



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 361 191**

51 Int. Cl.:
C12P 17/16 (2006.01)
C07D 333/20 (2006.01)
C07D 333/16 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **06830606 .7**
96 Fecha de presentación : **13.12.2006**
97 Número de publicación de la solicitud: **1969132**
97 Fecha de publicación de la solicitud: **17.09.2008**

54 Título: **Reducción enzimática para producir alcoholes ópticamente activos.**

30 Prioridad: **23.12.2005 DE 10 2005 062 662**

45 Fecha de publicación de la mención BOPI:
14.06.2011

45 Fecha de la publicación del folleto de la patente:
14.06.2011

73 Titular/es: **BASF SE**
67056 Ludwigshafen, DE

72 Inventor/es: **Stürmer, Rainer;**
Kessler, Maria;
Hauer, Bernhard;
Friedrich, Thomas;
Breuer, Michael y
Schröder, Hartwig

74 Agente: **Carvajal y Urquijo, Isabel**

ES 2 361 191 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

Reducción enzimática para producir alcoholes ópticamente activos

5 La presente invención se refiere a un método para producir alcanoles ópticamente activos mediante reducción enzimática de las cetonas correspondientes, en especial para la producción de (1S)-3-metilamino-1-(2-tienil)-propan-1-ol y (1S)-3-cloro-1-(2-tienil)-propan-1-ol.

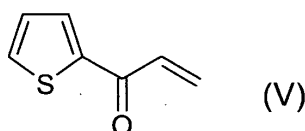
Estado de la técnica:

10 (1S)-3-metilamino-1-(2-tienil)-propan-1-ol ("alcohol de duloxetina") es un bloque de construcción en la síntesis de duloxetina. La Duloxetina® es un principio activo farmacéutico que en la actualidad se encuentra en la etapa de aprobación y está indicado para emplearse en la depresión y la incontinencia.

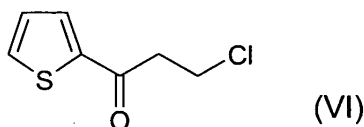
15 EP-B-0273658 describe un método para producir la base correspondiente de duloxetina mediante reacción de 2-acetiltiofeno en una reacción de Mannich con formaldehído y dimetilamina, reducción del grupo ceto de la base de Mannich obtenida en este caso para producir (S)-3-N,N-dimetilamino-1-(tien-2-il)propan-1-ol racémico, eterificación de la función alcohol con fluoruro de naftilo y finalmente conversión del grupo dimetilamino en una función metilamino. El enantiómero deseado del éter naftilo se obtiene mediante el empleo de materias primas quirales o mediante separación de racematos en la etapa del producto final, por ejemplo mediante sales con ácidos ópticamente activos o por cromatografía en una fase estacionaria.

20 US-5,362,886 describe un método análogo en el que se mezcla el propanol racémico, obtenido después de la reducción del grupo ceto, con ácido S-mandélico. El S-enantiómero del alcohol se emplea en las etapas siguientes de la reacción. EP-A-0457559 describe también un proceso análogo al de la EP-B-0273658. Aquí, el grupo ceto de la base de Mannich se reduce con el sistema de reducción asimétrico LAH-Icb (hidruro de litio aluminio -[(2R,2S)-(-)-4-dimetilamino-1,2-difenil-3-metil-2-butanol]) para producir alcohol en forma del S-enantiómero. La desventaja aquí, además de los costes, es la sensibilidad del sistema de reducción LAH-Icb que es estable solo unos pocos minutos.

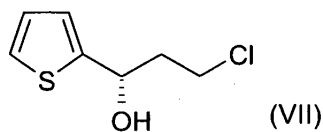
25 W. J. Wheeler y F. Kuo describen en el Journal of Labelled Compounds and Radiopharmaceuticals, volume XXXVI, No. 3, páginas 213 a 223, un método para la producción de duloxetina. Para este propósito se hace reaccionar cloruro de ácido tiofen-2-carboxílico en un acoplamiento de Stille con estanato de vinil-tri-n-butilo en presencia de cantidades catalíticas de bencilcloro-bis(trifenilfosfina) paladio (II) en DMPU (dimetilpropilenurea) para producir 1-(tien-2-il)-propanona de la fórmula (V),



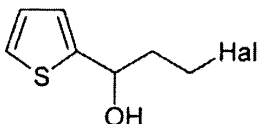
30 la cual se convierte a continuación mediante tratamiento con cloruro de hidrógeno en 3-cloro-1-(tien-2-il)-propan-1-ona de la fórmula (VI)



La cloropropanona así obtenida se reduce a continuación usando una oxazaborilidina quiral y BH_3 para producir (S)-3-cloro-1-(tien-2-il)-propan-1-ol de la fórmula (VII)



El alcohol obtenido de esta manera se convierte mediante reacción sucesiva con yoduro de sodio y se convierte a continuación con metilamina en (S)-3-metilamino-1-(tien-2-il)-propan-1-ol. Mediante reacción siguiente sucesiva con hidruro de sodio, 1-fluoronaftalina y cloruro de hidrógeno se obtiene en forma del cloruro de hidrógeno



5

Hal = Halógeno

WO 05/108590 describe dehidrogenasas para la producción de alcoholes ópticamente activos mediante reducción de las alcanonas correspondientes, en especial para la producción de 3-metilamino-1-(2-tienil)-propanona y 3-cloro-1-(2-tienil)-propanona.

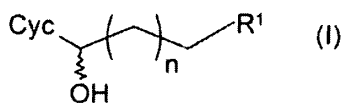
- 10 WO 05/033094 describe métodos enzimáticos y no enzimáticos para la producción de 3-metilamino-1-(tien-2-il)-propan-1-ol así como enzimas para la realización de estos métodos.

Descripción breve de la invención:

El objetivo en el que se fundamentó la invención fue encontrar un camino para la reducción estereoespecífico de alcanonas sustituidas, como la 3-metilamino-1-(2-tienil)-propanona y la 3-cloro-1-(2-tienil)-propanona.

- 15 Este objetivo se logra por el hallazgo sorprendente de las enzimas con actividad de dehidrogenasa que pueden producirse a partir de microorganismos del género *Lactobacillus*, en especial de la especie *Lactobacillus brevis*, que son capaces de la catálisis estereoespecífica de la reacción de arriba.

Un primer objeto de la invención se refiere a un método para producir alcoholes ópticamente activos de la fórmula I

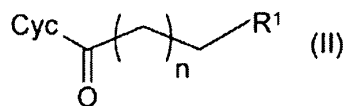


20 donde

n representa un valor de número entero de 0 a 5;

Cyc representa un anillo carbocíclico o heterocíclico, saturado o insaturado, opcionalmente sustituido, de uno o más núcleos, y

- 25 R¹ representa halógeno, SH, OH, NO₂, NR²R³ o NR²R³R⁴X', en cuyo caso R², R³ y R⁴, independientemente uno de otro, representan H o un residuo alquilo inferior de C₁-C₆ o alcoxi inferior de C₁-C₆ y X' representa un contraión, en cuyo caso, en un medio que contiene alcanona de la fórmula II

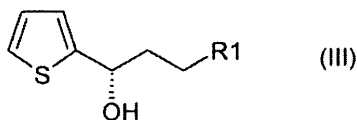


donde n, Cyc y R¹ tienen los significados arriba indicados, se incubó una enzima con una secuencia de polipéptido

(i) SEQ ID NO: 1 o

5 (ii) en la cual, en comparación con la SEQ ID NO: 1 se han modificado hasta un 25% de los residuos de aminoácidos mediante delección, inserción, sustitución o una combinación de éstas y posee además por lo menos un 50% de la actividad enzimática de SEQ ID NO:1,

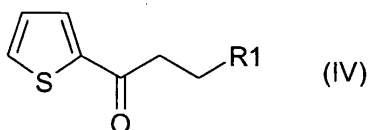
En cuyo caso el compuesto de la fórmula II se reduce enzimáticamente hasta el compuesto de la fórmula I, y se aísla el producto formado. En una forma particularmente preferida de realización, el método sirve para la preparación de derivados del 1-(2-tienil)-(S)-propanol de la fórmula III



10

donde R¹ = Cl o NHCH₃,

donde en un medio que contiene un derivado de la 1-(2-tienil)-propanona de la fórmula IV,



15

este compuesto se reduce enzimáticamente hasta el compuesto de la fórmula III, y el producto formado se aísla esencialmente en forma enantioméricamente pura.

En este método se usa preferentemente una enzima con actividad de dehidrogenasa, la cual puede prepararse a partir de microorganismos de los géneros Lactobacillus.

De manera particularmente preferida se usan dehidrogenasas de Lactobacillus brevis.

20

En una forma particularmente preferida de realización del método, la enzima con actividad de dehidrogenasa se selecciona entre enzimas que comprenden una secuencia de aminoácidos según SEQ ID NO: 1 o una secuencia derivada de la misma, en la cual hasta un 25%, preferible hasta un 20%, particularmente preferible hasta un 15%, en especial hasta un 10, un 9, un 8, un 7, un 6, un 5, un 4, un 3, un 2, un 1 % de los residuos de aminoácidos se han modificado mediante una delección, una sustitución, una inserción o una combinación de delección, sustitución e inserción, en cuyo caso las secuencias de polipéptidos modificadas frente a la SEQ ID NO: 1 además poseen al

25 menos un 50%, preferible un 65%, particularmente preferible un 80%, en especial más de 90 % de la actividad enzimática de la SEQ ID NO:1. En este contexto debe entenderse como actividad enzimática de SEQ ID NO:1 a la capacidad de reducir las cetonas de la fórmula (IV) con R¹ = Cl de manera enantioselectiva hasta el (S)-alcohol con la fórmula general (III).

30

De manera preferente, el método de la invención se realiza adicionando equivalentes de reducción (NADH o NADPH) o en las condiciones (bioquímicas o electroquímicas) que regeneran los equivalentes de reducción consumidos en la reacción. Para estas formas preferidas de realización se hace referencia a los ejemplos 5 y 6 que describen los sistemas de regeneración aplicables de manera general.

Otra forma adecuada de realización para una regeneración de co-factor en el método de la invención es el uso de un compuesto oxidable, inmiscible con agua, por ejemplo un alcohol superior, preferible hexanol, el cual puede presentarse en una segunda fase líquida en el medio de reacción.

- 5 Además, se prefiere hacer reaccionar el compuesto de la fórmula general II, como por ejemplo de la fórmula IV, en presencia de un microorganismo que se selecciona entre bacterias de las familias enterobacteriaceae, pseudomonadaceae, rhizobiaceae, streptomycetaceae y nocardiaceae. El microorganismo puede ser especialmente un microorganismo recombinante que se transforma con un constructo de ácido nucleico, el cual codifica para una enzima con actividad de dehidrogenasa según la definición de arriba.

Es objeto de la invención en especial que

- 10 - un microorganismo se produce una enzima con actividad de dehidrogenasa el cual se aísla de una fuente natural o se produce de manera recombinante,
 - este microorganismo se reproduce,
 - del microorganismo opcionalmente se aísla la enzima con actividad de dehidrogenasa o se produce una fracción de proteína que contiene esta enzima, y
- 15 - el microorganismo de la etapa b) o la enzima de la etapa c) se transfieren a un medio que contiene un compuesto de la fórmula I.

Son objeto de la invención enzimas con actividad de dehidrogenasa con una secuencia de polipéptido

(i) SEQ ID NO: 1 o

- 20 (ii) en la cual hasta un 25% de los residuos de aminoácido se han modificado frente a la SEQ ID NO:1 mediante delección, inserción, sustitución o una combinación de las mismas y la cual además posee al menos 50% de la actividad enzimática de SEQ ID NO:1.

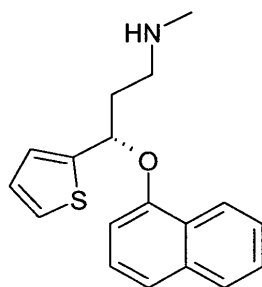
Además, son objeto de la invención secuencias de ácido nucleico codificantes que comprenden la secuencia codificante para un polipéptido según la definición de arriba.

- 25 Además, la invención se refiere a casetes de expresión que comprenden una secuencia de ácido nucleico codificante en conexión operativa con al menos una secuencia regulatoria de ácido nucleico.

Otro objeto de la invención son vectores recombinantes que comprenden al menos un casete de expresión así.

La invención se refiere también a huéspedes procarióticos o eucarióticos que se transforman con al menos un vector de la invención.

- 30 Un último objeto de la invención se refiere al uso de una enzima con actividad de hidrogenasa según la definición de arriba o de un microorganismo que produce esta enzima para la preparación de compuestos de las fórmulas I o II, y su post- procesamiento para la producción de duloxetina (fórmula VIII), por ejemplo.



(VIII)

Descripción detallada de la invención:

A. Términos generales y definiciones

Si no se hacen otras indicaciones, se aplican los siguientes significados generales:

"Halógeno" representa flúor, cloro, bromo o yodo, en especial flúor o cloro.

- 5 "Alquilo inferior" representa residuos alquilo de cadena recta o ramificada, de 1 a 6 átomos de C, como metilo, etilo, i- o n-propilo, n-, i-, sec.- o terc.-butilo, n-pentilo o 2-metilbutilo, n-hexilo, 2-metil-pentilo, 3-metil-pentilo, 2-etil-butilo.

"Alqueno inferior" representa los análogos insaturados, una o varias veces, preferentemente una o dos veces, de los residuos alquilo arriba nombrados con 2 a 6 átomos de carbono, en cuyo caso el enlace doble puede encontrarse en cualquier posición de la cadena de carbono.

- 10 "Alcoxi inferior" representa los análogos terminados en oxígeno de los residuos alquilo de arriba.

"Ariilo" representa un residuo aromático, opcionalmente sustituidos, mono- o polinuclear, preferentemente mono- o binuclear, en especial fenilo o representa un naftilo enlazado por una posición del anillo cualquiera, como 1- o 2-naftilo. Estos residuos ariilo pueden portar, opcionalmente, 1 o 2 sustituyentes iguales o diferentes, seleccionados de halógeno, alquilo inferior, alcoxi inferior según la definición de arriba o trifluorometilo.

- 15 Alcanonas sustituidas, (S)-alcoholes y derivados de las mismas

De acuerdo con la invención, los alcoholes que son accesibles por catálisis enzimática son aquellos de la fórmula de arriba (I) en la cual

n representa un valor de número entero de 0 a 5;

- 20 Cyc representa un anillo carbocíclico o heterocíclico, saturado o insaturado, mono- o polinuclear, opcionalmente sustituido, y

R^1 representa halógeno, SH, OH, NO_2 , NR^2R^3 o $NRZR^3R^4X$, en cuyo caso

R^2 , R^3 y R^4 , independientemente uno de otro, representan H o un residuo de alquilo inferior o alcoxi inferior y X representa un contra-ión.

- 25 Las alcanonas de la fórmula II de arriba, usadas para la síntesis enzimática, son compuestos conocidos de por sí y son accesibles usando métodos de síntesis orgánica conocidos en general (compárese, por ejemplo, EP-A- 0 273 658).

En los compuestos de arriba, n preferentemente representa 0, 1 o 2, en especial representa 1.

- 30 Como ejemplos de grupos carbo- y heterocíclicos Cyc pueden nombrarse en particular grupos mono- o binucleares, preferentemente mononucleares, con hasta 4, preferentemente 1 o 2 heteroátomos en el anillo, iguales o diferentes, seleccionados entre O, N y S:

- 35 Estos anillos carbo- o heterocíclicos comprenden en particular 3 a 12, preferentemente 4, 5 o 6 átomos de carbono en el anillo. Como ejemplos pueden nombrarse ciclopropilo, ciclobutilo, ciclopentilo, ciclohexilo, cicloheptilo, los análogos insaturados una o varias veces de los mismos, como ciclobutenilo, ciclopentenilo, ciclohexenilo, cicloheptenilo, ciclohexadienilo, cicloheptadienilo; así como residuos heterocíclicos de 5 a 7 miembros, saturados o mono- o poliinsaturados, con 1 a 4 heteroátomos que se seleccionan entre O, N y S, en cuyo caso el heterociclo puede condensarse opcionalmente con otro heterociclo o carbociclo. En especial pueden nombrarse residuos heterocíclicos derivados de pirrolidina, tetrahidrofurano, piperidina, morfolina, pirrol, furano, tiofeno, pirazol, imidazol, oxazol, tiazol, piridina, pirano, pirimidina, piridazina, pirazina, cumarona, indol y quinolina.

- 40 Los residuos Cyc en este caso pueden estar enlazadas a la alcanona por cualquier posición del anillo, preferentemente por un átomo de carbono del anillo.

Ejemplos de residuos Cyc adecuados son 2-tienilo, 3-tienilo; 2-furanilo, 3-furanilo; 2-piridilo, 3-piridilo o 4-piridilo; 2-tiazolilo, 4-tiazolilo o 5-tiazolilo; 4-metil-2-tienilo, 3-etil-2-tienilo, 2-metil-3-tienilo, 4-propil-3-tienilo, 5-n-butil-2-tienilo, 4-metil-3-tienilo, 3-metil-2-tienilo; 3-cloro-2-tienilo, 4-bromo-3-tienilo, 2-yodo-3-tienilo, 5-yodo-3-tienilo, 4-fluoro-2-tienilo, 2-bromo-3-tienilo, y 4-cloro-2-tienilo.

5 Los residuos Cyc pueden sustituirse además una o varias veces, como por ejemplo una o dos veces. Preferentemente, los sustituyentes se encuentran en un átomo de carbono del anillo. Ejemplos de sustituyentes adecuados son halógeno, alquilo inferior, alqueno inferior, alcoxi inferior, -OH, -SH, -NO₂ o NR²R³, en cuyo caso R² y R³ poseen los significados de arriba, preferible halógeno o alquilo inferior.

10 R¹ representa en especial halógeno, NR²R³ o NR²R³R⁴X⁻, en cuyo caso R², R³ o R², R³ y R⁴, independientemente uno de otro representan H o un residuo alquilo inferior o alcoxi inferior y X⁻ representa un contra-ión, y preferentemente uno de los residuos R², R³ y R⁴ representa H. Contraiones adecuados son, por ejemplo, aniones de ácido como se obtienen, por ejemplo, en la producción de una sal de adición de ácido. Un ejemplo de esto se nombra, por ejemplo, en la EP-A-0 273 658, a la cual se hace referencia por medio de la presente. Ejemplos preferidos de los residuos R¹ son en especial flúor o cloro, así como NR²R³ donde R² y R³ son iguales o diferentes y
15 representan H o metilo, etilo o n-propilo; particularmente preferible, R¹ representa cloro o -NHmetilo.

C. Enzimas adecuadas con actividad dehidrogenasa.

Enzimas preferidas con actividad de dehidrogenasa comprenden una secuencia de aminoácidos según SEQ ID NO: 1.

20 De acuerdo con la invención también están comprendidos "equivalentes funcionales" de las enzimas divulgadas concretamente con actividad de dehidrogenasa y el uso de estos en los métodos de acuerdo con la invención.

25 "Equivalentes funcionales" o análogos de las enzimas divulgadas concretamente son polipéptidos distintos de las mismas, los cuales poseen además la actividad biológica deseada, como por ejemplo especificidad de sustrato. Por ejemplo, de esta manera se entiende por "equivalentes funcionales" a las enzimas que reducen 3-cloro-1-(tien-2-il)-propan-1-ona hasta el S-alcohol y las cuales tienen 50 %, preferible 60 %, particularmente preferible 75 %, muy particularmente preferible 90 % de la actividad de una enzima con la secuencia de aminoácidos listada en la SEQ ID NO:1. Los equivalentes funcionales además son estables preferentemente entre pH 4 a 10 y poseen ventajosamente un óptimo de pH entre pH 5 y 8 así como un óptimo de temperatura en el rango de 20°C a 80°C.

30 Por "equivalentes funcionales" también se entienden, de acuerdo con la invención, en particular, los mutantes que tienen, en al menos una posición de secuencia, de las secuencias de aminoácidos nombradas arriba, otro aminoácido distinto del aminoácido nombrado concretamente pero que, sin embargo, poseen una de las actividades biológicas nombradas arriba. Los "equivalentes funcionales" comprenden de esta manera los mutantes obtenibles mediante una o varias adiciones, sustituciones, delecciones y/o inversiones de aminoácido, en cuyo caso las modificaciones nombradas pueden darse en cualquier posición de secuencia, en tanto conduzcan hacia un mutante que tenga el perfil de propiedades de acuerdo con la invención. La equivalencia funcional también se da cuando el
35 patrón de reactividad del mutante y el del polipéptido no modificado coinciden cualitativamente; es decir que, por ejemplo, sustratos iguales se convierten a velocidad diferente.

Ejemplos de sustituciones adecuadas de aminoácidos pueden deducirse de la siguiente tabla:

| | Residuo original | Ejemplos de la sustitución |
|----|------------------|----------------------------|
| | Ala | Ser |
| 40 | Arg | Lys |
| | Asn | Gln; His |
| | Asp | Glu |
| | Cis | Ser |
| | Gln | Asn |

| | | |
|----|-----|----------------|
| | Glu | Asp |
| | Gly | Pro |
| | His | Asn; Gln |
| | Ile | Leu; Val |
| 5 | Leu | Ile; Val |
| | Lys | Arg; Gln ; Glu |
| | Met | Leu; Ile |
| | Fe | Met; Leu; Tyr |
| | Ser | Thr |
| 10 | Thr | Ser |
| | Trp | Tyr |
| | Tyr | Trp; Fe |
| | Val | Ile; Leu |

15 De acuerdo con la invención, en especial también se entiende por "equivalentes funcionales" los mutantes que tienen un aminoácido distinto de los nombrados concretamente en al menos una posición de secuencia de las secuencias de aminoácidos arriba nombradas pero que, sin embargo, poseen una de las actividades biológicas nombradas arriba. "Equivalentes funcionales" comprenden de esta manera los mutantes obtenibles mediante una o varias adiciones, sustituciones, deleciones y/o inversiones de aminoácido, en cuyo caso las modificaciones nombradas pueden darse en cualquier posición de secuencia en tanto conduzcan hacia un mutante que tenga el perfil de propiedades de acuerdo con la invención. En especial también se da equivalencia funcional cuando el patrón de reactividad coincide cualitativamente entre los mutantes y el polipéptido no modificado; es decir que, por ejemplo, sustratos iguales se convierten a velocidad diferente.

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

25 "Precursores" son en este caso precursores naturales o sintéticos de los polipéptidos con o sin la actividad biológica deseada.

30 Por la expresión "sales" se entiende tanto sales de grupos carboxílicos como también sales de adición de ácido de las moléculas de proteína de acuerdo con la invención. Sales de grupos carboxílicos pueden prepararse de una manera conocida de por sí y comprenden sales inorgánicas como, por ejemplo, sales de sodio, calcio, amonio, hierro y cinc, así como sales con bases orgánicas como, por ejemplo, aminas, como trietanolamina, arginina, lisina, piperidina y similares. Sales de adición de ácido como, por ejemplo, sales con ácidos minerales, como ácido clorhídrico o ácido sulfúrico y sales con ácidos orgánicos como ácido acético y ácido oxálico también son objeto de la invención.

35 "Derivados funcionales" de polipéptidos de acuerdo con la invención también pueden prepararse con la ayuda de técnicas conocidas en los grupos laterales de aminoácidos funcionales o en sus extremos terminales de N o C. Los derivados de este tipo comprenden, por ejemplo, ésteres alifáticos de grupos carboxílicos, amidas de grupos carboxílicos que pueden obtenerse mediante reacción amoniaco o con una amina primaria o secundaria; derivados de N-acilo de grupos libres amino, preparados mediante reacción con grupos acilo; o derivados O-acilo de grupos hidroxilo, preparados mediante reacción con grupos acilo.

40 "Equivalentes funcionales" también comprenden naturalmente polipéptidos que se encuentran accesibles de otros organismos, así como variantes que tienen lugar naturalmente. A manera de ejemplo, pueden establecerse sectores

de regiones homólogas de secuencia mediante comparación de secuencia y pueden determinarse enzimas equivalentes de conformidad con las premisas concretas de la invención.

"Equivalentes funcionales" también comprenden fragmentos, preferentemente dominios individuales o motivos de secuencia de los polipéptidos de la invención que tienen, por ejemplo, la función biológica deseada.

5 "Equivalentes funcionales" son además proteínas de fusión que tienen una de las secuencias de polipéptido arriba mencionadas o equivalentes funcionales derivados de las mismas y al menos otra secuencia heteróloga funcionalmente distinta en conexión funcional con los terminales N o C (es decir, sin daño funcional esencial recíproco de las partes de proteína de fusión). Ejemplos no limitantes de secuencias heterólogas de este tipo son, por ejemplo, péptidos de señal o enzimas.

10 "Equivalentes funcionales" comprendidos de acuerdo con la invención son homólogos de las proteínas divulgadas concretamente. Estas poseen al menos 60 %, preferentemente al menos 75%, en especial al menos 85 %, como por ejemplo 90%, 95% o 99%, de homología con una de las secuencias divulgadas concretamente, calculada según el algoritmo de Pearson y Lipman, Proc. Natl. Acad. Sci. (USA) 85(8), 1988, 2444-2448. Una homología porcentual de un polipéptido homólogo de la invención significa especialmente identidad porcentual del residuo de aminoácido respecto de la longitud total de una de las secuencias de aminoácidos descritas aquí concretamente.

En el caso de una posible glicosilación de proteína, los "equivalentes funcionales" de acuerdo con la invención comprenden proteínas del tipo arriba designado en forma deglicosilada o en forma glicosilada, así como en formas derivadas obtenibles por modificación del patrón de glicosilación.

20 Homólogos de las proteínas o polipéptidos de acuerdo con la invención pueden generarse por mutagénesis, por ejemplo por mutación puntual o truncamiento de la proteína.

Homólogos de la proteína de acuerdo con la invención pueden identificarse mediante screening (tamizado) de los bancos combinatorios de mutantes, como por ejemplo mutantes de truncamiento. A manera de ejemplo, un banco abigarrado de variantes de proteína puede generarse por mutagénesis combinatoria a nivel de de ácido nucleico, como por ejemplo ligando de manera enzimática una mezcla de oligonucleótidos sintéticos. Hay una gran cantidad de métodos que pueden usarse para la producción de bancos de homólogos potenciales a partir de una secuencia degenerada de oligonucleótidos. La síntesis química de una secuencia degenerada puede realizarse en un sintetizador de ADN y el gen sintético puede ligarse luego en un vector adecuado de expresión. El uso de un conjunto génico degenerado hace posible suministrar todas las secuencias en una mezcla que codifican el conjunto deseado de secuencias potenciales de proteína. Para el experto en la materia son conocidos métodos para sintetizar oligonucleótidos degenerados (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 se conocen varias técnicas para tamizar productos génicos de bancos combinatorios que se han preparado mediante mutaciones puntuales o truncamiento y para tamizar bancos de ADNc para productos génicos con una propiedad seleccionada. Estas técnicas pueden ajustarse para tamizar rápido los bancos de datos que se han generado mediante mutagénesis combinatoria de los homólogos de la invención. Las técnicas usadas de manera más frecuente para tamizar grandes bancos de datos, que se someten a un análisis con alto rendimiento, comprenden la clonación del banco de datos en vectores de expresión replicables, transformación de las células adecuadas con el banco de datos resultante y expresión de los genes combinatorios en condiciones en las que la detección de la actividad deseada facilita el aislamiento del vector que codifica el gen cuyo producto se ha detectado. Recursive Ensemble Mutagenese (REM) (mutagénesis de ensamblaje recursiva), una técnica que aumenta la frecuencia de mutantes funcionales en los bancos, puede usarse en combinación con los ensayos de tamizado para identificar homólogos (Arkin y Yourvan (1992) PNAS 89:7811-7815; Delgrave et al. (1993) Protein Engineering 6(3):327-331).

D. Secuencias de ácido nucleico que codifican para dehidrogenasa

45 Objeto de la invención son especialmente secuencias de ácido nucleico (secuencias de ADN y ARN, monocatenario y bicatenario, como por ejemplo ADNc y ARNm), que codifican para una enzima con la actividad de dehidrogenasa de acuerdo con la invención. Se prefieren secuencias de ácido nucleico que codifican, por ejemplo, para secuencias de aminoácidos según SEQ ID NO:1 o secuencias parciales características de las mismas. Los ácidos nucleicos correspondientes pueden determinarse fácil mediante traducción inversa de la SEQ ID NO:1 según el código genético. También se prefiere la adaptación de la secuencia de ácido nucleico al "codon usage" del organismo huésped planificado en el que debe expresarse el ácido nucleico.

Todas las secuencias de ácido nucleico mencionadas aquí pueden producirse de manera conocida de por sí mediante síntesis química a partir de los bloques de construcción de nucleótidos, como por ejemplo mediante condensación de fragmentos de los bloques de construcción de ácido nucleico complementarios, que se solapan, de la hélice doble. La síntesis química de oligonucleótidos puede efectuarse de manera conocida, por ejemplo según el método de de fosfoamidita (Voet, Voet, 2. Edición, Wiley Press New York, páginas 896-897). La adición de oligonucleótidos sintéticos y el llenado de vacíos con ayuda del fragmento Klenow de la polimerasa-ADN y reacciones de ligación y métodos generales de clonación se describen en Sambrook et al. (1989), Molecular Cloning: A laboratory manual, Cold Spring Harbor Laboratory Press.

También son objeto de la invención secuencias de ácido nucleico (secuencias de ADN y ARN, monocatenario y bicatenario, como por ejemplo ADNc y ARNm), que codifican para uno de los polipéptidos de arriba y sus equivalentes funcionales que son accesibles, por ejemplo, usando análogos artificiales de nucleótido.

La invención se refiere tanto a moléculas aisladas de ácido nucleico que codifican para polipéptidos o proteínas o segmentos biológicamente activos de los mismos de acuerdo con la invención, como también fragmentos de ácido nucleico que pueden usarse, por ejemplo, como sondas de hibridación o cebadores para la identificación o amplificación de ácidos nucleicos codificantes de la invención.

Las moléculas de ácido nucleico de acuerdo con la invención pueden además contener secuencias no traducidas del extremo 3' y/o 5' de la región codificante del gen.

La invención comprende además las moléculas de ácido nucleico complementarias a las secuencias de nucleótidos descritas concretamente, o a un segmento de las mismas.

Las secuencias de nucleótidos de acuerdo con la invención hacen posible la generación de sondas y cebadores que pueden usarse para la identificación y/o clonación de secuencias homólogas en otros tipos de células y organismos. Tales sondas o cebadores comprenden habitualmente una región de secuencia de nucleótidos que hibrida en condiciones "stringent" (rigurosas) (véase abajo) en al menos aproximadamente 12, preferentemente al menos aproximadamente 25, como por ejemplo aproximadamente 40, 50 o 75 nucleótidos sucesivos uno tras otro, de una hebra sense de una secuencia de ácido nucleico de acuerdo con la invención o de una hebra antisense correspondiente.

Una molécula "aislada" de ácido nucleico se separa de otras moléculas de ácido nucleico que están presentes en la fuente natural del ácido nucleico y puede además estar esencialmente libre de otro material celular o medio de cultivo si se prepara mediante técnicas recombinantes o libre de precursores químicos u otros productos químicos, ci se sintetiza químicamente.

Una molécula de ácido nucleico de acuerdo con la invención puede aislarse mediante técnicas estándar de biología molecular y de la información de secuencias suministrada de acuerdo con la invención. A manera de ejemplo, el ADNc puede aislarse de un banco adecuado de ADNc usando una de las secuencias completas divulgadas concretamente o un segmento de la misma como sonda de hibridación y técnicas de hibridación estándar (como se describe, por ejemplo, en Sambrook, J., Fritsch, E.F. y Maniatis, T. Molecular Cloning: A Laboratory Manual. 2. Edición, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989). Además, mediante reacción en cadena de polimerasa puede aislarse una molécula de ácido nucleico que comprende una de las secuencias divulgadas o un segmento de la misma, en cuyo caso se usan los cebadores de oligonucleótidos que se prepararon con base en esta secuencia. El ácido nucleico amplificado puede clonarse a un vector adecuado y caracterizarse mediante análisis de secuencia de ADN. Los oligonucleótidos de la invención pueden prepararse, además, mediante procesos de síntesis estándar, por ejemplo con un sintetizador automático de ADN.

Las secuencias de ácido nucleico de la invención pueden identificarse y aislarse en principio a partir de todos los organismos del género *Lactobacillus*, en especial de *Lactobacillus brevis*. A partir de otros organismos las secuencias de ácido nucleico de acuerdo con la invención pueden aislarse, por ejemplo, con métodos usuales de hibridación o de la técnica PCR, por ejemplo por medio de bancos genómicos o de ADNc. Estas secuencias de ADN hibridan en condiciones estándar con las secuencias de acuerdo con la invención. Para la hibridación pueden usarse de manera ventajosa oligonucleótidos cortos de las regiones conservadas, por ejemplo, del centro activo, que pueden determinarse de manera conocida para el experto en la materia por medio de comparaciones con una con una dehidrogenasa de la invención. Pero para la hibridación también pueden usarse fragmentos más largos de los ácidos nucleicos de la invención o las secuencias completas. Dependiendo de los ácidos nucleicos usados (oligonucleótido, fragmento más largo o secuencia completa) o dependiendo de cuál tipo de ácido nucleico, ADN o ARN, se use para la hibridación, estas condiciones estándar varían. Así, las temperaturas de fusión para ADN :

híbridos de ADN se encuentran, por ejemplo, en cerca de 10 °C más bajo que para ADN : híbridos de ARN de igual longitud.

Por condiciones estándar pueden entenderse, por ejemplo según ácido nucleico, temperaturas entre 42 y 58 °C en una solución acuosa búfer con una concentración entre 0,1 a 5 x SSC (1 X SSC = 0,15 M NaCl, 15 mM de citrato de sodio, pH 7,2) o adicionalmente en presencia de 50% de formamida, como por ejemplo 42 °C en 5 x SSC, 50% de formamida. Las condiciones de hibridación para ADN : híbridos de ADN se encuentran ventajosamente en 0,1 x SSC y temperaturas entre aproximadamente 20 °C hasta 45 °C, preferible entre aproximadamente 30 °C hasta 45 °C. Para ADN : híbridos de ADN las condiciones de hibridación se encuentran ventajosamente en 0,1 x SSC y temperaturas entre aproximadamente 30 °C a 55 °C, preferiblemente entre aproximadamente 45 °C a 55 °C. Estas temperaturas indicadas para la hibridación son, por ejemplo, valores calculados de temperatura de fusión para un ácido nucleico con una longitud de cerca de nucleótidos y un contenido de G + C de 50 % en ausencia de formamida. Las condiciones experimentales para la hibridación de ADN se describen en manuales de consulta de genética, como por ejemplo Sambrook et al., "Molecular Cloning", Cold Spring Harbor Laboratory, 1989, y pueden calcularse de acuerdo con fórmulas conocidas para el experto en la materia, por ejemplo dependiendo de la longitud de los ácidos nucleicos, del tipo de los híbridos o del contenido de G + C. El experto técnico puede inferir más información de los siguientes manuales: Ausubel et al. (eds), 1985, Current Protocols in Molecular Biology, John Wiley & Sons, New York; Hames 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.

También son objeto de la invención derivados de las secuencias de ácido nucleico concretamente divulgadas o derivables.

De acuerdo con la invención también están comprendidas aquellas secuencias de ácido nucleico que comprenden las llamadas mutaciones silentes o que se han modificado en comparación con una secuencia nombrada concretamente, de manera correspondiente al uso de codón de un organismo especial de origen o huésped, como también variantes de las mismas que tienen lugar naturalmente, como por ejemplo variantes splice (de empalme) o variantes de alelos.

Así mismo son objeto las secuencias que pueden obtenerse por sustituciones conservativas de nucleótidos (es decir, el aminoácido en cuestión se reemplaza con un aminoácido de la misma carga, tamaño, polaridad y/o solubilidad).

También son objeto de la invención las moléculas derivadas de los ácidos nucleicos derivados concretamente mediante polimorfismos de secuencia. Estos polimorfismos pueden existir entre individuos de una población debido a la variación natural. Estas variaciones ocasionan usualmente una varianza de 1 a 5 % en la secuencia de nucleótido de un gen.

Por derivados de una secuencia de ácido nucleico de acuerdo con la invención pueden entenderse, por ejemplo, las variantes de alelos que tienen al menos 40 % de homología al nivel derivado de aminoácidos, preferible al menos 60 % de homología, muy particularmente preferible al menos 80, 85, 90, 93, 95 o 98 % de homología por toda la región de secuencia (respecto de la homología al nivel de aminoácidos se hace referencia a las explicaciones de arriba referentes a los polipéptidos). Por regiones parciales de las secuencias, las homologías pueden encontrarse ventajosamente más alto.

Además, por derivados también pueden entenderse homólogos de las secuencias de ácido nucleico de acuerdo con la invención, por ejemplo homólogos de hongos o de bacterias, secuencias truncadas, ADN monocatenario o ARN de la secuencia de ADN codificante y no codificante. Así, a nivel de ADN, por ejemplo, poseen una homología de al menos 40 %, preferible de al menos 60 %, particularmente preferible de al menos 70 %, muy particularmente preferible de al menos 80 % por toda la región indicada de ADN.

Además, por derivados pueden entenderse fusiones con promotores. Los promotores que se ubican antes de las secuencias indicadas de nucleótidos pueden modificarse mediante uno o varios intercambios de nucleótidos, inserciones, inversiones y/o deleciones, pero sin que se perjudique la funcionalidad o la efectividad de los promotores. Además, los promotores pueden incrementarse en su eficacia modificando su secuencia o intercambiarse completamente por promotores más efectivos, incluso de organismos de otras especies.

Por derivados también pueden entenderse variantes cuya secuencia de nucleótidos se hayan modificado en la región de -1 a -1000 bases corriente arriba del codón stop o 0 a 1000 bases corriente abajo después del codón stop, de tal manera que se modifique la expresión génica y/o la expresión de proteína, preferiblemente que se incremente.

Además, la invención también comprende secuencias de ácido nucleico que hibridan en "condiciones rigurosas" con las secuencias codificantes arriba nombradas. Estos polinucleótidos pueden encontrarse tamizando bancos genómicos o de ADNc y proliferarse opcionalmente a partir de éstos con cebadores adecuados por medio de PCR y, a continuación, aislarse con sondas adecuadas, por ejemplo. Además, los polinucleótidos de acuerdo con la invención también pueden sintetizarse de manera química. Por esta propiedad se entiende la capacidad de un poli-oligonucleótido de enlazarse a una secuencia casi complementaria en condiciones rigurosas, mientras que en estas condiciones no se tengan lugar enlaces no específicos entre las partes no complementarias. Para este propósito las secuencias deben ser complementarias en 70-100%, preferentemente en 90-100%. La propiedad de secuencias complementarias de poder enlazarse unas a otras de manera específica se utiliza, por ejemplo, en la técnica de Northern o Southern blot o en el enlazamiento de cebador en PCR o RT-PCR. Habitualmente se emplean para esto oligonucleótidos desde una longitud de 30 pares de bases. Por condiciones rigurosas se entiende, por ejemplo en la técnica Northern blot, el uso de una solución de lavado caliente a 50 - 70 °C, preferentemente a 60 - 65 °C, por ejemplo búfer 0,1x SSC con SDS de 0,1% (20x SSC: NaCl de 3M, citrato de Na de 0,3M, pH 7,0) para elución de sondas de ADNc u oligonucleótidos, hibridados de manera no específica. Tal como se mencionó arriba, en tal caso quedan enlazados uno al otro solo los ácidos nucleicos que son complementarios en gran medida. El establecimiento de condiciones rigurosas se conoce por parte del experto en la materia y se describe, por ejemplo, en Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6..

E. Diseños de constructos de acuerdo con la invención

Además, son objeto de la invención los constructos de expresión que contienen, bajo el control genético de secuencias regulatorias de ácido nucleico, una secuencia de ácido nucleico que codifica para un polipéptido de la invención; y vectores que comprenden al menos uno de estos constructos de expresión.

Tales constructos de la invención comprenden preferentemente un promotor 5'-corriente arriba de la secuencia codificante respectiva y una secuencia terminadora 3'-corriente abajo así como, opcionalmente, otros elementos regulatorios habituales, y que se conectan operativamente en cada caso con la secuencia codificante.

Por una "conexión operativa" se entiende la disposición secuencial de promotor, secuencia codificante, terminador y opcionalmente otros elementos regulatorios de tal manera que cada uno de los elementos regulatorios puede cumplir su función en la expresión de la secuencia codificante conforme a lo prescrito. Ejemplos de secuencias operativamente conectables son secuencias de targeting (selección de objetivo) y enhancers (promotores), señales de poliadenilación y similares. Otros elementos regulatorios comprenden marcadores seleccionables, señales de amplificación, orígenes de replicación y similares. Secuencias regulatorias adecuadas se describen, por ejemplo, en Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990).

Por un constructo de ácido nucleico de acuerdo con la invención pueden entenderse especialmente aquellos en los que para una dehidrogenasa de la invención se ha conectado de manera operativa o funcional el gen con una o varias señales de regulación para controlar, por ejemplo incrementar, la expresión génica.

Adicionalmente a estas secuencias de regulación, la regulación natural de estas secuencias puede aún estar presente corriente arriba de los genes estructurales propios y puede haberse modificado genéticamente de manera opcional, de tal modo que se haya eliminado la regulación natural y elevado la expresión de los genes. Pero el constructo de ácido nucleico también puede estar estructurado de una manera más sencilla, lo que significa que no se insertaron señales de regulación adicionales antes de la secuencia codificante y no retiró el promotor natural con su regulación. En lugar de esto, la secuencia natural de regulación se muta de tal modo que ya no se efectúa una regulación y se incrementa la expresión de gen.

Un constructo preferido de ácido nucleico también contiene ventajosamente uno o varios de las ya mencionadas secuencias "enhancer", conectadas funcionalmente con el promotor, las cuales hacen posible una expresión elevada de la secuencia de ácido nucleico. En el extremo 3' de las secuencias de ADN pueden insertarse secuencias ventajosas adicionales, como otros elementos regulatorios o terminadores. Los ácidos nucleicos de acuerdo con la invención pueden estar contenidos en una o varias copias en el constructo. En el constructo pueden estar contenidos aún más marcadores, como resistencias a antibióticos o genes que complementan auxotrofías, opcionalmente para la selección del constructo.

Secuencias ventajosas de regulación para el proceso de acuerdo con la invención se encuentran contenidas, por ejemplo, en promotores tales como promotores cos, tac, trp, tet, trp-tet, lpp, lac, lpp-lac, lacIq, T7, T5, T3, gal, trc, ara, rhaP (rhaP_{BAD})SP6, lambda-P_R o lambda-P_L, los cuales encuentran aplicación de manera ventajosa en bacterias gram-negativas. Otras secuencias ventajosas de regulación están contenidas, por ejemplo, en los promotores gram-positivos amy y SPO2, en los promotores de levadura o de hongo ADC1, MAlpha, AC, P-60, CYC1, GAPDH, TEF, rp28, ADH. En este contexto también son ventajosos los promotores de la piruvatodecarboxilasa y de la

metanoxidasa, por ejemplo a partir de *Hansenula*. También pueden usarse promotores artificiales para la regulación.

5 Para la expresión en un organismo huésped, el constructo de ácido nucleico se inserta de manera ventajosa en un vector como por ejemplo un plásmido o un fago que hacen posible una expresión óptima de los genes en el huésped. Por vectores también se entienden, aparte de plásmidos y fagos, todos los otros vectores conocidos para el experto en la materia; es decir, por ejemplo, virus como SV40, CMV, baculovirus y adenovirus, transposones, elementos IS, fásmidos, cósmidos y ADN lineal o circular. Estos vectores pueden replicarse de manera autónoma en el organismo huésped o replicarse de manera cromosómica. Estos vectores representan una modalidad más de la invención. Plásmidos adecuados son, por ejemplo, en *E. coli* pLG338, pACIC184, pBR322, pUC18, pUC19, pKC30, pRep4, pHS1, pKK223-3, pDHE19.2, pHS2, pPLc236, pMBL24, pLG200, pUR290, pIN-III¹³-B1, Igt11 o pBdCI, en *Streptomyces* pIJ101, pIJ364, pIJ702 o pIJ361, en *Bacillus* pUB110, pC194 o pBD214, en *Corynebacterium* pSA77 o pAJ667, en hongos pALS1, pIL2 o pBB116, en levaduras 2alphaM, pAG-1, YEp6, YEp13 o pEMBLYe23 o en vegetales pLGV23, pGHlac⁺, pBIN19, pAK2004 o pDH51. Los plásmidos nombrados representan una pequeña selección de los plásmidos posibles. Otros plásmidos son bien conocidos para el experto en la materia y pueden inferirse, por ejemplo, del libro *Cloning Vectors* (Eds. Pouwels P. H. et al. Elsevier, Amsterdam-New York-Oxford, 1985, ISBN 0 444 904018).

Para la expresión de los otros genes contenidos, el constructo de ácido nucleico también contiene ventajosamente y de manera adicional secuencias regulatorias 3'- y/o 5'-terminales para el aumento de la expresión, las cuales se seleccionan para una expresión óptima dependiendo del organismo huésped seleccionado y el gen o los genes.

20 Estas secuencias regulatorias deben hacer posible la expresión dirigida de los genes y de la expresión de proteína. Dependiendo del organismo huésped, esto puede significar, por ejemplo, que el gen se expresa o se sobre-expresa solo después de inducción o que se expresa y/o sobre-expresa inmediatamente.

25 Las secuencias o factores regulatorios pueden influir positivamente en este caso de manera preferente la expresión génica de los genes introducidos e incrementarla de esta manera. De esta manera puede efectuarse ventajosamente un refuerzo de los elementos regulatorios a nivel de transcripción, usando fuertes señales de transcripción como los promotores y/o "enhancers". Pero además también es posible un refuerzo de la traducción (translation) mejorando, por ejemplo, la estabilidad del ARNm.

30 En otra modalidad del vector, el vector que contiene el constructo de ácido nucleico de la invención o el ácido nucleico de la invención también puede introducirse a los microorganismos en forma de un ADN lineal e integrarse mediante recombinación heteróloga u homóloga al genoma del organismo huésped. Este ADN lineal puede componerse de un vector linealizado, como un plásmido, o solo del constructo de ácido nucleico o del ácido nucleico de la invención.

35 Para una expresión óptima de genes heterólogos en organismos es ventajoso modificar las secuencias de ácido nucleico de manera correspondiente con el uso específico de codón ("codon usage") empleado en el organismo. El "codon usage" puede determinarse fácilmente por medio de análisis de ordenador de otros genes conocidos del organismo en cuestión.

40 La preparación de un casete de expresión de acuerdo con la invención se efectúa mediante fusión de un producto adecuado con una secuencia codificante adecuada de nucleótidos así como con una señal de terminador y de poliadenilación. Para este propósito se usan técnicas corrientes de recombinación y de clonación, como se describen, por ejemplo, en T. Maniatis, E.F. Fritsch y J. Sambrook, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) así como en T.J. Silhavy, M.L. Berman y L.W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) y en Ausubel, F.M. et al., *Current Protocols in Molecular Biology*, Greene Publishing Assoc. and Wiley Interscience (1987).

45 Para la expresión en un organismo huésped, el constructo recombinante de ácido nucleico o el constructo del gen se insertan ventajosamente a un vector específico al huésped, el cual hace posible una expresión óptima de los genes en el huésped. Los vectores son bien conocidos para el experto en la materia y pueden inferirse, por ejemplo, de "Cloning Vectors" (Pouwels P. H. et al., editor Elsevier, Amsterdam-New York-Oxford, 1985).

F. Organismos huéspedes útiles de la invención

50 Con ayuda de los vectores o constructos de la invención pueden producirse microorganismos recombinantes, los cuales se transforman, por ejemplo, con al menos un vector de acuerdo con la invención y pueden emplearse para la producción de los polipéptidos de la invención. Ventajosamente los constructos recombinantes arriba descritos de

acuerdo con la invención se introducen a un sistema huésped adecuado y se expresa. En tal caso, métodos de clonación y de transfección corrientes conocidos para el experto en la materia, como por ejemplo co-precipitación, fusión de protoplastos, electroporación, transfección retroviral y similares, se usan preferentemente para hacer que los ácidos nucleicos nombrados se expresen en el sistema respectivo de expresión. Sistemas adecuados se describen, por ejemplo, en *Current Protocols in Molecular Biology*, F. Ausubel et al., Editor, Wiley Interscience, New York 1997, o Sambrook et al. *Molecular Cloning: A Laboratory Manual*. 2. Edición, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

De acuerdo con la invención, también pueden prepararse microorganismos recombinados homológamente. Para este propósito se prepara un vector que contiene al menos un segmento de un gen de la invención o de una secuencia codificante en donde se ha introducido opcionalmente al menos una delección, una adición o sustitución de aminoácido para modificar la secuencia de la invención, por ejemplo para desbaratar funcionalmente (vector "knockout"). La secuencia introducida también puede ser un homólogo de un microorganismo relacionado o de una fuente de mamífero, de levadura o de insecto, por ejemplo. El vector usado para la recombinación homóloga puede estar diseñado alternativamente de tal manera que el gen endógeno muta en el caso de recombinación homóloga o se modifica de otra manera, aunque aún codifica la proteína funcional (por ejemplo, la región regulatoria situada corriente arriba puede modificarse de tal manera que de esta manera se modifique la expresión de la proteína endógena). El segmento modificado del gen de la invención es un vector homólogo de recombinación. La construcción de vectores adecuados para la recombinación homóloga se describe, por ejemplo, en Thomas, K.R. y Capecchi, M.R. (1987) *Cell* 51:503.

Como organismos huéspedes recombinantes para el ácido nucleico de la invención o el constructo de ácido nucleico se consideran en principio todos los organismos procariontes o eucariontes. Ventajosamente se usan microorganismos como bacterias, hongos o levaduras en calidad de organismos huéspedes. De manera ventajosa se usan bacterias gram-positivas o gram-negativas, preferible bacterias de las familias Enterobacteriaceae, Pseudomonadaceae, Rhizobiaceae, Streptomycetaceae o Nocardiaceae, particularmente preferible bacterias de los géneros *Escherichia*, *Pseudomonas*, *Streptomyces*, *Nocardia*, *Burkholderia*, *Salmonella*, *Agrobacterium* o *Rhodococcus*. Muy particularmente se prefiere el género y la especie *Escherichia coli*. Otras bacterias ventajosas pueden encontrarse además en el grupo de las alfa-proteobacterias, beta-proteobacterias o gamma-proteobacterias.

El organismo huésped o los organismos huéspedes según la invención contienen en tal caso preferentemente al menos una de las secuencias de ácido nucleico, de los constructos de ácido nucleico o vectores descritos en esta invención, los cuales codifican para una enzima con la actividad de hidrogenasa de acuerdo con la invención.

Los organismos usados en el proceso de la invención se hacen crecer o se cultivan de una manera conocida para el experto en la materia dependiendo del organismo huésped. Los microorganismos regularmente se hacen crecer en un medio líquido que contiene una fuente de carbono la mayor parte en forma de azúcares, una fuente de nitrógeno la mayor parte en forma de fuentes orgánicas de nitrógeno tales como extracto de levadura o sales como sulfato de amonio, microelementos como sales de hierro, de manganeso, de magnesio y, opcionalmente, vitaminas, a temperaturas entre 0 °C y 100 °C, preferible entre 10 °C a 60 °C bajo exposición de gas nitrógeno. En tal caso, el pH del líquido nutriente puede mantenerse en valores fijos, lo que significa que durante el cultivo se regulan o no. El cultivo puede efectuarse a manera de "batch" (lote), "semi batch" (semicontinua) o continuamente. Los nutrientes pueden cargarse al inicio de la fermentación o de manera semicontinua o continua agregarse después. La cetona puede adicionarse directamente al cultivo o de manera ventajosa después del cultivo. Las enzimas pueden aislarse según el método descrito en los ejemplos a partir de los organismos o usarse como extracto crudo para la reacción.

G. Preparación recombinante de los polipéptidos de la invención

También son objeto de la invención métodos para la preparación recombinante de los polipéptidos de la invención o de los fragmentos funcionales, biológicamente activos, de los mismos, en cuyo caso se cultiva un microorganismo que produce polipéptidos, induce opcionalmente la expresión de los polipéptidos y los aísla del cultivo. Los polipéptidos también pueden producirse así a escala industrial si se desea.

El organismo recombinante puede cultivarse y fermentarse de acuerdo con métodos conocidos. Las bacterias pueden propagarse, por ejemplo, en medio TB o LB y a una temperatura de 20 a 40°C y un valor de pH de 6 a 9. En detalle se describen condiciones de cultivo adecuadas, por ejemplo, en T. Maniatis, E.F. Fritsch and J. Sambrook, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989).

Luego, las células se rompen su los polipéptidos no se secretan al medio de cultivo y el producto se obtiene según métodos conocidos de aislamiento de proteína a partir del lisado. Las células pueden abrirse a discreción mediante ultrasonido de alta frecuencia, mediante presión alta como, por ejemplo, en una célula de presión francesa, por

osmólisis, por acción de detergentes, enzimas líticas o solventes orgánicos, mediante homogeneizadores o por la combinación de varios de los métodos listados.

5 Puede lograrse una purificación de los polipéptidos con métodos cromatográficos conocidos, como cromatografía de tamices moleculares (filtración en gel), como cromatografía en Q-sefarsa, cromatografía de intercambio iónico y cromatografía hidrófoba, así como con otros métodos usuales como ultrafiltración, cristalización, precipitación de proteínas por acción de sal, diálisis y electroforesis nativa en gel. Métodos adecuados se describen, por ejemplo, en Cooper, F. G., Biochemische Arbeitsmethoden (Métodos bioquímicos de trabajo), Editorial Walter de Gruyter, Berlin, New York o en Scopes, R., Protein Purification, Editorial Springer Verlag, New York, Heidelberg, Berlin.

10 Puede ser ventajoso usar para el aislamiento de la proteína recombinante a sistemas vector que alargan el ADNc en determinadas secuencias de nucleótidos y de esta manera codifican para polipéptidos o proteínas de fusión modificados que sirven para una modificación más sencilla. Modificaciones adecuadas de este tipo son, por ejemplo, los así llamados "tags" que funcionan como anclas, como por ejemplo la modificación conocida como ancla de hexa-histidina o epítopes que pueden reconocerse como antígenos de anticuerpos (descritos por ejemplo en Harlow, E. y Lane, D., 1988, Antibodies: A Laboratory Manual. Cold Spring Harbor (N.Y.) Press). Estas anclas pueden servir para adherir las proteínas en un soporte sólido, como por ejemplo una matriz polimérica, que puede empacarse a una columna cromatográfica o puede usarse en una placa de microtitulado o en algún otro soporte.

15 Simultáneamente, estas anclas también pueden usarse para reconocer las proteínas. Además, para reconocer las proteínas pueden usarse marcadores habituales como colorantes de fluorescencia, marcadores de enzima que después de la reacción con un sustrato forman un producto de reacción detectable, o marcadores radioactivos, solos o en combinación con las anclas para derivatizar las proteínas.

20 H. Realización del método de la invención para preparar (S)-alcanoles

Las enzimas usadas de acuerdo con la invención que tienen actividad de dehidrogenasa pueden usarse en el método de la invención como enzima libre o inmovilizada.

25 El método de acuerdo con la invención se realiza ventajosamente a una temperatura entre 0 °C a 95 °C, p referible entre 10 °C a 85 °C, particularmente preferible entre 15 °C a 75 °C.

El valor de pH en el método de acuerdo con la invención se mantiene ventajosamente entre pH 4 y 12, preferible entre pH 4,5 y 9, particularmente preferible entre pH 5 y 8.

30 Por productos enantioméricos puros o quirales o alcoholes ópticamente activos en el método de la invención se entienden enantiómeros que muestran un enriquecimiento de enantiómero. En el método se logran preferiblemente unidades de enantiómero de al menos 70 %ee, preferible de mínimo 80 %ee, particularmente preferible de mínimo 90 %ee, muy particularmente preferible mínimo 98 %ee erreicht.

35 Para el método de la invención pueden usarse células crecientes que contienen los ácidos nucleicos, constructos de ácido nucleico o vectores de la invención. También pueden usarse células reposando o abiertas. Por células abiertas se entienden, por ejemplo, células que se han hecho permeables mediante un tratamiento con solventes, por ejemplo, o células que se han roto mediante un tratamiento enzimático, mediante un tratamiento mecánico (por ejemplo French Press o ultrasonido) o mediante otros métodos. Los extractos crudos obtenidos de esta manera son ventajosamente adecuados para el método de la invención. Para el método también pueden usarse enzimas purificadas o parcialmente purificadas. Así mismo son adecuados los microorganismos o enzimas inmovilizados que pueden encontrar aplicación ventajosamente en la reacción.

40 Si en el método de acuerdo con la invención se usan organismos o enzimas libres entonces éstos se separan antes de la extracción de manera conveniente, por ejemplo mediante una filtración o centrifugación.

45 El producto preparado en el método de acuerdo con la invención, por ejemplo (1S)-3-metilamino-1-(2-tienil)-propan-1-ol, puede obtenerse de manera ventajosa a partir de la solución acuosa de la reacción mediante extracción o destilación. La extracción puede repetirse varias veces para incrementar el rendimiento. Ejemplos de agentes de extracción son solventes como tolueno, cloruro de metileno, acetato de butilo, éter de diisopropilo, benceno, MTBE o acetato, sin limitarse a éstos.

Después de concentrar la fase orgánica, regularmente pueden obtenerse los productos en buenas purzas químicas, lo que significa una pureza química mayor a 80 %. Después de la extracción la fase orgánica con el producto solo puede concentrarse parcialmente y cristalizar el producto. Para este propósito se enfría la solución

5 ventajosamente a una temperatura de 0 °C a 10 °C. La cristalización también puede efectuarse directamente a partir de la solución orgánica o a partir de una solución acuosa. El producto cristalizado puede tomarse de nuevo para una nueva cristalización en el mismo solvente o en otro y cristalizarse una vez más. Si se requiere puede seguir aumentando la pureza enantiomérica del producto mediante la cristalización realizada a continuación al menos una vez.

10 En los tipos de procesamiento nombrados el producto del proceso de la invención puede aislarse en rendimientos de 60 a 100 %, preferible de 80 a 100 %, particularmente preferible de 90 a 100 %, respecto del sustrato empleado para la reacción, como por ejemplo de 3-metilamino-1-(2-tienil)-propan-1-ona. El producto aislado se caracteriza por una alta pureza química de > 90 %, preferible > 95 % particularmente preferible de > 98 %. Además, los productos tienen una alta pureza enantiomérica que, en caso de necesidad, puede seguir incrementándose mediante cristalización.

El método de la invención puede operarse en forma de lotes, en forma semi-continua o en forma continua.

La realización del método puede efectuarse de manera ventajosa en bioreactores, como se describe por ejemplo en Biotechnology, volumen 3, 2. Edición, Rehm et al editores, (1993) en especial el capítulo II.

15 La descripción de arriba y los ejemplos siguientes sirven solo para ilustrar la invención. Las numerosas variaciones posibles, obvias para el técnico en la materia, también están comprendidas de acuerdo con la invención.

Parte experimental:

Ejemplo 1

Análisis de 3-cloro-1-(tien-2-il)-propan-1-ona y 3-cloro-1-(tien-2-il)-propan-1-ol

20 La concentración de 3-cloro-1-(tien-2-il)-propan-1-ona y 3-cloro-1-(tien-2-il)-propan-1-ol pueden determinarse por medio de HPLC. Según la elección de las fases estacionaria y móvil, además de la concentración también puede determinarse el valor ee.

a) Análisis aquiral

La cuantificación de la reacción se realizó con el siguiente sistema:

| | | |
|----|-----------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| | Fase estacionaria: | Chromolith SpeedROD RP18, 50*4, 6Pm, Merck (Darmstadt) climatizada a 45°C |
| 25 | Fase móvil: | Eluente A: KH ₂ PO ₄ de 10mM, pH 2.5 Eluente B: acetonitrilo gradiente: 0-0.5 min, 35%B; 0.5-1.0 min 35 sobre 80% de B; 1.0-1.2 min 80% deB; 1.2-1.3 min 80% - 35%B; 1.3-2.0 min 35%B; |
| | Velocidad de flujo: | 1.5 ml/min |
| | Detección: | detección UV a 230 y 260nm |
| 30 | Tiempos de retención: | 3-cloro-1-(tien-2-il)-propan-1-ona: cerca de 1.6 min 3-Cloro-1-(tien-2-il)-propan-1-ol: cerca de 1.3 min |

Con material auténtico se produce una serie de calibración, por medio de la cual puede determinarse la concentración de las muestras desconocidas.

b) Análisis quiral

| | | |
|----|---------------------|-------------------------------------------------------------------|
| 35 | Fase estacionaria: | Chiracel OD-H, 250*4, 6µm, Daicel, calentado a 40°C |
| | Fase móvil: | Eluente A: n-Hexano Eluente B: iso-propanol isocrático con 2.5% B |
| | Velocidad de flujo: | 1.0ml/min |
| | Detección: | Detección UV a 230 y 260nm |

Tiempos de retención: 3-cloro-1-(tien-2-il)-propan-1-ona: cerca de 9.5 min (1S)-3-cloro-1-(tien-2-il)-propan-1-ol: cerca de 16.6min (1R)-3-cloro-1-(tien-2-il)-propan-1-ol: cerca de 18.3 min

Con material auténtico se produce una serie de calibración por medio de la cual puede determinarse la concentración de muestras desconocidas.

5 **Ejemplo 2:**

Suministro de glucosa-dehidrogenasa para regenerar co-factor

10 Para la regeneración de cofactor puede usarse glucosa dehidrogenasa. La enzima es accesible a partir de fuentes comerciales (por ejemplo, Jülich Fine Chemicals Order-No. 22.10 o 19.10) o también puede prepararse fácilmente basándose en la secuencia conocida de ADN. Se usó un clon de E. coli XL10 Gold que contiene el gen de glucosa dehidrogenasa en el plásmido pUC19 a partir de Bacillus subtilis (Genbank-Acc.No. M12276); este constructo lleva la denominación E. coli LU11293.

Para fermentar E. coli LU11293 se empleó el siguiente medio:

| | |
|-----------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------|
| 560 g | de extracto de levadura (65 %) |
| 448 g | de Trypton (Difco) |
| 42 g | de KH_2PO_4 |
| 84 g | de Na_2HPO_4 |
| 644 g | de glicerina (99%) |
| 100 mL | de solución SL4 (5 veces) |
| | |
| 1 g | de Tegosipon 3062 |
| | Rellenar el medio con agua hasta 13,5 L, ajustar el valor de pH a 7,0, |
| | sacar cerca de 300 mL para el precultivo, |
| | después esterilizar por 30 minutos 122 °C. |
| | |
| | Adicionar solución salina estéril* (previamente sacar la solución salina para los matraces agitados, véase Rapport). |
| *Solución salina: 2,1 g de $\text{CaCl}_2 \cdot 2 \text{H}_2\text{O}$ | |
| 3,5 g de $\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$ | |
| 14 g de NH_4Cl | |
| Disolver 14 mL de solución de ampicilina (100 mg/mL) en 500 mL de agua y filtrar de manera estéril. | |

15 150 mL de medio se esterilizaron respectivamente en dos matraces de Erlenmeyer de 1 L y se completó con 5 mL de solución salina estéril. Después de inocular una placa de LB-ampicilina-agar se incubaron los pre-cultivos por 12 horas a 37 °C y 200 rpm y se agregaron al medio de fermentación. La fermentación inició a 37°C, 0,1 bar de presión

5 interna, pH 7,0 (regulación con 20% de ácido fosfórico y 25% de NaOH) con una velocidad de aplicación de gas de 7,5 L/min y 300 rpm (regulación de pO_2 entre 20 y 50% con 10-20 L/min de alimentación de aire y 500-1500 rpm). Después de 2 h para la inducción se adicionó IPTG de 0,1 mM y se finalizó la fermentación después de un tiempo total de 13 h. Después de cosechar y lavar las células (1,3 kg) éstas se almacenaron hasta el uso (2-20 g/L en la mezcla de reacción) a $-20^{\circ}C$.

Ejemplo 3:

Regeneración de cofactor

10 La regeneración del cofactor también puede realizarse mediante la dehidrogenasa que causa la conversión de alcanonas (II) en alcoholes (I). En este caso no es necesaria la adición de una enzima de regeneración por separado. La dehidrogenasa acepta diferentes alcoholes simples como agentes reductores. Se oxidan hasta los compuestos correspondientes de carbonilo. Un alcohol simple que es adecuado para regenerar NADH con dehidrogenasa es iso-propanol. En lugar de la glucosedehidrogenasa y glucosa se usa 10% iso-propanol en la mezcla de reacción.

SEQUENCE LISTING

<110> BASF Aktiengesellschaft

<120> Nueva proteína para la síntesis de alcoholes ópticamente activos

<130> PF 57486

5 <160> 1

<170> PatentIn version 3.1

<210> 1

<211> 249

<212> PRT

10 <213> Lactobacillus brevis

<400> 1

ES 2 361 191 T3

Met Thr Asp Arg Leu Lys Asp Lys Val Ala Ile Ile Thr Gly Gly Val
 1 5 10 15

Ala Gly Ile Gly Leu Gly Ile Ala Glu Cys Tyr Val Arg Glu Gly Ala
 20 25 30

Lys Val Val Val Thr Ala Asn His Asn Val Asp Gly Gly His Ala Ala
 35 40 45

Val Ala Lys Phe Gly Asp Asp Val Ser Leu Phe Val Gln Gln Asp Val
 50 55 60

Ser Lys Glu Ala Asp Trp Gln Lys Val Ile Asp Ala Thr Ile Ala Lys
 65 70 75 80

Phe Gly Arg Val Asp Ile Leu Val Asn Asn Ala Gly Ile Gly Gly Val
 85 90 95

Asn Thr Ala Ile Glu Asp Leu Asp Leu Ala Asp Trp Gln Lys Val Ile
 100 105 110

Asp Val Asn Leu Thr Ala Asn Phe Leu Gly Glu Lys Ala Ala Ile Lys
 115 120 125

Ala Met Lys Gln Thr Ala Asp Ala Lys Gly Ser Ile Ile Asn Val Ser
 130 135 140

ES 2 361 191 T3

Ser Val Ala Gly Leu Val Gly Leu Pro Met Ala Pro Ala Tyr Ser Ala
145 150 155 160

Ser Lys Gly Gly Ser Arg Leu Leu Thr His Ala Thr Ala Leu Asn Leu
165 170 175

Ala Gln Arg Gly Ile Asp Ile Arg Val Asn Ser Val His Pro Gly Trp
180 185 190

Ile Asp Thr Ser Ile Val Pro Glu Ala Ala Arg Gln Gln Ile Ile Ala
195 200 205

Thr Ile Pro Val Gly His Met Gly Gln Pro Gln Asp Ile Gly Glu Ile
210 215 220

Cys Val Tyr Leu Gly Ser Asp Glu Ser Arg Phe Ala Asn Gly Ala Glu
225 230 235 240

Phe Thr Val Asp Gly Gly Gln Arg Ala
245

PF 57486

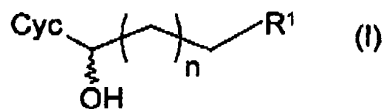
2

PF 57486

5 26.10.06 DP/SA

REIVINDICACIONES

1. Método para la preparación de alcoholes ópticamente activos de la fórmula I



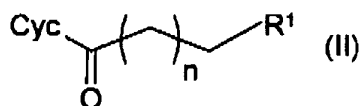
donde

5 n representa un valor de número entero desde 0 a 5;

Cyc representa un anillo carbocíclico o heterocíclico, saturado o insaturado, mono- o polinuclear, opcionalmente sustituido, y

R¹ representa halógeno, SH, OH, NO₂, NR²R³ o NR²R³R⁴X⁻, donde R², R³ y R⁴, independientemente unos de otros, representan H o un residuo alquilo inferior de C₁-C₈ o alcoxi inferior de C₁-C₈ y X⁻ representa un contra-ión,

10 en cuyo caso en un medio que contiene una alcanona de la fórmula II



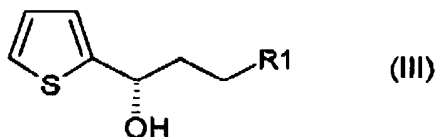
en la que n, Cyc y R¹ poseen los significados indicados arriba, se incuba una enzima con una secuencia de polipéptido

(i) SEQ ID NO: 1 o

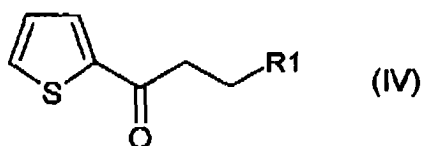
15 (ii) (ii) en la que hasta el 25% de los residuos de aminoácidos se han modificado frente a la secuencia SEQ ID NO:1 mediante delección, inserción, sustitución o una combinación de éstas y posee aún al menos el 50% de la actividad enzimática de la SEQ ID NO:1,

en cuyo caso el compuesto de la fórmula II se reduce de manera enzimática hasta el compuesto de la fórmula I y se aísla el producto formado.

20 2. Método según la reivindicación 1, para la preparación de derivados del 1-(2-tienil)-(S)-propanol de la fórmula III



donde R¹ = Cl o NHCH₃, en cuyo caso en un medio que contiene derivados de la 1-(2-tienil)-propanona de la fórmula IV



se reduce enzimáticamente este compuesto hasta el compuesto de la fórmula III, y se aísla el producto enantioméricamente puro con al menos 90%ee.

3. Método según una de las reivindicaciones precedentes, donde la enzima se selecciona entre enzimas de microorganismos del género de Lactobacillus.
- 5 4. Método según una de las reivindicaciones precedentes, donde la conversión se realiza adicionando equivalentes de reducción o en condiciones que regeneran los equivalentes de reducción consumidos durante la conversión.
5. Método según la reivindicación 4, **caracterizado porque** se realiza una regeneración de los cofactores con un alcohol de C₂-C₁₀.
- 10 6. Método según una de las reivindicaciones precedentes, donde la conversión del compuesto de la fórmula II se efectúa en presencia de un microorganismo que se selecciona entre bacterias de las familias de Enterobacteriaceae, Pseudomonadaceae, Rhizobiaceae, Streptomycetaceae y Nocardiaceae.
7. Método según una de las reivindicaciones precedentes, donde el microorganismo es un microorganismo recombinante que se transforma con un constructo de ácido nucleico, el cual codifica para una enzima según la definición en la reivindicación 1.
- 15 8. Enzima que tiene actividad de dehidrogenasa con un secuencia de polipéptido
- (i) SEQ ID NO: 1 o
- (ii) en la que hasta el 25% de los residuos de aminoácidos se ha modificado frente a SEQ ID NO:1 mediante delección, inserción, sustitución o una combinación de éstas y que posee aún al menos el 50% de la actividad enzimática de SEQ ID NO:1.
- 20 9. Uso de una enzima que tiene actividad dehidrogenasa con una secuencia de polipéptido
- (i) SEQ ID NO: 1 o
- (ii) en la que hasta el 25% de los residuos de aminoácidos se ha modificado frente a la SEQ ID NO:1 mediante delección, inserción, sustitución o una combinación de las mismas y que posee aún al menos el 50% de la actividad enzimática de SEQ ID NO:1 para la preparación de compuestos de las fórmulas I o III.
- 25 10. Uso según la reivindicación 9 en un método para la preparación de duloxetina.