His Val Phe Tyr Gly Ala Pro Asn Gly Asp
 85          90          95

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

Arg Asp Asp Glu Glu Pro Leu Arg Lys Ala Arg Asp Trp Ile Leu Ser
 115         120         125

Lys Gly Gly Leu Ala Asn Ile Arg Val Phe Thr Arg Tyr Trp Leu Ala
 130         135         140

Leu Ile Gly Glu Trp Pro Trp Glu Lys Thr Pro Asn Ile Leu Pro Glu
 145         150         155         160

Val Ile Trp Leu Pro Thr Trp Phe Pro Phe Ser Ile Tyr Asn Phe Ala
 165         170         175

Gln Trp Ala Arg Ala Thr Leu Met Pro Ile Ala Val Leu Ser Ala His
 180         185         190

Arg Pro Ser Arg Pro Leu Ala Pro Gln Asp Arg Leu Asp Ala Leu Phe
 195         200         205

Pro Gln Gly Arg Asp Ser Phe Asn Tyr Asp Leu Pro Ala Arg Leu Gly
 210         215         220
    
```

ES 2 703 770 T3

Ala Gly Val Trp Asp Val Ile Phe Arg Lys Ile Asp Thr Ile Leu His
 225 230 235 240

Arg Leu Gln Asp Trp Gly Ala Arg Arg Gly Pro His Gly Ile Met Arg
 245 250 255

Arg Gly Ala Ile Asp His Val Leu Gln Trp Ile Ile Arg His Gln Asp
 260 265 270

Tyr Asp Gly Ser Trp Gly Gly Ile Gln Pro Pro Trp Ile Tyr Gly Leu
 275 280 285

Met Ala Leu His Thr Glu Gly Tyr Ala Met Thr His Pro Val Met Ala
 290 295 300

Lys Ala Leu Asp Ala Leu Asn Glu Pro Gly Trp Arg Ile Asp Ile Gly
 305 310 315 320

Asp Ala Thr Phe Ile Gln Ala Thr Asn Ser Pro Val Trp Asp Thr Met
 325 330 335

Leu Ser Leu Leu Ala Phe Asp Asp Ala Gly Leu Gly Glu Arg Tyr Pro
 340 345 350

Glu Gln Val Glu Arg Ala Val Arg Trp Val Leu Lys Arg Gln Val Leu
 355 360 365

Val Pro Gly Asp Trp Ser Val Lys Leu Pro Asp Val Lys Pro Gly Gly
 370 375 380

Trp Ala Phe Glu Tyr Ala Asn Asn Phe Tyr Pro Asp Thr Asp Asp Thr
 385 390 395 400

Ser Val Ala Leu Met Ala Leu Ala Pro Phe Arg His Asp Pro Lys Trp
 405 410 415

Gln Ala Glu Gly Ile Glu Asp Ala Ile Gln Arg Gly Ile Asp Trp Leu
 420 425 430

Val Ala Met Gln Cys Lys Glu Gly Gly Trp Gly Ala Phe Asp Lys Asp
 435 440 445

Asn Asp Lys Lys Ile Leu Ala Lys Ile Pro Phe Cys Asp Phe Gly Glu
 450 455 460

Ala Leu Asp Pro Pro Ser Ala Asp Val Thr Ala His Ile Ile Glu Ala
 465 470 475 480

ES 2 703 770 T3

Phe Ala Lys Val Gly Leu Asp Arg Asn His Pro Ser Ile Val Arg Ala
 485 490 495

Leu Asp Tyr Leu Lys Arg Glu Gln Glu Pro Glu Gly Pro Trp Phe Gly
 500 505 510

Arg Trp Gly Val Asn Tyr Val Tyr Gly Thr Gly Ala Val Leu Pro Ala
 515 520 525

Leu Ala Ala Ile Gly Glu Asp Met Arg Gln Pro Tyr Ile Ala Arg Ala
 530 535 540

Cys Asp Trp Leu Ile Ala Arg Gln Gln Ala Asn Gly Gly Trp Gly Glu
 545 550 555 560

Ser Cys Val Ser Tyr Met Asp Ala Lys Gln Ala Gly Glu Gly Thr Ala
 565 570 575

Thr Ala Ser Gln Thr Ala Trp Ala Leu Met Ala Leu Ile Ala Ala Asp
 580 585 590

Arg Pro Gln Asp Arg Asp Ala Ile Glu Arg Gly Cys Leu Tyr Leu Thr
 595 600 605

Glu Thr Gln Arg Asp Gly Thr Trp Gln Glu Val His Tyr Thr Gly Thr
 610 615 620

Gly Phe Pro Gly Tyr Gly Val Gly Gln Thr Ile Lys Leu Asn Asp Pro
 625 630 635 640

Leu Leu Ser Lys Arg Leu Met Gln Gly Pro Glu Leu Ser Arg Ser Phe
 645 650 655

Met Leu Arg Tyr Asp Leu Tyr Arg His Tyr Phe Pro Met Met Ala Ile
 660 665 670

Gly Arg Val Leu Arg Gln Arg Gly Asp Arg Ser Gly His
 675 680 685

<210> 11
 <211> 680
 <212> PRT
 <213> Streptomyces coelicolor

5

<400> 11

Met Thr Ala Thr Thr Asp Gly Ser Thr Gly Ala Ser Leu Arg Pro Leu
 1 5 10 15

ES 2 703 770 T3

Ala Ala Ser Ala Ser Asp Thr Asp Ile Thr Ile Pro Ala Ala Ala Ala
 20 25 30

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

Leu Leu Ala Lys Gln Asp Ala Glu Gly Trp Trp Lys Gly Asp Leu Glu
 50 55 60

Thr Asn Val Thr Met Asp Ala Glu Asp Leu Leu Leu Arg Gln Phe Leu
 65 70 75 80

Gly Ile Gln Asp Glu Glu Thr Thr Arg Ala Ala Ala Leu Phe Ile Arg
 85 90 95

Gly Glu Gln Arg Glu Asp Gly Thr Trp Ala Thr Phe Tyr Gly Gly Pro
 100 105 110

Gly Glu Leu Ser Thr Thr Ile Glu Ala Tyr Val Ala Leu Arg Leu Ala
 115 120 125

Gly Asp Ser Pro Glu Ala Pro His Met Ala Arg Ala Ala Glu Trp Ile
 130 135 140

Arg Ser Arg Gly Gly Ile Ala Ser Ala Arg Val Phe Thr Arg Ile Trp
 145 150 155 160

Leu Ala Leu Phe Gly Trp Trp Lys Trp Asp Asp Leu Pro Glu Leu Pro
 165 170 175

Pro Glu Leu Ile Tyr Phe Pro Thr Trp Val Pro Leu Asn Ile Tyr Asp
 180 185 190

Phe Gly Cys Trp Ala Arg Gln Thr Ile Val Pro Leu Thr Ile Val Ser
 195 200 205

Ala Lys Arg Pro Val Arg Pro Ala Pro Phe Pro Leu Asp Glu Leu His
 210 215 220

Thr Asp Pro Ala Arg Pro Asn Pro Pro Arg Pro Leu Ala Pro Val Ala
 225 230 235 240

Ser Trp Asp Gly Ala Phe Gln Arg Ile Asp Lys Ala Leu His Ala Tyr
 245 250 255

Arg Lys Val Ala Pro Arg Arg Leu Arg Arg Ala Ala Met Asn Ser Ala

ES 2 703 770 T3

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

ES 2 703 770 T3

Trp Phe Gly Arg Trp Gly Val Asn Tyr Val Tyr Gly Thr Gly Ser Val
515 520 525

Ile Pro Ala Leu Thr Ala Ala Gly Leu Pro Thr Ser His Pro Ala Ile
530 535 540

Arg Arg Ala Val Arg Trp Leu Glu Ser Val Gln Asn Glu Asp Gly Gly
545 550 555 560

Trp Gly Glu Asp Leu Arg Ser Tyr Arg Tyr Val Arg Glu Trp Ser Gly
565 570 575

Arg Gly Ala Ser Thr Ala Ser Gln Thr Gly Trp Ala Leu Met Ala Leu
580 585 590

Leu Ala Ala Gly Glu Arg Asp Ser Lys Ala Val Glu Arg Gly Val Ala
595 600 605

Trp Leu Ala Ala Thr Gln Arg Glu Asp Gly Ser Trp Asp Glu Pro Tyr
610 615 620

Phe Thr Gly Thr Gly Phe Pro Trp Asp Phe Ser Ile Asn Tyr Asn Leu
625 630 635 640

Tyr Arg Gln Val Phe Pro Leu Thr Ala Leu Gly Arg Tyr Val His Gly
645 650 655

Glu Pro Phe Ala Lys Lys Pro Arg Ala Ala Asp Ala Pro Ala Glu Ala
660 665 670

Ala Pro Ala Glu Val Lys Gly Ser
675 680

REIVINDICACIONES

1. Procedimiento para la preparación de derivados de ambroxán, preferentemente ambroxán, de fórmula general (2) **caracterizado porque** se hacen reaccionar derivados de homofarnesol de fórmula general (1) de manera biocatalítica para dar los derivados de ambroxán correspondientes por medio de un polipéptido con la actividad de una homofarnesol-ambroxán ciclasa como enzima,



en donde la actividad del polipéptido es la reacción de homofarnesol para dar ambroxán con sustrato principal homofarnesol, **caracterizado porque** la enzima es un polipéptido codificado por una molécula de ácido nucleico que comprende al menos una molécula de ácido nucleico seleccionada del grupo que consiste en:

- 10 a) molécula de ácido nucleico que codifica un polipéptido que comprende la SEQ ID NO: 2, 5 o 10;
 b) molécula de ácido nucleico que comprende al menos un polinucleótido de la secuencia mostrada en SEQ ID NO: 1;
 c) molécula de ácido nucleico que codifica un polipéptido cuya secuencia presenta una identidad de al menos el 60 % con respecto a las secuencias SEQ ID NO: 2, 5 o 10.
- 15 2. Procedimiento según la reivindicación 1, **caracterizado porque** las homofarnesol-ambroxán ciclasas presentan una conversión en ambroxán en porcentaje en moles de 10 a 100.
3. Procedimiento según una de las reivindicaciones anteriores, **caracterizado porque** la enzima se encuentra en una forma seleccionada del grupo que consiste en:
- 20 a) polipéptido libre, dado el caso purificado o parcialmente purificado con la actividad de una homofarnesol-ambroxán ciclasa;
 b) polipéptido inmovilizado con la actividad de una homofarnesol-ambroxán ciclasa;
 c) polipéptido aislado a partir de células de acuerdo con a) o b);
 d) célula completa, dado el caso células quiescentes o disgregadas, que contiene al menos un polipéptido con la actividad de una homofarnesol-ambroxán ciclasa;
 25 e) lisado celular u homogeneizado celular de las células descritas en d).
4. Procedimiento según la reivindicación 1, **caracterizado porque** las células son microorganismos, preferentemente microorganismos transgénicos que expresan al menos una molécula heteróloga de ácido nucleico que codifica para un polipéptido con la actividad de una homofarnesol-ambroxán ciclasa.
- 30 5. Procedimiento según una de las reivindicaciones anteriores, **caracterizado porque** para la preparación de ambroxán como material de partida se emplea citral, que se hace reaccionar en varias etapas para dar homofarnesol.
6. Procedimiento según una de las reivindicaciones anteriores, **caracterizado porque** la preparación de ambroxán tiene lugar en sistemas acuosos de una sola fase o en sistemas de dos fases.
- 35 7. Procedimiento según una de las reivindicaciones anteriores, **caracterizado porque** la preparación de ambroxán tiene lugar en un modo de proceder discontinuo, de alimentación discontinua o continuo.
8. Procedimiento según una de las reivindicaciones anteriores, en el que la reacción de homofarnesol para dar ambroxán tiene lugar a una temperatura en el intervalo de 0 a 60 °C y/o un valor de pH en el intervalo de 4 a 8.
9. Procedimiento según la reivindicación 8, **caracterizado porque** la reacción de homofarnesol para dar ambroxán tiene lugar a una temperatura en el intervalo de 0 a 45 °C.
- 40 10. Procedimiento según una de las reivindicaciones anteriores, **caracterizado porque** la homofarnesol-ambroxán ciclasa se aisló a partir de un microorganismo seleccionado del grupo que consiste en *Zymomonas mobilis*, *Bradyrhizobium japonicum* y *Rhodospseudomonas palustris*.

11. Procedimiento según una de las reivindicaciones anteriores, **caracterizado porque** la homofarnesol-ambroxán ciclasa se aisló a partir de bacterias transgénicas de las especies *Escherichia coli*, *Pseudomonas putida*, *Burkholderia glumae*, *Streptomyces lividans*, *Streptomyces coelicolor* y *Methylococcus capsalatus*,
5 *Rhodopseudomonas palustris*, *Frankia spec.*, *Rhodopseudomonas palent*, *Frankia alni*, *Bacillus anthracis*, *Burkholderia ambifaria*, en particular *Zymomonas mobilis* y *Bradyrhizobium japonicum*, que sobreexpresan la homofarnesol-ambroxán ciclasa.
12. Uso de un polipéptido con la actividad de una homofarnesol-ambroxán ciclasa para la reacción biocatalítica de homofarnesol para dar ambroxán, **caracterizado porque** el polipéptido es codificado por una molécula de ácido nucleico que comprende al menos una molécula de ácido nucleico seleccionada del grupo que consiste en:
- 10 a) molécula de ácido nucleico que codifica un polipéptido que comprende las SEQ ID NO: 2, 5 o 10;
b) molécula de ácido nucleico que comprende al menos un polinucleótido de la secuencia mostrada en SEQ ID NO: 1;
c) molécula de ácido nucleico que codifica un polipéptido cuya secuencia presenta una identidad de al menos el 60 % con respecto a las secuencias SEQ ID NO: 2, 5 o 10.