(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 560 459**

(51) Int. Cl.:

C12N 9/02 (2006.01)
C12N 15/53 (2006.01)
C12P 7/00 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **27.08.2009** **E 09810573 (7)**

(97) Fecha y número de publicación de la concesión europea: **28.10.2015** **EP 2329013**

(54) Título: **Polipéptidos cetorreductasa para la producción de una 3-aril-3-hidroxipropanamina a partir de una 3-aril-3-cetopropanamina**

(30) Prioridad:

27.08.2008 US 92321 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
19.02.2016

(73) Titular/es:

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

(72) Inventor/es:

SAVILE, CHRISTOPHER;
GRUBER, JOHN M.;
MUNDORFF, EMILY;
HUISMAN, GJALT W. y
COLLIER, STEVEN J.

(74) Agente/Representante:

IZQUIERDO BLANCO, María Alicia

ES 2 560 459 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).

Polipéptidos cetorreductasa para la producción de una 3-aril-3-hidroxipropanamina a partir de una 3-aril-3-cetopropanamina

Descripción

CAMPO TÉCNICO

La presente descripción se refiere a polipéptidos cetorreductasa manipulados y a usos de los mismos para preparar 3-aril-3-hidroxipropanaminas a partir de las correspondientes 3-aril-3-cetopropanaminas.

ANTECEDENTES

Las enzimas que pertenecen a la clase cetorreductasa (KRED) o carbonil reductasa (EC 1.1.1.184) son útiles para la síntesis de alcoholes ópticamente activos a partir de los correspondientes sustratos cetonas pro-estereoisómeras y mediante reducción estereoespecífica de los correspondientes sustratos aldehído y cetona racémicos. Las KRED convierten por lo general sustratos cetona y aldehído en el correspondiente producto alcohol, pero también pueden catalizar la reacción inversa, la oxidación (deshidrogenación) de un sustrato alcohol al correspondiente producto cetona/aldehído. La reducción de cetonas y aldehídos, y la oxidación de alcoholes por enzimas tales como KRED requieren un cofactor, más comúnmente nicotinamida adenina dinucleótido en su forma reducida (NADH) o nicotinamida adenina dinucleótido fosfato en su forma reducida (NADPH), y nicotinamida adenina dinucleótido (NAD) o nicotinamida adenina dinucleótido fosfato (NADP) para la reacción de oxidación. NADH y NADPH hacen de donantes de electrones, mientras que NAD y NADP hacen de aceptores de electrones. Se observa con frecuencia que las cetorreductasas y las alcoholdeshidrogenasas aceptan tanto el cofactor fosforilado como el no fosforilado (en su estado oxidado y reducido).

Las enzimas KRED pueden encontrarse en gran diversidad de bacterias y levaduras (para revisiones: Kraus y Waldman, Enzyme catalysis in organic synthesis, vols. 1 y 2. VCH Weinheim, 1995; Faber, K, Biotransformations in organic chemistry, 4ª ed. Springer, Berlin Heidelberg Nueva York 2000; Hummel y Kula Eur. J. Biochem. 1989 184:1-13). Se han descrito varias secuencias enzimáticas y genes KRED, por ejemplo, *Candida magnoliae* (nº de registro del Genbank JC7338; GI:11360538), *Candida parapsilosis* (nº de registro del Genbank BAA24528.1; GI:2815409), *Sporobolomyces salmonicolor* (nº de registro del Genbank AF160799; GI:6539734).

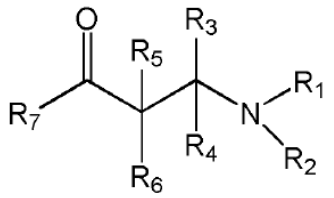
Con el fin de eludir muchos procedimientos de síntesis química para la producción de compuestos clave, se están empleando cada vez más las cetorreductasas para la conversión enzimática de diferentes sustratos ceto en productos quirales alcohol. Estas aplicaciones pueden emplear células enteras que expresan la cetorreductasa para reducciones de cetona biocatalíticas, o enzimas purificadas en aquellos casos en los que la presencia de múltiples cetorreductasas en células enteras influya en la pureza estereomérica y el rendimiento del producto deseado. Para las aplicaciones *in vitro*, puede utilizarse una enzima de regeneración de cofactor (NADH o NADPH) tal como la glucosa deshidrogenasa (GDH), la formiato deshidrogenasa, etc. junto con la cetorreductasa. Los ejemplos que utilizan cetorreductasas para generar compuestos químicos útiles incluyen la reducción asimétrica de ésteres de 4-cloroacetoacetato (Zhou, J. Am. Chem. Soc. 1983 105:5925-5926; Santaniello, J. Chem. Res. (S) 1984:132-133; la patente de EE.UU. Nº 5.559.030; la patente de EE.UU. Nº 5.700.670 y la patente de EE.UU. Nº 5.891.685), la reducción de ácidos dioxocarboxílicos (por ejemplo, la patente de EE.UU. Nº 6.399.339), la reducción de (S)-cloro-5-hidroxi-3-oxohexanoato de terc-butilo (por ejemplo, la patente de EE.UU. Nº 6.645.746 y el documento WO 01/40450), la reducción de compuestos a base de pirrolotriazina (por ejemplo, la solicitud de EE.UU. Nº 2006/0286646); la reducción de acetofenonas sustituidas (por ejemplo, la patente de EE.UU. Nº 6.800.477); y la reducción de cetotiolanos (WO 2005/054491).
Soni *et al*. 2007, Journal of industrial microbiology and biotechnology, 35:167-173 describen la producción de la carbonil reductasa de *Candida viswanathii* cepa MTCC 5158. Soni *et al*. 2007, Process biochemistry, 42:1632-1640 describen el aislamiento y la purificación de una carbonil reductasa capaz de llevar a cabo la reducción estereoselectiva de DKTP (N,N-dimetil-3-ceto-(2- tienil)-1-propanamina) para dar (S)-DHTP (N,N-dimetil-3-hidroxi-(2-tienil)-1-propanamina) con buen rendimiento con enantioselectividad casi absoluta, de *Candida viswanathii* MTCC 5158. En los documentos WO 2009/042984 y WO 2009/046153 se describen variantes de la cetorreductasa de *Lactobacillus*. En el documento US2006/195947 se describen cetorreductasas y variantes de *Candida magnolia*. El documento US6.531.308 se refiere a cetorreductasas de *Z. rouxii* y *S. cerevisiae*. En el documento US2004/023250 se describen 2-cetorreductasas de *Gluconobacter oxydans*. En el documento US7.301031 se describe la producción de ésteres de hidroxiaminoácidos a partir de beta-cetodiésteres sustituidos utilizando la batería de cribado de cetorreductasas KRE0-8000 disponible en el mercado. Kaluzna *et al*. 2005, Tetrahedron, 16:3682-3689 y Goldberg *et al*. 2007, Applied Microbiology and Biotechnology, 76:249-255 analizan las actividades enzimáticas de las cetorreductasas para la síntesis de alcoholes quirales. En el Genbank AA94029 se describe una alcohol deshidrogenasa R-específica dependiente de NADPH de *Lactobacillus kefir*. En el Genbank CAD66648 se describe una alcohol deshidrogenasa R-específica de *Lactobacillus brevis* DSM 20054. En el documento WO2005/033094 se describen métodos basados en cetorreductasa para reducir la 3-cloro-(2-tienil)-propan-1-ona.

Es deseable identificar otras enzimas cetorreductasa que puedan utilizarse para llevar a cabo la conversión de diversos sustratos ceto en sus correspondientes productos alcohol quiral.

RESUMEN DE LA INVENCIÓN

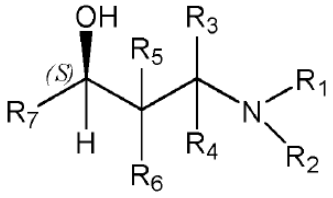
La invención proporciona un polipéptido cetorreductasa manipulado capaz de convertir el sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina en el producto (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina a una velocidad que está mejorada en comparación con la de un polipéptido de referencia que tiene la secuencia de aminoácidos de la SEQ ID NO: 2, en el que el polipéptido comprende una secuencia de aminoácidos que es idéntica en al menos un 85% a una secuencia de referencia basada en la SEQ ID NO: 2, en el que el polipéptido tiene las siguientes características: (a) el resto correspondiente al resto X94 es una glicina, (b) el resto correspondiente al resto X145 es un aminoácido aromático o leucina; y (c) el resto correspondiente al resto X190 es prolina; y (d) el resto en la posición X153 es treonina o valina, el resto en la posición X195 es metionina, el resto en la posición X206 es fenilalanina, triptófano o tirosina, y/o el resto en la posición X233 es glicina. La invención también proporciona un polinucleótido que codifica un polipéptido cetorreductasa manipulado según la invención. La invención también proporciona un vector de expresión que comprende el polinucleótido de la invención. La invención también proporciona una célula hospedadora que comprende el vector de expresión de la invención. La invención también proporciona un método para producir una *(S)*-3-aril-3-hidroxipropanamina, comprendiendo dicho método:

(a) proporcionar un sustrato 3-aril-3-cetopropanamina con la estructura de fórmula (I):



(I);

(b) poner en contacto el sustrato 3-aril-3-cetopropanamina con el polipéptido cetorreductasa manipulado de la invención en una mezcla de reacción en condiciones adecuadas para la reducción o conversión del sustrato en un producto (*S*)-3-aril-3-hidroxipropanamina que tiene la estructura de fórmula (II):



(II)

en el que para (I) y (II), cada uno de R_1 y R_2 está seleccionado independientemente del grupo que consiste en hidrógeno, un alquilo inferior opcionalmente sustituido, un cicloalquilo opcionalmente sustituido, un arilo opcionalmente sustituido, o como alternativa, en las que R_1 y R_2 juntos forman un cicloalquilo opcionalmente sustituido o un cicloarilo opcionalmente sustituido que tiene 3-7 átomos de carbono; cada uno de R_3, R_4, R_5 y R_6 está seleccionado independientemente del grupo que consiste en hidrógeno y un alquilo inferior opcionalmente sustituido; y R_7 es un arilo opcionalmente sustituido.

La invención también proporciona un método para la síntesis de *(3S)-N*-metil-3-naftalen-1-iloxi-3-tiofen-2-il-propan-1-amina que incluye como etapa la reducción estereoselectiva del sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina al producto (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina, en el que la etapa comprende poner en contacto el sustrato con el polipéptido cetorreductasa manipulado de la invención en condiciones de reacción adecuadas para la reducción o conversión del sustrato en el producto.
La invención también proporciona una composición que comprende el compuesto de fórmula estructural (I) y/o de fórmula estructural (II) y cualquiera de los polipéptidos cetorreductasa manipulados de la invención, comprendiendo adicionalmente de manera opcional un sistema de regeneración de cofactor seleccionado del grupo que consiste en glucosa deshidrogenasa y glucosa, formiato deshidrogenasa y formiato, isopropanol y una alcohol secundario deshidrogenasa, y fosfato y fosfato deshidrogenasa.

RESUMEN DE LA DESCRIPCIÓN

La presente descripción proporciona polipéptidos cetorreductasa que tienen la capacidad de reducir una 3-aril-3-cetopropanamina a una 3-aril-3-hidroxipropanamina, a polinucleótidos que codifican tales polipéptidos, y a métodos para preparar y utilizar los polipéptidos. Los polipéptidos cetorreductasa de la presente invención son particularmente útiles para preparar compuestos intermedios en la síntesis del fármaco duloxetina (es decir, *(3S)-N*-metil-3-naftalen-1-iloxi-3-tiofen-2-il-propan-1-amina). Los sustratos ejemplares incluyen, por ejemplo, *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina ("DMAK", el "sustrato dimetil amino cetona" o el "sustrato dimetilo") y *N*-metil-3-ceto-3-(2-tienil)-1-propanamina ("MMAK", el "sustrato monometil amino cetona" o el "sustrato monometilo"), que se reduce a (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina ("(S)-DMAA", el "producto dimetil amino alcohol" o el "producto dimetilo") y *(S)-N*-metil-3-hidroxi-3-(2-tienil)-1-propanamina ("(S)-MMAA"; el "producto monometil amino alcohol" o el "producto monometilo"), respectivamente. Los polipéptidos cetorreductasa de la presente invención presentan la capacidad de reducir el sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina al producto (*S*)-*N,N*-dimetil-3-hidroxi-3-(-2-tienil)-1-propanamina con una estereoselectividad especialmente alta.

En un aspecto, los polipéptidos cetorreductasa descritos en el presente documento tienen una secuencia de aminoácidos que tiene una o más diferencias de aminoácidos en comparación con una secuencia de aminoácidos de referencia de una cetorreductasa de tipo silvestre o una cetorreductasa manipulada que dan como resultado una propiedad mejorada de la enzima para el sustrato ceto definido. Generalmente, los polipéptidos cetorreductasa manipulados tienen una propiedad mejorada en comparación con las enzimas cetorreductasa de tipo silvestre naturales obtenidas a partir de *Lactobacillus kefir* ("*L. kefir*"; la SEQ ID NO: 2), *Lactobacillus brevis* ("*L. brevis*"; la SEQ ID NO: 4) o *Lactobacillus minor* ("*L. minor*", SEQ ID NO: 106). En algunas formas de realización, los polipéptidos de la descripción tienen una propiedad mejorada en comparación con otro polipéptido manipulado, tal como la SEQ ID NO: 6. Las mejoras en la propiedad de la enzima pueden incluir, entre otras, el aumento de la actividad enzimática, la estereoselectividad, la estereoespecificidad, la termoestabilidad, la estabilidad frente a disolventes o la inhibición por el producto reducido. En la presente descripción, los polipéptidos cetorreductasa tienen, en comparación con la secuencia de aminoácidos de las SEQ ID NO: 2, 4 ó 106, al menos las siguientes características: (1) el resto de aminoácido correspondiente a la posición de resto 94 es un resto no polar, (2) el resto de aminoácido correspondiente a la posición de resto 145 es un resto aromático o leucina, y (3) el resto de aminoácido correspondiente a la posición de resto 190 es una cisteína o un resto restringido. En algunas formas de realización, los polipéptidos cetorreductasa tienen, en comparación con las secuencias de las SEQ ID NO: 2, 4 ó 106, al menos las siguientes características: (1) el resto de aminoácido correspondiente a la posición de resto 94 es leucina, isoleucina, alanina, valina, metionina o glicina, (2) el resto de aminoácido correspondiente a la posición de resto 145 es tirosina, triptófano, fenilalanina o leucina, y (3) el resto de aminoácido correspondiente a la posición de resto 190 es cisteína o prolina. En algunas formas de realización, los polipéptidos de la descripción tienen, en comparación con las secuencias de las SEQ ID NO: 2, 4 ó 106, al menos las siguientes características: (1) el resto correspondiente a la posición de resto 94 es glicina, (2) el resto correspondiente a la posición de resto 145 es fenilalanina o leucina, y (3) el resto correspondiente a la posición de resto 190 es cisteína o prolina. En algunas formas de realización, los polipéptidos de la descripción tienen, en comparación con las secuencias de las SEQ ID NO: 2, 4 ó 106, al menos las siguientes características: (1) el resto correspondiente a la posición 94 es glicina, (2) el resto correspondiente a la posición 145 es fenilalanina, y (3) el resto correspondiente a la posición 190 es prolina.

En algunas formas de realización, las cetorreductasas manipuladas pueden tener una sola propiedad mejorada, o pueden tener dos o más propiedades mejoradas, en cualquier combinación o combinaciones. Por ejemplo, el polipéptido cetorreductasa manipulado puede tener una actividad enzimática aumentada en comparación con la enzima cetorreductasa de tipo silvestre para reducir el sustrato dimetilo al producto correspondiente. Las mejoras en la actividad enzimática pueden medirse comparando la actividad específica del polipéptido cetorreductasa con la de la enzima cetorreductasa de tipo silvestre utilizando ensayos enzimáticos convencionales. El grado de la mejora puede variar desde 1,5 veces la actividad enzimática de la correspondiente enzima cetorreductasa de tipo silvestre o la de una de referencia, hasta tanto como 2, 5, 10, 15 veces, o más, la mejora de la actividad enzimática. En formas de realización específicas, la enzima cetorreductasa manipulada presenta una actividad enzimática mejorada que es al menos 1,5 veces, 2 veces, 3 veces, 4 veces, 5 veces, 10 veces, 15 veces, 20 veces, 25 veces, 30 veces, 50 veces, o más, la de la enzima cetorreductasa de tipo silvestre o de la de referencia. Las mejoras en la actividad enzimática también incluyen aumentos de la estereoselectividad, la estereoespecificidad, la termoestabilidad, la estabilidad frente a disolventes o la inhibición por el producto reducido.

En algunas formas de realización, los polipéptidos cetorreductasa están mejorados en comparación con la SEQ ID NO: 2 y la SEQ ID NO: 6 con respecto a su tasa de actividad enzimática, es decir, su velocidad de conversión del sustrato dimetilo en el producto correspondiente. En algunas formas de realización, los polipéptidos cetorreductasa son capaces de convertir el sustrato dimetilo en el producto a una velocidad que es al menos 1,5 veces, 2 veces, 5 veces ó 10 veces la velocidad que muestran las cetorreductasas de la SEQ ID NO: 2 o la SEQ ID NO: 6.

En algunas formas de realización, el polipéptido cetorreductasa es capaz de convertir el sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina en (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina, a una velocidad que está mejorada en comparación con la de un polipéptido de referencia que tiene la secuencia de aminoácidos de la SEQ ID NO: 6. En otras formas de realización, tales polipéptidos cetorreductasa también son capaces de convertir el

sustrato dimetilo en el producto con un porcentaje de exceso estereomérico de al menos aproximadamente un 95%. En algunas formas de realización, tales polipéptidos cetorreductasa también son capaces de convertir el sustrato, *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina en el producto (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina con un porcentaje de exceso estereomérico de al menos aproximadamente un 99%. Los polipéptidos ejemplares con tales propiedades incluyen, pero no se limitan a, polipéptidos que comprenden las secuencias de aminoácidos correspondientes a las SEQ ID NO: 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76 y 78. Debido a que el polipéptido de referencia que tiene la secuencia de aminoácidos de la SEQ ID NO: 6 es capaz de convertir el sustrato dimetilo en el producto a una velocidad (por ejemplo, un 50% de 50 g/l de sustrato convertido en producto en 24 horas con aproximadamente 10 g/l de la variante de la enzima que tiene la secuencia de aminoácidos de la SEQ ID NO: 6) y con una estereoselectividad (99% de exceso estereomérico) que está mejorada en comparación con la de *Lactobacillus kefir* de tipo silvestre (SEQ ID NO: 2), los polipéptidos del presente documento que están mejorados en comparación con la SEQ ID NO: 6 también están mejorados en comparación con los de tipo silvestre.

En algunas formas de realización, el polipéptido cetorreductasa es capaz de convertir el sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina en el producto (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina, con un porcentaje de exceso estereomérico de al menos aproximadamente un 99% y a una velocidad que es al menos 1,5-2 veces la actividad del polipéptido de referencia con la secuencia de aminoácidos de la SEQ ID NO: 6. Los polipéptidos ejemplares con tales propiedades incluyen, pero no se limitan a, polipéptidos que comprenden las secuencias de aminoácidos correspondientes a las SEQ ID NO: 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76 y 78.

En algunas formas de realización, el polipéptido cetorreductasa es capaz de convertir el sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina en el producto (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina, con un porcentaje de exceso estereomérico de al menos aproximadamente un 99% y a una velocidad que está mejorada en al menos un 200%-1.000% en comparación con la de un polipéptido de referencia con la secuencia de aminoácidos de la SEQ ID NO: 6. Los polipéptidos ejemplares con tales propiedades incluyen, pero no se limitan a, polipéptidos que comprenden las secuencias de aminoácidos correspondientes a las SEQ ID NO: 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76 y 78.

En algunas formas de realización, el polipéptido cetorreductasa es capaz de convertir el sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina en el producto (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina, con un porcentaje de exceso estereomérico de al menos aproximadamente un 99% y a una velocidad que está mejorada en al menos 10 veces en comparación con la de un polipéptido de referencia con la secuencia de aminoácidos de la SEQ ID NO: 6. Los polipéptidos ejemplares con tales propiedades incluyen, pero no se limitan a, polipéptidos que comprenden las secuencias de aminoácidos correspondientes a las SEQ ID NO: 58, 60, 62, 64, 66, 68, 70, 72, 74, 76 y 78.

En algunas formas de realización, el polipéptido cetorreductasa mejorado es capaz de convertir al menos aproximadamente un 95% del sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina en el producto (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina en menos de aproximadamente 24 horas cuando se lleva a cabo con más de aproximadamente 100 g/l de sustrato y menos de aproximadamente 5 g/l del polipéptido. Los polipéptidos ejemplares que tienen esta capacidad incluyen, pero no se limitan a, polipéptidos que comprenden las secuencias de aminoácidos correspondientes a las SEQ ID NO: 60, 62, 64, 66, 68, 70, 72, 74, 76 y 78
.

En algunas formas de realización, los polipéptidos cetorreductasa son altamente estereoselectivos y pueden reducir el sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina al producto (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina en un exceso estereomérico superior a aproximadamente un 99%, 99,1%, 99,2%, 99,3%, 99,4%, 99,5%, 99,6%, 99,7%, 99,8% ó 99,9%. Los polipéptidos cetorreductasa ejemplares con una estereoselectividad tan alta incluyen, pero no se limitan a, los polipéptidos que comprenden las secuencias de aminoácidos correspondientes a las SEQ ID NO: 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76 y 78.

En algunas formas de realización, un polipéptido cetorreductasa mejorado comprende una secuencia de aminoácidos que corresponde a las fórmulas de secuencia proporcionadas en la SEQ ID NO: 107 o la SEQ ID NO: 108 o la SEQ ID NO: 109 (o una región de las mismas, tal como los restos 90-211). La SEQ ID NO: 107 se basa en la secuencia de aminoácidos de tipo silvestre de la cetorreductasa de *Lactobacillus kefir* (SEQ ID NO: 2); la SEQ ID NO: 108 se basa en la secuencia de aminoácidos de tipo silvestre de la cetorreductasa de *Lactobacillus brevis* (SEQ ID NO: 4); y la SEQ ID NO: 109 se basa en la secuencia de aminoácidos de tipo silvestre de la cetorreductasa de *Lactobacillus minor* (SEQ ID NO: 106). Además, las SEQ ID NO: 107, 108 y 109 especifican que el resto de aminoácido correspondiente a la posición 94 es un resto no polar, el resto de aminoácido correspondiente a la posición 145 es un resto aromático o leucina, y el resto de aminoácido correspondiente a la posición 190 es cisteína o un resto restringido.

En algunas formas de realización, un polipéptido cetorreductasa mejorado de la descripción se basa en las fórmulas de secuencia de las SEQ ID NO: 107, 108 ó 109 y puede comprender una secuencia de aminoácidos que

es idéntica en al menos aproximadamente un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ó 99% a una secuencia de referencia basada en las SEQ ID NO: 2, 4 ó 106, a condición de que (1) el resto de aminoácido correspondiente a la posición 94 sea un resto no polar, (2) el resto de aminoácido correspondiente a la posición 145 sea un resto aromático o leucina, y (3) el resto de aminoácido correspondiente a la posición 190 sea cisteína o un resto restringido. En algunas formas de realización, los polipéptidos cetorreductasa tienen la identidad de secuencia especificada con una secuencia de referencia basada en las SEQ ID NO: 2, 4 ó 106, a condición de que (1) el resto de aminoácido correspondiente a la posición 94 sea glicina, (2) el resto de aminoácido correspondiente a la posición 145 sea fenilalanina, tirosina o leucina, particularmente fenilalanina o leucina, y (3) el resto de aminoácido correspondiente a la posición 190 sea cisteína o prolina, en las que los polipéptidos cetorreductasa tienen al menos las condiciones especificadas. En algunas formas de realización, los polipéptidos cetorreductasa pueden tener adicionalmente una o más diferencias de resto de aminoácido en comparación con las secuencias de referencia anteriormente indicadas. Estas diferencias pueden ser inserciones, deleciones, sustituciones de aminoácidos, o cualquier combinación de tales cambios. En algunas formas de realización, las diferencias de secuencia de aminoácidos pueden comprender sustituciones de aminoácidos no conservadoras, conservadoras, así como una combinación de sustituciones de aminoácidos no conservadoras y conservadoras. En el presente documento se describen diversas posiciones de resto de aminoácido en las que pueden realizarse tales cambios.

En algunas formas de realización, un polipéptido cetorreductasa mejorado se basa en las fórmulas de secuencia de las SEQ ID NO: 107, 108 ó 109 y puede comprender una región con una secuencia de aminoácidos que es idéntica en al menos aproximadamente un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ó 99% a una región o dominio de las mismas, tal como los restos 90-211 de la SEQ ID NO: 2, 4 ó 106, a condición de que (1) el resto de aminoácido correspondiente a la posición 94 sea un resto no polar, (2) el resto de aminoácido correspondiente a la posición 145 sea un resto aromático o leucina, y (3) el resto de aminoácido correspondiente a la posición 190 sea cisteína o un resto restringido. En algunas formas de realización, los polipéptidos cetorreductasa tienen la identidad de secuencia especificada con una secuencia de referencia basada en las SEQ ID NO: 2, 4 ó 106, a condición de que (1) el resto de aminoácido correspondiente a la posición 94 sea glicina, (2) el resto de aminoácido correspondiente a la posición 145 sea fenilalanina, tirosina o leucina, y (3) el resto de aminoácido correspondiente a la posición 190 sea cisteína o prolina, y en los que los polipéptidos cetorreductasa tienen al menos las condiciones especificadas. En algunas formas de realización, los polipéptidos cetorreductasa pueden tener adicionalmente una o más diferencias de resto de aminoácido en la región o dominio definido. En algunas formas de realización, las diferencias de secuencia de aminoácido en el dominio pueden comprender sustituciones de aminoácidos no conservadoras, conservadoras, así como una combinación de sustituciones de aminoácidos no conservadoras y conservadoras. En el presente documento se describen diversas posiciones de resto de aminoácido en las que tales cambios pueden realizarse en la región definida.

En algunas formas de realización, una cetorreductasa mejorada tiene una secuencia de aminoácidos que es idéntica en al menos aproximadamente un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ó 99% a la secuencia de aminoácidos correspondiente a la SEQ ID NO: 6, 8, 10, 12, 14, 16, 18, 29, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76 ó 78, en la que la secuencia de aminoácidos del polipéptido cetorreductasa mejorado incluye cualquier conjunto de las combinaciones de sustitución de aminoácidos especificadas que se presentan en la Tabla 2. En algunas formas de realización, estos polipéptidos cetorreductasa pueden tener aproximadamente 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-25, 1-30, 1-35 ó 1-40 mutaciones en otros restos de aminoácidos. En algunas formas de realización, el número de mutaciones puede ser 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35 o aproximadamente 40 restos de aminoácidos. Las mutaciones pueden comprender inserciones, deleciones o sustituciones, o combinaciones de las mismas. En algunas formas de realización, las mutaciones son mutaciones conservadoras.

En algunas formas de realización, una cetorreductasa mejorada comprende una secuencia de aminoácidos correspondiente a la SEQ ID NO: 6, 8, 10, 12, 14, 16, 18, 29, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76 ó 78.

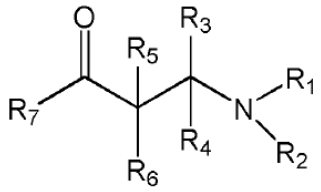
En otro aspecto, la presente descripción proporciona polinucleótidos que codifican las cetorreductasas manipuladas descritas en el presente documento o polinucleótidos que hibridan con tales polinucleótidos en condiciones muy rigurosas. El polinucleótido puede incluir promotores y otros elementos reguladores útiles para la expresión de la cetorreductasa manipulada codificada, y puede utilizar codones optimizados para los sistemas de expresión deseados específicos. Los polinucleótidos ejemplares incluyen, pero no se limitan a, la SEQ ID NO: 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75 y 77. Los polinucleótidos ejemplares también incluyen polinucleótidos que codifican polipéptidos que corresponden a las fórmulas de secuencia de las SEQ ID NO: 107, 108 ó 109.

En otro aspecto, la presente descripción proporciona células hospedadoras que comprenden los polinucleótidos y/o los vectores de expresión descritos en el presente documento. Las células hospedadoras pueden ser *L. kefir* o *L. brevis* o *L. minor*, o pueden ser un organismo diferente, tal como *E. coli*. Las células hospedadoras pueden utilizarse para la expresión y el aislamiento de las enzimas cetorreductasa manipuladas descritas en el

presente documento, o, como alternativa, pueden utilizarse directamente para la conversión del sustrato en el producto estereoisómero.

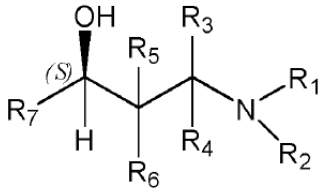
Ya se lleve a cabo el método con células enteras, extractos celulares o enzimas cetorreductasa purificadas, puede utilizarse una sola enzima cetorreductasa o, como alternativa, pueden utilizarse mezclas de dos o más enzimas cetorreductasa.

Las enzimas cetorreductasa descritas en el presente documento son útiles para catalizar la reacción de reducción del grupo ceto en la 3-aril-3-cetopropanamina que tiene la fórmula estructural (I):



(I)

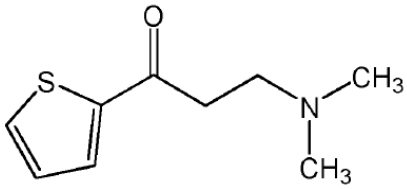
para dar la correspondiente *(S)*-3-aril-3-hidroxipropanamina con la fórmula estructural (II):



(II)

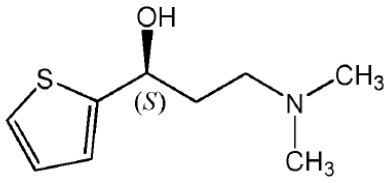
en las que para (I) y (II), cada uno de R_1 y R_2 está seleccionado independientemente del grupo que consiste en hidrógeno, un alquilo inferior opcionalmente sustituido, un cicloalquilo opcionalmente sustituido, un arilo opcionalmente sustituido, o como alternativa, en las que R_1 y R_2 juntos forman un cicloalquilo opcionalmente sustituido o un cicloarilo opcionalmente sustituido que tiene 3-7 átomos de carbono; cada uno de R_3, R_4, R_5 y R_6 está seleccionado independientemente del grupo que consiste en hidrógeno y un alquilo inferior opcionalmente sustituido; y R_7 es un arilo opcionalmente sustituido.

Los polipéptidos cetorreductasa descritos en el presente documento son particularmente capaces de reducir el sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina (III):



(III)

al producto (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina (IV):



(IV)

con alta estereoselectividad.

En algunas formas de realización, la presente descripción proporciona un método para reducir el sustrato con la fórmula estructural (I) al producto correspondiente de fórmula estructural (II), que comprende poner en contacto o incubar el sustrato con un polipéptido cetorreductasa descrito en el presente documento en condiciones de reacción adecuadas para la reducción o conversión del sustrato en el compuesto producto. Por ejemplo, en algunas formas de realización del presente método, el polipéptido cetorreductasa tiene, en comparación con las secuencias de KRED de *L. kefir* o *L. brevis* o *L. minor* de tipo silvestre de la SEQ ID NO: 2, 4 ó 106, al menos la siguiente características: (1) el resto correspondiente a X94 es un aminoácido no polar, (2) el resto correspondiente a X145 es un aminoácido aromático o leucina, y (3) el resto correspondiente a X190 es una cisteína o un aminoácido restringido. En algunas formas de realización del presente método, el polipéptido cetorreductasa utilizado en el método puede tener, en comparación con las secuencias de KRED de *L. kefir* o *L. brevis* o *L. minor* de tipo silvestre de la SEQ ID NO: 2, 4 y 106, respectivamente, al menos las siguientes características: (1) el resto correspondiente al resto X94 es glicina, (2) el resto correspondiente a la posición 145 es tirosina, fenilalanina o leucina, particularmente fenilalanina o leucina, y (3) el resto correspondiente a X190 es cisteína o prolina. En algunas formas de realización del presente método, los polipéptidos cetorreductasa pueden tener, en comparación con las secuencias de KRED de *L. kefir* o *L. brevis* o *L. minor* de la SEQ ID NO: 2, 4 ó 106, las siguientes características: (1) el resto correspondiente a X94 es glicina, (2) el resto correspondiente a X145 es fenilalanina y (3) el resto correspondiente a X190 es prolina.

En algunas formas de realización del método para reducir el sustrato al producto, el sustrato dimetilo (III) se reduce al producto dimetilo (IV) en un exceso estereomérico superior a aproximadamente un 99%, en las que el polipéptido cetorreductasa comprende una secuencia que corresponde a la SEQ ID NO: 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76 ó 78.

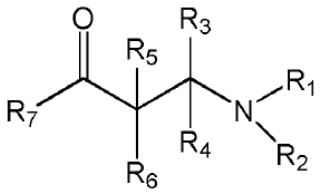
En algunas formas de realización del método para reducir el sustrato dimetilo (III) al producto dimetilo (IV), al menos aproximadamente un 95% del sustrato se convierte en el producto en menos de aproximadamente 24 horas cuando se lleva a cabo con más de aproximadamente 100 g/l de sustrato y menos de aproximadamente 5 g/l del polipéptido, en el que el polipéptido comprende una secuencia de aminoácidos correspondiente a la SEQ ID NO: 60, 62, 64, 66, 68, 70, 72, 74, 76 ó 78.

En algunas formas de realización del método para reducir el sustrato dimetilo (III) al producto dimetilo (IV), al menos aproximadamente un 10%-20% de 1 g/l de sustrato se convierte en el producto en menos de aproximadamente 24 horas con aproximadamente 10 g/l del polipéptido, en el que el polipéptido comprende una secuencia de aminoácidos correspondiente a la SEQ ID NO: 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76 ó 78.

En algunas formas de realización, el método se lleva a cabo en una condición de reacción en la que el pH es de aproximadamente 8 a 11. En algunas formas de realización, la condición de reacción tiene un pH de aproximadamente 9 a 10. En algunas formas de realización, la condición de reacción tiene un pH de aproximadamente 9,0. En algunas formas de realización, las condiciones de reacción comprenden ≥ 50% v/v de isopropanol. En algunas formas de realización, las condiciones de reacción comprenden ≥ 75% de isopropanol.

En algunas formas de realización, los métodos se refieren al uso de los polipéptidos cetorreductasa de la presente invención en la síntesis de una *(S)-N*-metil-3-hidroxi-3-(aril)-propanamina, comprendiendo el método:

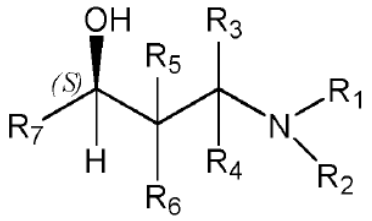
(a) proporcionar un sustrato 3-aril-3-cetopropanamina con la estructura de fórmula (I):



(I)

en el que cada uno de R_1 y R_2 es metilo, cada uno de R_3, R_4, R_5 y R_6 está seleccionado independientemente del grupo que consiste en hidrógeno y un alquilo inferior opcionalmente sustituido, y R_7 es un arilo opcionalmente sustituido;

(b) poner en contacto el sustrato 3-aril-3-cetopropanamina con uno o más polipéptidos cetorreductasa de cualquiera de los polipéptidos cetorreductasa descritos en el presente documento en una mezcla de reacción en condiciones adecuadas para la reducción o conversión del sustrato en un producto *(S)*-3-aril-3-hidroxipropanamina con la fórmula estructural (II):

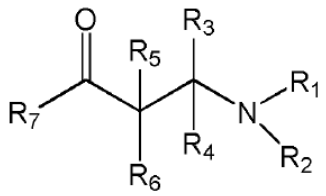


(II)

en el que cada uno de R_1 y R_2 es metilo, cada uno R_3, R_4, R_5 y R_6 está seleccionado independientemente del grupo que consiste en hidrógeno y un alquilo inferior opcionalmente sustituido, y R_7 es un arilo opcionalmente sustituido; (c) desmetilar el producto *(S)*-3-aril-3-hidroxipropanamina (es decir, *N,N*-dimetil-3-hidroxi-3-(aril)-propanamina) de la etapa (b) en una mezcla de reacción en condiciones adecuadas para producir una *(S)-N*-metil-3-hidroxi-3-(aril)-propanamina que tiene la fórmula estructural (II), en el que uno de R_1 y R_2 es metilo y el otro es hidrógeno, cada uno de R_3, R_4, R_5 y R_6 está seleccionado independientemente del grupo que consiste en hidrógeno y un alquilo inferior opcionalmente sustituido, y R_7 es un arilo opcionalmente sustituido.

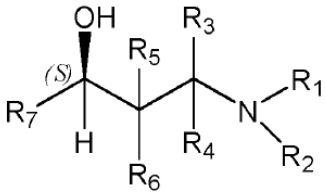
En algunas formas de realización, los métodos se refieren al uso de los polipéptidos cetorreductasa de la presente invención en la síntesis de una 3-ariloxi-3-arilpropanamina, comprendiendo dicho método:

(a) proporcionar una 3-aril-3-cetopropanamina con la estructura de fórmula (I):



(I)

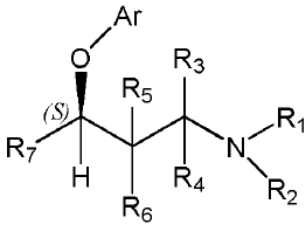
(b) poner en contacto la 3-aril-3-cetopropanamina con un polipéptido cetorreductasa de la presente invención en una mezcla de reacción en condiciones suficientes para producir una *(S)*-3-aril-3-hidroxipropanamina con la estructura de fórmula (II):



(II);

y

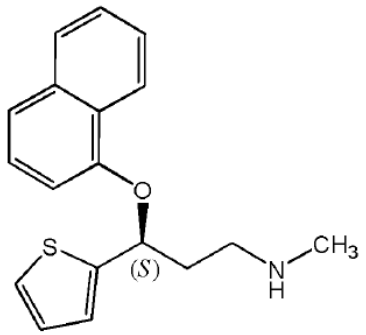
(c) poner en contacto la *(S)*-3-aril-3-hidroxipropanamina con un compuesto de arilo activado en una mezcla de reacción en condiciones suficientes para producir la *(S)*-3-ariloxi-3-(aril)-propanamina con la estructura de fórmula (VII):



(VII)

en las que para (I), (II) y (VII), cada uno de R_1 y R_2 está seleccionado independientemente del grupo que consiste en hidrógeno, un alquilo inferior opcionalmente sustituido, un cicloalquilo opcionalmente sustituido, un arilo opcionalmente sustituido, o como alternativa, en las que R_1 y R_2 juntos forman un cicloalquilo opcionalmente sustituido o un cicloarilo opcionalmente sustituido que tiene 3-7 átomos de carbono; cada uno de R_3, R_4, R_5 y R_6 está seleccionado independientemente del grupo que consiste en hidrógeno y un alquilo inferior opcionalmente sustituido; y R_7 es un arilo opcionalmente sustituido y, además, para (VII), Ar es un grupo arilo opcionalmente sustituido.

En algunas formas de realización, los métodos se refieren más concretamente al uso de polipéptidos cetorreductasa de la presente invención en la síntesis de duloxetina, *(*3S)-*N*-metil-3-naftalen-1-iloxi-3-tiofen-2-il-propan-1-amina, que tiene la estructura de fórmula (VIII):



(VIII)

y sales, hidratos y solvatos de la misma. Por lo tanto, en un método para la síntesis del compuesto de fórmula (VIII), una etapa en el método comprende reducir el sustrato de fórmula estructural (III) al producto correspondiente de fórmula (IV) poniendo en contacto o incubando el sustrato con un polipéptido cetorreductasa descrito en el presente documento en condiciones de reacción adecuadas para reducir el sustrato al compuesto producto.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

La FIG. 1 ilustra el papel de la cetorreductasa (KRED) en la conversión del compuesto sustrato de fórmula (III) en el correspondiente producto alcohol de fórmula (IV). La reducción utiliza un cofactor, tal como NADPH, que se puede regenerar mediante el uso de un sistema de regeneración de cofactor. En el sistema de reciclaje ilustrado, el $NADP^+$ es convertido en NADPH por una glucosa deshidrogenasa. La glucosa se convierte en ácido glucónico, que a su vez se convierte en su sal sódica con la adición de hidróxido sódico.

La FIG. 2 representa la ruta de síntesis para preparar duloxetina a partir del compuesto de fórmula estructural (III), *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina. La *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina se reduce a (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina (IV), que posteriormente se somete a naftilación hasta el compuesto intermedio (IX). A continuación, se desmetila el compuesto intermedio (IX) para generar duloxetina (*(3S)-N*-metil-3-naftalen-1-iloxi-3-tiofen-2-il-propan-1-amina).

DESCRIPCIÓN DETALLADA

Definiciones

Tal como se utilizan en el presente documento, los siguientes términos y expresiones pretenden tener los siguientes significados.

"Cetorreductasa" y "KRED" se utilizan indistintamente en el presente documento para referirse a un polipéptido que tiene la capacidad enzimática de reducir un grupo carbonilo a su alcohol correspondiente. Más concretamente, los polipéptidos cetorreductasa de la descripción son capaces de reducir estereoselectivamente el compuesto de fórmula (I), *supra*, al producto correspondiente de fórmula (II), *supra*. El polipéptido utiliza por lo general como agente reductor un cofactor nicotinamida adenina dinucleótido en su forma reducida (NADH) o nicotinamida adenina dinucleótido fosfato en su forma reducida (NADPH). Las cetorreductasas tal como se utilizan en el presente documento incluyen cetorreductasas naturales (de tipo silvestre), así como polipéptidos manipulados no naturales generados mediante manipulación humana.

"Secuencia codificante" se refiere a aquella porción de un ácido nucleico (por ejemplo, un gen) que codifica una secuencia de aminoácidos de una proteína.

"Natural" o "de tipo silvestre" se refiere a la forma que se encuentra en la naturaleza. Por ejemplo, una secuencia polipeptídica o polinucleotídica natural o de tipo silvestre es una secuencia presente en un organismo que puede aislarse a partir de una fuente en la naturaleza y que no ha sido modificada intencionadamente mediante manipulación humana.

"Recombinante" cuando se utiliza con referencia a, por ejemplo, una célula, un ácido nucleico o un polipéptido, se refiere a un material, o un material correspondiente a la forma natural o nativa del material, que ha sido modificado de tal manera que de lo contrario no existiría en la naturaleza, o que es idéntico al mismo pero producido o derivado a partir de materiales sintéticos y/o mediante manipulación utilizando técnicas recombinantes. Los ejemplos no limitativos incluyen, entre otros, células recombinantes que expresan genes que no se encuentran dentro de la forma nativa (no recombinante) de la célula o que expresan genes nativos que se expresan de otro modo a un nivel diferente.

"Porcentaje de identidad de secuencia" y "porcentaje de homología" se utilizan indistintamente en el presente documento para referirse a las comparaciones entre polinucleótidos y polipéptidos, y se determinan comparando dos secuencias óptimamente alineadas sobre una ventana de comparación, en las que la porción de la secuencia polinucleotídica o polipeptídica en la ventana de comparación puede comprender adiciones o deleciones (es decir, huecos) en comparación con la secuencia de referencia (que no comprende adiciones ni deleciones) para un alineamiento óptimo de las dos secuencias. El porcentaje puede calcularse determinando el número de posiciones en las que la base de ácido nucleico o resto de aminoácido idéntico aparece en ambas secuencias para dar el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes por el número total de posiciones en la ventana de comparación y multiplicando el resultado por 100 para dar el porcentaje de identidad de secuencia. Como alternativa, el porcentaje puede calcularse determinando el número de posiciones en las que, bien la base de ácido nucleico o resto de aminoácido idéntico aparece en ambas secuencias, o bien una base de ácido nucleico o resto de aminoácido se alinea con un hueco para dar el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes por el número total de posiciones en la ventana de comparación y multiplicando el resultado por 100 para dar el porcentaje de identidad de secuencia. Los expertos en la materia entenderán que existen muchos algoritmos establecidos disponibles para alinear dos secuencias. El alineamiento óptimo de secuencias para la comparación puede llevarse a cabo, por ejemplo, mediante el algoritmo de homología local de Smith y Waterman, 1981, Adv. Appl. Math. 2:482, mediante el algoritmo de alineamiento de homología de Needleman y Wunsch, 1970, J. Mol. Biol. 48:443, mediante el método de búsqueda de similitudes de Pearson y Lipman, 1988, Pro. Natl. Acad. Sci. EE.UU. 85:2444, mediante aplicaciones informatizadas de estos algoritmos (GAP, BESTFIT, FASTA y TFASTA en el GCG Wisconsin Software Package), o mediante inspección visual (véase en general, Current Protocols in Molecular Biology, Ausubel F.M. *et al.,* eds., Current Protocols, una empresa conjunta entre Greene Publishing Associates, Inc. y John Wiley & Sons, Inc., (suplemento de 1995) (Ausubel)). Los ejemplos de algoritmos que son adecuados para determinar el porcentaje de identidad de secuencia y la similitud de secuencia son los algoritmos BLAST y BLAST 2.0, que se describen en Altschul *et al.,* 1990, J. Mol. Biol. 215:403-410 y Altschul *et al.,* 1977, Nucleic Acids Res. 3389-3402, respectivamente. El software para realizar el análisis BLAST está a disposición del público a través de la página web del National Center for Biotechnology Information. Este algoritmo implica en primer lugar identificar pares de secuencia de alta puntuación (HSP) identificando palabras cortas de longitud W en la secuencia problema, que coinciden o satisfacen cierta puntuación umbral T de valor positivo cuando se alinean con una palabra de la misma longitud en una secuencia de base de datos. T se conoce como umbral de puntuación de la palabra vecina (Altschul *et al., supra*). Estos aciertos de palabra vecina iniciales actúan como semillas para iniciar búsquedas para encontrar HSP más largos que los contengan. A continuación, los aciertos de palabras se extienden en ambas direcciones a lo largo de cada secuencia tanto como pueda aumentarse la puntuación de alineamiento acumulativa. Las puntuaciones acumulativas se calculan utilizando, para secuencias nucleotídicas, los parámetros M (puntuación de recompensa para un par de restos coincidentes; siempre > 0) y N (puntuación de penalización para los restos no coincidentes; siempre < 0). Para las secuencias de aminoácidos, se utiliza una matriz de puntuación para calcular la puntuación acumulativa. La extensión de los aciertos de palabras en cada dirección se detiene cuando: la puntuación de alineamiento acumulativa disminuye en la cantidad X desde su valor máximo alcanzado; la puntuación acumulativa llega a cero o menos, debido a la acumulación de uno o más alineamientos de restos de puntuación negativa; o se alcanza el final de cualquiera de las secuencias. Los parámetros del algoritmo BLAST W, T y X determinan la sensibilidad y velocidad del alineamiento. El programa BLASTN (para las secuencias nucleotídicas) utiliza por defecto una longitud de palabra (W) de 11, una expectativa (E) de 10, M = 5, N = -4, y una comparación de ambas hebras. Para las secuencias de aminoácidos, el programa BLASTP utiliza por defecto una longitud de palabra (W) de 3, una expectativa (E) de 10, y la matriz de puntuación BLOSUM62 (véase Henikoff y Henikoff, 1989, Proc. Natl. Acad. Sci. EE.UU. 89:10915). La determinación ejemplar de alineamiento de secuencia y % de identidad de secuencia puede emplear los programas BESTFIT o GAP en el paquete GCG Wisconsin Software (Accelrys, Madison WI), utilizando los parámetros por defecto proporcionados.

"Secuencia de referencia" se refiere a una secuencia definida utilizada como base para una comparación de secuencias. Una secuencia de referencia puede ser un subconjunto de una secuencia mayor, por ejemplo, un segmento de un gen o secuencia polipeptídica de longitud completa. En general, una secuencia de referencia tiene al menos 20 nucleótidos o restos de aminoácidos de longitud, al menos 25 restos de longitud, al menos 50 restos de longitud, o la longitud completa del ácido nucleico o polipéptido. Dado que dos polinucleótidos o polipéptidos pueden

comprender cada uno (1) una secuencia (es decir, una porción de la secuencia completa) que es similar entre las dos secuencias, y (2) pueden comprender adicionalmente una secuencia que es divergente entre las dos secuencias, las comparaciones de secuencias entre dos (o más) polinucleótidos o polipéptidos se realizan por lo general comparando las secuencias de los dos polinucleótidos sobre una "ventana de comparación" para identificar y comparar regiones locales de similitud de secuencia.

En algunas formas de realización, una "secuencia de referencia" puede basarse en una secuencia primaria de aminoácidos, en la que la secuencia de referencia es una secuencia que puede tener uno o más cambios en la secuencia primaria. Por ejemplo, una secuencia de referencia "basada en la SEQ ID NO: 4 que tiene en el resto correspondiente a X94 una glicina" se refiere a una secuencia de referencia en la que el resto correspondiente a X94 en la SEQ ID NO: 4 se ha cambiado a una glicina.

"Ventana de comparación" se refiere a un segmento conceptual de al menos aproximadamente 20 posiciones contiguas de nucleótidos o restos de aminoácidos en el que una secuencia puede compararse con una secuencia de referencia de al menos 20 nucleótidos o aminoácidos contiguos y en el que la porción de la secuencia en la ventana de comparación puede comprender adiciones o deleciones (es decir, huecos) del 20 por ciento o menos en comparación con la secuencia de referencia (que no comprende adiciones ni deleciones) para un alineamiento óptimo de las dos secuencias. La ventana de comparación puede tener una longitud superior a 20 restos contiguos, e incluye, opcionalmente ventanas de longitud 30, 40, 50, 100, o mayor.

"Identidad sustancial" se refiere a una secuencia polinucleotídica o polipeptídica que tiene al menos un 80 por ciento de identidad de secuencia, al menos un 85 por ciento de identidad y de un 89 a un 95 por ciento de identidad de secuencia, más habitualmente al menos un 99 por ciento de identidad de secuencia comparada con una secuencia de referencia sobre una ventana de comparación de al menos 20 posiciones de restos, frecuentemente sobre una ventana de al menos 30-50 restos, en la que el porcentaje de identidad de secuencia se calcula comparando la secuencia de referencia con una secuencia que incluye deleciones o adiciones que suman el 20 por ciento o menos de la secuencia de referencia sobre la ventana de comparación. En las formas de realización específicas aplicadas a polipéptidos, la expresión "identidad sustancial" se refiere a que dos secuencias polipeptídicas, cuando se alinean óptimamente, tal como mediante los programas GAP o BESTFIT utilizando los pesos de hueco por defecto ("default gap weights"), comparten al menos un 80 por ciento de identidad de secuencia, preferentemente al menos un 89 por ciento de identidad de secuencia, al menos un 95 por ciento de identidad de secuencia o más (por ejemplo, un 99 por ciento de identidad de secuencia). Preferentemente, las posiciones de restos de aminoácidos que no son idénticas difieren en sustituciones de aminoácidos conservadoras.

"Correspondiente al", "con referencia a" o "con respecto a" cuando se utilizan en el contexto de la numeración de una secuencia de aminoácidos o polinucleotídica determinada se refieren a la numeración de los restos de una secuencia de referencia especificada cuando la secuencia de aminoácidos o polinucleotídica determinada se compara con la secuencia de referencia. En otras palabras, el número de resto o posición de resto de un polímero determinado se designa con respecto a la secuencia de referencia más que por la posición numérica real del resto dentro de la secuencia de aminoácidos o polinucleotídica determinada. Por ejemplo, una secuencia de aminoácidos determinada, tal como la de una cetorreductasa manipulada, puede alinearse con una secuencia de referencia introduciendo huecos para optimizar las coincidencias de restos entre las dos secuencias. En estos casos, aunque haya huecos presentes, la numeración del resto en la secuencia de aminoácidos o polinucleotídica determinada se hace con respecto a la secuencia de referencia con la que se ha alineado.

"Estereoselectividad" se refiere a la formación preferencial en una reacción química o enzimática de un estereoisómero frente a otro. La estereoselectividad puede ser parcial, cuando se favorece la formación de un estereoisómero frente al otro, o puede ser completa cuando sólo se forma un estereoisómero. Cuando los estereoisómeros son enantiómeros, la estereoselectividad se conoce como enantioselectividad, la fracción (por lo general presentada como porcentaje) de un enantiómero en la suma de ambos. Como alternativa, en la técnica se presenta comúnmente (por lo general como porcentaje) como el exceso enantiomérico (e.e.) calculado a partir del mismo según la fórmula [enantiómero principal - enantiómero minoritario]/[enantiómero principal + enantiómero minoritario]. Cuando los estereoisómeros son diastereoisómeros, la estereoselectividad se conoce como diastereoselectividad, la fracción (por lo general presentada como porcentaje) de un diastereómero en una mezcla de dos diastereómeros, como alternativa presentada comúnmente como el exceso diastereomérico (d.e.). El exceso enantiomérico y el exceso diastereomérico son tipos de exceso estereomérico.

"Altamente estereoselectivo" se refiere a un polipéptido cetorreductasa que es capaz de convertir o reducir el sustrato al correspondiente producto en la forma (S) con al menos aproximadamente un 99% de exceso estereomérico.

"Estereoespecificidad" se refiere a la conversión preferencial en una reacción química o enzimática de un estereoisómero frente a otro. La estereoespecificidad puede ser parcial, cuando se favorece la conversión de un estereoisómero frente al otro, o puede ser completa cuando sólo se convierte un estereoisómero.

"Quimioselectividad" se refiere a la formación preferencial en una reacción química o enzimática de un producto frente a otro.

"Propiedad de la enzima mejorada" se refiere a un polipéptido cetorreductasa que presenta una mejora en cualquier propiedad de la enzima en comparación con una cetorreductasa de referencia. Para los polipéptidos cetorreductasa manipulados descritos en el presente documento, la comparación se hace generalmente con la enzima cetorreductasa de tipo silvestre, aunque en algunas formas de realización, la cetorreductasa de referencia puede ser otra cetorreductasa manipulada mejorada. Las propiedades de la enzima para las que es deseable la mejora incluyen, pero no se limitan a, la actividad enzimática (que puede expresarse en términos de porcentaje de conversión del sustrato), la estabilidad térmica, el perfil de pH, los requisitos de cofactor, la refractariedad a inhibidores (por ejemplo, inhibición por el producto), la estereoespecificidad y la estereoselectividad (incluida la enantioselectividad).

"Actividad enzimática aumentada" se refiere a una propiedad mejorada de los polipéptidos cetorreductasa manipulados, que puede representarse por un aumento de la actividad específica (por ejemplo, producto producido/hora/peso de proteína) o un aumento del porcentaje de conversión del sustrato en el producto (por ejemplo, el porcentaje de conversión de la cantidad de sustrato de partida en producto en un período de tiempo especificado utilizando una cantidad especificada de KRED) en comparación con la enzima cetorreductasa de referencia. Los métodos ejemplares para determinar la actividad enzimática se proporcionan en los Ejemplos. Cualquier propiedad relacionada con la actividad enzimática puede verse afectada, incluidas las propiedades enzimáticas clásicas de K_m, $V_{máx}$ o k_{cat}, cuyos cambios pueden conducir a un aumento de la actividad enzimática. Las mejoras de la actividad enzimática pueden ser de aproximadamente una mejora del 100% en comparación con la actividad enzimática de la correspondiente enzima cetorreductasa de tipo silvestre, hasta incluso un 200%, 500%, 1.000%, 1.500%, o más, en comparación con la actividad enzimática de la cetorreductasa natural u otra cetorreductasa manipulada de la que se derivaron los polipéptidos cetorreductasa. En formas de realización específicas, la enzima cetorreductasa manipulada presenta una actividad enzimática mejorada en el intervalo de una mejora del 100% al 200%, del 200% al 1.000%, hasta el 1.500% y más, en comparación con la de la enzima cetorreductasa parental, de tipo silvestre, u otra enzima cetorreductasa de referencia. El experto en la materia entiende que la actividad de cualquier enzima está limitada por la difusión de manera que la velocidad de recambio catalítico no puede rebasar la velocidad de difusión del sustrato, incluido cualquier cofactor necesario. El máximo teórico del límite de difusión, o k_{cat}/K_m, es generalmente de aproximadamente 10^8 a 10^9 (M^{-1} s^{-1}). Por lo tanto, cualquier mejora de la actividad enzimática de la cetorreductasa tendrá un límite superior relacionado con la velocidad de difusión de los sustratos sobre los que actúa la enzima cetorreductasa. La actividad de la cetorreductasa puede medirse mediante cualquiera de los ensayos convencionales utilizados para medir la cetorreductasa, tales como una disminución de la absorbancia o fluorescencia del NADPH debida a su oxidación con la reducción concomitante de una cetona a un alcohol, o por el producto producido en un ensayo acoplado. Las comparaciones de las actividades enzimáticas se realizan utilizando una preparación definida de enzima, un ensayo definido en condiciones establecidas, y uno o más sustratos definidos, como se describe adicionalmente de manera detallada en el presente documento. Generalmente, cuando se comparan lisados, se determinan el número de células y la cantidad de proteína ensayada así como el uso de sistemas de expresión idénticos y células hospedadoras idénticas para minimizar las variaciones en la cantidad de enzima producida por las células hospedadoras y la cantidad presente en los lisados.

"Conversión" se refiere a la reducción enzimática del sustrato al producto correspondiente. "Porcentaje de conversión" se refiere al porcentaje del sustrato que se reduce al producto dentro de un período de tiempo en las condiciones especificadas. Por lo tanto, la "actividad enzimática" o "actividad" de un polipéptido cetorreductasa pueden expresarse como el "porcentaje de conversión" del sustrato en el producto.

"Termoestable" se refiere a un polipéptido cetorreductasa que mantiene una actividad similar (más del 60% al 80%, por ejemplo) tras la exposición a temperaturas elevadas (por ejemplo, 40°C-80°C) durante un período de tiempo (por ejemplo, 0,5-24 horas) en comparación con la enzima no tratada.

"Estable frente a disolventes" se refiere a un polipéptido cetorreductasa que mantiene una actividad similar (más de, por ejemplo, el 60% al 80%) tras la exposición a concentraciones variables (por ejemplo, del 0,5% al 99%) de disolvente (alcohol isopropílico, tetrahidrofurano, 2-metiltetrahidrofurano, acetona, tolueno, acetato de butilo, metil terc-butil éter, etc.) durante un período de tiempo (por ejemplo, 0,5-24 horas) en comparación con la enzima no tratada.

"Estable frente al pH" se refiere a un polipéptido cetorreductasa que mantiene una actividad similar (más de, por ejemplo, el 60% al 80%) tras la exposición a pH alto o bajo (por ejemplo, de 8 a 12 ó 4,5-6) durante un período de tiempo (por ejemplo, 0,5-24 horas) en comparación con la enzima no tratada.

"Termoestable y estable frente a disolventes" se refiere a un polipéptido cetorreductasa que es tanto termoestable como estable frente a disolventes.

"Derivado de" tal como se utiliza en el presente documento en el contexto de las enzimas cetorreductasa manipuladas, identifica la enzima cetorreductasa originaria, y/o el gen que codifica tal enzima cetorreductasa, en la que se basó la manipulación. Por ejemplo, la enzima cetorreductasa manipulada de la SEQ ID NO: 36 se obtuvo haciendo evolucionar artificialmente, durante múltiples generaciones, el gen que codifica la enzima cetorreductasa de *Lactobacillus kefir* de la SEQ ID NO: 2. Por lo tanto, esta enzima cetorreductasa manipulada se "deriva de" la cetorreductasa de tipo silvestre de la SEQ ID NO: 2.

"Aminoácido o resto hidrófilo" se refiere a un aminoácido o resto que tiene una cadena lateral que presenta una hidrofobicidad inferior a cero según la escala de hidrofobicidad consenso normalizada de Eisenberg *et al.*, 1984, J. Mol. Biol. 179:125-142. Los aminoácidos hidrófilos codificados genéticamente incluyen L-Thr (T), L-Ser (S), L-His (H), L-Glu (E), L-Asn (N), L-Gln (Q), L-Asp (D), L-Lys (K) y L-Arg (R).

"Aminoácido o resto ácido" se refiere a un aminoácido o resto hidrófilo que tiene una cadena lateral que presenta un valor de pK inferior a aproximadamente 6 cuando el aminoácido se encuentra incluido en un péptido o polipéptido. Los aminoácidos ácidos tienen por lo general cadenas laterales cargadas negativamente a pH fisiológico debido a la pérdida de un ion hidrógeno. Los aminoácidos ácidos codificados genéticamente incluyen L-Glu (E) y L-Asp (D).

"Aminoácido o resto básico" se refiere a un aminoácido o resto hidrófilo que tiene una cadena lateral que presenta un valor de pK superior a aproximadamente 6 cuando el aminoácido se encuentra incluido en un péptido o polipéptido. Los aminoácidos básicos tienen por lo general cadenas laterales cargadas positivamente a pH fisiológico debido a la asociación con un ion hidronio. Los aminoácidos básicos codificados genéticamente incluyen L-Arg (R) y L-Lys (K).

"Aminoácido o resto polar" se refiere a un aminoácido o resto hidrófilo que tiene una cadena lateral que no está cargada a pH fisiológico, pero que tiene al menos un enlace en el que el par de electrones compartidos en común por dos átomos se mantiene más cerca de uno de los átomos. Los aminoácidos polares codificados genéticamente incluyen L-Asn (N), L-Gln (Q), L-Ser (S) y L-Thr (T).

"Aminoácido o resto hidrófobo" se refiere a un aminoácido o resto que tiene una cadena lateral que presenta una hidrofobicidad superior a cero según la escala de hidrofobicidad consenso normalizada de Eisenberg *et al.*, 1984, J. Mol. Biol. 179:125-142. Los aminoácidos hidrófobos codificados genéticamente incluyen L-Pro (P), L-Ile (I), L-Phe (F), L-Val (V), L-Leu (L), L-Trp (W), L-Met (M), L-Ala (A) y L-Tyr (Y).

"Aminoácido o resto aromático" se refiere a un aminoácido o resto hidrófilo o hidrófobo que tiene una cadena lateral que incluye al menos un anillo aromático o heteroaromático. Los aminoácidos aromáticos codificados genéticamente incluyen L-Phe (F), L-Tyr (Y) y L-Trp (W). Aunque debido a la pKa de su átomo de nitrógeno heteroaromático L-His (H) se clasifica a veces como resto básico, o como resto aromático ya que su cadena lateral incluye un anillo heteroaromático, en el presente documento la histidina se clasifica como resto hidrófilo o como "resto restringido" (véase más adelante).

"Aminoácido o resto restringido" se refiere a un aminoácido o resto que tiene una geometría restringida. En el presente documento, los restos restringidos incluyen L-pro (P) y L-his (H). La histidina tiene una geometría restringida porque tiene un anillo imidazol relativamente pequeño. La prolina tiene una geometría restringida, ya que también tiene un anillo de cinco miembros.

"Aminoácido o resto no polar" se refiere a un aminoácido o resto hidrófobo que tiene una cadena lateral que no está cargada a pH fisiológico y que tiene enlaces en los que el par de electrones compartidos en común por dos átomos se mantiene generalmente a la misma distancia de cada uno de los dos átomos (es decir, la cadena lateral no es polar). Los aminoácidos no polares codificados genéticamente incluyen L-Gly (G), L-Leu (L), L-Val (V), L-Ile (I), L-Met (M) y L-Ala (A).

"Aminoácido o resto alifático" se refiere a un aminoácido o resto hidrófobo que tiene una cadena lateral hidrocarbonada alifática. Los aminoácidos alifáticos codificados genéticamente incluyen L-Ala (A), L-Val (V), L-Leu (L) y L-Ile (I).

"Cisteína". El aminoácido L-cisteína (C) es inusual en cuanto a que puede formar puentes disulfuro con otros aminoácidos L-Cys (C) u otros aminoácidos que contienen un grupo sulfanilo o sulfhidrilo. Los "restos tipo cisteína" incluyen cisteína y otros aminoácidos que contienen fracciones sulfhidrilo que están disponibles para la formación de puentes disulfuro. La capacidad de L-Cys (C) (y otros aminoácidos con cadenas laterales que contienen -SH) de existir en un péptido ya sea en la forma reducida con grupos -SH libres como en la oxidada con puentes disulfuro influye en si L-Cys (C) contribuye al carácter hidrófobo o hidrófilo final de un péptido. Mientras L-Cys (C) presenta una hidrofobicidad de 0,29 según la escala consenso normalizada de Eisenberg (Eisenberg *et al.*, 1984, *supra*), debe entenderse que para los fines de la presente descripción L-Cys (C) se clasifica en su propio grupo único.

"Aminoácido o resto pequeño" se refiere a un aminoácido o resto que tiene una cadena lateral que está compuesta en total por tres o menos átomos de carbono y/o heteroátomos (excepto el carbono α y los átomos de hidrógeno). Los aminoácidos o restos pequeños pueden clasificarse adicionalmente como aminoácidos o restos pequeños alifáticos, no polares, polares o ácidos, según las definiciones anteriores. Los aminoácidos pequeños codificados genéticamente incluyen L-Ala (A), L-Val (V), L-Cys (C), L-Asn (N), L-Ser (S), L-Thr (T) y L-Asp (D).

"Aminoácido o resto que contiene un grupo hidroxilo" se refiere a un aminoácido que contiene una fracción hidroxilo (-OH). Los aminoácidos que contienen grupos hidroxilo codificados genéticamente incluyen L-Ser (S) L-Thr (T) y L-Tyr (Y).

Sustituciones o mutaciones de aminoácidos "conservadoras" se refiere a la posibilidad de intercambiar restos que tienen cadenas laterales similares, y por lo tanto implica por lo general la sustitución del aminoácido en el polipéptido con aminoácidos dentro de la misma clase o una clase similar de aminoácidos definida. Sin embargo, tal como se utiliza en el presente documento, en algunas formas de realización, las mutaciones conservadoras no incluyen sustituciones de resto hidrófilo a hidrófilo, hidrófobo a hidrófobo, que contiene hidroxilo a que contiene hidroxilo, o pequeño a pequeño, si la mutación conservadora puede ser en su lugar una sustitución de resto alifático a alifático, no polar a no polar, polar a polar, ácido a ácido, básico a básico, aromático a aromático o restringido a restringido. Además, tal como se utiliza en el presente documento, A, V, L o I pueden mutarse de manera conservadora, ya sea a otro resto alifático o a otro resto no polar. La siguiente Tabla 1 muestra las sustituciones conservadoras ejemplares.

Tabla 1: Sustituciones conservadoras

Resto	Posibles mutaciones conservadoras
A, L, V, I	Otro alifático (A, L, V, I) Otro no polar (A, L, V, I, G, M)
G, M	Otro no polar (A, L, V, I, G, M)
D, E	Otro ácido (D, E)
K, R	Otro básico (K, R)
P, H	Otro restringido (P, H)
N, Q, S, T	Otro polar (N, Q, S, T)
Y, W, F	Otro aromático (Y, W, F)
C	Ninguno

"Sustitución no conservadora" se refiere a la sustitución o mutación de un aminoácido en el polipéptido con un aminoácido con propiedades significativamente diferentes de la cadena lateral. Las sustituciones no conservadoras pueden utilizar aminoácidos entre, más que dentro de, los grupos definidos mencionadas anteriormente. En algunas formas de realización, una mutación no conservadora influye en (a) la estructura del esqueleto del péptido en la zona de la sustitución (por ejemplo, glicina por prolina) (b) la carga o hidrofobicidad, o (c) la mayor parte de la cadena lateral.

"Deleción" se refiere a la modificación del polipéptido mediante la eliminación de uno o más aminoácidos del polipéptido de referencia. Las deleciones en una secuencia de aminoácidos pueden comprender la eliminación de 1 o más aminoácidos, 2 o más aminoácidos, 3 o más aminoácidos, 4 o más aminoácidos, 5 o más aminoácidos, 6 o más aminoácidos, 8 o más aminoácidos, 10 o más aminoácidos, 15 o más aminoácidos, ó 20 o más aminoácidos, hasta el 10% del número total de aminoácidos, hasta el 15% del número total de aminoácidos, o hasta el 20% del número total de aminoácidos que componen la enzima de referencia al tiempo que se conserva la actividad enzimática y/o se conservan las propiedades mejoradas de una enzima cetorreductasa manipulada. Las deleciones pueden dirigirse a las porciones internas y/o a las porciones terminales del polipéptido. En diversas formas de realización, la deleción puede comprender un segmento continuo o puede ser discontinua.

"Inserción" se refiere a la modificación del polipéptido añadiendo uno o más aminoácidos al polipéptido de referencia. En algunas formas de realización, las enzimas cetorreductasa manipuladas mejoradas comprenden inserciones de uno o más aminoácidos en el polipéptido cetorreductasa natural, así como inserciones de uno o más aminoácidos en otros polipéptidos cetorreductasa mejorados. Las inserciones pueden estar en las porciones internas del polipéptido, o en el extremo carboxilo o amino terminal. Las inserciones, tal como se utiliza en el presente documento, incluyen proteínas de fusión como se conoce en la técnica. La inserción puede ser un segmento contiguo de aminoácidos o estar separada por uno o más de los aminoácidos del polipéptido natural.

"Fragmento" tal como se utiliza en el presente documento se refiere a un polipéptido que tiene una deleción amino terminal y/o carboxilo terminal, pero en el que la secuencia de aminoácidos restante es idéntica a las posiciones correspondientes en la secuencia. Los fragmentos pueden tener al menos 14 aminoácidos de longitud, al menos 20 aminoácidos de longitud, al menos 50 aminoácidos de longitud o más, y hasta el 70%, 80%, 90%, 95%, 98% y 99% del polipéptido cetorreductasa de longitud completa.

"Polipéptido aislado" se refiere a un polipéptido que está sustancialmente separado de otros contaminantes que lo acompañan de forma natural, por ejemplo, proteínas, lípidos y polinucleótidos. La expresión abarca polipéptidos que se han extraído o purificado a partir de su sistema de expresión o entorno natural (por ejemplo, célula hospedadora o síntesis *in vitro*). Las enzimas cetorreductasa pueden estar presentes dentro de una célula, presentes en el medio celular o prepararse de diversas formas, tal como lisados o preparaciones aisladas. Como tal, en algunas formas de realización, la enzima cetorreductasa mejorada puede ser un polipéptido aislado.

"Polipéptido sustancialmente puro" se refiere a una composición en la que la especie polipeptídica es la especie predominante presente (es decir, en base a la molaridad o al peso, es más abundante que cualquier otra especie macromolecular individual en la composición), y es generalmente una composición sustancialmente purificada cuando la especie objetivo comprende al menos aproximadamente un 50 por ciento de la especie macromolecular presente en moles o % en peso. Generalmente, una composición de cetorreductasa sustancialmente pura comprenderá aproximadamente un 60% o más, aproximadamente un 70% o más, aproximadamente un 80% o más, aproximadamente un 90% o más, aproximadamente un 95% o más, y aproximadamente un 98% o más de todas las especies macromoleculares en moles o % en peso presentes en la composición. En algunas formas de realización, la especie objetivo se purifica hasta la homogeneidad esencial (es decir, no pueden detectarse especies contaminantes en la composición mediante métodos de detección convencionales) en la que la composición consiste esencialmente en una única especie macromolecular. Las especies disolventes, las moléculas pequeñas (< 500 daltons) y las especies iones elementales no se consideran especies macromoleculares. En algunas formas de realización, el polipéptido cetorreductasa mejorado aislado es una composición polipeptídica sustancialmente pura.

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

"Rigurosidad de hibridación" se refiere a tales condiciones de lavado de ácidos nucleicos. Generalmente, las reacciones de hibridación se realizan en condiciones de baja rigurosidad, seguidas de lavados de rigurosidad variable, pero mayor. La expresión "hibridación moderadamente rigurosa" se refiere a condiciones que permiten que el ADN diana se una a un ácido nucleico complementario que tiene aproximadamente un 60% de identidad, preferentemente de aproximadamente un 75% de identidad, aproximadamente un 85% de identidad con el ADN diana, teniendo una identidad superior a aproximadamente un 90% con el polinucleótido diana. Las condiciones moderadamente rigurosas ejemplares son condiciones equivalentes a la hibridación en formamida al 50%, 5x solución de Denhart, 5xSSPE, SDS al 0,2% a 42°C, seguida de lavado en 0,2xSSPE, SDS al 0,2%, a 42°C. "Hibridación en condiciones de alta rigurosidad" se refiere generalmente a condiciones que tienen una diferencia de aproximadamente 10°C o menos con la temperatura de fusión térmica T_m tal como se determina en condición de solución para una secuencia polinucleotídica definida. En algunas formas de realización, una condición de alta rigurosidad se refiere a condiciones que permiten la hibridación de solamente aquellas secuencias de ácidos nucleicos que forman híbridos estables en NaCl 0,018 M a 65°C (es decir, si un híbrido no es estable en NaCl 0,018 M a 65°C, no será estable en condiciones de alta rigurosidad, tal como se contempla en el presente documento). Las condiciones de alta rigurosidad pueden proporcionarse, por ejemplo, por hibridación en condiciones equivalentes a formamida al 50%, 5x solución de Denhart, 5xSSPE, SDS al 0,2% a 42°C, seguida de lavado en 0,1xSSPE, y SDS al 0,1% a 65°C. Otra condición de alta rigurosidad es hibridar en condiciones equivalentes a la hibridación en 5x SSC que contiene SDS al 0,1% (p:v) a 65°C y lavado en 0,1x SSC que contiene SDS al 0,1% a 65°C. Otras condiciones de hibridación de alta rigurosidad, así como condiciones moderadamente rigurosas, se describen en las referencias citadas anteriormente.

Polinucleótido "heterólogo" se refiere a cualquier polinucleótido que se introduce en una célula hospedadora mediante técnicas de laboratorio, e incluye polinucleótidos que se extraen de una célula hospedadora, se someten a manipulación de laboratorio, y a continuación se reintroducen en una célula hospedadora.

"Con codones optimizados" se refiere a cambios en los codones del polinucleótido que codifica una proteína en relación a los utilizados preferentemente en un organismo particular, de manera que la proteína codificada se exprese de manera eficaz en el organismo de interés. Aunque el código genético está degenerado en cuanto a que la mayoría de los aminoácidos están representados por varios codones, llamados "sinónimos" o

“codones sinónimos", es bien sabido que el uso de codones por organismos particulares no es aleatorio y existe preferencia por tripletes de codones concretos. Esta preferencia de uso de codones puede ser mayor en relación con un determinado gen, genes de función común u origen ancestral, proteínas altamente expresadas frente a proteínas con bajo número de copias, y las regiones codificantes de proteínas globales del genoma de un organismo. En algunas formas de realización, los polinucleótidos que codifican las enzimas cetorreductasas pueden tener codones optimizados para la producción óptima a partir del organismo hospedador seleccionado para la expresión.

"Codones de elevada preferencia codónica, óptimos, preferentes" se refiere indistintamente a codones que se utilizan con mayor frecuencia en las regiones codificantes de proteínas que otros codones que codifican el mismo aminoácido. Los codones preferentes pueden determinarse con relación al uso de codones en un solo gen, un conjunto de genes de origen o función común, genes altamente expresados, la frecuencia de codones en las regiones codificantes de proteínas globales de todo el organismo, la frecuencia de codones en las regiones codificantes de proteínas globales de organismos relacionados, o combinaciones de los mismos. Los codones cuya frecuencia aumenta en función del nivel de expresión génica son por lo general codones óptimos para la expresión. Se conocen diversos métodos para determinar la frecuencia de codones (por ejemplo, el uso de codones, el uso relativo de codones sinónimos) y la preferencia codónica en organismos específicos, incluido el análisis multivariante, por ejemplo, utilizando análisis por conglomerados o análisis de correspondencias, y el número efectivo de codones utilizados en un gen (véase GCG CodonPreference, Genetics Computer Group Wisconsin Package; CodonW, John Peden, Universidad de Nottingham; McInerney, J. O, 1998, Bioinformatics 14:372-73; Stenico *et al.,* 1994, Nucleic Acids Res. 222437-46; Wright, F, 1990, Gene 87:23-29). Se dispone de tablas de uso de codones para una lista cada vez mayor de organismos (véase, por ejemplo, Wada *et al.,* 1992, Nucleic Acids Res. 20:2111-2118; Nakamura *et al.,* 2000, Nucl. Acids Res. 28:292; Duret, *et al., supra*; Henaut y Danchin, "*Escherichia coli* and Salmonella”, 1996, Neidhardt, *et al*., eds., ASM Press, Washington D.C., p. 2047-2066). La fuente de datos para obtener el uso de codones puede basarse en cualquier secuencia nucleotídica disponible capaz de codificar una proteína. Estos conjuntos de datos incluyen secuencias de ácidos nucleicos que se sabe realmente codifican proteínas expresadas (por ejemplo, secuencias codificantes de proteínas completas - CDS), marcadores de secuencias expresadas (ESTS), o regiones codificantes de secuencias genómicas predichas (véase, por ejemplo, Mount, D., Bioinformatics: Sequence and Genome Analysis, capítulo 8, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York, 2001; Uberbacher, E.C., 1996, Methods Enzymol. 266:259-281; Tiwari *et al.,* 1997, Comput. Appl. Biosci. 13:263-270).

"Secuencia de control" se define en el presente documento para que incluya todos los componentes que son necesarios o ventajosos para la expresión de un polipéptido de la presente descripción. Cada secuencia de control puede ser nativa o extraña a la secuencia de ácido nucleico que codifica el polipéptido. Tales secuencias de control incluyen, pero no se limitan a, una secuencia líder, secuencia de poliadenilación, secuencia polipeptídica, promotor, secuencia de péptido señal, y terminador de la transcripción. Como mínimo, las secuencias de control incluyen un promotor, y señales de terminación de la transcripción y de la traducción. Las secuencias de control pueden estar provistas de conectores con el fin de introducir sitios de restricción específicos que faciliten la ligación de las secuencias de control con la región codificante de la secuencia de ácido nucleico que codifica un polipéptido.

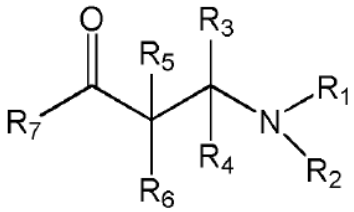
"Unido operativamente" se define en el presente documento como una configuración en la que una secuencia de control está situada apropiadamente en una posición con respecto a la secuencia codificante de la secuencia de ADN de manera que la secuencia de control dirija o regule la expresión de un polinucleótido y/o polipéptido.

"Secuencia promotora" es una secuencia de ácido nucleico que es reconocida por una célula hospedadora para la expresión de la región codificante. La secuencia de control puede comprender una secuencia promotora apropiada. La secuencia promotora contiene secuencias de control de la transcripción, que intervienen en la expresión del polipéptido. El promotor puede ser cualquier secuencia de ácido nucleico que muestre actividad transcripcional en la célula hospedadora de elección incluidos promotores mutantes, truncados e híbridos, y puede obtenerse a partir de genes que codifican polipéptidos extracelulares o intracelulares ya sean homólogos o heterólogos a la célula hospedadora.

"Sistema de regeneración de cofactor" se refiere a un conjunto de reaccionantes que participan en una reacción que reduce la forma oxidada del cofactor (por ejemplo, $NADP^+$ a NADPH). Los cofactores oxidados por la reducción catalizada por cetorreductasa del sustrato ceto son regenerados en forma reducida por el sistema de regeneración de cofactor. Los sistemas de regeneración de cofactor comprenden un reductor estequiométrico que es una fuente de equivalentes de hidrógeno reductores y es capaz de reducir la forma oxidada del cofactor. El sistema de regeneración de cofactor puede comprender adicionalmente un catalizador, por ejemplo un catalizador enzimático, que cataliza la reducción de la forma oxidada del cofactor por el reductor. En la técnica se conocen sistemas de regeneración de cofactor para regenerar NADH o NADPH a partir de NAD^+ o $NADP^+$, respectivamente, y pueden utilizarse en los métodos que se describen en el presente documento.

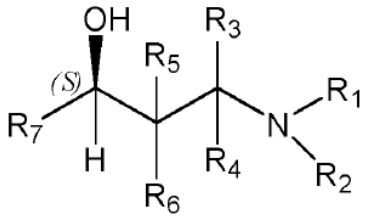
Polipéptidos cetorreductasa

La presente descripción proporciona enzimas cetorreductasa manipuladas ("KRED") que son útiles para reducir o convertir un sustrato 3-aril-3-cetopropanamina que tiene la fórmula estructural (I):



(I)

a la correspondiente (S)-3-aril-3-hidroxipropanamina que tiene la fórmula estructural (II):



(II)

en la que para (I) y (II), cada uno de R_1 y R_2 está seleccionado independientemente del grupo que consiste en hidrógeno, un alquilo inferior opcionalmente sustituido, un cicloalquilo opcionalmente sustituido, un arilo opcionalmente sustituido, o como alternativa, en las que R_1 y R_2 juntos forman un cicloalquilo opcionalmente sustituido o un cicloarilo opcionalmente sustituido que tiene 3-7 átomos de carbono; cada uno de R_3, R_4, R_5 y R_6 está seleccionado independientemente del grupo que consiste en hidrógeno y un alquilo inferior opcionalmente sustituido; y R_7 es un arilo opcionalmente sustituido.

Tal como se ha descrito anteriormente, los sustratos adecuados para los polipéptidos cetorreductasa de la presente invención incluyen aquellos con los sustituyentes descritos para la fórmula (I) que pueden estar opcionalmente sustituidos. La expresión "opcionalmente sustituido" se refiere a la sustitución de hidrógeno con un radical monovalente o divalente. Los grupos de sustitución adecuados incluyen, por ejemplo, hidroxilo, nitro, amino, imino, ciano, halo, tio, tioamido, amidino, imidino, oxo, oxamidino, metoxamidino, imidino, guanidina, sulfonamida, carboxilo, formilo, alquilo inferior, alcoxi inferior, y similares. Los productos que tienen la estructura de fórmula (II) tendrán los mismos sustituyentes opcionalmente sustituidos descritos anteriormente.

Tal como se utiliza en el presente documento, la expresión "alquilo inferior" se refiere a grupos alquilo de cadena ramificada o lineal que comprenden de uno a diez átomos de carbono que están sustituidos o no sustituidos, por ejemplo, con uno o más grupos halógeno, hidroxilo, y similares. La expresión "alcoxi inferior" se refiere en el presente documento a RO-, en el que R es alquilo inferior. Los ejemplos representativos de grupos alcoxi inferior incluyen metoxi, etoxi, t-butoxi, trifluorometoxi, y similares.

"Arilo" se refiere en el presente documento a grupos aromáticos monocíclicos y policíclicos que tienen de 3 a 14 átomos de carbono o heteroátomos en el esqueleto, e incluye grupos arilo carbocíclico y grupos arilo heterocíclico. Los grupos arilo carbocíclico son grupos arilo en los que todos los átomos del anillo en el anillo aromático son carbono. La expresión "arilo heterocíclico" se refiere en el presente documento a grupos arilo que tienen de 1 a 4 heteroátomos como átomos del anillo en un anillo aromático siendo los demás átomos del anillo átomos de carbono. Cuando se utiliza en relación a sustituyentes arilo, el término "policíclico" se refiere en el presente documento a estructuras cíclicas condensadas y no condensadas en las que al menos una estructura cíclica es aromática. Las fracciones arilo ejemplares empleadas como sustituyentes en los compuestos de la presente invención incluyen fenilo, naftilo, tiofenilo, y similares. Los sustituyentes arilo sustituidos ejemplares incluyen bencilo, y similares.

"Cicloalquilo" se refiere en el presente documento a un sustituyente alquilo monocíclico o policíclico, heterocíclico o carbocíclico. Los sustituyentes cicloalquilo típicos tienen de 3 a 10 átomos en el esqueleto (es decir, anillo) en los que cada átomo del esqueleto es un carbono o un heteroátomo. Los heteroátomos adecuados incluyen nitrógeno, oxígeno y azufre.

"Cicloarilo" se refiere en el presente documento a un sustituyente alquilo monocíclico o policíclico, heterocíclico o carbocíclico. Los sustituyentes cicloarilo típicos tienen de 3 a 10 átomos en el esqueleto en el que cada átomo del esqueleto es un carbono o un heteroátomo. Los heteroátomos adecuados incluyen nitrógeno, oxígeno y azufre.

Los sustratos típicos tienen la estructura de fórmula (I) en la que cada uno de R_1 y R_2 es independientemente hidrógeno, un alquilo inferior de uno a diez o de uno a seis átomos de carbono, o un fenilo. Por lo general, uno de R_1 y R_2 es hidrógeno y el otro es metilo, o ambos R_1 y R_2 son metilo. Cada uno de R_3, R_4, R_5 y R_6 es por lo general hidrógeno o un alquilo inferior de uno a diez o de uno a seis átomos de carbono, y más generalmente cada uno de R_3, R_4, R_5 y R_6 es hidrógeno. R_7 es por lo general un tiofenilo opcionalmente sustituido (tal como, por ejemplo, un 2-tienilo opcionalmente sustituido o un 3-tienilo opcionalmente sustituido) o un fenilo opcionalmente sustituido. Más generalmente, R_7 es 2-tienilo o fenilo. Los sustratos útiles tienen la combinación de sustituyentes: R_1 es hidrógeno, R_2 es metilo, cada uno de R_3-R_6 es hidrógeno, R_7 está seleccionado del grupo que consiste en 2-tienilo y fenilo; y R_1 y R_2 son metilo, cada uno de R_3-R_6 es hidrógeno, R_7 está seleccionado del grupo que consiste en 2-tienilo y fenilo. La reducción de estos sustratos produce un producto 3-aril-3-hidroxipropanamina correspondiente con la estructura de fórmula (II) siendo R1-R_7 como se ha descrito anteriormente.

Los polipéptidos cetorreductasa de la presente descripción son particularmente útiles para la reducción o conversión estereoselectiva del sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina ("el sustrato dimetilo", es decir, una compuesto sustrato con la fórmula estructural (I) en la que R_1 y R_2 son ambos metilo, cada uno de R_3-R_6 es hidrógeno, y R_7 es 2-tienilo) en el producto (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina ("el producto dimetilo", es decir, un compuesto producto con la fórmula estructural (II) en la que R_1 y R_2 son ambos metilo, cada uno de R_3-R_6 es hidrógeno, y R_7 es 2-tienilo) y que tiene una propiedad mejorada cuando se compara con la enzima KRED de tipo silvestre natural obtenida a partir de *Lactobacillus kefir* (SEQ ID NO: 2), *Lactobacillus brevis* (SEQ ID NO: 4) o *Lactobacillus minor* (SEQ ID NO: 106), o cuando se compara con otras enzimas cetorreductasa manipuladas. Las propiedades de la enzima para las que es deseable la mejora incluyen, pero no se limitan a, la actividad enzimática, la estabilidad térmica, el perfil de pH, los requisitos de cofactor, la refractariedad a inhibidores (por ejemplo, inhibición por el producto), la estereoespecificidad, la estereoselectividad, y la estabilidad frente a disolventes. Las mejoras pueden estar relacionadas con una sola propiedad de la enzima, tal como la actividad enzimática, o una combinación de diferentes propiedades de la enzima, tal como la actividad enzimática y la estereoselectividad.

Para los polipéptidos descritos en el presente documento, la secuencia de aminoácidos de los polipéptidos cetorreductasa tienen, en comparación con la secuencia de referencia de la SEQ ID NO: 2, 4 ó 106, el requisito de que: (1) el resto de aminoácido correspondiente a la posición de resto 94 es un aminoácido no polar, (2) el resto de aminoácido correspondiente a la posición de resto 145 es un aminoácido aromático o leucina, y (3) el resto de aminoácido correspondiente a la posición de resto 190 es una cisteína o un aminoácido restringido. En algunas formas de realización, los polipéptidos cetorreductasa tienen, en comparación con las secuencias de KRED de *L. kefir* o *L. brevis* o *L. minor* de tipo silvestre de la SEQ ID NO: 2, 4 y 106, respectivamente, al menos las siguientes características: (1) el resto de aminoácido correspondiente a la posición 94 es leucina, isoleucina, alanina, valina, metionina o glicina, (2) el resto de aminoácido correspondiente a la posición 145 es tirosina, triptófano, fenilalanina o leucina, y (3) el resto de aminoácido correspondiente a la posición 190 es cisteína o prolina. En algunas formas de realización, los polipéptidos cetorreductasa tienen, en comparación con las secuencias de KRED de *L. kefir L.* o *brevis* o *L. minor* de la SEQ ID NO: 2, 4 y 106, al menos las siguientes características: (1) el resto de aminoácido correspondiente a la posición 94 es glicina, (2) el resto de aminoácido correspondiente a la posición 145 es fenilalanina, tirosina o leucina, particularmente fenilalanina o leucina, y (3) el aminoácido correspondiente a la posición 190 es prolina o cisteína, particularmente prolina. En algunas formas de realización, los polipéptidos cetorreductasa tienen, en comparación con las secuencias de KRED de *L. kefir* o *L. brevis* o *L. minor* de la SEQ ID NO: 2, 4 y 106, al menos las siguientes características: (1) el resto de aminoácido correspondiente a la posición 94 es glicina, (2) el resto de aminoácido correspondiente a la posición 145 es fenilalanina, y (3) el aminoácido correspondiente a la posición 190 es prolina.

En algunas formas de realización, tal como se ha indicado anteriormente, la cetorreductasa manipulada con actividad enzimática mejorada se describe con referencia a la cetorreductasa de *Lactobacillus kefir* de la SEQ ID NO: 2, la cetorreductasa de *Lactobacillus brevis* de la SEQ ID NO: 4, la cetorreductasa de *Lactobacillus minor* de la SEQ ID NO: 106, o una cetorreductasa manipulada. La posición de resto de aminoácido se determina en estas cetorreductasas a partir del resto metionina iniciador (M) (es decir, M representa la posición de resto 1), aunque el experto en la materia entenderá que este resto metionina iniciador puede ser eliminado por la maquinaria biológica de procesamiento, tal como en una célula hospedadora o en un sistema de traducción *in vitro*, para generar una proteína madura que carece del resto metionina iniciador. La posición de resto de aminoácido en la que un aminoácido o cambio de aminoácido particular está presente en una secuencia de aminoácidos se describe a veces en el presente documento con los términos "*Xn*", o "posición *n*", en los que *n* se refiere a la posición del resto. Cuando los restos de aminoácidos en la misma posición de resto difieren entre las cetorreductasas, los diferentes restos se indican mediante un símbolo "/" siendo la disposición, por ejemplo, "resto de kefir/resto de brevis/minor". En algunas formas de realización, una mutación por sustitución, que es un reemplazo de un resto de aminoácido en un resto correspondiente de una secuencia de referencia, por ejemplo, la cetorreductasa de tipo silvestre de la

SEQ ID NO: 2, SEQ ID NO: 4, o SEQ ID NO: 106 con un resto de aminoácido diferente se indica mediante el símbolo "→".

En el presente documento, las mutaciones se describen a veces como una mutación "a un" tipo de aminoácido. Por ejemplo, el resto 211 puede mutarse "a un" resto básico. Pero el uso de la expresión "a un" no excluye las mutaciones de un aminoácido de una clase a otro aminoácido de la misma clase. Por ejemplo, el resto 211 puede mutarse de una lisina a una arginina.

La secuencia polinucleotídica que codifica la cetorreductasa natural de *Lactobacillus kefir* y *Lactobacillus brevis* (también denominada "alcohol deshidrogenasa" o "ADH"), y por lo tanto las correspondientes secuencias de aminoácidos, están disponibles en el nº de registro del GenBank AAP94029 GI:33112056 para *Lactobacillus kefir*; nº de registro del GenBank CAD66648 GI:28400789 para *Lactobacillus brevis*; y la solicitud de patente de EE.UU. Nº 20040265978 o la SEQ ID NO: 106 para *Lactobacillus minor*.

En algunas formas de realización, los polipéptidos cetorreductasa del presente documento puede tener varias modificaciones con respecto a la secuencia de referencia (por ejemplo, un polipéptido natural o un polipéptido manipulado) que dan como resultado una propiedad mejorada de la cetorreductasa. En tales formas de realización, el número de modificaciones en la secuencia de aminoácidos puede comprender uno o más aminoácidos, 2 o más aminoácidos, 3 o más aminoácidos, 4 o más aminoácidos, 5 o más aminoácidos, 6 o más aminoácidos, 8 o más aminoácidos, 10 o más aminoácidos, 15 o más aminoácidos, ó 20 o más aminoácidos, hasta el 10% del número total de aminoácidos, hasta el 15% del número total de aminoácidos, hasta el 20% del número total de aminoácidos, o hasta el 30% del número total de aminoácidos de la secuencia polipeptídica de referencia. En algunas formas de realización, el número de modificaciones con respecto al polipéptido natural o un polipéptido manipulado que produce una propiedad mejorada de la cetorreductasa puede comprender de aproximadamente 1-2, 1-3, 1-4, 1-5, 1-6, 1- 7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-26, 1-30, 1-35 o aproximadamente 1-40 modificaciones de la secuencia de referencia. En algunas formas de realización, el número de modificaciones puede ser 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35 o aproximadamente 40 restos de aminoácidos. Las modificaciones pueden comprender inserciones, deleciones, sustituciones, o combinaciones de las mismas.

En algunas formas de realización, las modificaciones comprenden sustituciones de aminoácidos con respecto a la secuencia de referencia. Las sustituciones que pueden producir una propiedad mejorada de la cetorreductasa pueden estar en uno o más aminoácidos, 2 o más aminoácidos, 3 o más aminoácidos, 4 o más aminoácidos, 5 o más aminoácidos, 6 o más aminoácidos, 8 o más aminoácidos, 10 o más aminoácidos, 15 o más aminoácidos, ó 20 o más aminoácidos, hasta el 10% del número total de aminoácidos, hasta el 20% del número total de aminoácidos, o hasta el 30% del número total de aminoácidos de la secuencia de la enzima de referencia. En algunas formas de realización, el número de sustituciones en el polipéptido natural o un polipéptido manipulado que produce una propiedad mejorada de la cetorreductasa puede comprender de aproximadamente 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-26, 1-30 1-35 o aproximadamente 1-40 sustituciones de aminoácidos de la secuencia de referencia. En algunas formas de realización, el número de sustituciones puede ser 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35 o aproximadamente 40 restos de aminoácidos.

En algunas formas de realización, la propiedad mejorada (en comparación con el polipéptido de tipo silvestre u otro polipéptido manipulado) del polipéptido cetorreductasa es con respecto a un aumento de su estereoselectividad para reducir o convertir el sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina en el producto (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina. En algunas formas de realización, la propiedad mejorada del polipéptido cetorreductasa es con respecto a un aumento en la estereoselectividad, es decir, en el presente documento, un aumento del exceso estereomérico del producto. En algunas formas de realización, la propiedad mejorada del polipéptido cetorreductasa es con respecto a un aumento de su capacidad de reducir o convertir un mayor porcentaje del sustrato en producto. En algunas formas de realización, la propiedad mejorada del polipéptido cetorreductasa es con respecto a un aumento de su velocidad de conversión del sustrato en el producto. Esta mejora de la actividad enzimática puede manifestarse por la capacidad para utilizar menos polipéptido mejorado en comparación con el de tipo silvestre u otra secuencia de referencia (por ejemplo, SEQ ID NO: 6) para reducir o convertir la misma cantidad de sustrato. En algunas formas de realización, la propiedad mejorada del polipéptido cetorreductasa es con respecto a su estabilidad o termoestabilidad. En algunas formas de realización, el polipéptido cetorreductasa tiene más de una propiedad mejorada.

En algunas formas de realización, el polipéptido cetorreductasa de la descripción es capaz de convertir el sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina en el producto (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina, con un porcentaje de exceso estereomérico de al menos aproximadamente un 99% y a una velocidad que está mejorada en comparación con la de un polipéptido de referencia con la secuencia de aminoácidos de la SEQ ID NO: 6. Los polipéptidos ejemplares con tales propiedades incluyen, pero no se limitan a, polipéptidos que comprenden las secuencias de aminoácidos correspondientes a las SEQ ID NO: 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76 y 78. Debido a que el polipéptido de referencia con la secuencia de aminoácidos de la SEQ ID NO: 6 es capaz de convertir el sustrato

dimetilo en el producto dimetilo a una velocidad (por ejemplo, 50% de 50 g/l de sustrato convertidos en producto en 24 horas con aproximadamente 5 g/l de la KRED) y con una estereoselectividad (99% de exceso estereomérico) que está mejorada en comparación con la de tipo silvestre (por ejemplo, SEQ ID NO: 2), los polipéptidos del presente documento que están mejorados en comparación con la SEQ ID NO: 6 también están mejorados en comparación con los de tipo silvestre.

En algunas formas de realización, el polipéptido cetorreductasa es capaz de convertir el sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina en el producto (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina, con un porcentaje de exceso estereomérico de al menos aproximadamente un 99% y a una velocidad que está al menos 1,5 a 2 veces mejorada en comparación con la de un polipéptido de referencia con la secuencia de aminoácidos de la SEQ ID NO: 6. Los polipéptidos ejemplares con tales propiedades incluyen, pero no se limitan a, polipéptidos que comprenden las secuencias de aminoácidos correspondientes a las SEQ ID NO: 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76 y 78.

En algunas formas de realización, el polipéptido cetorreductasa es capaz de convertir el sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina en el producto (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina, con un porcentaje de exceso estereomérico de al menos aproximadamente un 99% y a una velocidad que está al menos 2-10 veces mejorada en comparación con la de un polipéptido de referencia con la secuencia de aminoácidos de la SEQ ID NO: 6. Los polipéptidos ejemplares con tales propiedades incluyen, pero no se limitan a, polipéptidos que comprenden las secuencias de aminoácidos correspondientes a las SEQ ID NO: 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76 y 78.

En algunas formas de realización, el polipéptido cetorreductasa es capaz de convertir el sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina en el producto (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina, con un porcentaje de exceso estereomérico de al menos aproximadamente un 99% y a una velocidad que está mejorada más de 10 veces en comparación con la de un polipéptido de referencia con la secuencia de aminoácidos de la SEQ ID NO: 6. Los polipéptidos ejemplares con tales propiedades incluyen, pero no se limitan a, polipéptidos que comprenden las secuencias de aminoácidos correspondientes a las SEQ ID NO: 60, 62, 64, 66, 68, 70, 72, 74, 76 y 78.

En algunas formas de realización, el polipéptido cetorreductasa es capaz de convertir al menos aproximadamente un 95% del sustrato dimetilo en el producto dimetilo en menos de aproximadamente 24 horas cuando se lleva a cabo con más de aproximadamente 100 g/l de sustrato y menos de aproximadamente 5 g/l del polipéptido. Los polipéptidos ejemplares que tienen esta capacidad incluyen, pero no se limitan a, los polipéptidos que comprenden las secuencias de aminoácidos correspondientes a las SEQ ID NO: 60, 62, 64, 66, 68, 70, 72, 74, 76 y 78.

En algunas formas de realización, los polipéptidos cetorreductasa de la descripción son altamente estereoselectivos y pueden reducir el sustrato dimetilo al producto dimetilo en más de aproximadamente un 99%, 99,1%, 99,2%, 99,3%, 99,4%, 99,5%, 99,6%, 99,7%, 99,8% ó 99,9% de exceso estereomérico. Los polipéptidos cetorreductasa ejemplares con una estereoselectividad tan alta incluyen, pero no se limitan a, los polipéptidos que comprenden las secuencias de aminoácidos correspondientes a las SEQ ID NO: 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76 y 78.

La siguiente Tabla 2 proporciona una lista de las SEQ ID NO descritas en el presente documento con los niveles de actividad asociados con respecto a la reducción de *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina a (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina. Las siguientes secuencias se basan en las secuencias cetorreductasa de *L. kefir* de tipo silvestre (SEQ ID NO: 1 y 2) a menos que se especifique otra cosa. En la siguiente Tabla 2, en cada fila aparecen dos SEQ ID NO, en las que el número impar se refiere a la secuencia nucleotídica que codifica la secuencia de aminoácidos proporcionada por el número par. La columna donde aparece el número de mutaciones (es decir, los cambios de restos) es con respecto al número de sustituciones de aminoácidos en comparación con la secuencia de aminoácidos de KRED de *L. kefir* de la SEQ ID NO: 2. En la columna de actividad, un signo "+" indica que la actividad de la enzima es 1,5-3 veces superior a la de la enzima de la SEQ ID NO: 6. Del mismo modo, un símbolo de dos signos más "++" indica que el polipéptido tiene una actividad aproximadamente 3-10 veces superior a la de la enzima de referencia de la SEQ ID NO: 6, y el símbolo de tres signos más "+++" indica que el polipéptido tiene una actividad más de 10 veces superior a la de la enzima de referencia de la SEQ ID NO: 6.

Tabla 2: Lista de secuencias y mejora de actividad correspondiente con respecto a la reducción de *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina ("sustrato dimetilo"):

SEQ ID NO (Polinucleótido/ aminoácido)	Cambios de residuos con respecto la SEQ ID NO: 2	Número de cambios con respecto a la secuencia de *L. kefir*	Actividad sustrato dimetilo
5/6	G7H; D66E; A94G; S96I; E145L; F147L; Y190P; L195R; V196L; A202W; K211L; Y249W	12	"Control"
7/8	A94G; S96V; E145L; Y190P; V196L; I226V; Y249W	7	+
9/10	A94G; S96V; E145F; Y190P; V196L; I226V; Y249W	7	+
11/12	A94G; S96V; E145L; F147L; Y190P; V196L; I226V; Y249W	8	+
13/14	A94G; S96V; E145L; F 147M; Y190P; V196L; I226V; Y249W	8	+
15/16	A94G S96V E145L F147V Y190P V196L I226V Y249W	8	+
17/18	A94G; S96V; E145L; F147I; Y190P; V196L; I226V; Y249W	8	+
19/20	A94G; S96V; E145F; F147L; Y190P; V196L; I226V; Y249W	8	+
21/22	A94G; S96V; E145F; F147M; Y190P; V196L; I226V; Y249W	8	+
23/24	A94G; S96V; E145L; F147I; Y190P; V196L; I226V; Y249W	8	+
25/26	H40R A94G S96V E145F F147M Y190P V196L M206Y I226V Y249W	10	++
27/28	H40R A94G S96V E145F F147M Y190P V196L M206F I226V Y249W	10	++
29/30	H40R; A94G; S96V; E145L; F147M; Y190P; V196L; M206W; I226V; Y249W	10	++
31/32	H40R A94G S96V E145F F147M Y190P V196L I226V Y249W	9	++
33/34	A94G S96V E145F F147M L153T Y190P L195M V196L L199Y I226V Y249W	11	++
35/36	H40R; A94G; S96V; E145F; F147M; Y190P; L195M; V196L; M206W; I226V; D233G; Y249W	12	++
37/38	H40R A94G S96V E145F F147M L153T Y190P L195M V196L M206W I226V Y249W	12	++
39/40	H40R A94G S96V E145F F147M Y190P L195M V196L M206W I226V D233G Y249W	12	++
41/42	A94G S96V E145F F147M T152S L153T Y190P V196L I226V Y249W	10	++
43/44	H40R A94G S96V E145F F147M T152S Y190P V196L I226V Y249W	10	++
45/46	H40R; A94G; S96V; E145F; F147M; Y190P; V196L; M206W; I226V; Y249W	10	++
47/48	H40R A94G S96V E145F F147M Y190P V196L L199W I226V Y249W	10	++
49/50	H40R A94G S96V E145F F147M L153T Y190P V196L I226V Y249W	10	++
51/52	H40R; A94G; S96V; E145Y; F147M; Y190P; L195M; V196L; M206W; I226V; D233G; Y249W	12	++
53/54	H40R A94G S96V E145L F147M Y190P L195M V196L M206W I226V D233G Y249W	12	++
55/56	H40R; A94G; S96V; E145F; F147M; L153V; Y190P; L195M; V196L; M206W; I226V; D233G; 249W	13	++
57/58	H40R; A94G; S96V; E145F; F147M; Y190P; L195M; V196L; L199V; M206W; I226V; D233G; Y249W	13	++

SEQ ID NO (Polinucleótido/ aminoácido)	Cambios de residuos con respecto la SEQ ID NO: 2	Número de cambios con respecto a la secuencia de *L. kefir*	Actividad sustrato dimetilo
59/60	H40R; K46R; A94G; S96V; E145F; F147M; L153V; Y190P; L195M; V196L; L199V; M206W; I226V; 233G; V245I; Y249W	16	+++
61/62	H40R; K46R; A94G; S96V; E145F; F147M; L153V; T185S; Y190P; L195M; V196L; L199V; M206W; 226V; D233G; V245I; Y249W	17	+++
63/64	H40R K72T A94G S96V R108H L134M M141V E145L F147M T185S Y190P L195M V196L L199V 206W I226V D233G Y249W	18	+++
65/66	H40R; A94G; S96V; E145F; F147M; S169C; Y190P; L195M; V196L; L199V; M206W; I226V; 233A; Y249W	14	+++
67/68	H40R; A94G; S96V; I144V; E145F; F147M; Y190P; L195M; V196L; L199V; M206W; I226V; D233G; Y249W	14	+++
69/70	H40R; A94G; S96V; R108H; E145F; F147M; L153V; Y190P; L195M; V196L; L199V; M206W; I226V; D233G; V245I; Y249W	16	+++
71/72	H40R; A94G; S96V; E145F; F147M; L153V; V186I; Y190P; L195M; V196L; L199V; M206W; I226V; D233G; Y249W	15	+++
73/74	H40R; A94G; S96V; R108H; E145F; F147M; L153V; T185S; Y190P; L195M; V196L; L199V; M206W; I226V; D233G; V245I; Y249W	17	+++
75/76	H40R; K46R; V60I; K72T; A94G; S96V; E145F; F147M; L153V; Y190P; L195M; V196L; L199V; M206W; I226V; D233G; V245I; Y249W	18	+++
77/78	H40R; K46R; A94G; S96V; R108H; E145F; F147M; L153V; Y190P; L195M; V196L; L199V; M206W; I226V; D233G; V245I; Y249W	17	+++

En algunas formas de realización, los polipéptidos cetorreductasa del presente documento comprenden una secuencia de aminoácidos que es idéntica en al menos aproximadamente un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más, en comparación con una secuencia de referencia basada en las SEQ ID NO: 2, 4 ó 106 que tiene las siguientes características descritas en el presente documento, por ejemplo: (1) el resto de aminoácido correspondiente a la posición 94 es glicina, (2) el resto de aminoácido correspondiente a la posición 145 es fenilalanina, tirosina o leucina, particularmente fenilalanina o leucina, y (3) el aminoácido correspondiente a la posición 190 es cisteína o prolina, particularmente prolina, en las que los polipéptidos cetorreductasa tienen al menos las características anteriores. En algunas formas de realización, estos polipéptidos cetorreductasa pueden tener una o más modificaciones con respecto a la secuencia de aminoácidos de referencia. Las modificaciones pueden incluir sustituciones, deleciones e inserciones. Las sustituciones pueden ser sustituciones no conservadoras, sustituciones conservadoras, o una combinación de sustituciones no conservadoras y conservadoras. En algunas formas de realización, estos polipéptidos cetorreductasa pueden tener opcionalmente de aproximadamente 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-26, 1-30, 1-35 o aproximadamente 1-40 mutaciones en otros restos de aminoácidos. En algunas formas de realización, el número de mutaciones puede ser 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35 o aproximadamente 40 mutaciones en otros restos de aminoácidos.

En algunas formas de realización, el polipéptido cetorreductasa del presente documento comprende una secuencia de aminoácidos que corresponde a las fórmulas de secuencia que se exponen en la SEQ ID NO: 107 o la SEQ ID NO: 108 o la SEQ ID NO: 109, o una región de las mismas, tal como los restos 90-211. La fórmula de secuencia de la SEQ ID NO: 107 se basa en la secuencia de aminoácidos de la cetorreductasa de *Lactobacillus kefir* (SEQ ID NO: 2). La fórmula de secuencia de la SEQ ID NO: 108 se basa en la secuencia de aminoácidos de la cetorreductasa de *Lactobacillus brevis* (SEQ ID NO: 4), y la fórmula de secuencia de la SEQ ID NO: 109 se basa en la secuencia de aminoácidos de la cetorreductasa de *Lactobacillus minor* (SEQ ID NO: 106). El polipéptido cetorreductasa basado en la fórmula de secuencia de las SEQ ID NO: 107, 108 y 109 especifica que el resto correspondiente a X94 es un resto de aminoácido no polar, el resto correspondiente a X145 es leucina o un resto de aminoácido aromático; y el resto correspondiente a X190 es una cisteína o un resto de aminoácido restringido. En algunas formas de realización, el polipéptido cetorreductasa basado en las fórmulas de secuencia de las SEQ ID NO: 107, 108 ó 109 puede comprender una secuencia de aminoácidos que es idéntica en al menos

aproximadamente un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ó 99% a una secuencia de aminoácidos de referencia basada en la SEQ ID NO: 2, 4 ó 106 que tiene al menos las características descritas en el presente documento para los restos de aminoácidos X94, X145, y X190, a condición de que los polipéptidos cetorreductasa tengan al menos las características especificadas. En algunas formas de realización, los polipéptidos cetorreductasa tienen al menos las siguientes características: (1) el resto de aminoácido correspondiente a X94 es glicina, (2) el resto de aminoácido correspondiente a X145 es fenilalanina, tirosina o leucina, particularmente fenilalanina o leucina y (3) el aminoácido correspondiente a X190 es cisteína o prolina, particularmente prolina, a condición de que los polipéptidos cetorreductasa tengan también al menos las características especificadas.

En algunas formas de realización, los polipéptidos que comprenden una secuencia de aminoácidos basada en las fórmulas de secuencia proporcionadas en las SEQ ID NO: 107, 108 ó 109, o una región de las mismas, tal como los restos 90-211, pueden tener una o más mutaciones conservadoras en comparación con las secuencias de aminoácidos de la SEQ ID NO: 2, 4 ó 106. Las mutaciones conservadoras ejemplares incluyen sustituciones de aminoácidos tales como, pero no limitadas a: sustitución del resto correspondiente a X46 lisina (K) con otro resto hidrófilo o básico, por ejemplo, arginina; sustitución del resto correspondiente a X60 valina (V) con otro resto alifático o no polar, por ejemplo, isoleucina (I); sustitución del resto correspondiente a X66 ácido aspártico (D) con otro resto hidrófilo o ácido, por ejemplo, ácido glutámico (E); sustitución del resto correspondiente a X134 con otro resto no polar o alifático, por ejemplo, metionina (M); sustitución del resto correspondiente a X141 con otro resto no polar o alifático, por ejemplo, valina (V); sustitución del resto correspondiente a X144 isoleucina (I) con otro resto no polar o alifático, por ejemplo, valina (V); sustitución del resto correspondiente a X147 con otro resto hidrófobo, por ejemplo, metionina (M), valina (V), isoleucina (I) o leucina (L); sustitución del resto correspondiente a X152 (S) con otro resto polar, por ejemplo, serina; sustitución del resto correspondiente a X153 (L) con otro resto no polar o alifático, por ejemplo, valina; sustitución del resto correspondiente a X185 treonina (T) con otro resto polar, por ejemplo, serina (S); sustitución del resto X186 valina (V) con otro resto no polar o alifático, por ejemplo, isoleucina (I); sustitución del resto X195 leucina (L) con otro resto alifático o no polar, por ejemplo, metionina (M); sustitución del resto X196 valina (V) con otro resto no polar o alifático, por ejemplo, leucina (L), isoleucina (I); sustitución del resto X199 leucina (L) con otro resto alifático o hidrófobo, por ejemplo, valina (V); sustitución del resto X226 isoleucina (I) con otro resto no polar o alifático, por ejemplo, valina (V); sustitución del resto X245 valina (V) con otro resto no polar o alifático, por ejemplo, isoleucina (I); y la sustitución del resto X249 tirosina (Y) con otro resto hidrófobo o aromático, por ejemplo, triptófano (W).

En algunas formas de realización, los polipéptidos cetorreductasa pueden comprender una secuencia de aminoácidos basada en las fórmulas de secuencia de las SEQ ID NO: 107, 108 ó 109, o una región de las mismas, tal como los restos 90-211, que tiene las características especificadas para los restos X94, X145 y X190 como se describe en el presente documento, y en las que la secuencia de aminoácidos puede incluir adicionalmente una o más características seleccionadas de entre las siguientes: el resto correspondiente a X7 es glicina, metionina, alanina, valina, leucina, isoleucina, prolina o histidina, particularmente histidina; el resto correspondiente a X40 es treonina, serina, histidina, ácido glutámico, asparagina, glutamina, ácido aspártico, lisina o arginina, particularmente arginina; el resto correspondiente a X46 es arginina o lisina; el resto correspondiente a X60 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente isoleucina; el resto correspondiente a X66 es ácido glutámico o ácido aspártico; el resto correspondiente a X72 es treonina, serina, histidina, ácido glutámico, asparagina, glutamina, ácido aspártico, lisina y arginina, particularmente treonina; el resto correspondiente a X96 es prolina, serina, treonina, asparagina, glutamina, glicina, metionina, alanina, valina, leucina o isoleucina, particularmente valina; el resto correspondiente a X108 es treonina, serina, histidina, ácido glutámico, asparagina, glutamina, ácido aspártico, lisina, particularmente histidina; el resto correspondiente a X134 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente metionina; el resto correspondiente a X141 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente metionina, particularmente valina; el resto correspondiente a X144 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente valina; el resto correspondiente a X147 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina y tirosina, particularmente isoleucina, metionina, valina o leucina; el resto correspondiente a X152 es serina, treonina, asparagina o glutamina, particularmente serina; el resto correspondiente a X153 es alanina, valina, leucina, isoleucina, serina, treonina, asparagina o glutamina, particularmente treonina; el resto correspondiente a X169 es cisteína, serina, treonina, asparagina, glutamina, particularmente cisteína; el resto correspondiente a X185 es serina, treonina, asparagina o glutamina, particularmente treonina; el resto correspondiente a X186 es glicina, metionina, alanina, valina, leucina, isoleucina, particularmente isoleucina; el resto correspondiente a X195 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente metionina; el resto correspondiente a X196 es glicina, metionina, alanina, valina, leucina, isoleucina, particularmente leucina; el resto correspondiente a X199 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina o tirosina, particularmente tirosina, triptófano o valina; el resto correspondiente a X206 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina o tirosina, particularmente tirosina, triptófano o fenilalanina; el resto correspondiente a X211 es lisina, arginina, glicina, metionina, alanina, valina, leucina o isoleucina, particularmente leucina; el resto correspondiente a X226 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente valina; el resto correspondiente a X233 es glicina, metionina, alanina, valina, leucina, isoleucina, ácido aspártico o ácido glutámico, particularmente glicina; el resto correspondiente a X245 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente isoleucina; y el resto correspondiente a X249 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina o tirosina, particularmente triptófano. En

algunas de las formas de realización anteriores, el resto de aminoácido correspondiente a X94 es glicina; el resto de aminoácido correspondiente a X145 es fenilalanina o leucina, particularmente fenilalanina; y el aminoácido correspondiente a X190 es cisteína o prolina, particularmente prolina. En algunas formas de realización, los polipéptidos que comprenden una secuencia de aminoácidos que corresponde a las fórmulas de secuencia proporcionadas en las SEQ ID NO: 107, 108 ó 109 (o región de las mismas) pueden tener adicionalmente uno o más de los restos no especificados mediante una X a mutar en comparación con la secuencia de referencia de la SEQ ID NO: 2, 4 ó 106. En algunas formas de realización, las mutaciones pueden ser de aproximadamente 1-2, 1-3, 1-4, 1-5, 1-6, 1- 7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-26, 1-30, 1-35 o aproximadamente 1-40 mutaciones en otros restos de aminoácidos no definidos mediante una X anteriormente. En algunas formas de realización, el número de mutaciones puede ser 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35 ó aproximadamente 40 mutaciones en otros restos de aminoácidos. En algunas formas de realización, las mutaciones comprenden mutaciones conservadoras.

En algunas formas de realización, los polipéptidos cetorreductasa pueden comprender una secuencia de aminoácidos basada en las fórmulas de secuencia de las SEQ ID NO: 107, 108 ó 109, o una región de las mismas, tal como los restos 90-211, en las que la secuencia de aminoácidos tiene las características especificadas para los restos X94, X145 y X190 como se describe en el presente documento, y en las que el polipéptido puede incluir adicionalmente una o más o al menos todas las características seleccionadas de entre las siguientes: el resto correspondiente a X40 es treonina, serina, histidina, ácido glutámico, asparagina, glutamina, ácido aspártico, lisina o arginina, particularmente arginina; el resto correspondiente a X96 es serina, treonina, asparagina, glutamina, glicina, metionina, alanina, valina, leucina o isoleucina, particularmente valina; el resto correspondiente a X147 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina y tirosina, particularmente isoleucina, metionina, valina o leucina; el resto correspondiente a X195 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente metionina; el resto correspondiente a X196 es glicina, metionina, alanina, valina, leucina, isoleucina, particularmente leucina; el resto correspondiente a X199 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina o tirosina, particularmente tirosina, triptófano o valina; el resto correspondiente a X206 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina o tirosina, particularmente tirosina, triptófano o fenilalanina; el resto correspondiente a X226 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente valina; el resto correspondiente a X233 es glicina, metionina, alanina, valina, leucina, isoleucina, ácido aspártico o ácido glutámico, particularmente glicina; y el resto correspondiente a X249 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina o tirosina, particularmente triptófano. En algunas de las formas de realización anteriores, el resto de aminoácido correspondiente a X94 es glicina; el resto de aminoácido correspondiente a X145 es fenilalanina, tirosina o leucina, particularmente fenilalanina o leucina; y el aminoácido correspondiente a X190 es cisteína o prolina, particularmente prolina. En algunas formas de realización, los polipéptidos cetorreductasa pueden tener adicionalmente 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-26, 1-30, 1-35 o aproximadamente 1-40 diferencias de restos en otros restos de aminoácidos en comparación con la secuencia de referencia de la SEQ ID NO: 2, 4 ó 106. En algunas formas de realización, el número de diferencias puede ser 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35 o aproximadamente 40 diferencias de restos en otros restos de aminoácidos. En algunas formas de realización, las diferencias comprenden mutaciones conservadoras. En algunas formas de realización, el polipéptido cetorreductasa comprende una secuencia de aminoácidos con al menos las características anteriores, y en el que la secuencia de aminoácidos tiene una identidad de al menos un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ó 99% en comparación con una secuencia de referencia basada en las SEQ ID NO: 2, 4 ó 106 con las características anteriores.

En algunas formas de realización, los polipéptidos cetorreductasa pueden comprender una secuencia de aminoácidos basada en las fórmulas de secuencia de las SEQ ID NO: 107, 108 ó 109, o una región de las mismas, tal como los restos 90-211, en las que la secuencia de aminoácidos tiene las características especificadas para los restos X94, X145 y X190 tal como se describe en el presente documento, y en las que el polipéptido puede incluir adicionalmente una o más o al menos todas las características seleccionadas de entre las siguientes: el resto correspondiente a X40 es treonina, serina, histidina, ácido glutámico, asparagina, glutamina, ácido aspártico, lisina o arginina, particularmente arginina; el resto correspondiente a X96 es serina, treonina, asparagina, glutamina, glicina, metionina, alanina, valina, leucina o isoleucina, particularmente valina; el resto correspondiente a X147 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina y tirosina, particularmente isoleucina, metionina, valina o leucina; el resto correspondiente a X153 es alanina, valina, leucina, isoleucina, serina, treonina, asparagina o glutamina, particularmente treonina; el resto correspondiente a X195 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente metionina; el resto correspondiente a X196 es glicina, metionina, alanina, valina, leucina, isoleucina, particularmente leucina; el resto correspondiente a X199 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina o tirosina, particularmente tirosina, triptófano o valina; el resto correspondiente a X206 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina o tirosina, particularmente tirosina, triptófano o fenilalanina; el resto correspondiente a X226 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente valina; el resto correspondiente a X233 es glicina, metionina, alanina, valina, leucina, isoleucina, ácido aspártico o ácido glutámico, particularmente glicina; y el resto correspondiente a X249 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina o tirosina, particularmente triptófano. En algunas de las formas de realización anteriores, el resto de aminoácido correspondiente a X94 es glicina; el resto de aminoácido correspondiente a X145 es fenilalanina, tirosina o leucina, particularmente fenilalanina o leucina; y el aminoácido

correspondiente a X190 es cisteína o prolina, particularmente prolina. En algunas formas de realización, los polipéptidos cetorreductasa pueden tener adicionalmente 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-26, 1-30, 1-35 o aproximadamente 1-40 diferencias de restos en otros restos de aminoácidos en comparación con la secuencia de referencia de la SEQ ID NO: 2, 4 ó 106. En algunas formas de realización, el número de diferencias puede ser 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35 o aproximadamente 40 diferencias de restos en otros restos de aminoácidos. En algunas formas de realización, las diferencias comprenden mutaciones conservadoras. En algunas formas de realización, el polipéptido cetorreductasa comprende una secuencia de aminoácidos con al menos las características anteriores, y en el que la secuencia de aminoácidos tiene una identidad de al menos un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ó 99% en comparación con una secuencia de referencia basada en las SEQ ID NO: 2, 4 ó 106 con las características anteriores.

En algunas formas de realización, los polipéptidos cetorreductasa pueden comprender una secuencia de aminoácidos basada en las fórmulas de secuencia de las SEQ ID NO: 107, 108 ó 109, o una región de las mismas, tal como los restos 90-211, en las que la secuencia de aminoácidos tiene las características especificadas para los restos X94, X145 y X190 como se describe en el presente documento, y en las que el polipéptido puede incluir adicionalmente una o más o al menos todas las características seleccionadas de entre las siguientes: el resto correspondiente a X40 es treonina, serina, histidina, ácido glutámico, asparagina, glutamina, ácido aspártico, lisina o arginina, particularmente arginina; el resto correspondiente a X96 es serina, treonina, asparagina, glutamina, glicina, metionina, alanina, valina, leucina o isoleucina, particularmente valina; el resto correspondiente a X147 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina y tirosina, particularmente isoleucina, metionina, valina o leucina; el resto correspondiente a X195 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente metionina; el resto correspondiente a X196 es glicina, metionina, alanina, valina, leucina, isoleucina, particularmente leucina; el resto correspondiente a X199 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina o tirosina, particularmente tirosina, triptófano o valina; el resto correspondiente a X206 s isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina o tirosina, particularmente tirosina, triptófano o fenilalanina; el resto correspondiente a X226 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente valina; el resto correspondiente a X233 es glicina, metionina, alanina, valina, leucina, isoleucina, ácido aspártico o ácido glutámico, particularmente glicina; el resto correspondiente a X245 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente isoleucina; y el resto correspondiente a X249 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina o tirosina, particularmente triptófano. En algunas de las formas de realización anteriores, el resto de aminoácido correspondiente a X94 es glicina; el resto de aminoácido correspondiente a X145 es fenilalanina, tirosina o leucina, particularmente fenilalanina o leucina; y el aminoácido correspondiente a X190 es cisteína o prolina, particularmente prolina. En algunas formas de realización, los polipéptidos cetorreductasa pueden tener adicionalmente 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-26, 1-30, 1-35 o aproximadamente 1-40 diferencias de restos en otros restos de aminoácidos en comparación con la secuencia de referencia de la SEQ ID NO: 2, 4 ó 106. En algunas formas de realización, el número de diferencias puede ser 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35 o aproximadamente 40 diferencias de restos en otros restos de aminoácidos. En algunas formas de realización, las diferencias comprenden mutaciones conservadoras. En algunas formas de realización, el polipéptido cetorreductasa comprende una secuencia de aminoácidos con al menos las características anteriores, y en el que la secuencia de aminoácidos tiene una identidad de al menos un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ó 99% en comparación con una secuencia de referencia basada en las SEQ ID NO: 2, 4 ó 106 con las características anteriores.

En algunas formas de realización, los polipéptidos cetorreductasa pueden comprender una secuencia de aminoácidos basada en las fórmulas de secuencia de las SEQ ID NO: 107, 108 ó 109, o una región de las mismas, tal como los restos 90-211, en las que la secuencia de aminoácidos tiene las características especificadas para los restos X94, X145 y X190 como se describe en el presente documento, y en las que el polipéptido puede incluir adicionalmente una o más o al menos todas las características seleccionadas de entre las siguientes: el resto correspondiente a X40 es treonina, serina, histidina, ácido glutámico, asparagina, glutamina, ácido aspártico, lisina o arginina, particularmente arginina; el resto correspondiente a X96 es serina, treonina, asparagina, glutamina, glicina, metionina, alanina, valina, leucina o isoleucina, particularmente valina; el resto correspondiente a X147 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina y tirosina, particularmente isoleucina, metionina, valina o leucina; el resto correspondiente a X153 es alanina, valina, leucina, isoleucina, serina, treonina, asparagina o glutamina, particularmente treonina; el resto correspondiente a X195 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente metionina; el resto correspondiente a X196 es glicina, metionina, alanina, valina, leucina, isoleucina, particularmente leucina; el resto correspondiente a X199 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina o tirosina, particularmente tirosina, triptófano o valina; el resto correspondiente a X206 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina o tirosina, particularmente tirosina, triptófano o fenilalanina; el resto correspondiente a X226 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente valina; el resto correspondiente a X233 es glicina, metionina, alanina, valina, leucina, isoleucina, ácido aspártico o ácido glutámico, particularmente glicina; el resto correspondiente a X245 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente isoleucina; y el resto correspondiente a X249 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina o tirosina, particularmente triptófano. En algunas de las formas de realización anteriores, el resto de aminoácido correspondiente a X94 es glicina; el resto de aminoácido

correspondiente a X145 es fenilalanina, tirosina o leucina, particularmente fenilalanina o leucina; y el aminoácido correspondiente a X190 es cisteína o prolina, particularmente prolina. En algunas formas de realización, los polipéptidos cetorreductasa pueden tener adicionalmente 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-26, 1-30, 1-35 o aproximadamente 1-40 diferencias de restos en otros restos de aminoácidos en comparación con la secuencia de referencia de la SEQ ID NO: 2, 4 ó 106. En algunas formas de realización, el número de diferencias puede ser 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35 o aproximadamente 40 diferencias de restos en otros restos de aminoácidos. En algunas formas de realización, las diferencias comprenden mutaciones conservadoras. En algunas formas de realización, el polipéptido cetorreductasa comprende una secuencia de aminoácidos con al menos las características anteriores, y en el que la secuencia de aminoácidos tiene una identidad de al menos un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ó 99% en comparación con una secuencia de referencia basada en la SEQ ID NO: 2, 4 ó 106 con las características anteriores.

En algunas formas de realización, una cetorreductasa mejorada comprende una secuencia de aminoácidos basada en las fórmulas de secuencia de las SEQ ID NO: 107, 108 ó 109, o una región de las mismas, tal como los restos 90-211, en las que la secuencia de aminoácidos tiene al menos las siguientes características: el resto correspondiente a X94 es glicina; el resto correspondiente a X145 es leucina o un resto aromático, particularmente fenilalanina o leucina; el resto correspondiente a X190 es cisteína o un resto restringido, particularmente cisteína o prolina; el resto correspondiente a X40 es arginina; el resto correspondiente a X96 es valina; el resto correspondiente a X147 es metionina; el resto correspondiente a X153 es valina, el resto correspondiente a X186 es isoleucina, el resto correspondiente a X195 es metionina, el resto correspondiente a X196 es leucina; el resto correspondiente a X199 es valina; el resto correspondiente a X206 es triptófano; el resto correspondiente a X233 es glicina, y el resto correspondiente a X249 es triptófano. En algunas formas de realización, los polipéptidos cetorreductasa pueden tener adicionalmente 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-25, 1-30, 1-35 o aproximadamente 1-40 diferencias de restos en otras posiciones de resto en comparación con una secuencia de referencia de la SEQ ID NO: 72. En algunas formas de realización, el número de diferencias puede ser 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35 o aproximadamente 40 diferencias de restos en otros restos de aminoácidos. En algunas formas de realización, las diferencias comprenden mutaciones conservadoras. En algunas formas de realización, el polipéptido cetorreductasa comprende una secuencia de aminoácidos con las características anteriores, y en el que la secuencia de aminoácidos tiene una identidad de al menos un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ó 99% con la SEQ ID NO: 72.

En algunas formas de realización, los polipéptidos cetorreductasa con una región correspondiente a los restos 90-211 y que tienen las características especificadas para los restos X94, X145 y X190 como se describe en el presente documento, pueden incluir adicionalmente una o más características seleccionadas de entre las siguientes: el resto correspondiente a X96 es serina, treonina, asparagina, glutamina, glicina, metionina, alanina, valina, leucina o isoleucina, particularmente valina; el resto correspondiente a X108 es treonina, serina, histidina, ácido glutámico, asparagina, glutamina, ácido aspártico, lisina, particularmente histidina; el resto correspondiente a X134 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente metionina; el resto correspondiente a X141 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente metionina, particularmente valina; el resto correspondiente a X144 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente valina; el resto correspondiente a X147 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina y tirosina, particularmente isoleucina, metionina, valina o leucina; el resto correspondiente a X152 es serina, treonina, asparagina o glutamina, particularmente serina; el resto correspondiente a X153 es alanina, valina, leucina, isoleucina, serina, treonina, asparagina o glutamina, particularmente treonina; el resto correspondiente a X169 es cisteína, serina, treonina, asparagina, glutamina, particularmente cisteína; el resto correspondiente a X185 es serina, treonina, asparagina o glutamina, particularmente treonina; el resto correspondiente a X186 es glicina, metionina, alanina, valina, leucina, isoleucina, particularmente isoleucina; el resto correspondiente a X195 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente metionina; el resto correspondiente a X196 es glicina, metionina, alanina, valina, leucina, isoleucina, particularmente leucina; el resto correspondiente a X199 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina o tirosina, particularmente tirosina, triptófano o valina; el resto correspondiente a X206 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina o tirosina, particularmente tirosina, triptófano o fenilalanina; y el resto correspondiente a X211 es lisina, arginina, glicina, metionina, alanina, valina, leucina o isoleucina, particularmente leucina. En algunas de las formas de realización anteriores, el resto de aminoácido correspondiente a X94 es glicina; el resto de aminoácido correspondiente a X145 es fenilalanina, tirosina o leucina, particularmente fenilalanina o leucina; y el aminoácido correspondiente a X190 es cisteína o prolina, particularmente prolina. En algunas formas de realización, la región o dominio correspondiente a los restos 90-211 pueden tener adicionalmente 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18 ó 1-20 diferencias de restos en otros restos de aminoácidos en comparación con el dominio de una secuencia de referencia basada en la SEQ ID NO: 2, 4 ó 106. En algunas formas de realización, el número de diferencias puede ser 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, o aproximadamente 20 diferencias de restos en otros restos de aminoácidos en el dominio. En algunas formas de realización, las diferencias comprenden mutaciones conservadoras.

En algunas formas de realización, una cetorreductasa mejorada comprende una secuencia de aminoácidos que tiene una región o dominio correspondiente a los restos 90-211 de la fórmula de secuencia de las SEQ ID NO: 107, 108 ó 109, que tiene las características especificadas para los restos X94, X145 y X190 como se describe en el presente documento, y en las que el polipéptido puede incluir adicionalmente una o más características seleccionadas de entre las siguientes: el resto correspondiente a X96 es serina, treonina, asparagina, glutamina, glicina, metionina, alanina, valina, leucina o isoleucina, particularmente valina; el resto correspondiente a X147 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina y tirosina, particularmente isoleucina, metionina, valina o leucina; el resto correspondiente a X153 es alanina, valina, leucina, isoleucina, serina, treonina, asparagina o glutamina, particularmente treonina; el resto correspondiente a X195 es glicina, metionina, alanina, valina, leucina o isoleucina, particularmente metionina; el resto correspondiente a X196 es glicina, metionina, alanina, valina, leucina, isoleucina, particularmente leucina; el resto correspondiente a X199 es isoleucina, fenilalanina, valina, leucina, triptófano, metionina, alanina o tirosina, particularmente tirosina, triptófano o valina; y el resto correspondiente a X206 es fenilalanina, tirosina, o triptófano. En algunas de las formas de realización anteriores, el resto de aminoácido correspondiente a X94 es glicina; el resto de aminoácido correspondiente a X145 es fenilalanina, tirosina o leucina, particularmente fenilalanina o leucina; y el aminoácido correspondiente a X190 es cisteína o prolina, particularmente prolina. En algunas formas de realización, las mutaciones comprenden mutaciones conservadoras. En algunas formas de realización, la región o dominio correspondiente a los restos 90-211 pueden tener adicionalmente 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18 ó 1-20 diferencias de restos en otros restos de aminoácidos en comparación con el dominio de una secuencia de referencia basada en la SEQ ID NO: 2, 4 ó 106. En algunas formas de realización, el número de diferencias puede ser 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, o aproximadamente 20 diferencias de restos en otros restos de aminoácidos en el dominio. En algunas formas de realización, las diferencias comprenden mutaciones conservadoras. En algunas formas de realización, las diferencias comprenden mutaciones conservadoras. En algunas formas de realización, el polipéptido cetorreductasa comprende una secuencia de aminoácidos con al menos las características anteriores, y en el que la secuencia de aminoácidos tiene una identidad de al menos un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ó 99% en comparación con la secuencia de aminoácidos correspondiente a los restos 90-211 de una secuencia de referencia basada en la SEQ ID NO: 2, 4 ó 106 con las características anteriores.

En algunas formas de realización, los polipéptidos cetorreductasa pueden comprender una secuencia de aminoácidos que es idéntica en al menos aproximadamente un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ó 99% a una secuencia de referencia basada en la SEQ ID NO: 2, 4 ó 106, o un dominio de la misma, tal como los restos 90-211, a condición de que el resto correspondiente a la posición 94 de la SEQ ID NO: 2, 4 ó 106 sea glicina; el resto correspondiente a la posición 145 de la SEQ ID NO: 2, 4 ó 106 sea fenilalanina o leucina, y el resto correspondiente a la posición 190 de la SEQ ID NO: 2, 4 ó 106 sea cisteína o prolina, en las que el polipéptido cetorreductasa tiene al menos las características anteriores y puede tener opcionalmente una o más de las siguientes sustituciones de manera que el polipéptido esté adicionalmente mejorado (con respecto a la estereoselectividad, la actividad enzimática y/o la termoestabilidad) en comparación con la cetorreductasa de *kefir* de tipo silvestre u otra cetorreductasa manipulada (tal como la SEQ ID NO: 6): 7 → H, S; 40 → R; 46 → R; 60 → I; 66 → E; 72 → T; 96 → V, I, L; 108 → H; 134 → M; 141 → V; 144 → V; 152 → S; 153 → T, V; 169 → C; 185 → S; 186 → I; 195 → R, M; 196 → L; 199 → L, W, F, Y, V; 206 → F, W, Y; 211 → L; 226 → V; 233 → G, A; 245 → I; y 249 → W.

En algunas formas de realización, los polipéptidos cetorreductasa pueden comprender una secuencia de aminoácidos que es idéntica en al menos aproximadamente un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ó 99% a una secuencia de referencia basada en la SEQ ID NO: 2, 4 ó 106, o una región o dominio de la misma, tal como los restos 90-211, a condición de que el resto correspondiente a la posición 94 de la SEQ ID NO: 2, 4 ó 106 sea glicina, el resto correspondiente a la posición 145 de la SEQ ID NO: 2, 4 ó 106 sea fenilalanina y el resto correspondiente a la posición 190 de la SEQ ID NO: 2, 4 ó 106 sea prolina, en las que el polipéptido cetorreductasa tiene al menos las características anteriores y, puede tener opcionalmente una o más de las siguientes sustituciones de manera que el polipéptido esté mejorado adicionalmente (con respecto a la estereoselectividad, la actividad enzimática y/o la termoestabilidad) en comparación con la cetorreductasa de *kefir* de tipo silvestre u otra cetorreductasa manipulada (tal como la SEQ ID NO: 8): 7 → H, S; 40 → R; 46 → R; 60 → I; 66 → E; 72 → T; 96 → V, I, L; 108 → H; 144 → V; 152 → S; 153 → T, V; 169 → C; 185 → S; 186 → I; 195 → R, M; 198 → N; 199 → L, W, F, Y, V; 206 → F, W, Y; 211 → L; 226 → V; 233 → G, A; 242 → I; 245 → I; y 249 → W.

En algunas formas de realización, los polipéptidos cetorreductasa de la descripción pueden comprender una región con una secuencia de aminoácidos que es idéntica en al menos aproximadamente un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ó 99% a una secuencia de referencia basada en la SEQ ID NO: 2, 4 ó 106, o de una región o dominio de la misma, tal como los restos 90-211, a condición de que el resto correspondiente a la posición 94 de la SEQ ID NO: 2, 4 ó 106 sea glicina, el resto correspondiente a la posición 145 de la SEQ ID NO: 2, 4 ó 106 sea fenilalanina o leucina, el resto correspondiente a la posición 190 de la SEQ ID NO: 2, 4 ó 106 sea prolina o cisteína, y el resto correspondiente a la posición 40 de la SEQ ID NO: 2, 4 ó 106 sea arginina, y, además, tenga una o más de las siguientes sustituciones de manera que el polipéptido esté mejorado adicionalmente (con respecto a la estereoselectividad, la actividad enzimática y/o la termoestabilidad) en comparación con la cetorreductasa de *kefir* de tipo silvestre u otra cetorreductasa manipulada (tal como la

SEQ ID NO: 8): 46 → R; 96 → V; 108 → H; 152 → N; 153 → V; 195 → M; 199 → V, E, R; 206 → W; 226 → V; 233 → G; 245 → I; y 249 → W.

En algunas formas de realización, los polipéptidos cetorreductasa de la descripción pueden comprender una región con una secuencia de aminoácidos que es idéntica en al menos aproximadamente un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ó 99% a una secuencia de referencia basada en la SEQ ID NO: 2, 4 ó 106 (o una región o dominio de la misma, tal como los restos 90-211) a condición de que el resto correspondiente a la posición 94 de la SEQ ID NO: 2, 4 ó 106 sea glicina, el resto correspondiente a la posición 145 de la SEQ ID NO: 2, 4 ó 106 sea fenilalanina, el resto correspondiente a la posición 190 de la SEQ ID NO: 2, 4 ó 106 sea prolina, el resto correspondiente a la posición 40 de la SEQ ID NO: 2, 4 ó 106 sea arginina, en las que los polipéptidos cetorreductasa tienen al menos las características anteriores y, opcionalmente, una o más de la siguientes sustituciones de manera que el polipéptido esté mejorado adicionalmente (con respecto a la estereoselectividad, la actividad enzimática y/o la termoestabilidad) en comparación con la cetorreductasa de *kefir* de tipo silvestre u otra cetorreductasa manipulada (tal como la SEQ ID NO: 6): 46 → R; 96 → V; 108 → H; 152 → N; 153 → V; 195 → M; 199 → Y, W, V; 206 → W; 226 → V; 233 → G; 245 → I; y 249 → W.

En algunas formas de realización, cada una de las enzimas cetorreductasa manipuladas mejoradas descritas en el presente documento puede comprender deleciones de los polipéptidos descritos en el presente documento. Por lo tanto, para todas y cada una de las formas de realización de los polipéptidos cetorreductasa de la descripción, las deleciones, que pueden ser deleciones internas o externas (es decir, truncamientos C-terminales y/o N-terminales), pueden comprender uno o más aminoácidos, 2 o más aminoácidos, 3 o más aminoácidos, 4 o más aminoácidos, 5 o más aminoácidos, 6 o más aminoácidos, 8 o más aminoácidos, 10 o más aminoácidos, 15 o más aminoácidos, o 20 o más aminoácidos ácidos, hasta el 10% del número total de aminoácidos, hasta el 20% del número total de aminoácidos, o hasta el 30% del número total de aminoácidos de los polipéptidos cetorreductasa, siempre que se mantenga la actividad funcional de la actividad de cetorreductasa. En algunas formas de realización, las deleciones pueden comprender, 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-25, 1-30, 1-35 o aproximadamente 1-40 aminoácidos. En algunas formas de realización, las deleciones pueden comprender deleciones de 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10 ó 1-20 restos de aminoácidos. En algunas formas de realización, el número de deleciones puede ser 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35 o aproximadamente 40 aminoácidos. En algunas formas de realización, las deleciones pueden comprender deleciones de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, ó 20 restos de aminoácidos.

En algunas formas de realización, la secuencia de aminoácidos de las cetorreductasas del presente documento está sujeta a una o más de las siguientes condiciones: (1) el resto correspondiente a X153 debe ser una mutación conservadora en comparación con el resto correspondiente de la SEQ ID NO: 2, 4 ó 106; (3) el resto correspondiente a X206 es una fenilalanina (F); (4) el resto correspondiente a X233 debe ser un resto no polar o alifático; y (5) las cetorreductasas de la descripción no incluyen las secuencias polipeptídicas específicas de la SEQ ID NO: 6, 50, 52 ó 110.

La presente descripción también proporciona fragmentos de las enzimas cetorreductasa manipuladas mejoradas anteriormente descritas que presentan actividad cetorreductasa en los que la secuencia de aminoácidos del fragmento comprende las siguientes características: (a) el resto correspondiente al resto X94 es una glicina, (b) el resto correspondiente al resto X145 es un aminoácido aromático o leucina; y (c) el resto correspondiente al resto X190 es una cisteína o un aminoácido restringido, en el que la posición del aminoácido viene determinada por el alineamiento óptimo con una secuencia de referencia seleccionada de entre la SEQ ID NO: 2, 4 ó 106.

Como entenderá el experto en la materia, los polipéptidos descritos en el presente documento no se limitan a los aminoácidos codificados genéticamente. Además de los aminoácidos codificados genéticamente, los polipéptidos descritos en el presente documento pueden estar compuestos, ya sea en su totalidad o en parte, por aminoácidos no codificados naturales y/o sintéticos. Determinados aminoácidos no codificados que se encuentran comúnmente de los cuales los polipéptidos descritos en el presente documento pueden estar compuestos incluyen, pero no se limitan a: los D-estereómeros de los aminoácidos codificados genéticamente; ácido 2,3-diaminopropiónico (Dpr); ácido α-aminoisobutírico (Aib); ácido ε-aminohexanoico (Aha); ácido δ-aminovalérico (Ava); N-metilglicina o sarcosina (MeGly o Sar); ornitina (Orn); citrulina (Cit); t-butilalanina (Bua); t-butilglicina (Bug); N-metilisoleucina (MeIle); fenilglicina (Phg); ciclohexilalanina (Cha); norleucina (Nle); naftilalanina (Nal); 2-clorofenilalanina (Ocf); 3-clorofenilalanina (Mcf); 4-clorofenilalanina (Pcf); 2-fluorofenilalanina (Off); 3-fluorofenilalanina (Mff); 4-fluorofenilalanina (Pff); 2-bromofenilalanina (Obf); 3-bromofenilalanina (Mbf); 4-bromofenilalanina (Pbf); 2-metilfenilalanina (Omf); 3-metilfenilalanina (Mmf); 4-metilfenilalanina (Pmf); 2-nitrofenilalanina (Onf); 3-nitrofenilalanina (Mnf); 4-nitrofenilalanina (Pnf); 2-cianofenilalanina (Ocf); 3-cianofenilalanina (Mcf); 4-cianofenilalanina (Pcf); 2-trifluorometilfenilalanina (Otf); 3-trifluorometilfenilalanina (Mtf); 4-trifluorometilfenilalanina (Ptf); 4-aminofenilalanina (Paf); 4-yodofenilalanina (Pif); 4-aminometilfenilalanina (Pamf); 2,4-diclorofenilalanina (Opef); 3,4-diclorofenilalanina (Mpcf); 2,4-difluorofenilalanina (Opff); 3,4-difluorofenilalanina (Mpff); pirid-2-ilalanina (2pAla); pirid-3-ilalanina (3pAla); pirid-4-ilalanina (4pAla); naft-1-ilalanina (1nAla); naft-2-ilalanina (2nAla); tiazolilalanina (taAla); benzotienilalanina (bAla); tienilalanina (tAla); furilalanina (fAla); homofenilalanina (hPhe); homotirosina (hTyr); homotriptófano (hTrp); pentafluorofenilalanina (5ff); estirilalanina

(sAla); autrilalanina (aAla); 3,3-difenilalanina (Dfa); ácido 3-amino-5-fenilpentanoico (Afp); penicilamina (Pen); ácido 1,2,3,4-tetrahidroisoquinolina-3-carboxílico (Tic); β-2-tienilalanina (Thi); sulfóxido de metionina (Mso); N(w)-nitroarginina (nArg); homolisina (hLys); fosfonometilfenilalanina (pmPhe); fosfoserina (pSer); fosfotreonina (pThr); ácido homoaspártico (hAsp); ácido homoglutámico (hGlu); ácido 1-aminociclopent-(2 ó 3)-eno-4-carboxílico; ácido pipecólico (PA), ácido azetidina-3-carboxílico (ACA); ácido 1-aminociclopentano-3-carboxílico; alilglicina (aOly); propargilglicina (pgGly); homoalanina (hAla); norvalina (nVal); homoleucina (hLeu), homovalina (Hval); homoisoleucina (hIle); homoarginina (hArg); N-acetil lisina (AcLys); ácido 2,4-diaminobutírico (Dbu); ácido 2,3-diaminobutírico (Dab); N-metilvalina (MeVal); homocisteína (hCys); homoserina (hSer); hidroxiprolina (Hyp) y homoprolina (hPro). Los aminoácidos no codificados adicionales de los cuales los polipéptidos descritos en el presente documento pueden estar compuestos resultarán evidentes para los expertos en la materia (véanse, por ejemplo, los diversos aminoácidos proporcionados en Fasman, 1989, CRC Practical Handbook of Biochemistry and Molecular Biology, CRC Press, Boca Raton, FL, en las pág. 3-70 y las referencias citadas en el mismo). Estos aminoácidos pueden estar en la configuración L o D.

Los expertos en la materia reconocerán que los aminoácidos o restos que portan grupos protectores de la cadena lateral pueden comprender también los polipéptidos descritos en el presente documento. Los ejemplos no limitativos de tales aminoácidos protegidos, que en este caso pertenecen a la categoría de aromáticos, incluyen (los grupos protectores que aparecen entre paréntesis), pero no se limitan a: Arg(tos), Cys(metilbencilo), Cys(nitropiridinasulfenilo), Glu(δ-benciléster), Gln(xantilo), Asn(N-δ-xantilo), His(bom), His(bencilo), His(tos), Lys(fmoc), Lys(tos), Ser(O-bencilo), Thr(O-bencilo) y Tyr(O-bencilo).

Los aminoácidos no codificantes que están conformacionalmente restringidos de los cuales los polipéptidos descritos en el presente documento puede estar compuesto incluyen, pero no se limitan a, N-metilaminoácidos (configuración L); ácido 1-aminociclopent-(2 ó 3)-ene-4-carboxílico; ácido pipecólico; ácido azetidina-3-carboxílico; homoprolina (hPro); y ácido 1-aminociclopentano-3-carboxílico.

Los restos en el dominio catalítico de *L. kefir* son S143, Y156, K160 y N114. Los restos del dominio bolsillo de unión están situados en las posiciones 94, 96, 153, 150, 144, 145, 190, 195, 196, 199, 202, 206, 211 y 249 en la KRED de *L. kefir*. Q252 se encuentra en el dominio de unión a $Mg2^{+}$. Los restos en el dominio de unión a NADP son los restos 14-20, 37-40, 62-64, 90-93, 113, 141, 188-191, 193, y 195. Los análisis de secuencia-actividad indicaron que las sustituciones específicas descritas en el presente documento en las posiciones 94, 145 y 190 eran importantes con respecto a la capacidad de los polipéptidos cetorreductasa manipulados descritos en el presente documento para reducir estereoselectivamente el sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina a (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-propanamina con un alto porcentaje de exceso estereomérico. Estas posiciones interactúan directamente con el sustrato en el bolsillo de unión de la cetorreductasa.

Polinucleótidos que codifican las cetorreductasas manipuladas

En otro aspecto, la presente descripción proporciona polinucleótidos que codifican las enzimas cetorreductasa manipuladas descritas en el presente documento. Los polinucleótidos pueden estar unidos operativamente a una o más secuencias reguladoras heterólogas que controlan la expresión génica para crear un polinucleótido recombinante capaz de expresar el polipéptido. Pueden introducirse constructos de expresión que contienen un polinucleótido heterólogo que codifica la cetorreductasa manipulada en células hospedadoras apropiadas que expresen el polipéptido cetorreductasa correspondiente.

Debido a los conocimientos de los codones correspondientes a los diversos aminoácidos, la disponibilidad de una secuencia proteica proporciona una descripción de todos los polinucleótidos capaces de codificar el polipéptido de la invención. La degeneración del código genético, en la que los mismos aminoácidos son codificados por codones alternativos o sinónimos, permite generar un número extremadamente grande de ácidos nucleicos, todos los cuales codifican las enzimas cetorreductasa mejoradas descritas en el presente documento. Por lo tanto, después de haber identificado una secuencia de aminoácidos particular, los expertos en la materia podrían generar multitud de diferentes ácidos nucleicos simplemente modificando la secuencia de uno o más codones de manera que no cambie la secuencia de aminoácidos de la proteína. En este sentido, la presente descripción contempla específicamente cada una de, y todas, las posibles variaciones de polinucleótidos que podrían generarse seleccionando combinaciones basadas en las posibles elecciones de codones, y todas estas variaciones deben considerarse descritas específicamente para cualquier polipéptido descrito en el presente documento, incluidas las secuencia de aminoácidos que se presentan en la Tabla 2. En diversas formas de realización, los codones se seleccionan preferentemente para que se adapten a la célula hospedadora en la que se produce la proteína. Por ejemplo, se utilizan los codones preferentes utilizados en bacterias para expresar el gen en bacterias; se utilizan los codones preferentes utilizados en levadura para la expresión en levaduras; y se utilizan los codones preferentes utilizados en los mamíferos para la expresión en células de mamífero. A modo de ejemplo, el polinucleótido de la SEQ ID NO: 1 se ha sometido a optimización de codones para la expresión en *E. coli*, pero por lo demás codifica la cetorreductasa natural de *Lactobacillus kefir*.

En algunas formas de realización, el polinucleótido comprende una secuencia nucleotídica que codifica un polipéptido cetorreductasa con una secuencia de aminoácidos que tiene una identidad de secuencia de al menos

aproximadamente un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ó 99%, o más, con cualquiera de los polipéptidos cetorreductasa manipulados de referencia descritos en el presente documento, en las que el polipéptido cetorreductasa codificado comprende una secuencia de aminoácidos que tiene al menos las siguientes características: (1) el resto de aminoácido correspondiente a la posición de resto 94 de la SEQ ID NO: 2, 4 ó 106 es un aminoácido no polar, (2) el resto de aminoácido correspondiente a la posición de resto 145 de la SEQ ID NO: 2, 4 ó 106 es un aminoácido aromático o leucina, y (3) el resto de aminoácido correspondiente a la posición de resto 190 de la SEQ ID NO: 2, 4 ó 106 es un aminoácido restringido. En algunas formas de realización, los polinucleótidos con la identidad de secuencia anteriormente especificada codifican polipéptidos cetorreductasa que tienen al menos las siguientes características: (1) el resto de aminoácido correspondiente a la posición 94 es leucina, isoleucina, alanina, valina, metionina o glicina, (2) el resto de aminoácido correspondiente a la posición 145 es tirosina, triptófano, fenilalanina o leucina, y (3) el resto de aminoácido correspondiente a la posición 190 es cisteína o prolina. En algunas formas de realización, los polinucleótidos con la identidad de secuencia anteriormente especificada codifican polipéptidos cetorreductasa que tienen al menos las siguientes características: (1) el resto de aminoácido correspondiente a la posición 94 es glicina, (2) el resto de aminoácido correspondiente a la posición 145 es fenilalanina, tirosina o leucina, particularmente fenilalanina o leucina, y (3) el aminoácido correspondiente a la posición 190 es cisteína o prolina, particularmente prolina. En algunas formas de realización, los polinucleótidos codifican un polipéptido cetorreductasa manipulado que comprende una secuencia de aminoácidos seleccionada de entre las SEQ ID NO: 6, 10, 12, 14, 16, 18, 29, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76 y 78.

En algunas formas de realización, los polinucleótidos que codifican las cetorreductasas manipuladas están seleccionados de entre la SEQ ID NO: 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75 y 77. En algunas formas de realización, los polinucleótidos son capaces de hibridar en condiciones muy rigurosas con un polinucleótido que comprende la SEQ ID NO: 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75 y 77, en las que el polinucleótido que hibrida en condiciones muy rigurosas codifica una cetorreductasa funcional capaz de convertir el sustrato de fórmula estructural (I) en el producto de fórmula estructural (II), incluido, por ejemplo, el sustrato con la fórmula estructural (III) en el producto con la fórmula estructural (IV).

En algunas formas de realización, los polinucleótidos comprenden polinucleótidos que codifican los polipéptidos descritos en el presente documento, pero que tienen una identidad de secuencia de aproximadamente un 80% o más, una identidad de secuencia de aproximadamente un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ó 99% o más, a nivel de nucleótidos con un polinucleótido de referencia que codifica una cetorreductasa manipulada. En algunas formas de realización, el polinucleótido de referencia está seleccionado de entre las secuencias polinucleotídicas representadas por la SEQ ID NO: 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75 y 77.

En algunas formas de realización, los polinucleótidos descritos en el presente documento están sujetos a la condición de que no incluyan las secuencias nucleotídicas de la SEQ ID NO: 5, 49 ni 51.

Un polinucleótido aislado que codifica un polipéptido cetorreductasa mejorado puede manipularse de diversas maneras para proporcionar la expresión del polipéptido. Puede ser deseable o necesario manipular el polinucleótido aislado antes de insertarlo en un vector, dependiendo del vector de expresión. Las técnicas para modificar los polinucleótidos y las secuencias de ácidos nucleicos utilizando métodos de ADN recombinante son bien conocidas en la técnica. Se proporciona orientación en Sambrook *et al.,* 2001, Molecular Cloning: A Laboratory Manual, 3ª ed., Cold Spring Harbor Laboratory Press; y Current Protocols in Molecular Biology, Ausubel. F. ed., Greene Pub. Associates, 1998, actualizaciones hasta 2006.

Para las células hospedadoras bacterianas, los promotores adecuados para dirigir la transcripción de los constructos de ácidos nucleicos de la presente descripción incluyen los promotores obtenidos del operón lac de *E. coli*, el gen de la agarasa de *Streptomyces coelicolor* (dagA), el gen de la levansucrasa de *Bacillus subtilis* (sacB), el gen de la alfa-amilasa de *Bacillus licheniformis* (amyL), el gen de la amilasa maltogénica de *Bacillus stearotermophilus* (amyM), el gen de la alfa-amilasa de *Bacillus amyloliquefaciens* (amyQ), el gen de la penicilinasa de *Bacillus licheniformis* (penP), los genes xylA y xylB de *Bacillus subtilis*, y el gen de la β-lactamasa procariota (véase, Villa-Kamaroff *et al.,* 1978, Proc. Natl. Acad. Sci. EE.UU. 75:3727-3731), así como el promotor *tac* (véase, DeBoer *et al.,* 1983, Proc. Natl. Acad. Sci. EE.UU. 80:21-25). Se describen promotores adicionales en Sambrook *et al., supra*.

Para las células de hongos filamentosos hospedadoras, los promotores adecuados para dirigir la transcripción de los constructos de ácidos nucleicos de la presente descripción incluyen promotores obtenidos de los genes para la TAKA-amilasa de *Aspergillus oryzae*, la proteinasa aspártica de *Rhizomucor miehei*, la alfa-amilasa neutra de *Aspergillus niger*, la alfa-amilasa estable frente a ácidos de *Aspergillus niger*, la glucoamilasa de *Aspergillus niger* o *Aspergillus awamori* (glaA), la lipasa de *Rhizomucor miehei*, la proteasa alcalina de *Aspergillus oryzae*, la triosa-fosfato isomerasa de *Aspergillus oryzae*, la acetamidasa de *Aspergillus nidulans*, y la proteasa tipo tripsina de *Fusarium oxysporum* (WO 96/00787), así como el promotor NA2-tpi (un híbrido de los promotores de los

genes para la alfa-amilasa neutra de *Aspergillus niger* y la triosa-fosfato isomerasa de *Aspergillus oryzae*), y promotores mutantes, truncados e híbridos de los mismos.

En un hospedador levadura, los promotores útiles pueden ser de los genes de la enolasa de *Saccharomyces cerevisiae* (ENO-1), la galactoquinasa de *Saccharomyces cerevisiae* (GAL1), la alcohol deshidrogenasa/gliceraldehido-3-fosfato deshidrogenasa de *Saccharomyces cerevisiae* (ADH2/GAP), y la 3-fosfoglicerato quinasa de *Saccharomyces cervisiae*. Se describen otros promotores útiles para células de levadura hospedadoras en Romanos *et al.,* 1992, Yeast 8:423-488.

La secuencia de control también puede ser una secuencia terminadora de la transcripción adecuada, una secuencia reconocida por una célula hospedadora para terminar la transcripción. La secuencia terminadora está unida operativamente al extremo 3' terminal de la secuencia de ácido nucleico que codifica el polipéptido. Puede utilizarse en la presente invención cualquier terminador que sea funcional en la célula hospedadora de elección.

Por ejemplo, pueden obtenerse terminadores de la transcripción ejemplares para células de hongos filamentosos hospedadoras de los genes de la TAKA-amilasa de *Aspergillus oryzae*, la glucoamilasa de *Aspergillus niger*, la antranilato sintasa de *Aspergillus nidulans*, la alfa-glucosidasa de *Aspergillus niger*, y la proteasa tipo tripsina de *Fusarium oxysporum*.

Pueden obtenerse terminadores ejemplares para células de levadura hospedadoras de los genes de la enolasa de *Saccharomyces cerevisiae*, la citocromo C de *Saccharomyces cerevisiae* (CYC1), y la gliceraldehído-3-fosfato deshidrogenasa de *Saccharomyces cerevisiae*. Se describen otros terminadores útiles para células de levadura hospedadoras en Romanos *et al.,* 1992, *supra*.

La secuencia de control también puede ser una secuencia líder adecuada, una región no traducida de un ARNm que sea importante para la traducción por la célula hospedadora. La secuencia líder está unida operativamente al extremo 5' terminal de la secuencia de ácido nucleico que codifica el polipéptido. Puede utilizarse cualquier secuencia líder que sea funcional en la célula hospedadora de elección. Se obtienen líderes ejemplares para células de hongos filamentosos hospedadoras de los genes de la TAKA-amilasa de *Aspergillus oryzae* y la triosa-fosfato isomerasa de *Aspergillus nidulans*. Se obtienen líderes adecuadas para células de levadura hospedadoras de los genes de la enolasa de *Saccharomyces cerevisiae* (ENO-1), la 3-fosfoglicerato quinasa de *Saccharomyces cerevisiae*, el factor alfa de *Saccharomyces cerevisiae*, y la alcohol deshidrogenasa/gliceraldehído-3-fosfato deshidrogenasa de *Saccharomyces cerevisiae* (ADH2/GAP).

La secuencia de control también puede ser una secuencia de poliadenilación, una secuencia unida operativamente al extremo 3' terminal de la secuencia de ácido nucleico y que, cuando se transcribe, es reconocida por la célula hospedadora como una señal para añadir restos de poliadenosina al ARNm transcrito. Puede utilizarse en la presente invención cualquier secuencia de poliadenilación que sea funcional en la célula hospedadora de elección. Las secuencias de poliadenilación ejemplares para células de hongos filamentosos hospedadoras pueden ser de los genes de la TAKA-amilasa de *Aspergillus oryzae*, la glucoamilasa de *Aspergillus niger*, la antranilato sintasa de *Aspergillus nidulans*, la proteasa tipo tripsina de *Fusarium oxysporum*, y la alfa-glucosidasa de *Aspergillus niger*. Se describen secuencias de poliadenilación útiles para células de levadura hospedadoras en Guo y Sherman, 1.995, Mol Cell Bio 15:5983-5990.

La secuencia de control también puede ser una región codificante del péptido señal que codifica una secuencia de aminoácidos unida al extremo amino terminal de un polipéptido y dirige el polipéptido codificado a la vía secretora de la célula. El extremo 5' de la secuencia codificante de la secuencia de ácido nucleico puede contener intrínsecamente una región codificante del péptido señal naturalmente unida en el marco de lectura de la traducción con el segmento de la región codificante que codifica el polipéptido secretado. Como alternativa, el extremo 5' de la secuencia codificante puede contener una región codificante del péptido señal que sea extraña a la secuencia codificante. La región codificante del péptido señal extraña puede ser necesaria cuando la secuencia codificante no contenga naturalmente una región codificante del péptido señal.

Como alternativa, la región codificante del péptido señal extraña puede simplemente reemplazar la región codificante del péptido señal natural con el fin de potenciar la secreción del polipéptido. Sin embargo, puede utilizarse en la presente invención cualquier región codificante del péptido señal que dirija el polipéptido expresado a la vía secretora de una célula hospedadora de elección.

Las regiones codificantes del péptido señal eficaces para células hospedadoras bacterianas son las regiones codificantes del péptido señal obtenidas de los genes de la amilasa maltogénica de *Bacillus* NC1B 11837, la alfa-amilasa de *Bacillus stearotermophilus*, la subtilisina de *Bacillus licheniformis*, la beta-lactamasa de *Bacillus licheniformis*, las proteasas neutras de *Bacillus stearotermophilus* (nprT, nprS, nprM), y prsA de *Bacillus subtilis*. Se describen péptidos señal adicionales en Simonen y Palva, 1993, Microbiol. Rev. 57:109-137.

Las regiones codificantes del péptido señal eficaces para las células de hongos filamentosos hospedadoras pueden ser las regiones codificantes del péptido señal obtenidas de los genes de la TAKA-amilasa de *Aspergillus*

oryzae, la amilasa neutra de *Aspergillus niger*, la glucoamilasa de *Aspergillus niger*, la proteinasa aspártica de *Rhizomucor miehei*, la celulasa de *Humicola insolens*, y la lipasa de *Humicola lanuginosa*.

Los péptidos señal útiles para células de levadura hospedadoras pueden ser de los genes del factor alfa de *Saccharomyces cerevisiae* y la invertasa de *Saccharomyces cerevisiae*. Se describen otras regiones codificantes del péptido señal útiles en Romanos *et al.,* 1992, *supra*.

La secuencia de control también puede ser una región codificante del propéptido que codifica una secuencia de aminoácidos situada en el extremo amino terminal de un polipéptido. El polipéptido resultante se conoce como proenzima o propolipéptido (o zimógeno en algunos casos). Un propolipéptido es generalmente inactivo y puede convertirse en un polipéptido activo maduro por escisión catalítica o autocatalítica del propéptido a partir del propolipéptido. La región codificante del propéptido puede obtenerse de los genes de la proteasa alcalina de *Bacillus subtilis* (aprE), la proteasa neutra de *Bacillus subtilis* (nprT), el factor alfa de *Saccharomyces cerevisiae*, la proteinasa aspártica de *Rhizomucor miehei*, y la lactasa de *Myceliophthora termophila* (véase, el documento WO 95/33836).

Cuando las regiones tanto del péptido señal como del propéptido están presentes en el extremo amino terminal de un polipéptido, la región del propéptido se sitúa junto al extremo amino terminal de un polipéptido y la región del péptido señal se sitúa junto al extremo amino terminal de la región del propéptido.

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

Otros ejemplos de secuencias reguladoras son aquellas que permiten la amplificación génica. En los sistemas eucariotas, estos incluyen el gen de la dihidrofolato reductasa, que se amplifica en presencia de metotrexato, y los genes de la metalotioneína, que se amplifican con metales pesados. En estos casos, la secuencia de ácido nucleico que codifica el polipéptido KRED de la presente invención estaría unida operativamente con la secuencia reguladora.

Por lo tanto, en otra forma de realización, la presente descripción también se refiere a un vector de expresión recombinante que comprende un polinucleótido que codifica un polipéptido cetorreductasa manipulado o una variante del mismo, y una o más regiones reguladoras de la expresión tales como un promotor y un terminador, un origen de replicación, etc., dependiendo del tipo de hospedadores en los que se vaya a introducir. Las diversas secuencias de ácidos nucleicos y de control descritas anteriormente pueden juntarse para producir un vector de expresión recombinante que puede incluir uno o más sitios de restricción convenientes para permitir la inserción o sustitución de la secuencia de ácido nucleico que codifica el polipéptido en tales sitios. Como alternativa, la secuencia de ácido nucleico de la presente descripción puede expresarse insertando la secuencia de ácido nucleico o un constructo de ácido nucleico que comprenda la secuencia en un vector apropiado para la expresión. Al crear el vector de expresión, la secuencia codificante se sitúa en el vector de manera que la secuencia codificante quede unida operativamente con las secuencias de control apropiadas para la expresión.

El vector de expresión recombinante puede ser cualquier vector (por ejemplo, un plásmido o virus), que pueda someterse oportunamente a procedimientos de ADN recombinante y pueda producir la expresión de la secuencia polinucleotídica. La elección del vector dependerá por lo general de la compatibilidad del vector con la célula hospedadora en la que va a introducirse el vector. Los vectores pueden ser plásmidos circulares lineales o cerrados.

El vector de expresión puede ser un vector de replicación autónoma, es decir, un vector que existe como una entidad extracromosómica, cuya replicación es independiente de la replicación cromosómica, por ejemplo, un plásmido, un elemento extracromosómico, un minicromosoma o un cromosoma artificial. El vector puede contener cualquier medio para asegurar la autorreplicación. Como alternativa, el vector puede ser uno que, cuando se introduce en la célula hospedadora, se integra en el genoma y se replica junto con el cromosoma o cromosomas en los que se ha integrado. Además, pueden utilizarse un único vector o plásmido o dos o más vectores o plásmidos que juntos contengan el ADN total a introducir en el genoma de la célula hospedadora, o un transposón.

El vector de expresión de la presente invención contiene preferentemente uno o más marcadores seleccionables, que permiten una selección fácil de las células transformadas. Un marcador seleccionable es un gen cuyo producto proporciona resistencia biocida o vírica, resistencia a metales pesados, prototrofía a auxótrofos, y similares. Los ejemplos de marcadores seleccionables bacterianos son los genes *dal* de *Bacillus subtilis* o *Bacillus licheniformis*, o marcadores, que confieren resistencia a antibióticos tal como resistencia a la ampicilina, la

kanamicina, el cloranfenicol (Ejemplo 1) o la tetraciclina. Los marcadores adecuados para células de levadura hospedadoras son ADE2, HIS3, LEU2, LYS2, MET3, TRP1 y URA3.

Los marcadores seleccionables para su uso en una célula de hongo filamentoso hospedadora incluyen, pero no se limitan a, amdS (acetamidasa), argB (ornitina carbamoiltransferasa), bar (fosfinotricina acetiltransferasa), hph (higromicina fosfotransferasa), niaD (nitrato reductasa), pyrG (orotidina-5'-fosfato descarboxilasa), sC (sulfato adeniltransferasa) y trpC (antranilato sintasa), así como equivalentes de las mismas. Las formas de realización para el uso en una célula de *Aspergillus* incluyen los genes amdS y pyrG de *Aspergillus nidulans* o *Aspergillus oryzae* y el gen *bar* de *Streptomyces hygroscopicus*.

Los vectores de expresión de la presente invención contienen preferentemente un elemento o elementos que permiten la integración del vector en el genoma de la célula hospedadora o la replicación autónoma del vector en la célula con independencia del genoma. Para la integración en el genoma de la célula hospedadora, el vector puede depender de la secuencia de ácido nucleico que codifica el polipéptido o cualquier otro elemento del vector para la integración del vector en el genoma mediante recombinación homóloga o no homóloga.

Como alternativa, el vector de expresión puede contener secuencias de ácidos nucleicos adicionales para dirigir la integración mediante recombinación homóloga en el genoma de la célula hospedadora. Las secuencias de ácidos nucleicos adicionales permiten que el vector se integre en el genoma de la célula hospedadora en una ubicación o ubicaciones precisas en el cromosoma o cromosomas. Para aumentar la probabilidad de integración en una ubicación precisa, los elementos de integración deberían contener preferentemente un número suficiente de ácidos nucleicos, tal como de 100 a 10.000 pares de bases, preferentemente de 400 a 10.000 pares de bases, y lo más preferentemente de 800 a 10.000 pares de bases, que presentan una gran homología con la secuencia diana correspondiente para potenciar la probabilidad de recombinación homóloga. Los elementos de integración pueden ser cualquier secuencia que presente homología con la secuencia diana en el genoma de la célula hospedadora. Además, los elementos de integración pueden ser secuencias de ácidos nucleicos codificantes o no codificantes. Por otro lado, el vector puede integrarse en el genoma de la célula hospedadora mediante recombinación no homóloga.

Para la replicación autónoma, el vector puede comprender adicionalmente un origen de replicación que permita al vector replicarse de manera autónoma en la célula hospedadora en cuestión. Los ejemplos de orígenes de replicación bacterianos son P15A ori (como se muestra en el plásmido de la FIG. 5) o los orígenes de replicación de los plásmidos pBR322, pUC19, pACYC177 (cuyo plásmido tienen el P15A ori), o pACYC184 que permite la replicación en *E. coli*, y pUB110, pE194, pTA1060 o pAMβ1 que permiten la replicación en *Bacillus*. Los ejemplos de orígenes de replicación para su uso en una célula de levadura hospedadora son el origen de replicación de 2 micrómetros, ARS1, ARS4, la combinación de ARS1 y CEN3, y la combinación de ARS4 y CEN6. El origen de replicación puede ser uno que tenga una mutación que haga su funcionamiento sensible a la temperatura en la célula hospedadora (véase, por ejemplo, Ehrlich, 1978, Proc. Natl. Acad. Sci. EE.UU. 75:1433).

Puede insertarse en la célula hospedadora más de una copia de una secuencia de ácido nucleico de la presente invención para aumentar la producción del producto del gen. Puede obtenerse un aumento del número de copias de la secuencia de ácido nucleico integrando al menos una copia adicional de la secuencia en el genoma de la célula hospedadora o incluyendo un gen marcador seleccionable amplificable con la secuencia de ácido nucleico en el que pueden seleccionarse las células que contienen copias amplificadas del gen marcador seleccionable, y de ese modo copias adicionales de la secuencia de ácido nucleico, cultivando las células en presencia del agente seleccionable apropiado.

Muchos de los vectores de expresión para su uso en la presente invención están disponibles en el mercado. Los vectores de expresión comerciales adecuados incluyen los vectores de expresión p3xFLAGTM™ de Sigma-Aldrich Chemicals, St. Louis MO, que incluye un promotor de CMV y el sitio de poliadenilación de hGH para la expresión en células hospedadoras de mamífero y un origen de replicación de pBR322 y marcadores de resistencia a la ampicilina para la amplificación en *E. coli*. Otros vectores de expresión adecuados son pBlue-scriptII SK(-) y pBK-CMV, que están disponibles en el mercado en Stratagene, La Jolla CA, y plásmidos que se derivan de pBR322 (Gibco BRL), pUC (Gibco BRL), pREP4, pCEP4 (Invitrogen) o pPoly (véase Lathe *et al.,* 1987, Gene 57:193-201).

<u>Células hospedadoras para la expresión de los polipéptidos cetorreductasa</u>

En otro aspecto, la presente descripción proporciona una célula hospedadora que comprende un polinucleótido que codifica un polipéptido cetorreductasa mejorado de la presente descripción, estando el polinucleótido unido operativamente a una o más secuencias de control para la expresión de la enzima cetorreductasa en la célula hospedadora. Las células hospedadoras para su uso en la expresión de los polipéptidos KRED codificados por los vectores de expresión de la presente invención son bien conocidas en la técnica e incluyen pero no se limitan a, células bacterianas, tales como células de *E. coli*, *Lactobacillus kefir*, *Lactobacillus brevis*, *Lactobacillus minor*, *Streptomyces* y *Salmonella typhimurium*; células fúngicas, tales como células de levadura (por ejemplo, *Saccharomyces cerevisiae* o *Pichia pastoris* (nº de registro de la ATCC 201178)); células de insecto tales como células de *Drosophila* S2 y *Spodoptera* Sf9; células animales tales como CHO, COS, BHK, 293 y

células de melanoma de Bowes; y células vegetales. Los medios de cultivo y las condiciones de crecimiento apropiadas para las células hospedadoras descritas anteriormente son bien conocidos en la técnica.

Los polinucleótidos para la expresión de la cetorreductasa pueden introducirse en las células mediante diversos métodos conocidos en la técnica. Las técnicas incluyen, entre otras, electroporación, bombardeo de partículas de biobalística, transfección mediada por liposomas, transfección con cloruro cálcico, y fusión de protoplastos. Para el experto en la materia resultarán evidentes diversos métodos para introducir polinucleótidos en células.

Como se describe en el Ejemplo 1, una célula hospedadora ejemplar es *Escherichia coli* W3110. En el Ejemplo 1, el vector de expresión se creó uniendo operativamente un polinucleótido que codifica una cetorreductasa mejorada en el plásmido pCK110900 unido operativamente al promotor *lac* bajo el control del represor *lac*. El vector de expresión también contenía el origen de replicación p15A y el gen de resistencia a cloranfenicol. Las células que contenían el polinucleótido de la invención en *Escherichia coli* W3110 se aislaron sometiendo las células a selección con cloranfenicol.

Métodos de generación de los polipéptidos cetorreductasa manipulados

En algunas formas de realización, para generar los polipéptidos y polinucleótidos KRED mejorados de la presente descripción, la enzima cetorreductasa natural que cataliza la reacción de reducción se obtiene (o deriva) de *Lactobacillus kefir* o *Lactobacillus brevis* o *Lactobacillus minor*. En algunas formas de realización, la secuencia polinucleotídica parental se somete a optimización de codones para potenciar la expresión de la cetorreductasa en una célula hospedadora específica. Como ilustración, la secuencia polinucleotídica parental que codifica el polipéptido KRED de tipo silvestre de *Lactobacillus kefir* se construyó a partir de oligonucleótidos preparados basándose en la secuencia polipeptídica conocida de la secuencia de KRED de *Lactobacillus kefir* disponible en la base de datos del GenBank (nº de registro del GenBank AAP94029 GI:33112056). La secuencia polinucleotídica parental, designada SEQ ID NO: 1, se sometió a optimización de codones para la expresión en *E. coli* y el polinucleótido con codones optimizados se clonó en un vector de expresión, poniendo la expresión del gen de cetorreductasa bajo el control del promotor lac y el gen represor lacI. Se identificaron los clones que expresaban la cetorreductasa activa en *E. coli* y se secuenciaron los genes para confirmar su identidad. La secuencia, designada SEQ ID NO: 1, fue la secuencia parental utilizada como punto de partida para la mayoría de los experimentos y la construcción de la biblioteca de cetorreductasas manipuladas evolucionadas a partir de la cetorreductasa de *Lactobacillus kefir*.

Las cetorreductasas manipuladas pueden obtenerse sometiendo el polinucleótido que codifica la cetorreductasa natural a mutagénesis y/o a métodos de evolución dirigida, como se ha analizado anteriormente. Una técnica de evolución dirigida ejemplar es la mutagénesis y/o el barajado de ADN como se describe en Stemmer, 1994, Proc. Natl. Acad. Sci EE.UU. 91:10747-10751; los documentos WO 95/22625, WO 97/0078, WO 97/35966, WO 98/27230, WO 00/42651, WO 01/75767 y la patente de EE.UU. 6.537.746. Otros procedimientos de evolución dirigida que pueden utilizarse incluyen, entre otros, el proceso de extensión escalonada (StEP), la recombinación *in vitro* (Zhao *et al.,* 1998, Nat. Biotechnol. 16:258-261), PCR mutagénica (Caldwell *et al.,* 1994, PCR Methods Appl. 3:S136-S140), y mutagénesis con casete (Black *et al.,* 1996, Proc. Natl. Acad. Sci EE.UU. 93:3525-3529).

Los clones obtenidos tras el tratamiento de mutagénesis se criban para seleccionar las cetorreductasas manipuladas que tengan una propiedad de la enzima mejora deseada. La medición de la actividad enzimática de las bibliotecas de expresión puede realizarse utilizando la técnica bioquímica convencional de seguimiento de la velocidad de disminución (por una disminución de la absorbancia o fluorescencia) de la concentración de NADH o NADPH, a medida que se convierte en NAD^+ o $NADP^+$. En esta reacción, el NADH o el NADPH son consumidos (oxidados) por la cetorreductasa a medida que la cetorreductasa reduce un sustrato cetona al grupo hidroxilo correspondiente. La velocidad de disminución de la concentración de NADH o NADPH, medida por la disminución de la absorbancia o fluorescencia, por unidad de tiempo indica la actividad (enzimática) relativa del polipéptido KRED en una cantidad fija del lisado (o un polvo liofilizado hecho del mismo). Cuando la propiedad de la enzima mejorada deseada es la estabilidad térmica, la actividad enzimática puede medirse después de someter las preparaciones enzimáticas a una temperatura definida y midiendo la cantidad de actividad enzimática que queda después de los tratamientos térmicos. A continuación, se aíslan los clones que contienen un polinucleótido que codifica una cetorreductasa, se secuencian para identificar los cambios de secuencias nucleotídicas (de haberlos), y se utilizan para expresar la enzima en una célula hospedadora.

Cuando se conoce la secuencia del polipéptido manipulado, los polinucleótidos que codifican la enzima pueden prepararse mediante métodos en fase sólida convencionales, según métodos de síntesis conocidos. En algunas formas de realización, pueden sintetizarse individualmente fragmentos de hasta aproximadamente 100 bases, a continuación unirse (por ejemplo, mediante métodos de ligación enzimática o química, o métodos dependientes de polimerasa) para formar cualquier secuencia continua deseada. Por ejemplo, los polinucleótidos y oligonucleótidos de la invención pueden prepararse mediante síntesis química utilizando, por ejemplo, el método de la fosforamidita clásico descrito por Beaucage *et al.,* 1981, Tet. Lett. 22: 1859-69, o el método descrito por Matthes *et al.,* 1984, EMBO J. 3:801-05, por ejemplo, como se practica por lo general en métodos de síntesis automatizados.

Según el método de la fosforamidita, los oligonucleótidos se sintetizan, por ejemplo, en un sintetizador de ADN automático, se purifican, se hibridan, se ligan y se clonan en vectores apropiados. Además, puede obtenerse esencialmente cualquier ácido nucleico a partir de cualquiera de diversas fuentes comerciales, tales como The Midland Certified Reagent Company, Midland, TX, The Great American Gene Company, Ramona, CA, ExpressGen Inc. Chicago, IL, Operon Technologies Inc, Alameda, CA, y muchas otras.

Las enzimas cetorreductasa manipuladas expresadas en una célula hospedadora pueden recuperarse de las células y/o del medio de cultivo utilizando una cualquiera o más de las técnicas bien conocidas para la purificación de proteínas, incluidas, entre otras, tratamiento con lisozima, sonicación, filtración, precipitación por adición de sal, ultracentrifugación y cromatografía. Se dispone en el mercado de soluciones adecuadas para la lisis y la extracción de alta eficiencia de proteínas a partir de bacterias, tales como *E. coli*, con el nombre comercial CelLytic B™ de Sigma-Aldrich de St. Louis MO.

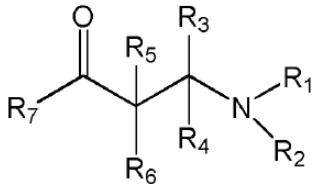
Las técnicas cromatográficas para el aislamiento del polipéptido cetorreductasa incluyen, entre otras, cromatografía en fase inversa, cromatografía de líquidos de alto rendimiento, cromatografía de intercambio iónico, electroforesis en gel, y cromatografía de afinidad. Las condiciones para purificar una enzima particular dependerán, en parte, de factores tales como la carga neta, la hidrofobicidad, la hidrofilicidad, el peso molecular, la forma molecular, etc., y resultarán evidentes para los expertos en la materia.

En algunas formas de realización, pueden utilizarse técnicas de afinidad para aislar las enzimas cetorreductasa mejoradas. Para la purificación por cromatografía de afinidad, puede utilizarse cualquier anticuerpo que se una específicamente al polipéptido cetorreductasa. Para la producción de anticuerpos, puede inmunizarse a diversos animales hospedadores, incluidos pero no limitados a conejos, ratones, ratas, etc., inyectándoles un polipéptido. El polipéptido puede estar fijado a un vehículo adecuado, tal como BSA, por medio de un grupo funcional de cadena lateral o conectores fijados a un grupo funcional de cadena lateral. Pueden utilizarse diversos adyuvantes para aumentar la respuesta inmunitaria, dependiendo de la especie hospedadora, incluidos pero no limitados a adyuvante (completo e incompleto) de Freund, geles minerales tales como hidróxido de aluminio, sustancias tensioactivas tales como lisolecitina, polioles plurónicos, polianiones, péptidos, emulsiones de aceite, hemocianina de lapa californiana, dinitrofenol, y adyuvantes para uso humano potencialmente útiles tales como BCG (bacilo de Calmette Guerin) y *Corynebacterium parvum*.

Métodos de uso de las enzimas cetorreductasa manipuladas y compuestos preparados con las mismas

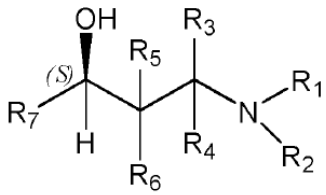
En algunas formas de realización, la invención proporciona un método para producir una *(S)*-3-aril-3-hidroxipropanamina, en las que el método comprende:

(a) proporcionar un sustrato 3-aril-3-cetopropanamina que tiene la estructura de fórmula (I):



(I);

(b) poner en contacto o incubar el sustrato 3-aril-3-cetopropanamina con uno (o más) de los polipéptidos cetorreductasa descritos en el presente documento en condiciones de reacción adecuadas para la reducción o conversión del sustrato en el producto *(S)*-3-aril-3-hidroxipropanamina que tiene la fórmula estructural (II):



(II)

en las que para (I) y (II), cada uno de R_1 y R_2 está seleccionado independientemente del grupo que consiste en hidrógeno, un alquilo inferior opcionalmente sustituido, un cicloalquilo opcionalmente sustituido, un arilo opcionalmente sustituido, o como alternativa, en las que R_1 y R_2 juntos forman un cicloalquilo opcionalmente sustituido o un cicloarilo opcionalmente sustituido que tiene 3-7 átomos de carbono; cada uno de R_3, R_4, R_5 y R_6

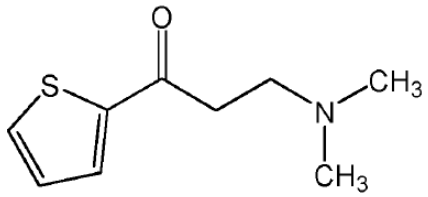
está seleccionado independientemente del grupo que consiste en hidrógeno y un alquilo inferior opcionalmente sustituido; y R_7 es un arilo opcionalmente sustituido.

La *(S)*-3-aril-3-hidroxipropanamina resultante puede recuperarse de la etapa (b) y, opcionalmente, purificarse utilizando métodos conocidos.

Por lo general, para cada uno de (I) y (II), cada uno de R_1 y R_2 es independientemente hidrógeno, un alquilo inferior de uno a diez o de uno a seis átomos de carbono, o un fenilo. Por lo general, uno de R_1 y R_2 es hidrógeno y el otro es metilo o ambos R_1 y R_2 son metilo. Cada uno de R_3, R_4, R_5 y R_6 es por lo general hidrógeno o un alquilo inferior de uno a seis átomos de carbono, y más generalmente cada uno de R_3, R_4, R_5 y R_6 es hidrógeno. R_7 es por lo general un tiofenilo opcionalmente sustituido (tal como, por ejemplo, un 2-tienilo opcionalmente sustituido o un 3-tienilo opcionalmente sustituido) o un fenilo opcionalmente sustituido. Más generalmente, R_7 es 2-tienilo o fenilo. Por lo general, R_7 es 2-tienilo. En determinados sustratos/productos ejemplares, R_1 es hidrógeno, R_2 es metilo, cada uno de R_3-R_6 es hidrógeno, R_7 está seleccionado del grupo que consiste en 2-tienilo y fenilo; y R_1 y R_2 son metilo, cada uno de R_3-R_6 es hidrógeno, R_7 está seleccionado del grupo que consiste en 2-tienilo y fenilo.

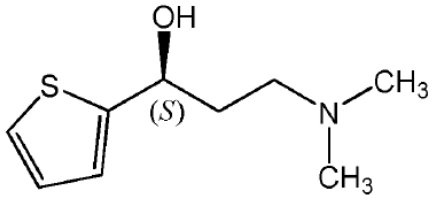
En algunas formas de realización, puede utilizarse cualquiera de los polipéptidos cetorreductasa proporcionados en el presente documento en la producción de productos intermedios para la síntesis de duloxetina (es decir, (3S)-*N*-metil-3-naftalen-1-iloxi-3-tiofen-2-il-propan-1-amina), un fármaco para el tratamiento de la depresión. En determinados enfoques para la síntesis de duloxetina, una etapa importante es la conversión de determinados compuestos de fórmula (I) en los correspondientes compuestos de fórmula (II).

Más concretamente, las enzimas cetorreductasa descritas en el presente documento son particularmente útiles para catalizar la reducción del compuesto sustrato de fórmula estructural (III), *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina ("el sustrato dimetilo", es decir, con respecto a la fórmula (I), cada uno de R_1 y R_2 es metilo, cada uno de R_3, R_4, R_5 y R_6 es hidrógeno, y R_7 es 2-tienilo):



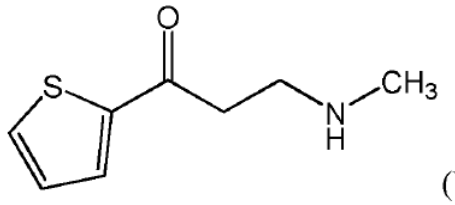
(III)

al producto alcohol estereoisómero de fórmula estructural (IV), (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina ("el producto dimetilo", es decir, con respecto a la fórmula (II), cada uno de R_1 y R_2 es metilo, cada uno de R_3, R_4, R_5 y R_6 es hidrógeno, y R_7 es 2-tienilo):



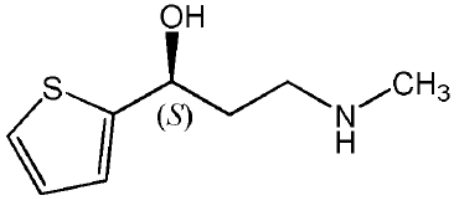
(IV)

Los polipéptidos cetorreductasa descritos en el presente documento también son útiles para catalizar la reducción del compuesto sustrato de fórmula estructural (V), *N*-metil-3-ceto-3-(2-tienil)-1-propanamina ("el sustrato monometilo", es decir, con respecto a la fórmula (I), R_1 es metilo, R_2 es hidrógeno, cada uno de R_3, R_4, R_5 y R_6 es hidrógeno, y R_7 es 2-tienilo):



(V)

al correspondiente producto alcohol estereoisómero de fórmula estructural (VI), *(S)-N*-metil-3-hidroxi-3-(2-tienil)-1-propanamina ("el producto monometilo", es decir, con respecto a la fórmula (II), R_1 es metilo, R_2 es hidrógeno, cada uno de R_3, R_4, R_5 y R_6 es hidrógeno, y R_7 es 2-tienilo):



(VI)

Ambos productos (IV) y (VI) son útiles como productos intermedios en la síntesis de duloxetina.

Por ejemplo, en algunas formas de realización del presente método para reducir el sustrato al correspondiente compuesto producto, los polipéptidos cetorreductasa tienen, en comparación con las secuencias de KRED de *L. kefir* o *L. brevis* o *L. minor* de tipo silvestre de la SEQ ID NO: 2, 4 y 106, respectivamente, al menos las siguientes características: (1) el resto correspondiente a X94 es un resto de aminoácido no polar, (2) el resto correspondiente a X145 es un resto de aminoácido aromático o leucina, y (3) el resto correspondiente a X190 es una cisteína o un resto de aminoácido restringido. En algunas formas de realización del presente método, los polipéptidos cetorreductasa tienen, en comparación con las secuencias de KRED de *L. kefir* o *L. brevis* o *L. minor* de tipo silvestre de la SEQ ID NO: 2, 4 y 106, respectivamente, al menos las siguientes características : (1) el resto correspondiente a X94 es glicina, (2) el resto correspondiente a X145 es tirosina, fenilalanina o leucina y (3) el resto correspondiente a X190 es cisteína o prolina. En algunas formas de realización del presente método, los polipéptidos de la invención tienen, en comparación con las secuencias de KRED de *L. kefir* o *L. brevis* o *L. minor* de tipo silvestre de la SEQ ID NO: 2, 4 y 106, respectivamente, al menos las siguientes características: (1) el resto correspondiente a X94 es glicina, (2) el resto correspondiente a X145 fenilalanina y (3) el resto correspondiente a X190 es prolina. Como se ha indicado en el presente documento, el polipéptido cetorreductasa puede tener, además, una o más mutaciones en otros restos de aminoácidos en comparación con las secuencias de referencia de las SEQ ID NO: 2, 4 ó 106.

En algunas formas de realización del presente método, el polipéptido cetorreductasa comprende, a modo de ejemplo y no de limitación, una secuencia de aminoácidos que es idéntica en al menos un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o más, a una secuencia de referencia basada en la SEQ ID NO: 2, 4 ó 106, a condición de que el resto X94 sea glicina, el resto X145 sea fenilalanina, y el resto X190 sea prolina, en las que los polipéptidos cetorreductasa tienen al menos las características anteriores.

En algunas formas de realización del presente método para reducir el sustrato al producto, el sustrato se reduce al producto en un exceso estereomérico superior a aproximadamente un 99%, en las que el polipéptido cetorreductasa comprende una secuencia que corresponde a la SEQ ID NO: 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76 ó 78.

En algunas formas de realización del presente método para reducir el sustrato 3-sustituido-3-cetopropanamina al correspondiente producto 3-sustituido-3-hidroxipropanamina, al menos aproximadamente un 95% del sustrato se convierte en el producto en menos de aproximadamente 24 horas cuando se lleva a cabo con más de aproximadamente 100 g/l de sustrato y menos de aproximadamente 5 g/l del polipéptido, en las que el polipéptido comprende una secuencia de aminoácidos correspondiente a la SEQ ID NO: 60, 62, 64, 66, 68, 70, 72, 74, 76 ó 78.

En algunas formas de realización del presente método, al menos aproximadamente el 10%-20% de 1 g/l de sustrato se convierte en el correspondiente producto dimetilo en menos de aproximadamente 24 horas con

aproximadamente 10 g/l del polipéptido, en las que el polipéptido comprende una secuencia de aminoácidos correspondiente a la SEQ ID NO: 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76 ó 78.

En algunas formas de realización, el método para catalizar la reducción del compuesto sustrato 3-aril-3-cetopropanamina de fórmula estructural (I), al producto alcohol estereoisómero de fórmula estructural (II) comprende poner en contacto el polipéptido cetorreductasa con el sustrato en condiciones de reacción en las que el pH es aproximadamente 11 o inferior. En algunas formas de realización, la condición de reacción es un pH de aproximadamente 8 a 11. En algunas formas de realización, la condición de reacción es un pH de aproximadamente 9 a 10. En algunas formas de realización, la condición de reacción es un pH de aproximadamente 9.

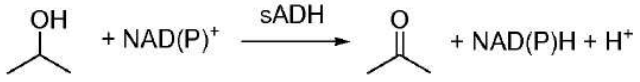
En algunas formas de realización, la reducción del compuesto sustrato 3-sustituido-3-cetopropanamina de fórmula estructural (I) al correspondiente producto alcohol de fórmula estructural (II) se lleva a cabo en presencia de alcohol isopropílico (IPA). En algunas formas de realización, el IPA está presente en ≥ 50% v/v. En algunas formas de realización, el IPA en la reacción está presente en al menos aproximadamente un 75% v/v.

En algunas formas de realización en las que el sustrato es *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina, la condición de reacción se elige para reducir o minimizar la formación de productos secundarios, tales como 1-(tiofen-2-il)prop-2-en-1-ona); 1-(tiofen-2-il)propan-1-ona; 1-(tiofen-2-il)propan-1-ol; y 1-(tiofen-2-il)prop-2-en-1-ol. En algunas formas de realización, la condición de reacción es un pH de aproximadamente 9 a aproximadamente 10 y tiene ≥ 75% v/v de IPA.

Como saben los expertos en la materia, las reacciones de reducción catalizadas por cetorreductasa requieren por lo general un cofactor. Las reacciones de reducción catalizadas por las enzimas cetorreductasa manipuladas descritas en el presente documento también requieren por lo general un cofactor, aunque muchas formas de realización de las cetorreductasas manipuladas requieren mucho menos cofactor que las reacciones catalizadas con enzimas cetorreductasa de tipo silvestre. Tal como se utiliza en el presente documento, el término "cofactor" se refiere a un compuesto no proteico que opera en combinación con una enzima cetorreductasa. Los cofactores adecuados para utilizarse con las enzimas cetorreductasa manipuladas descritas en el presente documento incluyen, pero no se limitan a, $NADP^+$ (nicotinamida adenina dinucleótido fosfato), NADPH (la forma reducida de $NADP^+$), NAD^+ (nicotinamida adenina dinucleótido) y NADH (la forma reducida de NAD^+). Generalmente, se añade a la mezcla de reacción la forma reducida del cofactor. La forma NAD(P)H reducida puede regenerarse opcionalmente a partir de la forma $NAD(P)^+$ oxidada utilizando un sistema de regeneración de cofactor.

La expresión "sistema de regeneración de cofactor" se refiere a un conjunto de reaccionantes que participan en una reacción que reduce la forma oxidada del cofactor (por ejemplo, $NADP^+$ a NADPH). Los cofactores oxidados por la reducción catalizada por cetorreductasa del sustrato ceto son regenerados en forma reducida por el sistema de regeneración de cofactor. Los sistemas de regeneración de cofactor comprenden un reductor estequiométrico que es una fuente de equivalentes de hidrógeno reductores y es capaz de reducir la forma oxidada del cofactor. El sistema de regeneración de cofactor puede comprender adicionalmente un catalizador, por ejemplo un catalizador enzimático que cataliza la reducción de la forma oxidada del cofactor por el reductor. En la técnica se conocen sistemas de regeneración de cofactor para regenerar NADH o NADPH a partir de NAD^+ o $NADP^+$, respectivamente, y pueden utilizarse en los métodos que se describen en el presente documento.

Los sistemas de regeneración de cofactor adecuados para utilizarse en la práctica de la presente invención incluyen aquellos que utilizan una alcohol secundario deshidrogenasa. La expresión "alcohol secundario deshidrogenasa" y el término "sADH" se utilizan indistintamente en el presente documento para referirse a una enzima dependiente de NAD^+ o $NADP^+$ que cataliza la conversión de un alcohol secundario y NAD^+ o $NADP^+$ a una cetona y NADH o NADPH, respectivamente. El siguiente esquema (1) describe la reducción de NAD^+ o $NADP^+$ mediante un alcohol secundario, ilustrado mediante isopropanol. Por lo tanto, por ejemplo, la cetorreducción puede llevarse a cabo en presencia de un alcohol, tal como, por ejemplo, isopropanol, para proporcionar un sustrato para la reacción inversa (deshidrogenación del alcohol). De esta manera, el NADH/NADPH consumido en la reacción de cetorreducción se regenera mediante la reacción de oxidación inversa tal como se muestra en el Esquema (1):



Esquema (1)

Las alcohol secundario deshidrogenasas que son adecuadas para su uso como sistemas de regeneración de cofactor en las reacciones de reducción catalizadas por cetorreductasa descritas en el presente documento incluyen alcohol secundario deshidrogenasas naturales, así como alcohol secundario deshidrogenasas no naturales. Las alcohol secundario deshidrogenasas naturales incluyen alcohol deshidrogenasas conocidas de

Termoanaerobium brockii, *Rhodococcus erythropolis*, *Lactobacillus kefir*, *Lactobacillus minor* y *Lactobacillus brevis*, y las alcohol secundario deshidrogenasas no naturales incluyen alcohol deshidrogenasas manipuladas derivadas de las mismas. Las alcohol secundario deshidrogenasas empleadas en los métodos que se describen en el presente documento, sean naturales o no naturales, pueden presentar una actividad de al menos aproximadamente 1 µmol/min/mg, a veces al menos aproximadamente 10 µmol/min/mg, o al menos aproximadamente 10^2 µmol/min/mg, hasta aproximadamente 10^3 µmol/min/mg o mayor.

Los alcoholes secundarios adecuados incluyen alcoholes secundarios inferiores y carbinoles de arilalquilo. Los ejemplos de alcoholes secundarios inferiores incluyen isopropanol, 2-butanol, 3-metil-2-butanol, 2-pentanol, 3-pentanol, 3,3-dimetil-2-butanol, y similares. En algunas formas de realización el alcohol secundario es isopropanol. Los carbinoles de arilalquilo adecuados incluyen 1-ariletanoles sustituidos y no sustituidos.

Cuando se emplean como sistema de regeneración de cofactor un alcohol secundario y una alcohol secundario deshidrogenasa, el NAD^+ o $NADP^+$ resultante es reducido mediante la oxidación acoplada del alcohol secundario a la cetona por la alcohol secundario deshidrogenasa. Algunas cetorreductasas manipuladas también tienen actividad para deshidrogenar un reductor alcohol secundario. En algunas formas de realización que utilizan como reductor un alcohol secundario, la cetorreductasa manipulada y la alcohol secundario deshidrogenasa son la misma enzima.

La acumulación de acetona generada por la reducción de isopropanol por las enzimas KRED de la invención puede impedir que la reacción deseada (por ejemplo, la reducción de *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina a (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina o la reducción de *N*-metil-3-ceto-3-(2-tienil)-1-propanamina a *(S)-N*-metil-3-hidroxi-3-(2-tienil)-1-propanamina)) llegue a su término. Por consiguiente, en determinadas formas de realización, la reacción de reducción deseada se lleva a cabo a presión reducida (por ejemplo a 100 Torr) para eliminar la acetona de la mezcla de reacción por destilación. En tales formas de realización, puede añadirse a la reacción en curso isopropanol para reponer el que se consume por reducción a acetona, así como el que se pierde por destilación a presión reducida.

Los sistemas de regeneración de cofactor ejemplares adecuados que pueden emplearse pueden incluir, pero no se limitan a, glucosa y glucosa deshidrogenasa, formiato y formiato deshidrogenasa, glucosa-6-fosfato y glucosa-6-fosfato deshidrogenasa, un alcohol secundario (por ejemplo, isopropanol) y alcohol secundario deshidrogenasa, fosfito y fosfito deshidrogenasa, hidrógeno molecular e hidrogenasa, y similares. Estos sistemas pueden utilizarse en combinación con cualquiera de $NADP^+$/NADPH o NAD^+/NADH como cofactor. También puede utilizarse como sistema de regeneración de cofactor la regeneración electroquímica mediante hidrogenasa. Véanse, por ejemplo, las patentes de EE.UU. Nº 5.538.867 y 6.495.023. También son adecuados los sistemas químicos de regeneración de cofactor que comprenden un catalizador metálico y un agente reductor (por ejemplo, hidrógeno molecular o formiato). Véase, por ejemplo, la publicación PCT WO 2000/053731.

La expresión "glucosa deshidrogenasa" y el término "GDH" se utilizan indistintamente en el presente documento para referirse a una enzima dependiente de NAD^+ o $NADP^+$ que cataliza la conversión de D-glucosa y NAD^+ o $NADP^+$ en ácido glucónico y NADH o NADPH, respectivamente. La siguiente ecuación (1) describe la reducción catalizada por la glucosa deshidrogenasa de NAD^+ o $NADP^+$ mediante glucosa:

$$\text{glucosa} + NAD(P)^+ + H_2O \xrightarrow{GDH} \text{ácido glucónico} + NAD(P)H + H^+$$

Las glucosa deshidrogenasas que son adecuadas para su uso en la práctica de los métodos que se describen en el presente documento incluyen glucosas deshidrogenasas naturales, así como glucosa deshidrogenasas no naturales. Los genes que codifican la glucosa deshidrogenasa natural han sido descritos en la literatura. Por ejemplo, se expresó el gen de GDH 61297 de *Bacillus subtilis* en *E. coli* y se informó que presentaba las mismas propiedades fisicoquímicas que la enzima producida en su hospedador nativo (véase, Vasantha *et al.,* 1983, Proc. Natl. Acad. Sci. EE.UU. 80:785). La secuencia génica del gen GDH de *B. subtilis*, que corresponde al nº de registro del GenBank M12276, fue descrita por Lampel *et al.,* 1986, J. Bacteriol. 166:238-243, y en forma corregida por Yamane *et al.,* 1.996, Microbiology 142:3047-3056 como el nº de registro del GenBank D50453. Los genes de GDH naturales también incluyen los que codifican la GDH de *B. cereus* ATCC 14579 (Nature, 2003, 423:87-91; nº de registro del Genbank AE017013) y *B. megaterium* (Eur. J. Biochem., 1988, 174:485-490, nº de registro del Genbank X12370; J. Ferment. Bioeng., 1990, 70:363-369, nº de registro del Genbank GI216270). Las glucosa deshidrogenasas de *Bacillus sp*. se proporcionan en la publicación PCT WO 2005/018579 como las SEQ ID NO: 10 y 12 (codificadas por secuencias polinucleotídicas correspondientes a las SEQ ID NO: 9 y 11, respectivamente, de la publicación PCT). La Figura 1 representa la conversión del compuesto sustrato de fórmula (III) en el correspondiente producto alcohol de fórmula (IV) mediante los cofactores NAD(P)H/$NAD(P)^+$ y un sistema de reciclaje de cofactor glucosa/glucosa deshidrogenasa.

Pueden generarse glucosa deshidrogenasas no naturales utilizando métodos conocidos, tales como, por ejemplo, mutagénesis, evolución dirigida, y similares. Las enzimas GDH con actividad adecuada, sean naturales o no naturales, pueden identificarse fácilmente utilizando el ensayo descrito en el Ejemplo 4 de la publicación PCT WO 2005/018579. Se proporcionan glucosa deshidrogenasas no naturales ejemplares en la publicación PCT WO 2005/018579 como las SEQ ID NO: 62, 64, 66, 68, 122, 124 y 126. Las secuencias polinucleotídicas que las codifican se proporcionan en la publicación PCT WO 2005/018579 como las SEQ ID NO: 61, 63, 65, 67, 121, 123 y 125, respectivamente. Las glucosa deshidrogenasas adicionales no naturales que son adecuadas para su uso en las reacciones de reducción catalizadas por cetorreductasa descritas en el presente documento se proporcionan en la publicación de solicitud de EE.UU. Nº 2005/0095619 y Nº 2005/0153417.

Las glucosa deshidrogenasas empleadas en las reacciones de reducción catalizadas por cetorreductasa descritas en el presente documento pueden presentar una actividad de al menos aproximadamente 10 µmol/min/mg y algunas veces al menos aproximadamente 10^2 µmol/min/mg o aproximadamente 10^3 µmol/min/mg, hasta aproximadamente 10^4 µmol/min/mg o superior en el ensayo descrito en el Ejemplo 4 de la publicación PCT WO 2005/018579.

Las reacciones de reducción catalizadas por cetorreductasa descritas en el presente documento se llevan a cabo generalmente en un disolvente. Los disolventes adecuados incluyen agua, disolventes orgánicos (por ejemplo, acetato de etilo, acetato de butilo, 1-octacnol, heptano, octano, metil t-butil éter (MTBE), tolueno, y similares), líquidos iónicos (por ejemplo, tetrafluoroborato de 1-etil-4-metilimidazolio, tetrafluoroborato de 1-butil-3-metilimidazolio, hexafluorofosfato de 1-butil-3-metilimidazolio, y similares). En algunas formas de realización, se utilizan disolventes acuosos, incluida el agua y sistemas de codisolventes acuosos.

Los sistemas de codisolventes acuosos ejemplares tienen agua y uno o más disolventes orgánicos. En general, se selecciona un componente disolvente orgánico de un sistema de codisolventes acuoso de manera que no inactive completamente la enzima cetorreductasa. Los sistemas de codisolventes apropiados pueden identificarse fácilmente midiendo la actividad enzimática de la enzima cetorreductasa manipulada especificada con un sustrato de interés definido en el sistema disolvente candidato, utilizando un ensayo de actividad enzimática, tal como los descritos en el presente documento.

El componente disolvente orgánico de un sistema de codisolventes acuoso puede ser miscible con el componente acuoso, proporcionando una única fase líquida, o puede ser parcialmente miscible o inmiscible con el componente acuoso, proporcionando dos fases líquidas. Generalmente, cuando se emplea un sistema de codisolventes acuoso, éste se selecciona para que sea bifásico, dispersándose el agua en un disolvente orgánico, o viceversa. Generalmente, cuando se utiliza un sistema de codisolventes acuoso, es deseable seleccionar un disolvente orgánico que pueda separarse fácilmente de la fase acuosa. En general, la relación entre agua y disolvente orgánico en el sistema de codisolventes se encuentra por lo general en el intervalo comprendido entre aproximadamente 90:10 y aproximadamente 10:90 (v/v) entre disolvente orgánico y agua, y entre 80:20 y 20:80 (v/v) entre disolvente orgánico y agua. El sistema de codisolventes puede formarse previamente antes de añadirse a la mezcla de reacción, o puede formarse *in situ* en el recipiente de reacción.

El disolvente acuoso (agua o el sistema de codisolventes acuoso) puede estar tamponado o no tamponado. Generalmente, la reducción puede llevarse a cabo a un pH de aproximadamente 11 o inferior, normalmente en el intervalo comprendido entre aproximadamente 8 y aproximadamente 11. En algunas formas de realización, la reducción se lleva a cabo a un pH de aproximadamente 9 o superior, normalmente en el intervalo comprendido entre aproximadamente 9 y aproximadamente 10. En algunas formas de realización, la reducción se lleva a cabo a un pH de aproximadamente 9.

A lo largo de las reacciones de reducción, el pH de la mezcla de reacción puede cambiar. El pH de la mezcla de reacción puede mantenerse a un pH deseado o dentro de un intervalo de pH deseado añadiendo un ácido o una base a lo largo de la reacción. Como alternativa, el pH puede controlarse utilizando un disolvente acuoso que comprenda un tampón. Los tampones adecuados para mantener los intervalos de pH deseados se conocen en la técnica e incluyen, por ejemplo, tampón fosfato, tampón trietanolamina, y similares. También pueden utilizarse combinaciones de tamponamiento y adición de ácido o base.

Cuando se emplea un sistema de regeneración de cofactor de glucosa/glucosa deshidrogenasa, la coproducción de ácido glucónico (pKa = 3,6), tal como se representa en la ecuación (1) puede hacer que disminuya el pH de la mezcla de reacción si no se neutraliza el ácido glucónico acuoso resultante. El pH de la mezcla de reacción puede mantenerse en el nivel deseado mediante técnicas convencionales de tamponamiento, en las que el tampón neutraliza el ácido glucónico hasta la capacidad de tamponamiento proporcionada, o añadiendo una base simultáneamente al transcurso de la conversión. También pueden utilizarse combinaciones de tamponamiento y adición de base. Los tampones adecuados para mantener los intervalos de pH deseados se han descrito anteriormente. Las bases adecuadas para la neutralización de ácido glucónico son bases orgánicas, por ejemplo aminas, alcóxidos y similares, y bases inorgánicas, por ejemplo, sales de hidróxido (por ejemplo, NaOH), sales de carbonato (por ejemplo, $NaHCO_3$), sales de bicarbonato (por ejemplo, K_2CO_3), sales de fosfato básico (por ejemplo, K_2HPO_4, Na_3PO_4), y similares. La adición de una base simultáneamente al transcurso de la conversión puede

hacerse manualmente mientras se supervisa el pH de la mezcla de reacción o, más oportunamente, utilizando un titulador automático tal como un pH-stat. También puede utilizarse una combinación de capacidad de tamponamiento parcial y adición de base para el control del proceso.

Cuando se emplea la adición de base para neutralizar el ácido glucónico liberado durante una reacción de reducción catalizada por cetorreductasa, puede supervisarse el avance de la conversión mediante la cantidad de base añadida para mantener el pH. Por lo general, las bases añadidas a las mezclas de reacción no tamponadas o parcialmente tamponadas a lo largo de la reducción se añaden en soluciones acuosas.

En algunas formas de realización, el sistema de regeneración de cofactor puede comprender una formiato deshidrogenasa. La expresión "formiato deshidrogenasa" y el término "FDH" se utilizan indistintamente en el presente documento para referirse a una enzima dependiente de NAD^+ o $NADP^+$ que cataliza la conversión de formiato y NAD^+ o $NADP^+$ en dióxido de carbono y NADH o NADPH, respectivamente. Las formiato deshidrogenasas que pueden ser adecuadas para utilizarse como sistemas de regeneración de cofactor en las reacciones de reducción catalizadas por cetorreductasa descritas en el presente documento incluyen formiato deshidrogenasas naturales, así como formiato deshidrogenasas no naturales. Las formiato deshidrogenasas incluyen las correspondientes a las SEQ ID NO: 70 (*Pseudomonas sp*.) Y 72 (*Candida boidinii*) de la publicación PCT WO 2005/018579, que son codificadas por las secuencias polinucleotídicas correspondientes a las SEQ ID NO: 69 y 71, respectivamente, de la publicación PCT 2005/018579. Las formiato deshidrogenasas que pueden emplearse en los métodos que se describen en el presente documento, sean naturales o no naturales, pueden presentar una actividad de al menos aproximadamente 1 µmol/min/mg, a veces al menos aproximadamente 10 µmol/min/mg, o al menos aproximadamente 10^2 µmol/min/mg, hasta aproximadamente 10^3 µmol/min/mg o superior, y pueden cribarse fácilmente para determinar su actividad en el ensayo que se describe en el Ejemplo 4 de la publicación PCT WO 2005/018579.

Tal como se utiliza en el presente documento, el término "formiato" se refiere al anión formiato (HCO_2^-), al ácido fórmico (HCO_2H), y a mezclas de los mismos. El formiato puede proporcionarse en forma de sal, por lo general un álcali o sal de amonio (por ejemplo, HCO_2Na HCO_2NH_4, y similares), en forma de ácido fórmico, por lo general ácido fórmico acuoso, o mezclas de los mismos. El ácido fórmico es moderadamente ácido. En soluciones acuosas dentro de varias unidades de pH de su pKa (pKa = 3,7 en agua) el formiato está presente como HCO_2^- y como HCO_2H en concentraciones de equilibrio. A valores de pH por encima de aproximadamente pH 4, el formiato está predominantemente presente como HCO_2^-. Cuando el formiato se proporciona como ácido fórmico, la mezcla de reacción por lo general se tampona o se hace menos ácida añadiendo una base para proporcionar el pH deseado, por lo general de aproximadamente pH 5 o superior. Las bases adecuadas para neutralizar el ácido fórmico incluyen, pero no se limitan a, bases orgánicas, por ejemplo aminas, alcóxidos y similares, y bases inorgánicas, por ejemplo, sales de hidróxido (por ejemplo, NaOH), sales de carbonato (por ejemplo, $NaHCO_3$), sales de bicarbonato (por ejemplo, K_2CO_3), sales de fosfato básico (por ejemplo, K_2HPO_4, Na_3PO_4), y similares.

Para valores de pH superiores a aproximadamente pH 5, en los que el formiato está predominantemente presente como HCO_2^-, la siguiente Ecuación (2) describe la reducción catalizada por formiato deshidrogenasa de NAD^+ o $NADP^+$ mediante formiato:

$$HCO_2^- + NAD(P)^+ \xrightarrow{FDH} CO_2 + NAD(P)H$$

Cuando se emplean como sistema de regeneración de cofactor formiato y formiato deshidrogenasa, el pH de la mezcla de reacción puede mantenerse en el nivel deseado mediante técnicas de tamponamiento convencionales, en las que el tampón libera protones hasta la capacidad de tamponamiento proporcionada, o añadiendo un ácido simultáneamente al transcurso de la reacción. Los ácidos adecuados a añadir a lo largo de la reacción para mantener el pH incluyen ácidos orgánicos, por ejemplo ácidos carboxílicos, ácidos sulfónicos, ácidos fosfónicos, y similares, ácidos minerales, por ejemplo ácidos halohídricos (tal como ácido clorhídrico), ácido sulfúrico, ácido fosfórico y similares, sales ácidas, por ejemplo, sales dihidrogenofosfato (por ejemplo, KH_2PO_4), sales bisulfato (por ejemplo, $NaHSO_4$) y similares. Algunas formas de realización utilizan ácido fórmico, de manera que se mantengan tanto la concentración de formiato como el pH de la solución.

Cuando se emplea la adición de ácido para mantener el pH durante una reacción de reducción utilizando el sistema de regeneración de cofactor formiato/formiato deshidrogenasa, puede supervisarse el avance de la conversión mediante la cantidad de ácido añadido para mantener el pH. Por lo general, los ácidos añadidos a las mezclas de reacción no tamponadas o parcialmente tamponadas a lo largo de la conversión se añaden en soluciones acuosas.

Cuando se llevan a cabo las formas de realización de las reacciones de reducción catalizadas por las cetorreductasa descritas en el presente documento empleando un sistema de regeneración de cofactor, puede

proporcionarse inicialmente la forma oxidada o reducida del cofactor. Como se ha descrito anteriormente, el sistema de regeneración de cofactor convierte el cofactor oxidado en su forma reducida, que a continuación se utiliza en la reducción del sustrato de la cetorreductasa.

En algunas formas de realización, no se utilizan sistemas de regeneración de cofactor. Para las reacciones de reducción llevadas a cabo sin utilizar un sistema de regeneración de cofactor, el cofactor se añade a la mezcla de reacción en forma reducida.

En algunas formas de realización, cuando el proceso se lleva a cabo utilizando células enteras del organismo hospedador, la célula entera puede proporcionar el cofactor de forma nativa. Como alternativa o en combinación, la célula puede proporcionar de forma nativa o recombinante la glucosa deshidrogenasa.

Cuando se llevan a cabo las reacciones de reducción estereoselectivas descritas en el presente documento, la enzima cetorreductasa manipulada, y cualesquiera enzimas que comprendan el sistema de regeneración de cofactor opcional, pueden añadirse a la mezcla de reacción en forma de enzimas purificadas, células enteras transformadas con el gen o genes que codifican las enzimas, y/o extractos celulares y/o lisados de tales células. El gen o genes que codifican la enzima cetorreductasa manipulada y las enzimas de regeneración de cofactor opcionales pueden transformarse en células hospedadoras por separado o juntas en la misma célula hospedadora. Por ejemplo, en algunas formas de realización, puede transformarse un conjunto de células hospedadoras con el gen o genes que codifican la enzima cetorreductasa manipulada y puede transformarse otro conjunto con el gen o genes que codifican las enzimas de regeneración de cofactor. Ambos conjuntos de células transformadas pueden utilizarse juntos en la mezcla de reacción en forma de células enteras, o en forma de lisados o extractos derivados de las mismas. En otras formas de realización, una célula hospedadora puede transformarse con el gen o genes que codifican tanto la enzima cetorreductasa manipulada como las enzimas de regeneración de cofactor.

Las células enteras transformadas con el gen o genes que codifican la enzima cetorreductasa manipulada y/o las enzimas de regeneración de cofactor opcionales, o los extractos celulares y/o lisados de las mismas, pueden emplearse de diversas formas diferentes, incluidas sólida (por ejemplo, liofilizada, secada por aspersión, y similares) o semisólida (por ejemplo, una pasta bruta).

Los extractos celulares o lisados celulares pueden purificarse parcialmente mediante precipitación (sulfato de amonio, polietilenimina, tratamiento térmico o similar, seguido de un procedimiento de desalado antes de la liofilización (por ejemplo, ultrafiltración, diálisis, y similares). Cualquiera de las preparaciones celulares pueden estabilizarse por reticulación utilizando agentes de reticulación conocidos, tales como, por ejemplo, glutaraldehído o inmovilización en una fase sólida (por ejemplo, Eupergit C, y similares).

Pueden proporcionarse a la reacción reaccionantes sólidos (por ejemplo, enzima, sales, etc.) de diversas formas diferentes, incluidas en polvo (por ejemplo, liofilizada, secada por aspersión, y similares), solución, emulsión, suspensión, y similares. Los reaccionantes pueden liofilizarse o secarse por aspersión fácilmente utilizando métodos y equipos conocidos por los expertos en la materia. Por ejemplo, la solución de proteínas puede congelarse a -80°C en pequeñas alícuotas, añadirse a continuación a una cámara de liofilización previamente enfriada, seguido de la aplicación de vacío. Después de extraer el agua de las muestras, la temperatura se eleva a por lo general a 4°C durante dos horas antes de liberar el vacío y recuperar las muestras liofilizadas.

Las cantidades de reaccionantes utilizados en la reacción de reducción variarán generalmente en función de las cantidades de producto deseadas, y de manera concomitante la cantidad de sustrato de la cetorreductasa empleado. Pueden utilizarse las siguientes directrices para determinar las cantidades de cetorreductasa, cofactor, y sistema de regeneración de cofactor opcional a utilizar. En general, los sustratos ceto pueden emplearse a una concentración de aproximadamente 20 a 300 gramos/litro utilizando de aproximadamente 50 mg a aproximadamente 5 g de cetorreductasa y de aproximadamente 10 mg a aproximadamente 150 mg de cofactor. Los expertos en la técnica entenderán fácilmente cómo variar estas cantidades para adaptarlas al nivel de productividad y a la escala de producción deseados. Las cantidades apropiadas del sistema de regeneración de cofactor opcional pueden determinarse fácilmente mediante experimentación rutinaria en base a la cantidad de cofactor y/o cetorreductasa utilizados. En general, el reductor (por ejemplo, glucosa, formiato, isopropanol) se utiliza a niveles por encima del nivel equimolar del sustrato de la cetorreductasa para conseguir la conversión esencialmente completa o casi completa del sustrato de la cetorreductasa.

Para mejorar la eficacia de mezcla cuando se utiliza un sistema de codisolventes acuoso, el sistema de regeneración de cofactor, cetorreductasa, y el cofactor pueden añadirse y mezclarse primero en la fase acuosa. A continuación, puede añadirse y mezclarse la fase orgánica, seguido de la adición del sustrato de la cetorreductasa. Como alternativa, el sustrato de la cetorreductasa puede mezclarse previamente en la fase orgánica, antes de añadirse a la fase acuosa.

Las condiciones adecuadas para llevar a cabo las reacciones de reducción catalizadas por cetorreductasa descritas en el presente documento incluyen gran diversidad de condiciones que pueden optimizarse fácilmente

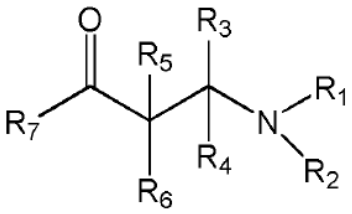
mediante experimentación rutinaria que incluye, pero no se limita a, poner en contacto la enzima cetorreductasa manipulada y el substrato a un pH y temperatura experimentales, y detectar el producto, por ejemplo, utilizando los métodos descritos en los Ejemplos que se proporcionan en el presente documento.

La reducción catalizada por cetorreductasa se lleva a cabo por lo general a una temperatura en el intervalo comprendido entre aproximadamente 15°C y aproximadamente 75°C. Para algunas formas de realización, la reacción se lleva a cabo a una temperatura en el intervalo comprendido entre aproximadamente 20°C y aproximadamente 55°C. En todavía otras formas de realización, se lleva a cabo a una temperatura en el intervalo comprendido entre aproximadamente 20°C y aproximadamente 45°C. La reacción también puede llevarse a cabo en condiciones ambientales.

Generalmente, se deja que la reacción de reducción avance hasta llegar básicamente a la finalización, o casi la finalización, de la reducción del sustrato. La reducción del sustrato a producto puede supervisarse utilizando métodos conocidos detectando el sustrato y/o el producto. Los métodos adecuados incluyen cromatografía de gases, HPLC, y similares. Los rendimientos de conversión del producto de reducción alcohólica generados en la mezcla de reacción son generalmente superiores a aproximadamente un 50%, también pueden ser superiores a aproximadamente un 60%, también pueden ser superiores a aproximadamente un 70%, también pueden ser superiores a aproximadamente un 80%, también pueden ser superiores a un 90%, y son con frecuencia superiores a aproximadamente un 97%.

La presente invención proporciona un método para preparar una *(S)-N*-metil-3-hidroxi-3-(aril)-propanamina, comprendiendo el método:

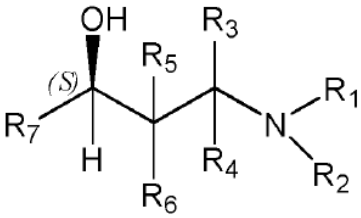
(a) proporcionar un sustrato 3-aril-3-cetopropanamina que tiene la estructura de fórmula (I):



(I)

en el que cada uno de R_1 y R_2 es metilo, cada uno de R_3, R_4, R_5 y R_6 está seleccionado independientemente del grupo que consiste en hidrógeno y un alquilo inferior opcionalmente sustituido, y R_7 es un arilo opcionalmente sustituido;

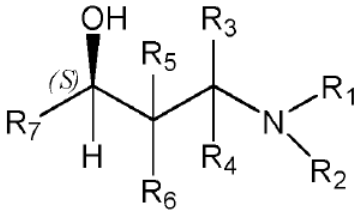
(b) poner en contacto o incubar el sustrato 3-aril-3-cetopropanamina con uno o más polipéptidos cetorreductasa de la presente invención en una mezcla de reacción en condiciones adecuadas para la reducción o conversión del sustrato en un producto *(S)*-3-aril-3 -hidroxipropanamina que tiene la fórmula estructural (II):



(II)

en el que cada uno de R_1 y R_2 es metilo, cada uno de R_3, R_4, R_5 y R_6 está seleccionado independientemente del grupo que consiste en hidrógeno y un alquilo inferior opcionalmente sustituido, y R_7 es un arilo opcionalmente sustituido;

(c) desmetilar el producto *(S)*-3-aril-3-hidroxipropanamina (es decir, *N,N*-dimetil-3-hidroxi-3-(aril)-propanamina) de la etapa (b) en una mezcla de reacción en condiciones adecuadas para producir una *(S)-N*-metil-3-hidroxi-3-(aril)-propanamina que tiene la fórmula estructural (II):



(II)

en la que uno de R_1 y R_2 es metilo y el otro es hidrógeno, cada uno de R_3, R_4, R_5 y R_6 está seleccionado independientemente del grupo que consiste en hidrógeno y un alquilo inferior opcionalmente sustituido, y R_7 es un arilo opcionalmente sustituido.

Los sustituyentes R_3-R_7 son como se ha descrito anteriormente en el presente documento para los compuestos que tienen la estructura de fórmula (I) y (II). Por lo general, R_3, R_4, R_5 y R_6 son todos hidrógenos y R_7 es un arilo seleccionado del grupo que consiste en fenilo y tienilo. Por lo general, R_7 es 2-tienilo. Se ha descubierto que el producto desmetilado, es decir, *(S)-N*-dimetil-3-hidroxi-3-(aril)-propanamina, producido mediante el método anteriormente descrito se genera con una pureza enantiomérica muy alta del 99% o más.

Las condiciones adecuadas para producir la *(S)-N*-metil-3-hidroxi-3-(aril)-propanamina en la etapa (b) incluyen poner en contacto el producto *N,N*-dimetil-3-hidroxi-3-(aril)-propanamina de la etapa (b) con un agente de desmetilación en presencia de una base para formar un producto intermedio que se somete a hidrólisis adicional para producir el producto *(S)-N*-metil-3-hidroxi-3-(aril)-propanamina de la etapa (c). La expresión "agente de desmetilación" se refiere en el presente documento a un compuesto que facilita la eliminación de un grupo metilo de una amina. Los agentes de desmetilación adecuados incluyen cloroformiato y sus derivados, fosgeno y sus derivados, y otros compuestos adecuados que son bien conocidos en la técnica. Los derivados de cloroformiato ejemplares incluyen cloroformiato de etilo, cloroformiato de metilo, cloroformiato de propilo, cloroformiato de butilo, cloroformiato de i-butilo, cloroformiato de fenilo, cloroformiato de 2,2,2-tricloroetilo, cloroformiato de 2-cloroetilo, cloroformiato de 2-yodoetilo, cloroformiato de bencilo, cloroformiato de nitrobencilo, cloroformiato de 1-cloroetilo, cloroformiato de 2,2-dicloroetilo, cloroformiato de 1,1-dimetil-2,2,2-tricloroetilo, cloroformiato de 1,1-dimetil-2-cloroetilo, cloroformiato de 1,1-dimetil-2-bromoetilo, y similares. Por lo general, el derivado de cloroformiato es cloroformiato de i-butilo o cloroformiato de fenilo. Los derivados de fosgeno adecuados para utilizarse en la práctica de la invención incluyen difosgeno (es decir, cloroformiato de triclorometilo), trifosgeno (es decir, bis(carbonato de triclorometilo), y similares.

Las bases adecuadas para utilizar junto con el agente de desmetilación incluyen aminas (por ejemplo, trialquilaminas tales como trietilamina, trimetilamina, dialquilaminas, y similares), hidróxidos de un metal alcalino o sus sales con un ácido débil (tal como, por ejemplo, hidróxido sódico, hidróxido potásico, carbonato sódico, hidrogenocarbonato sódico, carbonato potásico, hidrogenocarbonato potásico, carbonato de cesio, fosfato tripotásico, fosfato tripotásico dihidratado, y similares), hidróxidos de amonio cuaternario o sus sales con un ácido débil, y similares. Por lo general, la base es una amina, tal como, por ejemplo, trietilamina.

El producto intermedio se prepara en una mezcla de reacción del agente de desmetilación, base, y un disolvente, tal como, por ejemplo, una pirrolidona, una cetona (por ejemplo, acetona, etil metil cetona, y similares), un disolvente aprótico dipolar (por ejemplo, dimetilsulfóxido (DMSO), *N*-metil-2-pirrolidona, dimetilformamida, dimetilacetamida, y similares), un hidrocarburo aromático (por ejemplo, benceno, tolueno, xileno, mesitileno, y similares), un nitrilo (por ejemplo, acetonitrilo, y similares), un éter (por ejemplo, metil t-butil éter, diisopropil éter, tetrahidrofurano, 1,4-dioxano, 1,2-dimetoxietano, anisol, y similares), una amida y similares, así como combinaciones de cualesquiera dos o más de los mismos. Por lo general, el disolvente es un hidrocarburo aromático, tal como, por ejemplo, tolueno. Si el disolvente no se sustituye para la posterior etapa de hidrólisis, el disolvente es por lo general un hidrocarburo aromático, tal como, por ejemplo, benceno, tolueno, xileno y mesitileno, un éter, tal como, por ejemplo, t-butil metil éter, diisopropil éter, tetrahidrofurano, 1,4-dioxano, 1,2-dimetoxietano y anisol, o cualquier combinación de dos o más de los mismos.

La reacción se lleva a cabo por lo general a una temperatura en el intervalo comprendido entre aproximadamente -30°C y aproximadamente 100°C. Por lo general, la reacción se lleva a cabo a una temperatura en el intervalo comprendido entre aproximadamente 0°C y aproximadamente 90°C, y normalmente a una temperatura en el intervalo comprendido entre aproximadamente 25°C ó 30°C y aproximadamente 80°C, y con frecuencia a una temperatura de aproximadamente 75°C.

La posterior reacción de hidrólisis se lleva a cabo en presencia de una base y un disolvente, tal como, por ejemplo, un alcohol (por ejemplo, metanol, etanol, propanol, isopropanol, butanol, y similares), una amida, una pirrolidona, un disolvente aprótico dipolar (por ejemplo, DMSO, *N*-metil-2-pirrolidona, dimetilformamida, dimetilacetamida, y similares), una cetona (por ejemplo, acetona, etil metil cetona, y similares), un hidrocarburo

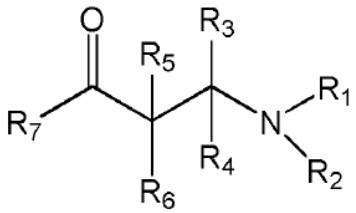
aromático (por ejemplo, benceno, tolueno, xileno, mesitileno, y similares), y éter (por ejemplo, butil metil éter, diisopropil éter, tetrahidrofurano, 1,4-dioxano, 1,2-dimetoxietano, anisol, y similares), agua, y similares, así como cualquier combinación de dos o más de los mismos. Los sistemas de codisolventes acuosos que son adecuados para utilizarse en el sistema de hidrólisis incluyen agua y un alcohol en una mezcla de 25:75, 50:50 ó 75:25 de agua:alcohol. Las bases adecuadas para utilizarse en la etapa de hidrólisis incluyen un hidróxido de un metal alcalino o sal del mismo con un ácido débil, un hidróxido de amonio cuaternario o sal del mismo con un ácido débil, y similares. Las bases ejemplares incluyen hidróxido sódico, hidróxido potásico, carbonato sódico, hidrogenocarbonato sódico, carbonato potásico, hidrogenocarbonato potásico, carbonato de cesio, fosfato tripotásico y fosfato tripotásico dihidratado. Por lo general, la hidrólisis se lleva a cabo utilizando un hidróxido de un metal alcalino, tal como, por ejemplo, potasio o sodio, en un sistema de codisolventes de agua-alcohol, tal como por ejemplo agua-metanol, agua-isopropanol, y similares. Las condiciones adecuadas para llevar a cabo la desmetilación se describen en la publicación de patente de EE.UU. 2006/0167278, la patente de EE.UU. Nº 5.023.269, la patente de EE.UU. Nº 5.491.243, la patente de EE.UU. Nº 6.541.668, el documento WO 2007/095200, el documento EP 0 457 559A2 y el documento EP 0 273 658A1.

La (*S*)-*N,N*-dimetil-3-hidroxi-3-(aril)-propanamina puede recuperarse opcionalmente después de llevar a cabo la etapa (b), antes de la etapa de desmetilación (c). El producto desmetilado, *(S)-N*-metil-3-hidroxi-3-(aril)-propanamina, puede recuperarse posteriormente y, opcionalmente, purificarse adicionalmente después de llevar a cabo la etapa (c), utilizando métodos que son bien conocidos en la técnica.

La presente invención incluye *(S)-N*-metil-3-hidroxi-3-(aril)-propanamina producida mediante los métodos que se describen en el presente documento. La *(S)-N*-metil-3-hidroxi-3-(aril)-propanamina producida de esta manera tiene una pureza enantiomérica del 99% o más.

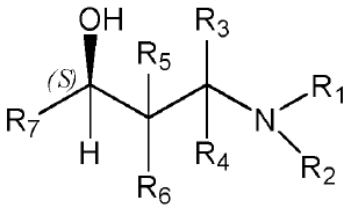
En una forma de realización adicional, la presente invención proporciona un método de producción de una (S)-3-ariloxi-3-(aril)-propanamina, comprendiendo el método:

(a) proporcionar una 3-aril-3-cetopropanamina que tiene la estructura de fórmula (I):



(I)

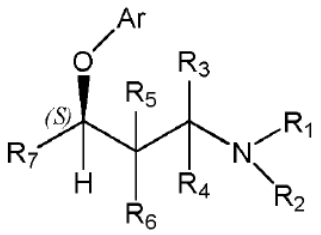
(b) poner en contacto la 3-aril-3-cetopropanamina con un polipéptido cetorreductasa de la presente invención en una mezcla de reacción en condiciones suficientes para producir una *(S)*-3-aril-3-hidroxipropanamina que tiene la estructura de fórmula (II):



(II)

y

(c) poner en contacto la *(S)*-3-aril-3-hidroxipropanamina con un compuesto de arilo activado en una mezcla de reacción en condiciones suficientes para producir la *(S)*-3-ariloxi-3-arilpropanamina que tiene la estructura de fórmula (VII)



(VII)

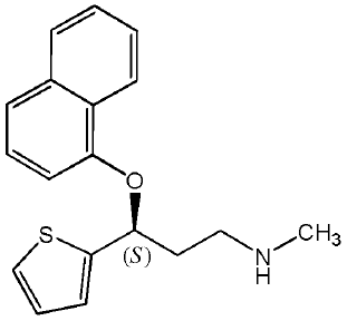
en la que para (I), (II) y (VII), cada uno de R_1 y R_2 está seleccionado independientemente del grupo que consiste en hidrógeno, un alquilo inferior opcionalmente sustituido, un cicloalquilo opcionalmente sustituido, un arilo opcionalmente sustituido, o como alternativa, en las que R_1 y R_2 juntos forman un cicloalquilo opcionalmente sustituido o un cicloarilo opcionalmente sustituido que tiene 3-7 átomos de carbono; cada uno de R_3, R_4, R_5 y R_6 está seleccionado independientemente del grupo que consiste en hidrógeno y un alquilo inferior opcionalmente sustituido; y R_7 es un arilo opcionalmente sustituido y, además, para (VII), Ar es un grupo arilo opcionalmente sustituido. Por lo general, R1-R_7 son sustituyentes tal como se ha descrito anteriormente con respecto a las fórmulas (I) y (II). Por lo general, Ar es un arilo seleccionado del grupo que consiste en 1-naftilo, fenilo, 4-trifluorometilfenilo, 2-metilfenilo, 2-metoxifenilo y 2-tiometoxifenilo.

El método puede incluir opcionalmente una etapa de desmetilación si R_1 y R_2 de la 3-aril-3-cetopropanamina son ambos metilo como se ha descrito anteriormente en el presente documento. La etapa de desmetilación puede producirse antes o después de la etapa (c). Preferentemente, la etapa de desmetilación se produce antes de la etapa (c).

Tal como se utiliza en el presente documento, la expresión "compuesto de arilo activado" se refiere a un compuesto de arilo que está sustituido con un grupo saliente. Los grupos salientes adecuados incluyen haluros (tales como, por ejemplo, fluoro, cloro, y similares), pseudohaluros (tales como, por ejemplo, sulfonatos, tales como triflato, tosilato, mesilato, y similares), y similares. Las condiciones de reacción suficientes para producir la 3-ariloxi-3-(aril)-propanamina con la estructura de fórmula (VII) incluyen poner en contacto la 3-aril-3-hidroxipropanamina con el compuesto de arilo activado en presencia de un hidruro de metal alcalino, tal como, por ejemplo, hidruro sódico, hidruro potásico, y similares, o un hidróxido de metal alcalino, tal como, por ejemplo, hidróxido potásico, hidróxido sódico, y similares. La 3-aril-3-hidroxipropanamina también puede ponerse en contacto con un compuesto de arilo activado con un pseudohaluro (tal como, por ejemplo, triflato, tosilato, y similares) o un haluro en presencia de un catalizador metal de transición (tal como, por ejemplo, paladio (Pd), cobre (Cu), y similares). La reacción se lleva a cabo en un disolvente, tal como, por ejemplo, *N,N*-dimetilacetamida, dimetilsulfóxido, dimetilformamida, piridina, y similares. La reacción se lleva a cabo por lo general a una temperatura en el intervalo comprendido entre aproximadamente 10°C hasta la temperatura de reflujo del disolvente utilizado. Los métodos para preparar una 3-ariloxi-3-(aril)-propanamina a partir de una 3-aril-3-hidroxipropanamina se describen, por ejemplo, en la patente de EE.UU. Nº 5.023.269, la patente de EE.UU. Nº 5.491.243, la patente de EE.UU. Nº 6.541.668, el documento WO 2007/095200, el documento EP 0 457 559A2 y el documento EP 0 273 658A1.

Los métodos para preparar una 3-ariloxi-3-(aril)-propanamina incluyen opcionalmente la etapa adicional de preparar una sal farmacéuticamente aceptable de la 3-ariloxi-3-(aril)-propanamina. Los métodos para preparar una sal farmacéuticamente aceptable son bien conocidos en la técnica. Véase, por ejemplo, la patente de EE.UU. Nº 5.023.269, la patente de EE.UU. Nº 5.491.243, la patente de EE.UU. Nº 6.541.668, el documento WO 2007/095200, el documento EP 0 457 559A2 y el documento EP 0 273 658A1. Las sales de adición de ácido farmacéuticamente aceptables se preparan por lo general haciendo reaccionar una 3-ariloxi-3-(aril)-propanamina con una cantidad equimolar o en exceso de ácido. Los ácidos adecuados incluyen ácidos tanto inorgánicos como orgánicos. Los ácidos inorgánicos ejemplares incluyen ácido clorhídrico, bromhídrico, yodhídrico, nítrico, sulfúrico, fosfórico, metafosfórico, pirofosfórico, y similares. Los ácidos orgánicos ejemplares incluyen ácidos mono y dicarboxílicos alifáticos, ácidos alcanoicos sustituidos (por ejemplo, sustituidos con fenilo, sustituidos con hidroxilo, y similares), ácidos aromáticos, ácidos sulfónicos alifáticos y aromáticos, y similares.

En algunas formas de realización, los métodos se refieren más concretamente al uso de los polipéptidos cetorreductasa proporcionados en el presente documento en la síntesis de duloxetina (es decir, *(3S)-N*-metil-3-naftalen-1-iloxi-3-tiofen-2-il-propan-1-amina), que tiene la fórmula estructural (VIII):



(VIII)

y sales, hidratos y solvatos de la misma. Por consiguiente, en un método para la síntesis del compuesto de fórmula (VIII), una etapa en el método puede comprender la reducción estereoselectiva del sustrato de fórmula estructural (III) o (V) al correspondiente producto alcohol de fórmula estructural (IV) o (VI), respectivamente, poniendo en contacto o incubando el sustrato con uno cualquiera o más de los polipéptidos cetorreductasa de la descripción en condiciones de reacción adecuadas para la reducción o conversión del sustrato en el producto. La ruta de síntesis para preparar duloxetina a partir del compuesto de fórmula estructural (III) se ilustra en la Figura 2. Los métodos para preparar duloxetina a partir de una 3-aril-3-hidroxipropanamina se describen, por ejemplo, en la patente de EE.UU. Nº 5.023.269, la patente de EE.UU. Nº 5.491.243, la patente de EE.UU. Nº 6.541.668, el documento WO 2007/095200, el documento EP 0 457 559A2 y el documento EP 0 273 658A1. La Figura 2 representa una ruta de síntesis para producir duloxetina. En la Figura 2, el compuesto de fórmula estructural (III), *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina se reduce a (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina (IV), que se somete posteriormente a naftilación para dar el compuesto intermedio (IX). El producto intermedio (IX) naftilado se desmetila posteriormente para producir duloxetina, *(3S)-N*-metil-3-naftalen-1-iloxi-3-tiofen-2-il-propan-1-amina.

Por lo tanto, la presente invención proporciona métodos para preparar duloxetina, es decir, *(3S)-N*-metil-3-naftalen-1-iloxi-3-tiofen-2-il-propan-1-amina (Fórmula VIII). En una forma de realización, el método comprende:

(a) proporcionar *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina;

(b) poner en contacto la *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina con un polipéptido cetorreductasa de la presente invención en una mezcla de reacción en condiciones suficientes para producir (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina;

(c) desmetilar la (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina para producir *(S)-N*-metil-3-hidroxi-3-(2-tienil)-1-propanamina; y

(d) poner en contacto la *(S)-N*-metil-3-hidroxi-3-(2-tienil)-1-propanamina con un naftaleno activado en una mezcla de reacción en condiciones suficientes para producir *(3S)-N*-metil-3 -naftalen-1-iloxi-3-tiofen-2-il-propan-1-amina (duloxetina) que tiene la estructura de fórmula (VIII); y

(e) recuperar la *(3S)-N*-metil-3-naftalen-1-iloxi-3-tiofen-2-il-propan-1-amina de la mezcla de reacción.

En otra forma de realización para preparar duloxetina, el método comprende:

(a) proporcionar *N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina;

(b) poner en contacto la *N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina en una mezcla de reacción en condiciones suficientes para producir *(S)-N*-metil-3-hidroxi-3-(2-tienil)-1-propanamina; y

(c) poner en contacto la *(S)-N*-metil-3-hidroxi-3-(2-tienil)-1-propanamina con un naftaleno activado en una mezcla de reacción en condiciones suficientes para producir *(3S)-N*-metil-3-naftalen-1-iloxi-3-tiofen-2-il-propan-1-amina (duloxetina) que tiene la estructura de fórmula (VIII); y

(d) recuperar la *(3S)-N*-metil-3-naftalen-1-iloxi-3-tiofen-2-il-propan-1-amina de la mezcla de reacción.

Tal como se utiliza en el presente documento, la expresión "naftaleno activado" se refiere a un arilo activado como se ha descrito anteriormente, en el que el grupo arilo es 1-naftilo. Las condiciones de reacción para llevar a cabo la naftilación y desmetilación son como se han descrito anteriormente en la descripción del método para preparar una 3-ariloxi-3-arilpropanamina a partir de una *(S)*-3-aril-3-hidroxipropanamina. Los agentes de desmetilación adecuados para utilizarse en el método para preparar duloxetina a partir de (*S*)-*N,N*-dimetil-3-naftalen-1-iloxi-3-tiofen-2-il-propan-1-diamina incluyen los descritos anteriormente en el presente documento. Los métodos

para preparar una 3-ariloxi-3-arilpropanamina a partir de una 3-aril-3-hidroxipropanamina se describen, por ejemplo, en la patente de EE.UU. Nº 5.023.269, la patente de EE.UU. Nº 5.491.243 y la patente de EE.UU. Nº 6.541.668, el documento WO 2007/095200, el documento EP 0 457 449A2 y el documento EP 0 273 658A1.

En los siguientes ejemplos representativos se ilustran diversas características y formas de realización de la descripción, los cuales pretenden ser ilustrativos, y no limitativos.

EJEMPLOS

Ejemplo 1: Adquisición de genes de cetorreductasa de tipo silvestre y construcción de vectores de expresión.

Se diseñaron genes que codifican cetorreductasa (KRED) para la expresión en *E. coli* en base a la secuencia de aminoácidos descrita de la cetorreductasa y utilizando métodos convencionales de optimización de codones. El software convencional de optimización de codones se revisa, por ejemplo, en "OPTIMIZER: a web server for optimizing the codon usage of DNA sequences", Puigbò *et al.,* Nucleic Acids Res. (julio de 2007); 35 (publicado en el servidor WEB): W126-31, Epub 16 de abril de 2007. Los genes se sintetizaron utilizando oligonucleótidos, compuestos generalmente por 42 nucleótidos, que se clonaron en el vector de expresión pCK110900, representado como FIG. 3 en la publicación de solicitud de patente de los Estados Unidos 20060195947, bajo el control de un promotor *lac*. Este vector de expresión también contiene el origen de replicación p15A y el gen de resistencia a cloranfenicol. Los plásmidos resultantes se transformaron en *E. coli* W3110 utilizando métodos convencionales. Los genes con codones optimizados y también los polipéptidos codificantes aparecen en las Tablas 2 y 3, y sus secuencias se proporcionan en la SEQ ID NO: 5-78.

Las cetorreductasas útiles para el desarrollo de enzimas capaces de reducir el compuesto de Fórmula (III) al compuesto de Fórmula (IV) se proporcionan en la siguiente Tabla 3.

Tabla 3: Abreviaturas, fuente y referencias para las cetorreductasas

Cetorreductasa	Microorganismo a partir del que se identificó originalmente la enzima	Número de registro del Genbank	Número GI	SEQ ID NO del polinucleótido	fuente o número de SEQ ID del polipéptido
ADH-CM	*Candida magnolia*	AB036927.1	12657576	SEQ ID NO:79	SEQ ID NO:80
YDL	*Saccharomyces cervisiae*	NP_010159.1	6320079	SEQ ID NO:85	SEQ ID NO:88
ADH-LB	*Lactobacillus brevis*	1NXQ A	30749782	SEQ ID NO:3	SEQ ID NO:4
ADH-RE	*Rhodococcus erythropolis*	AAN73270.1	34776951	SEQ ID NO:81	SEQ ID NO:82
YGL	*Saccharomyces cerevisiae*	NP_011476	6321399	SEQ ID NO:83	SEQ ID NO:84
YPR	*Saccharomyces cerevisiae*	NP_010656.1	6320576	SEQ ID NO:87	SEQ ID NO:88
GRE	*Saccharomyces cerevisiae*	NP_014490.1	6324421	SEQ ID NO:89	SEQ ID NO:90
ADH-LK	*Lactobacillus kefir*	AAP94029.1	33112056	SEQ ID NO:1	SEQ ID NO:2
ADH-SB	*Sporobolomyces salmonicolor*	Q9UUN9	30315955	SEQ ID NO:93	SEQ ID NO:94
ADH-SC	*Streptomyces coeliclor*	NP_631415.1	21225636	SEQ ID NO:91	SEQ ID NO:92
ADH-TB	*Thermoanaerobium brockii*	X64841.1	1771790	SEQ ID NO:101	SEQ ID NO:102

ADH-CP	*Candida parapsilosis*	BAA24528	2815409	-	Julich Chiral Solutions, nº cat. 03.11
DR-LB	diacetil reductasa de *Lactobacillus brevis*	ABJ63353.1	116098204	-	Julich Chiral Solutions, nº cat. 8.1
ADH-HE	Hígado de caballo	DEHOAL	625197	SEQ ID NO:103	SEQ ID NO:104
ADH-CB	*Candida boidinii*	CAD66648	28400789	-	Julich Chiral Solutions, nº cat. 02.10
LDH-LL	*Lactobacillus leichmannii*	-	-	-	Fluka, nº cat. 61306
ADH-AF	*Aspergillus flavus*	P41747	1168346	SEQ ID NO:95	SEQ ID NO:96
ADH-OO1	*Oenococcus oeni*	ZP_00318704.1	48864831	SEQ ID NO:97	SEQ ID NO:98
ADH-RU	*Ralstonia eutropha*	ZP_00202558.1	46131317	SEQ ID NO:99	SEQ ID NO:100

Como se ha indicado anteriormente, los polinucleótidos que codifican las cetorreductasas manipuladas de la presente invención también pueden clonarse en el vector pCK1 10900 para la expresión en *E. coli* W3110.

Ejemplo 2: Producción de polvos de cetorreductasa - Método de agitación en matraz.

Se inoculó una sola colonia microbiana de *E. coli* que contenía un plásmido que codificaba una cetorreductasa de interés en 50 ml de caldo Luria Bertani que contenía 30 µg/ml de cloranfenicol y glucosa al 1%. Las células se cultivaron durante la noche (al menos 16 horas) en una incubadora a 30°C con agitación a 250 rpm. El cultivo se diluyó en 250 ml de Terrific Broth (12 g/l de bacto-triptona, 24 g/l de extracto de levadura, 4 ml/l de glicerol, fosfato potásico 65 mM, pH 7,0, $MgSO_4$ 1 mM) que contenía 30 µg/ml de cloranfenicol, en un matraz de 1 litro a una densidad óptica a 600 nm (DO_{600}) de 0,2 y se dejaron crecer a 30°C. La expresión del gen de cetorreductasa se indujo añadiendo iso-propil-β-D-tiogalactósido ("IPTG") a una concentración final de 1 mM cuando la DO_{600} del cultivo era de 0,6 a 0,8 y, a continuación, se continuó la incubación durante la noche (al menos 16 horas). Las células se recogieron por centrifugación (5000 rpm, 15 min., 4°C) y se descartó el sobrenadante. El sedimento celular se resuspendió con un volumen igual de tampón (cloruro de) trietanolamina 100 mM frío (4°C), pH 7,0 (que incluía $MgSO_4$ 2 mM en el caso de ADH-LK (SEQ ID NO: 4) y ADH-LB (SEQ ID NO: 2) y las cetorreductasas manipuladas derivadas de las mismas), y se recolectaron por centrifugación como anteriormente. Las células lavadas se volvieron a suspender en dos volúmenes de tampón (cloruro de) trietanolamina frío y se hicieron pasar por una prensa francesa dos veces a 12.000 psi, mientras que se mantenían a 4°C. Los restos celulares se eliminaron por centrifugación (9.000 rpm, 45 min., 4°C). El sobrenadante del lisado transparente se recogió y se almacenó a -20°C. La liofilización del lisado transparente congelado proporcionó un polvo seco de la enzima cetorreductasa bruta. Como alternativa, el sedimento celular (antes o después del lavado) se almacenó a 4°C ó -80°C.

Ejemplo 3: Producción de cetorreductasas - Procedimiento de fermentación.

Se realizaron fermentaciones a escala experimental a 30°C en un fermentador de 15 litros, agitado y aireado, utilizando 6,0 litros de medio de crecimiento (0,88 g/l de sulfato de amonio, 0,98 g/l de citrato sódico; 12,5 g/l de hidrogenofosfato dipotásico trihidratado, 6,25 g/l de dihidrogenofosfato potásico, 6,2 g/l de extracto de levadura Tastone-154, 0,083 g/l de citrato de amonio férrico, y 8,3 ml/l de una solución de elementos traza que contenía 2 g/l de cloruro cálcico dihidratado, 2,2 g/l de sulfato de zinc heptahidratado, 0,5 g/l de sulfato de manganeso monohidratado, 1 g/l de sulfato cuproso heptahidratado, 0,1 g/l de molibdato de amonio tetrahidratado y 0,02 g/l de tetraborato sódico). El fermentador se inoculó con un cultivo exponencial tardío de *E. coli* W3110 que contenía un plásmido que codificaba el gen de cetorreductasa de interés (cultivadas en un matraz de agitación como se describe en el Ejemplo 2) a una DO_{600} de partida de 0,5 a 2,0. El fermentador se agitó a 500-1.500 rpm y se suministró aire al recipiente de fermentación a 1,0-15,0 l/min para mantener un nivel de oxígeno disuelto con una saturación del 30% o superior. El pH del cultivo se mantuvo a 7,0 añadiendo hidróxido de amonio al 20% v/v. El crecimiento del cultivo se mantuvo añadiendo una solución de alimentación que contenía 500 g/l de cerelosa, 12 g/l de cloruro de amonio y 10,4 g/l de sulfato de magnesio heptahidratado. Una vez que el cultivo alcanzó una DO_{600} de 50, se indujo la expresión de cetorreductasa añadiendo isopropil-β-D-tiogalactósido (IPTG) a una concentración final de 1 mM y la fermentación continuó durante otras 14 horas. A continuación, se enfrió el cultivo a 4°C y se mantuvo a esa temperatura hasta la recolección. Las células se recogieron por centrifugación a 5.000 G durante 40 minutos en una centrífuga Sorval RC12BP a 4°C. Las células recolectadas se utilizaron directamente en el siguiente proceso de recuperación corriente abajo o pueden almacenarse a 4°C o congelarse a -80°C hasta tal uso.

El sedimento celular se resuspendió en 2 volúmenes de tampón (cloruro de) trietanolamina 100 mM, pH 6,8, a 4°C por cada volumen de pasta celular húmeda. La cetorreductasa intracelular se liberó de las células haciendo pasar la suspensión por un homogeneizador equipado con un conjunto de válvula de homogeneización de doble etapa utilizando una presión de 12.000 psig. El homogeneizado celular se enfrió a 4°C inmediatamente después de la rotura. Se añadió al lisado una solución de polietilenimina al 10% p/v, pH 7,2, a una concentración final del 0,5% p/v y se agitó durante 30 minutos. La suspensión resultante se aclaró por centrifugación a 5.000 G en una centrífuga de laboratorio convencional durante 30 minutos. El sobrenadante transparente se decantó y se concentró diez veces utilizando una membrana de ultrafiltración de celulosa con un peso molecular límite de 30 kD. El concentrado final se distribuyó en recipientes poco profundos, se congeló a -20°C y se liofilizó hasta un polvo. El polvo de cetorreductasa se almacenó a -80°C.

Ejemplo 4: Métodos analíticos: Conversión de *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina ("DMAK") en (*S*)-*N,N*-dimetil-3-hidroxi-3 -(2-tienil)-1-propanamina ("(S)-DMAA")

Método de HPLC aquiral para determinar la conversión de *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina en (*S*)-*N,N*-dimetil-3-hidroxi-3-(2 tienil)-1-propanamina ("DMAK"): se determinó la reducción de DMAK (preparada como se describe en el documento EP0273658) a (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina ("(*S*)-DMAA") utilizando un Agilent 1100 HPLC equipado con una columna Agilent Zorbax 5 µm SB-AQ (15 cm de longitud, 2,1 mm de diámetro, eluyente: 8:2 de NH_4Ac/MeCN 40 mM, caudal 0,4 ml/min, y temperatura de la columna de 50°C). Tiempos de retención: (*S*)-DMAA: 2,0 min.; DMAK: 2,8 min. El producto ((*S*)-DMAA) se detectó como el área del pico a 235 nm y mientras que la del sustrato DMAK se determinó a 290 nm.

Método de HPLC quiral para determinar la pureza estereomérica del producto DMAA: Se determinó la pureza estereomérica de DMAA utilizando un Agilent 1100 HPLC equipado con una columna Chiralcel OJ-H (15 cm de longitud, 2,1 mm de diámetro) utilizando como eluyente 95:5:0,1 de heptano/etanol/dietilamina a un caudal de 0,5 ml/min; y a una temperatura de la columna de 40°C). Tiempos de retención: (*S*)-DMAA: 2,2 min., (*R*)-DMAA: 2,6 min., DMAK: 2,0 min.

Ejemplo 5: Selección celular, crecimiento y expresión inducida de enzimas variantes de ADH-LK

Se realizó una recolección robotizada de colonias individuales con un instrumento Q-Bot™ (Genetix, USA Inc., Boston, MA) a 180 µl de caldo Luria-Bertani (LB) que contenía glucosa al 1% y 30 µg/ml de cloranfenicol (CAM) en una placa de 96 pocillos NUNC® (Nalge Nunc International, Rochester NY). La placa (la "placa maestra") se selló con cinta AirPore (Qiagen, Valencia CA), y se incubó durante la noche a 30°C a 250 rpm (recorrido de 2 pulgadas) a una humedad relativa del 85%. Las placas maestras se subcultivaron inoculando una alícuota de 10 µl de cada pocillo en un pocillo de una placa de pocillos profundos Costar® (Corning®, Acton MA) que contenía 390 µl de Terrific Broth, pH 7,0, complementado con $MgSO_4$ 1 mM y 30 µg/ml de cloranfenicol (CAM). Las placas de pocillos profundos Costar® inoculadas se incubaron durante tres horas a 30°C, una humedad relativa del 85%, a 250 rpm en un agitador con un recorrido de 2 pulgadas. A continuación, se añadió a cada pocillo el inductor IPTG a una concentración final de 1 mM y la incubación continuó durante 18 horas más. Las células se recolectaron centrifugando las placas de pocillos profundos Costar® a 4.000 rpm durante 10 minutos, desechando el sobrenadante. Generalmente, los sedimentos se congelaron durante una hora antes de la lisis.

Se añadió glicerol a los pocillos de la placa maestra a una concentración final del 20%. A continuación, las placas maestras se almacenaron a -80°C.

Ejemplo 6: Cribado previo para la selección de cetorreductasas capaces de reducir el isopropanol en presencia de $NADP^+$ produciendo NADPH y acetona

Se recolectaron colonias recombinantes de *E. coli* que portaban un gen que codificaba ADH-LK o una variante de la misma utilizando un recolector de colonias robotizado Q-bot® (Genetix USA, Inc, Boston, MA) en placas de microtitulación de pocillos poco profundos de 96 pocillos que contenían 180 µl de Terrific Broth (TB), glucosa al 1% y 30 µg/ml de cloranfenicol (CAM). Las células se cultivaron durante la noche a 30°C con agitación a 200 rpm. A continuación, se transfirió una alícuota de 10 µl de este cultivo a placas de pocillos profundos de 96 pocillos que contenían 390 µl de Terrific Broth (TB), $MgSO_4$ 1 mM y 30 µg/ml de CAM. Después incubar las placas de pocillos profundos a 30°C con agitación a 250 rpm durante 2 a 3 horas, se indujo la expresión de los genes recombinantes dentro de las células cultivadas añadiendo IPTG a una concentración final de 1 mM. A continuación, las placas se incubaron a 30°C con agitación a 250 rpm durante 18 horas.

Se observó que el sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina experimenta una reacción de eliminación (produciendo, por ejemplo, 1-(tiofen-2-il)prop-2-en-1-ona) a pH 8 que puede ir seguida de reducción enzimática para proporcionar varios productos adicionales, por ejemplo, 1-(tiofen-2-il)propan-1-ona, 1-(tiofen-2-il)propan-1-ol, y 1-(tiophen-2-il)prop-2-en-1-ol. Sin embargo, una alta concentración de isopropanol (> 75%) reduce la eliminación y elevar el pH a los intervalos de 9 a 11 evita la reacción de eliminación.

Las células se sedimentaron por centrifugación (4.000 rpm, 10 min., 4°C), se resuspendieron en 400 µl de tampón de lisis y se lisaron mediante agitación a temperatura ambiente durante 2 horas. Cuando el cribado se realizaba a pH 11, las células se lisaron añadiendo a cada pocillo 400 µl de tampón de lisis (tampón (cloruro de) trietilamina 100 mM, pH 11, 1 mg/ml de lisozima, 500 µg/ml de sulfato de polimixina B ("PMBS") y $MgSO_4$ 1 mM). Con variantes de ADH-LK más activas, el cribado se realizó a pH 9, las células se lisaron añadiendo a cada pocillo 400 µl de tampón de lisis ($MgSO_4$ 1 mM, 500 µg/ml de sulfato de polimixina B y 1 mg/ml de lisozima en (cloruro de) trietanolamina 100 mM pH 9). Después de sellar las placas con cinta termosellante laminada de aluminio/polipropileno (Velocity 11 (Menlo Park, CA), nº cat. 06643-001), se agitaron enérgicamente durante 2 horas a temperatura ambiente. Los restos celulares se recogieron por centrifugación (4.000 rpm, 10 min., 4°C) y el sobrenadante transparente se ensayó directamente o se almacenó a 4°C hasta su uso.

En este ensayo, se añadieron 20 µl de muestra (diluida en tampón (cloruro de) trietilamina o (cloruro de) trietanolamina 100 mM, al mismo pH que el tampón de lisis, y $MgSO_4$ 1 mM) a 180 µl de una mezcla de ensayo en placas de microtitulación de color negro de 96 pocillos. El tampón de ensayo consiste en tampón (cloruro de) trietilamina 100 mM (pH 11) o tampón (cloruro de) trietanolamina 100 mM (pH 9), de alcohol isopropílico (IPA) al 50%, $MgSO_4$ 1 mM y 222 µM de $NADP^+$. Se hizo seguimiento de la reacción midiendo la reducción de la fluorescencia del $NADP^+$, a medida que se convertía en NADPH utilizando un instrumento Flexstation® (Molecular Devices, Sunnyvale, CA). Se midió fluorescencia del NADPH a 445 nm tras excitación a 330 nm. Para identificar las variantes de ADH-LK termoestables, se preincubaron muestras de los lisados durante 18 a 24 horas a 25°C-40°C en presencia o ausencia de IPA al 50% antes de añadirse a la mezcla de ensayo.

Ejemplo 7: Evaluación de cetorreductasas de tipo silvestre para la reducción de *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina a (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina

Se cribaron las KRED descritas en la Tabla 3 del Ejemplo 1 utilizando como cofactores NADH y NADPH, y glucosa deshidrogenasa/glucosa o alcohol isopropílico ("IPA") como sistema de regeneración de cofactor. Se añadieron 100 µl de lisado celular a una placa de pocillos profundos (Costar # 3960) que contenía 25 µl 5 mg/ml de Na-NADP (Oriental Yeast) y $MgSO_4$ 2 mM en (cloruro de) trietilamina 100 mM (pH 11,0), y 125 µl de alcohol isopropílico que contenía 2 g/l de *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina. Después de sellar las placas con cinta termosellante laminada de aluminio/polipropileno (Velocity 11 (Menlo Park, CA), nº cat. 06643-001), las reacciones se llevaron a cabo durante al menos 16 horas a temperatura ambiente. Al final de la reacción, se añadió a cada pocillo NaOH 6N para ajustar el pH a > 12 seguido de adición de 1 ml de MTBE por pocillo. Las placas se volvieron a sellar, se agitaron durante 20 minutos, y se centrifugaron (4.000 rpm, 10 min., 4°C). Se transfirieron 200 µl de la capa orgánica a una nueva placa de microtitulación de pocillos poco profundos para el análisis como se describe en el Ejemplo 4. Ninguna de las enzimas ensayadas mostró una conversión > 1% de *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina.

Este ejemplo demuestra que las cetorreductasas de tipo silvestre tienen muy poca o ninguna actividad sobre la *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina.

Ejemplo 8: Identificación de una enzima ADH-LK variante capaz de convertir la *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina en (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina con > 90% de exceso enantiomérico ("ee").

Los experimentos iniciales indicaron que la cetorreductasa/alcohol deshidrogenasa (ADH-LK) "de tipo silvestre" (SEQ ID NO: 4) no reducía la *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina a (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina de forma significativa. Sin embargo, el cribado de diversas variantes de ADH-LK como se describe en la solicitud de EE.UU. con nº de serie 60/957.974 (pub. de pat. de EE.UU. 2009/0093031A, que corresponde a USSN 12/197.286), nº de serie 60/972.058 (pub. de pat. de EE.UU. 2009/0191605, que corresponde a USSN 12/210.195), y nº de serie 60/976.345 (pub. de pat. de EE.UU. 2009/0155863, que corresponde a USSN 12/240.986) identificó una variante de esa enzima que portaba 12 sustituciones de aminoácidos (es decir, G7H; D66E; A94G; S96I; E145L; F147L; Y190P; L195R; V196L; A202W; K211L; y Y249W; (SEQ ID NO: 6)) capaz de convertir la *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina en (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina a una velocidad de 0,5 g/l de producto al día por gramo de enzima, teniendo > 90% de ese producto la configuración (*S*).

Ejemplo 9: Identificación de variantes de ADH-LK con una actividad y estereoselectividad mejoradas para la reducción de *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina a (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina

Lisis celular: Los sedimentos celulares (recogidos en los pocillos de una placa de microtitulación) se lisaron añadiendo a cada pocillo 400 µl de tampón de lisis ($MgSO_4$ 1 mM, 0,5 mg/ml de PMBS, 1 mg/ml de lisozima, trietilamina 100 mM (pH 11)). Las placas se sellaron con cinta termosellante laminada de aluminio/polipropileno (Velocity 11 (Menlo Park, CA), nº cat. 06643-001), se agitaron enérgicamente durante dos horas a temperatura ambiente, y, a continuación, se centrifugaron a 4.000 rpm durante 20 minutos a 4°C. Los sobrenadantes se recuperaron y se almacenaron a 4°C hasta el ensayo.

Reacción de reducción enzimática: Se añadió, mediante robot, una alícuota (300 µl) de mezcla de reacción concentrada a cada pocillo de una placa de pocillos profundos Costar® utilizando un instrumento Multidrop (MTX Lab Systems, Viena VA), seguido de la adición robotizada de 200 µl del sobrenadante de lisado recuperado utilizando un dispensador de Multimek™ (Multimek, Inc. Santa Clara CA), para proporcionar una reacción que comprendía 1 mg/ml de sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina, 0,2 mg/ml de $NADP^+$, trietilamina 25 mM pH 11, y isopropanol al 35% (v/v). Las placas se termosellaron con cinta termosellante laminada de aluminio/polipropileno (Velocity 11 (Menlo Park, CA), nº cat. 06643-001) a 180°C durante 2,5 segundos y, a continuación, se agitaron durante la noche (al menos 16 horas) a temperatura ambiente.

Las reacciones se interrumpieron añadiendo 1 ml de metil-*t*-butil éter ("MTBE") y se volvieron a sellar las placas. Las placas selladas se agitaron durante 20-30 minutos, y las capas orgánicas y acuosas se separaron por centrifugación (4.000 rpm, 5 min., a temperatura ambiente). Se transfirió una alícuota de 250 µml de la capa orgánica de cada pocillo a los pocillos de una nueva placa de polipropileno (Costar # 3365) (Corning®, Acton MA) y después de volver a sellar, las muestras de la placa se sometieron a análisis HPLC como se describe, por ejemplo, en el Ejemplo 4.

Para un cribado de alto rendimiento a un pH de 9, se transfirieron 50 µl de lisado celular a una placa de pocillos profundos (Costar # 3960; Corning®, Acton MA) que contenía 200 µl de mezcla de ensayo (por 100 ml: 37,5 ml de (cloruro de) trietanolamina 100 mM (pH 9), 12,5 mg Na-NADP (Oriental Yeast, Andover, MA), 15,1 g de DMAK, y 62,5 ml de alcohol isopropílico). Después de sellar las placas con cinta termosellante laminada de aluminio/polipropileno (Velocity 11 (Menlo Park, CA), nº cat. 06643-001), las reacciones se llevaron a cabo durante 4-24 horas a temperaturas que iban de temperatura ambiente a 40°C. Las reacciones se interrumpieron añadiendo a cada pocillo 100 µl de NaOH 1 N, seguido de 1 ml de MTBE. Las placas se volvieron a sellar, se agitaron durante 5 minutos, y las capas orgánica y acuosa se separaron por centrifugación (4.000 rpm, 5 min., a temperatura ambiente). Se transfirieron 25 µl de la capa orgánica de cada pocillo a los pocillos de una nueva placa de polipropileno de pocillos poco profundos (Costar # 3365; Corning®, Acton MA) que contenía cada uno 225 µl de MTBE. Después de volver a sellar las placas, las muestras se sometieron a análisis HPLC como se describe en el ejemplo 4.

Las variantes de la cetorreductasa de *Lactobacillus kefir* (ADH-LK) capaces de convertir la *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina en (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina pueden identificarse utilizando los métodos anteriormente indicados. Se utilizó la iteración de estos procedimientos a través de varias rondas, en las que se utilizaron como material de partida uno o más aislados mejorados de una ronda para las siguientes rondas de mutagénesis y cribado, para desarrollar o "evolucionar" las variantes de cetorreductasa (ADH-LK) de *Lactobacillus kefir* con una capacidad mejorada para reducir la *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina de manera estereoespecífica a (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina, por ejemplo, una enzima capaz de convertir más del 99% del sustrato (a una carga de 100 g/l) en el correspondiente alcohol quiral de más del 99% de exceso enantiomérico (ee).

Ejemplo 10: Reducción estereoselectiva de *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina ("DMAK") utilizando alcohol isopropílico para la regeneración de cofactor mediante cetorreductasas manipuladas derivadas de ADH-LK

Se evaluaron a escala preparativa cetorreductasas mejoradas derivadas de variantes de ADH-LK para la reducción de DMAK de la siguiente manera. Se añadieron 200 µl de solución de una variante de ADH-LK a ensayar (de 5 mg/ml a 25 mg/ml) y NADP-Na (0,5 mg/ml) en (cloruro de) trietanolamina 100 mM pH 9, a un vial de reacción de 5 ml equipado con una barra de agitación magnética. Posteriormente, se añadieron a la solución de enzima/NADP-Na 800 µl de una solución de DMAK (base libre; 187,5 mg/ml) en alcohol isopropílico al 62,5% y (cloruro de) trietanolamina 100 mM al 37,5% pH 9 (ajustado después de la adición de DMAK), y se incubaron con agitación a 40°C. La reacción se muestreó periódicamente y se analizó con los métodos de HPLC descritos anteriormente. En la Tabla 2 se resumen las características de las variantes de ADH-LK identificadas y aisladas mediante los métodos que se describen en el presente documento. En la Tabla 2 se proporciona la SEQ ID NO correspondiente a las variantes de cetorreductasa, el número de mutaciones de aminoácidos a partir de la ADH-LK de tipo silvestre, y la actividad reductora de DMAK de cada una, con respecto a la de la enzima con la secuencia de aminoácidos de la SEQ ID NO: 6.

Este ejemplo ilustra que las cetorreductasas manipuladas derivadas de la cetorreductasa ADH-LK proporcionan actividades mejoradas en comparación con la cetorreductasa ADH-LK de tipo silvestre para la reducción de DMAK.

Ejemplo 11: Producción de DMAA a escala preparativa

Se añadió una solución de NaOH al 25% (40 ml), IPA (15 ml) y clorhidrato de *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina (21,6 g) a un matraz de tres bocas de fondo redondo de 250 ml equipado con un agitador magnético, septo y sistema de destilación de disolvente. La suspensión resultante se agitó hasta que quedaron dos fases transparentes. La capa inferior se aspiró con pipeta y se añadió agua (40 ml) a la capa superior, seguido de ajuste del pH a 9 por adición de H_2SO_4 concentrado. A continuación, se añadieron 40 ml de IPA y una solución de $NADP^+$ (12 mg) y la cetorreductasa de la SEQ ID NO: 60 (120 mg) preparada en 15 ml de (cloruro de) trietanolamina

100 mM pH 9 para iniciar la reacción. El contenido de la reacción se calentó a 38ºC-40ºC con un bloque térmico OptiTherm (ajuste a 60°C). Durante el calentamiento, se aplicó vacío (~ 100 mmHg) para destilar el IPA/acetona a 38°C-40°C. El bloque térmico se ajustó a aproximadamente 52ºC-55ºC para destilar lentamente el IPA/acetona (~ 10 ml/hr.). Se añadió periódicamente IPA (total de 100 ml - 150 ml, IPA al 90%, agua al 10%) para reemplazar el que se destilaba. Después de 9 horas, se detuvo el calentamiento, se liberó el vacío y se dejó que la reacción continuase durante otras 10-20 horas a temperatura ambiente. El IPA restante (55 ml - 65 ml) se retiró por destilación durante 2-3 horas a una temperatura del recipiente de 52°C con el mismo ajuste de vacío utilizado para la reacción. Durante la destilación, se añadieron 10 ml de agua que contenía NaOH al 50% (1 g). La mezcla también se sembró con cristales de producto en agua en ese momento. Cuando la temperatura del recipiente alcanzó 52°C, se liberó el vacío y la mezcla se enfrió a 25°C. La suspensión resultante se volvió a calentar brevemente a 40°C y, a continuación, se enfrió lentamente (30 minutos a temperatura ambiente y 30 minutos en un baño de hielo). La suspensión resultante se filtró utilizando un embudo Buchner de vidrio sinterizado de porosidad gruesa (40µ -60µ) y la torta del filtro se lavó con agua fría. El sólido granulado blanco se secó hasta peso constante a vacío a 40°C durante 5 horas, lo que produjo un sólido blanco (16,4 g, rendimiento del 90%) de una pureza del 99,3% de área según lo determinado mediante HPLC (detección a 235 nm) encontrándose presente cetona al 0,7%. La cetona puede eliminarse por recristalización. Por ejemplo, se disolvieron 20,0 g de producto bruto que contenía cetona al 3% en 200 ml de EtOAc/heptano 1:1 mediante calentamiento a 40°C. La suspensión se filtró a través de un lecho de celite y, a continuación, se destiló para eliminar el EtOAc (temperatura de cabeza de 80°C al final de la destilación). La solución transparente resultante se dejó enfriar y el producto se cristalizó a aproximadamente 50°C. El enfriamiento adicional (baño de hielo, 30 minutos) y la filtración proporcionaron 17,14 g (86% de rendimiento recuperado, p.f. 76,0°C-77,5°C) de producto puro (HPLC, ^{1}H-RMN).

Ejemplo 12: Reducción enzimática de *N,N*-dimetil-3-ceto-3-fenil-1-propanamina

Se añadió una solución 200 µl de la cetorreductasa de la SEQ ID NO: 60 (5 mg/ml) y NADP-Na (0,5 mg/ml) en (cloruro de) trietanolamina 100 mM pH 9, a un vial de reacción de 5 ml equipado con una barra de agitación magnética. Posteriormente, se añadió a la solución de enzima/NADP-Na una solución de 800 µl de la sal de clorhidrato del sustrato (62,5 mg/ml) en alcohol isopropílico al 62,5% y TEA-cloruro 100 mM al 37,5% pH 9, ajustado después de la adición del sustrato, y se agitó a temperatura ambiente. En estas condiciones de reacción, la cetorreductasa de la SEQ ID NO: 60 fue capaz de reducir la *N,N*-dimetil-3-ceto-3-fenil-1-propanamina, es decir, después de 72 horas a temperatura ambiente, la conversión en alcohol era detectable a pesar de que podía ser < 3%.

Este ejemplo ilustra que las cetorreductasas manipuladas derivadas de la cetorreductasa ADH-LK son capaces de reducir la *N,N*-dimetil-3-ceto-3-fenil-1-propanamina.

LISTADO DE SECUENCIAS

<110> SAVILE, Christopher GRUBER, John M. MUNDORFF, Emily HUISMAN, Gjalt Collier, Steven

<120> POLIPÉPTIDOS CETORREDUCTASA PARA LA PRODUCCIÓN DE UNA 3-ARIL-3-HIDROXIPROPANAMINA A PARTIR DE UNA 3-ARIL-3-CETOPROPANAMINA

<130> CX2-024

<150> 61/092.321

<151> 27-08-2008

<160> 110

<170> PatentIn versión 3.5

<210> 1
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Secuencia de *L. kefir* con optimización de codones

<400> 1

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggactct gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cagtttccaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggtt cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggctat atcaagaccc cgctggtcga tgatctggaa 600
ggtgctgagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga                        759
```

<210> 2
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Traducción de secuencia de *L. kefir* con optimización de codones

<400> 2

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


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

<210> 3
<211> 762
<212> ADN
<213> Secuencia artificial

<220>

<223> Secuencia de *L. brevis* con optimización de codones

<400> 3

```
atgtctaacc gtctggatgg caaagtagcc atcattaccg gcgggactct gggtatcggt 60
ttggcaatcg ccacgaaatt tgtagaggag ggtgcgaaag taatgattac tggtcgtcac 120
tccgatgtag gtgaaaaggc cgccaaatca gtaggcactc cggatcagat tcagtttttt 180
cagcacgatt catccgatga agatggctgg acgaaactgt tcgacgccac cgagaaagca 240
ttcggcccgg ttagcacctt agtgaacaat gcagggattg cagttaacaa aagcgttgaa 300
gaaactacca cggccgaatg gcgtaaactg ctggccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggtt cgtaggcgat ccgagcctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggctat atcaagaccc cgctggtcga tgatctgccg 600
ggtgctgagg aagcgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catatatctg tgtgtacctg gcatctaatg aatcgaaatt tgcgacgggt 720
tccgaatttg tggtcgacgg cgggtatacc gcacagtaat ga                    762
```

<210> 4
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Traducción de secuencia de *L. brevis* con optimización de codones

<400> 4

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

            100                 105                 110
Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg
        115                 120                 125
Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
    130                 135                 140
Glu Gly Phe Val Gly Asp Pro Ser Leu Gly Ala Tyr Asn Ala Ser Lys
145                 150                 155                 160
Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu
                165                 170                 175
Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Tyr Ile Lys
            180                 185                 190
Thr Pro Leu Val Asp Asp Leu Pro Gly Ala Glu Glu Ala Met Ser Gln
        195                 200                 205
Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
    210                 215                 220
Tyr Ile Cys Val Tyr Leu Ala Ser Asn Glu Ser Lys Phe Ala Thr Gly
225                 230                 235                 240
Ser Glu Phe Val Val Asp Gly Gly Tyr Thr Ala Gln
                245                 250
```

<210> 5
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 5

```
atgaccgatc gtctgaagca taaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgaaga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttattaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttctggggct cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgcggctcga tgatctggaa 600
ggttgggagg aaatgatgtc acagcgtacg ttaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 6
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 6

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

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

<210> 7
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 7

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttctggggtt cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggtgctgagg aaatgatgtc acagcgtacg aaaacccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                         759
```

<210> 8
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 8

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


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

<210> 9
<211> 759
<212> ADN
<213> Secuencia artificial

<220>

<223> Variante de *L. kefir*

<400> 9

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta tttttgggtt cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggtgctgagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 10
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 10

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


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

<210> 11
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 11

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttctggggct ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggtgctgagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 12
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 12

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


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

<210> 13
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 13

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttctggggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggtgctgagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 14
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 14

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

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

<210> 15
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 15

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttctgggggt tgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggtgctgagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 16
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 16

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


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

<210> 17
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 17

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttctggggat tgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggtgctgagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 18
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 18

```
<400> 18
Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
1               5                   10                  15
Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
            20                  25                  30
Lys Val Val Ile Thr Gly Arg His Ala Asp Val Gly Glu Lys Ala Ala
        35                  40                  45
Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
    50                  55                  60
Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
65                  70                  75                  80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Gly Val Val
                85                  90                  95
Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser

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

<210> 19
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 19

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggct ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggtgctgagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 20
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 20

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


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

<210> 21
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 21

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggtgctgagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 22
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 22

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


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

<210> 23
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 23

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat tgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggtgctgagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 24
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 24

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


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

<210> 25
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 25

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggtgctgagg aaatgtattc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 26
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 26

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


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

<210> 27
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 27

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcgc 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggtgctgagg aaatgttttc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 28
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 28

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


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

<210> 29
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 29

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttggggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggtgctgagg aaatgtggtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 30
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 30

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


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

<210> 31
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 31

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggtgctgagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 32
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 32

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


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

<210> 33
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 33

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgactaccg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgattatgaa 600
ggtgctgagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 34
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 34

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

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

<210> 35
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 35

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgatctggaa 600
ggtgctgagg aaatgtggtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctggtg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 36
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 36

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


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

<210> 37
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 37

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgatctggaa 600
ggtgctgagg aaatgtggtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 38
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 38

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


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

<210> 39
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 39

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgactctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgatctggaa 600
ggtgctgagg aaatgtattc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctggtg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 40
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 40

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


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

<210> 41
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 41

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgtctaccg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggtgctgagg aaatgatgtc acagcgtacg aaaacccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 42
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 42

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


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

<210> 43
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 43

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgtctctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggtgctgagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 44
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 44

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


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

<210> 45
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 45

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcgc 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggtgctgagg aaatgtggtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 46
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 46

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


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

<210> 47
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 47

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgattgggaa 600
ggtgctgagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 48
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 48

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


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

<210> 49
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 49

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcgc 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgactaccg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgctgctcga tgatctggaa 600
ggtgctgagg aaatgatgtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 50
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 50

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


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

<210> 51
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 51

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta tttacgggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgatctggaa 600
ggtgctgagg aaatgtggtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctggtg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 52
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 52

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


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

<210> 53
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 53

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttggggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgatctggaa 600
ggtgctgagg aaatgtggtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctggtg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 54
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 54

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


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

<210> 55
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 55

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgacggtgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgatctggaa 600
ggtgctgagg aaatgtggtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctggtg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                         759
```

<210> 56
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 56

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


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

<210> 57
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 57

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggccgtcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgatgtggaa 600
ggtgctgagg aaatgtggtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctggtg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 58
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 58

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


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

<210> 59
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 59

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggccgtcgt 120
gcggatgtag gtgaacgtgc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgacggttg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgatgtggaa 600
ggtgctgagg aaatgtggtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctggtg aatcgaaatt tgcgacgggt 720
gcagaatttg tgattgacgg cgggtggacc gcacagtga                        759
```

<210> 60
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 60

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


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

<210> 61
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 61

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggccgtcgt 120
gcggatgtag gtgaacgtgc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgccag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgacggttg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acagcgtcca tccgggcccc atcaagaccc cgatgctcga tgatgtggaa 600
ggtgctgagg aaatgtggtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctggtg aatcgaaatt tgcgacgggt 720
gcagaatttg tgattgacgg cgggtggacc gcacagtga                        759
```

<210> 62
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 62

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


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

<210> 63
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 63

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggccgtcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgactctgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcacaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggca tgggcgctag catcatcaat 420
gtgagcagta ttttcgggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acagcgtcca tccgggcccc atcaagaccc cgatgctcga tgatgtggaa 600
ggtgctgagg aaatgtggtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctggtg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 64
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 64

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


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

<210> 65
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 65

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggccgtcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaatgcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgatgtggaa 600
ggtgctgagg aaatgtggtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgctg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 66
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 66

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


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

<210> 67
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 67

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggccgtcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagtg ttttcgggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgcttga tgatgtggaa 600
ggtgctgagg aaatgtggtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctggtg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 68
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 68

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


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

<210> 69
<211> 759
<212> ADN
<213> Secuencia artificial

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggccgtcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcacaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgacggttg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgatgtggaa 600
ggtgctgagg aaatgtggtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctggtg aatcgaaatt tgcgacgggt 720
gcagaatttg tgattgacgg cgggtggacc gcacagtga                        759
```

<210> 70
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 70

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


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

<210> 71
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 71

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggccgtcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgacggttg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acaccatcca tccgggcccc atcaagaccc cgatgctcga tgatgtggaa 600
ggtgctgagg aaatgtggtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctggtg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga                        759
```

<210> 72
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 72

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


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

<210> 73
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 73

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggccgtcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcacaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgacggttg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acagcgtcca tccgggcccc atcaagaccc cgatgctcga tgatgtggaa 600
ggtgctgagg aaatgtggtc acagcgtacg aaaacccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctggtg aatcgaaatt tgcgacgggt 720
gcagaatttg tgattgacgg cgggtggacc gcacagtga                         759
```

<210> 74
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 74

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


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

<210> 75
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 75

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggccgtcgt 120
gcggatgtag gtgaacgtgc cgccaaatca atcggcggca ctgatgttat tcgctttatt 180
cagcacgatg cgtccgatga agcaggctgg acgactctgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgacggttg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgatgtggaa 600
ggtgctgagg aaatgtggtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctggtg aatcgaaatt tgcgacgggt 720
gcagaatttg tgattgacgg cgggtggacc gcacagtga                        759
```

<210> 76
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 76

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


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

<210> 77
<211> 759
<212> ADN
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 77

```
atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggtatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggccgtcgt 120
gcggatgtag gtgaacgtgc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcacaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgacggttg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagaccc cgatgctcga tgatgtggaa 600
ggtgctgagg aaatgtggtc acagcgtacg aaaaccccta tgggccacat tggcgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctggtg aatcgaaatt tgcgacgggt 720
gcagaatttg tgattgacgg cgggtggacc gcacagtga                        759
```

<210> 78
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 78

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


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

<210> 79
<211> 852
<212> ADN
<213> Secuencia artificial

<220>
<223> ADH de *Candida magnoliae*

<400> 79

```
atggcaaaga attttagcaa tgtagagtat cccgcacccc cccccgcaca tacaaagaat 60
gagagcttac aagtattaga tttatttaag ttaaatggaa aagtagcaag cataacagga 120
agcagcagcg gaataggata tgcattagca gaggcttttg cacaagtcgg agcagatgta 180
gcaatatggt ataatagcca tgatgcaaca ggaaaagcag aggcattagc aaagaagtat 240
ggagtaaagg taaaggcata taaagcaaat gtaagcagca gcgatgcagt caagcaaaca 300
atagagcaac aaataaagga ttttggacat ttagatatag tagtagcaaa tgcaggaata 360
ccctggacaa agggagcata tatagatcaa gatgatgaca agcattttga ccaagtagta 420
gatgtagact taaagggagt aggatacgta gcaaagcatg caggaaggca ttttagggaa 480
aggtttgaga aagagggaaa aaagggagca ttagtattta cagcaagcat gagcggacat 540
atagtaaatg tcccccaatt ccaagcaaca tataatgcag caaaggcagg agtaaggcat 600
tttgcaaaga gcttagcagt cgagtttgca ccctttgcaa gggtaaatag cgtaagcccc 660
ggatatataa atacagagat aagcgatttc gtcccccaag agacacaaaa taagtggtgg 720
agcttagtcc ccttaggaag gggaggagag acagcagagt tagtaggagc atatttattc 780
ttagcaagcg atgcaggaag ctatgcaaca ggaacagata taatagtaga tggaggatat 840
acattaccct aa                                                     852
```

<210> 80
<211> 283
<212> PRT
<213> Secuencia artificial

<220>
<223> ADH de *Candida magnoliae*

<400> 80

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


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

<210> 81
<211> 1047

<212> ADN
<213> Secuencia artificial

<220>
<223> ADH de *Rhodococcus erythropolis*

<400> 81

```
atgaaagcca ttcagtacac tcgtatcggt gcggaaccag aactgactga aatcccgaag 60
ccggaaccgg gcccgggcga agtactgctg gaagtcacgg cagctggcgt gtgccattcc 120
gatgatttca ttatgtctct gccggaagaa cagtacacct acggcctgcc gctgaccctg 180
ggtcatgaag gtgctggtaa agttgccgca gttggcgaag gtgttgaagg gttggatatt 240
ggcaccaatg tggttgtgta cggcccatgg ggttgtggca actgttggca ttgcagtcag 300
ggcctggaga actattgctc ccgtgcgcag gaactgggta ttaacccgcc tggtctgggt 360
gctccggggg ctttggcaga atttatgatt gtcgactcac cacgtcattt ggtcccgatt 420
ggcgatttag accctgttaa aactgttccg ttgactgatg cgggcctgac cccataccat 480
gcaatcaaac gctccctgcc gaaactgcgc ggcggctctt atgcagtagt gatcggtacg 540
ggtggcctgg gccacgtggc tatccaactg ctgcgtcatt tatctgctgc aacggtgatc 600
gccttggacg tttctgccga taaactggaa ctggctacca aagtcggcgc acatgaagta 660
gtcctgtctg ataaagatgc agcggagaat gtgcgtaaaa ttactggtag ccaaggtgca 720
gctttggtgt tggattttgt gggctatcag cctaccattg acaccgccat ggcagtggcg 780
ggcgtgggct ctgacgtcac cattgttggt atcggtgatg gccaggcaca tgcgaaagtt 840
ggtttcttcc agagtcctta tgaggcatcg gttacggtac cttattgggg cgctcgtaat 900
gaactgatcg aattgatcga tctggcgcat gctggtattt tcgacattgc cgttgagacc 960
ttctctttgg ataatggtgc agaggcctat cgtcgcctgg ctgcgggcac actgtcaggc 1020
cgtgcggtag tcgtcccggg cctgtaa                                     1047
```

<210> 82
<211> 348
<212> PRT
<213> Secuencia artificial

<220>
<223> ADH de *Rhodococcus erythropolis*

<400> 82

```
Met Lys Ala Ile Gln Tyr Thr Arg Ile Gly Ala Glu Pro Glu Leu Thr
1               5                   10                  15
Glu Ile Pro Lys Pro Glu Pro Gly Pro Gly Glu Val Leu Leu Glu Val

            20                  25                  30
Thr Ala Ala Gly Val Cys His Ser Asp Asp Phe Ile Met Ser Leu Pro
        35                  40                  45
Glu Glu Gln Tyr Thr Tyr Gly Leu Pro Leu Thr Leu Gly His Glu Gly
    50                  55                  60
Ala Gly Lys Val Ala Ala Val Gly Glu Gly Val Glu Gly Leu Asp Ile
65                  70                  75                  80
Gly Thr Asn Val Val Val Tyr Gly Pro Trp Gly Cys Gly Asn Cys Trp
                85                  90                  95
His Cys Ser Gln Gly Leu Glu Asn Tyr Cys Ser Arg Ala Gln Glu Leu
            100                 105                 110
Gly Ile Asn Pro Pro Gly Leu Gly Ala Pro Gly Ala Leu Ala Glu Phe
        115                 120                 125
Met Ile Val Asp Ser Pro Arg His Leu Val Pro Ile Gly Asp Leu Asp
    130                 135                 140
Pro Val Lys Thr Val Pro Leu Thr Asp Ala Gly Leu Thr Pro Tyr His
145                 150                 155                 160
Ala Ile Lys Arg Ser Leu Pro Lys Leu Arg Gly Gly Ser Tyr Ala Val
                165                 170                 175
Val Ile Gly Thr Gly Gly Leu Gly His Val Ala Ile Gln Leu Leu Arg
            180                 185                 190
His Leu Ser Ala Ala Thr Val Ile Ala Leu Asp Val Ser Ala Asp Lys
        195                 200                 205
Leu Glu Leu Ala Thr Lys Val Gly Ala His Glu Val Val Leu Ser Asp
    210                 215                 220
Lys Asp Ala Ala Glu Asn Val Arg Lys Ile Thr Gly Ser Gln Gly Ala
225                 230                 235                 240
Ala Leu Val Leu Asp Phe Val Gly Tyr Gln Pro Thr Ile Asp Thr Ala
                245                 250                 255
Met Ala Val Ala Gly Val Gly Ser Asp Val Thr Ile Val Gly Ile Gly
            260                 265                 270
Asp Gly Gln Ala His Ala Lys Val Gly Phe Phe Gln Ser Pro Tyr Glu
        275                 280                 285
Ala Ser Val Thr Val Pro Tyr Trp Gly Ala Arg Asn Glu Leu Ile Glu
    290                 295                 300
Leu Ile Asp Leu Ala His Ala Gly Ile Phe Asp Ile Ala Val Glu Thr
305                 310                 315                 320
Phe Ser Leu Asp Asn Gly Ala Glu Ala Tyr Arg Arg Leu Ala Ala Gly
                325                 330                 335
Thr Leu Ser Gly Arg Ala Val Val Val Pro Gly Leu
            340                 345
```

<210> 83
<211> 1047
<212> ADN
<213> Secuencia artificial

<220>
<223> YGL ADH de *Saccharomyces cervisiae*

<400> 83

```
atgaccacgg aaaaaaccgt agtgttcgtg tcaggcgcga ccggttttat tgctctgcac 60
gtggtagatg acctgctgaa aactggttac aaagtaattg gttccggtcg ttctcaggaa 120
aaaaatgacg gtttgctgaa gaagttcaag tccaacccga atctgagcat ggaaattgtg 180
gaagatattg cggcaccaaa cgccttcgat aaagtattcc agaaacatgg taaagaaatt 240
aaagtggtcc tgcatatcgc gtccccggtc catttcaaca ctaccgattt cgaaaaagac 300
ttactgatcc cggcggtaaa cggtaccaaa tctattttgg aagcaattaa gaactatgcc 360
gcagacaccg tggaaaaagt ggttattact tcatctgttg ccgcgttggc ctctccgggt 420
gatatgaaag ataccagctt cgtggttaac gaagaatcct ggaataaaga cacctgggaa 480
tcgtgtcagg cgaatgctgt gtccgcttat tgcggttcta aaaaattcgc agagaaaacg 540
gcgtgggact tcttggaaga aaaccagagc agcattaaat ttactctgtc cacgattaac 600
ccaggcttcg tttttggtcc gcagctgttc gccgactcct tgcgcaatgg tattaactct 660
agcagtgcga ttattgcgaa cctggtgtcg tataaattag gggataactt ctacaattat 720
agcggcccgt ttatcgacgt ccgtgacgtt tccaaagctc atctgctggc atttgagaaa 780
cctgaatgcg ccggtcagcg cctgtttctg tgcgaggata tgttctgttc ccaggaagcc 840
ctggacattc tgaacgaaga atttccacag ctgaagggca agatcgcaac gggcgaacct 900

ggcagcggct cgaccttcct gactaaaaat tgttgcaaat gcgacaatcg taaaactaaa 960
aacttgctgg gcttccagtt caacaaattc cgcgactgca ttgtcgatac tgcgtcccag 1020
ttgctggaag tgcaaagcaa aagctaa                                     1047
```

<210> 84
<211> 348
<212> PRT
<213> Secuencia artificial

<220>
<223> YGL ADH de *Saccharomyces cervisiae*

<400> 84

```
Met Thr Thr Glu Lys Thr Val Val Phe Val Ser Gly Ala Thr Gly Phe
1               5                   10                  15
Ile Ala Leu His Val Val Asp Asp Leu Leu Lys Thr Gly Tyr Lys Val
            20                  25                  30
Ile Gly Ser Gly Arg Ser Gln Glu Lys Asn Asp Gly Leu Leu Lys Lys
        35                  40                  45
Phe Lys Ser Asn Pro Asn Leu Ser Met Glu Ile Val Glu Asp Ile Ala
    50                  55                  60
Ala Pro Asn Ala Phe Asp Lys Val Phe Gln Lys His Gly Lys Glu Ile
65                  70                  75                  80
Lys Val Val Leu His Ile Ala Ser Pro Val His Phe Asn Thr Thr Asp
                85                  90                  95
Phe Glu Lys Asp Leu Leu Ile Pro Ala Val Asn Gly Thr Lys Ser Ile
            100                 105                 110
Leu Glu Ala Ile Lys Asn Tyr Ala Ala Asp Thr Val Glu Lys Val Val
        115                 120                 125
Ile Thr Ser Ser Val Ala Ala Leu Ala Ser Pro Gly Asp Met Lys Asp
    130                 135                 140
Thr Ser Phe Val Val Asn Glu Glu Ser Trp Asn Lys Asp Thr Trp Glu
145                 150                 155                 160
Ser Cys Gln Ala Asn Ala Val Ser Ala Tyr Cys Gly Ser Lys Lys Phe
                165                 170                 175
Ala Glu Lys Thr Ala Trp Asp Phe Leu Glu Glu Asn Gln Ser Ser Ile
            180                 185                 190
Lys Phe Thr Leu Ser Thr Ile Asn Pro Gly Phe Val Phe Gly Pro Gln
        195                 200                 205
Leu Phe Ala Asp Ser Leu Arg Asn Gly Ile Asn Ser Ser Ser Ala Ile
    210                 215                 220
Ile Ala Asn Leu Val Ser Tyr Lys Leu Gly Asp Asn Phe Tyr Asn Tyr
225                 230                 235                 240
Ser Gly Pro Phe Ile Asp Val Arg Asp Val Ser Lys Ala His Leu Leu
                245                 250                 255
Ala Phe Glu Lys Pro Glu Cys Ala Gly Gln Arg Leu Phe Leu Cys Glu
            260                 265                 270
Asp Met Phe Cys Ser Gln Glu Ala Leu Asp Ile Leu Asn Glu Glu Phe
        275                 280                 285
Pro Gln Leu Lys Gly Lys Ile Ala Thr Gly Glu Pro Gly Ser Gly Ser
    290                 295                 300
Thr Phe Leu Thr Lys Asn Cys Cys Lys Cys Asp Asn Arg Lys Thr Lys
305                 310                 315                 320
Asn Leu Leu Gly Phe Gln Phe Asn Lys Phe Arg Asp Cys Ile Val Asp
                325                 330                 335
Thr Ala Ser Gln Leu Leu Glu Val Gln Ser Lys Ser
            340                 345
```

<210> 85
<211> 939
<212> ADN
<213> Secuencia artificial

<220>
<223> YDL ADH de *Saccharomyces cervisiae*

<400> 85

```
atgtcttttc accaacagtt tttcacgctg aacaacggca ataaaatccc ggcgattgcc 60
atcatcggca ctggtacacg ttggtataaa aatgaagaaa ctgacgcgac cttctccaat 120
agtctggttg aacaaatcgt gtatgcgttg aaactgccgg ggattatcca catcgacgcc 180
gcggagattt atcgcaccta cccggaagtg ggtaaagcac tgtccctgac cgaaaagcct 240
cgtaacgcga tttttctgac ggataaatat tctccgcaga ttaaaatgag tgactcccct 300
gcggacggtc tggatttagc attgaagaaa atgggtacag attatgttga tttatatctg 360
ttacattccc cgtttgtttc gaaggaagtg aatggcttaa gcttagaaga ggcttggaaa 420
gatatggagc agttatacaa aagtggtaaa gctaaaaaca tcggggtttc caatttcgca 480
gtggaagacc tgcaacgtat cctgaaagtc gctgaagtta aacctcaggt caaccagatt 540
gagttctctc cgttcctgca aaaccaaaca ccaggcattt ataaattctg tcaggagcac 600
gatatcctgg tggaagcata ttctccgctg ggcccgctgc agaagaaaac cgcgcaggat 660
gacagccaac catttttga gtacgtcaaa gaattgagcg aaaaatacat caaatccgag 720
gcccagatca tcctgcgctg ggtcactaaa cgcggtgtgc tgccagttac cacctcttca 780
aagcctcagc gcattagcga tgctcagaac ctgttttcct tcgacctgac agcggaagag 840
gttgataaaa tcacggagct gggtctggaa catgaaccgc tgcgcctgta ctggaataaa 900
ttgtatggca aatataacta cgccgcccag aaagtgtaa                        939
```

<210> 86
<211> 312
<212> PRT
<213> Secuencia artificial

<220>
<223> YDL ADH de *Saccharomyces cervisiae*

<400> 86

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


                    Tyr Asn Tyr Ala Ala Gln Lys Val
                    305                 310
```

<210> 87
<211> 939
<212> ADN
<213> Secuencia artificial

<220>
<223> YPR ADH de *Saccharomyces cervisiae*

<400> 87

```
atgccggcaa cgttaaaaaa cagcagtgct accttaaaat taaacacagg tgcgagcatt 60
cctgtcctgg ggttcggcac ctggcgctct gtcgataaca acggctatca tagtgtaatt 120
gcggcgctga aagcggggta ccgtcatatc gatgctgcgg ccatctatct gaatgaagaa 180
gaagtcggcc gtgcgatcaa ggactccggt gttcctcgtg aagaaatttt tattaccacc 240
aaactgtggg gcaccgaaca acgcgatcca gaagcagccc tgaacaaatc tctgaaacgt 300
ctgggtctgg actatgtgga cctgtatctg atgcactggc cggtccctct gaaaacagac 360
cgtgtaactg acggtaacgt cctgtgcatc ccgaccctgg aagatggcac cgtggacatc 420
gataccaaag agtggaattt tattaaaacc tgggaactga tgcaggaatt gccgaaaact 480
ggtaagacca aagccgtcgg tgtgtccaat ttttccatca acaatatcaa agaactgctg 540
gaatcgccaa ataacaaggt cgttccagca accaatcaga tcgagattca tccgttgctg 600
ccgcaggatg aattaatcgc cttttgtaaa gaaaaaggca ttgtggtcga agcatatagc 660
ccattcggct ccgctaacgc cccgctgctg aaagaacagg cgattatcga tatggccaaa 720
aagcacggcg tcgaaccggc gcaactgatt atcagctggt cgattcagcg cggttatgtg 780
gtattggcca agtccgtaaa tccggagcgt atcgtgtcga actttaagat ttttaccctg 840
ccagaggatg atttcaaaac catctctaac ctgagcaaag tgcacggtac caaacgtgtc 900
gttgacatga aatggggctc atttccgatt tttcaataa                        939
```

<210> 88
<211> 312
<212> PRT
<213> Secuencia artificial

<220>
<223> YPR ADH de *Saccharomyces cervisiae*

<400> 88

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

```
    210                 215                 220
Ala Asn Ala Pro Leu Leu Lys Glu Gln Ala Ile Ile Asp Met Ala Lys
225                 230                 235                 240
Lys His Gly Val Glu Pro Ala Gln Leu Ile Ile Ser Trp Ser Ile Gln
                245                 250                 255
Arg Gly Tyr Val Val Leu Ala Lys Ser Val Asn Pro Glu Arg Ile Val
            260                 265                 270
Ser Asn Phe Lys Ile Phe Thr Leu Pro Glu Asp Asp Phe Lys Thr Ile
        275                 280                 285
Ser Asn Leu Ser Lys Val His Gly Thr Lys Arg Val Val Asp Met Lys
    290                 295                 300
Trp Gly Ser Phe Pro Ile Phe Gln
305                 310
```

<210> 89
<211> 1029
<212> ADN
<213> Secuencia artificial

<220>
<223> GRE ADH de *Saccharomyces cervisiae*

<400> 89

```
atgtctgtgt tcgtgtcagg cgcgaatggt tttattgctc agcacatcgt agatctgctg 60
ctgaaagaag attacaaagt aattggttcc gcacgttctc aggaaaaagc tgaaaatttg 120
accgaagcct tcggtaacaa cccgaaattt agcatggaag tggtgcctga tattagcaaa 180
ctggatgcct tcgatcatgt attccagaaa catggtaaag atattaaaat cgtcctgcat 240
accgcgtccc cgttttgttt cgatattacc gattccgaac gtgacttact gatcccggcg 300
gtaaacggtg tcaaaggtat tttgcacagt attaagaaat atgccgcaga cagcgtggaa 360
cgcgtggttc tgacttcatc ttacgccgcg gtatttgata tggcgaagga aaacgataag 420
agcctgacct tcaacgaaga atcctggaat ccggcgacct gggaatcgtg tcagagtgat 480
ccggtgaacg cttattgcgg ttctaaaaaa ttcgcagaga aagcagcgtg ggaattcttg 540
gaagaaaacc gtgatagcgt gaaatttgag ctgacagcgg tcaacccagt ttacgttttt 600
ggtccgcaga tgttcgataa agatgttaaa aaacacttga acaccagctg cgaactggtg 660
aactctctga tgcatctgag ccctgaagat aaaattccgg aactgtttgg cggttacatc 720
gacgtccgtg acgttgcgaa agctcatctg gttgcatttc agaaacgtga aacaatcggt 780
cagcgcctga tcgtgtcgga ggcacgtttc acgatgcagg acgttctgga cattctgaac 840
gaagactttc cagtactgaa gggcaatatc ccggtcggca agcctggcag cggcgccacc 900
cataatactc tgggcgccac cctggacaat aaaaaaagca aaaaattgct gggcttcaaa 960
ttccgtaatc tgaaagagac tattgacgat actgcgtccc agatcctgaa attcgaaggt 1020
cgcatttaa                                                         1029
```

<210> 90
<211> 342
<212> PRT
<213> Secuencia artificial

<220>
<223> GRE ADH de *Saccharomyces cervisiae*

<400> 90

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

Lys Tyr Ala Ala Asp Ser Val Glu Arg Val Val Leu Thr Ser Ser Tyr
        115                 120                 125
Ala Ala Val Phe Asp Met Ala Lys Glu Asn Asp Lys Ser Leu Thr Phe
    130                 135                 140
Asn Glu Glu Ser Trp Asn Pro Ala Thr Trp Glu Ser Cys Gln Ser Asp
145                 150                 155                 160
Pro Val Asn Ala Tyr Cys Gly Ser Lys Lys Phe Ala Glu Lys Ala Ala
                165                 170                 175
Trp Glu Phe Leu Glu Glu Asn Arg Asp Ser Val Lys Phe Glu Leu Thr
            180                 185                 190
Ala Val Asn Pro Val Tyr Val Phe Gly Pro Gln Met Phe Asp Lys Asp
        195                 200                 205
Val Lys Lys His Leu Asn Thr Ser Cys Glu Leu Val Asn Ser Leu Met
    210                 215                 220
His Leu Ser Pro Glu Asp Lys Ile Pro Glu Leu Phe Gly Gly Tyr Ile
225                 230                 235                 240
Asp Val Arg Asp Val Ala Lys Ala His Leu Val Ala Phe Gln Lys Arg
                245                 250                 255
Glu Thr Ile Gly Gln Arg Leu Ile Val Ser Glu Ala Arg Phe Thr Met
            260                 265                 270
Gln Asp Val Leu Asp Ile Leu Asn Glu Asp Phe Pro Val Leu Lys Gly
        275                 280                 285
Asn Ile Pro Val Gly Lys Pro Gly Ser Gly Ala Thr His Asn Thr Leu
    290                 295                 300
Gly Ala Thr Leu Asp Asn Lys Lys Ser Lys Lys Leu Leu Gly Phe Lys
305                 310                 315                 320
Phe Arg Asn Leu Lys Glu Thr Ile Asp Asp Thr Ala Ser Gln Ile Leu
                325                 330                 335
Lys Phe Glu Gly Arg Ile
            340
```

<210> 91
<211> 1041
<212> ADN
<213> Secuencia artificial

<220>
<223> ADH de *Streptomyces coelicolor*

<400> 91

```
atgaaggcac tgcagtaccg caccatcggc gccccgcccg aggtcgtcac cgtcccggac 60
ccggagccgg gccccggcca ggtgctgttg aaggtgaccg cggccggagt ctgccactcc 120
gacatcgcgg tgatgagctg gcccgccgag ggcttcccgt acgagctgcc gctcaccctc 180
ggccacgagg gcgtcggcac cgtggccgcg ctcggcgccg gggtgacggg gctcgccgag 240
ggcgacgcgg tcgccgtgta cgggccctgg ggctgcggca cctgcgccaa gtgcgcggag 300
ggcaaggaga actactgcct gcgcgccgac gagctgggca tccgtccgcc ggggctcggg 360
cgtccggggt ccatggccga gtacctgctg atcgacgacc cccggcacct ggtcccgctg 420
gacgggctcg acccggtcgc ggcggtgccg ctcaccgacg ccggactgac gccgtaccac 480
gcgatcaagc ggtcgctgcc caagctggtc cccggctcca ccgcggtggt catcggcacc 540
ggtggtctcg gccacgtcgc catccagctg ctgcgcgccc tgacgtccgc ccgggtggtc 600
gccctggacg tcagcgagga gaagctgcgc ctcgcccgtg cggtgggcgc gcacgaggcg 660
gtgctgtcgg acgcgaaggc cgcggacgcg gtgcgcgaga tcaccggcgg tctcggtgcc 720
gaggccgtgt tcgacttcgt cggcgtggcg cccaccgtgc agaccgccgg agccgtcgcg 780
gccgtcgagg gcgatgtcac cctggtcggc atcggcggcg gatcgctgcc cgtcggcttc 840
ggcatgctgc cgttcgaggt gtcggtcaac gccccctact ggggcagccg cagcgagctg 900
accgaggtgc tgaacctggc ccgctccggt gccgtgtcgg tgcacaccga gacgtactcc 960
ctggacgacg ccccgctcgc ctacgagcgg ctgcacgagg gcagggtcaa cggccgcgcg 1020
gtgatcctgc cccacggctg a                                           1041
```

<210> 92
<211> 346
<212> PRT
<213> Secuencia artificial

<220>
<223> ADH de *Streptomyces coelicolor*

<400> 92

```
Met Lys Ala Leu Gln Tyr Arg Thr Ile Gly Ala Pro Pro Glu Val Val
1               5                   10                  15
Thr Val Pro Asp Pro Glu Pro Gly Pro Gly Gln Val Leu Leu Lys Val
            20                  25                  30
Thr Ala Ala Gly Val Cys His Ser Asp Ile Ala Val Met Ser Trp Pro
        35                  40                  45
Ala Glu Gly Phe Pro Tyr Glu Leu Pro Leu Thr Leu Gly His Glu Gly
    50                  55                  60
Val Gly Thr Val Ala Ala Leu Gly Ala Gly Val Thr Gly Leu Ala Glu
65                  70                  75                  80
Gly Asp Ala Val Ala Val Tyr Gly Pro Trp Gly Cys Gly Thr Cys Ala
                85                  90                  95
Lys Cys Ala Glu Gly Lys Glu Asn Tyr Cys Leu Arg Ala Asp Glu Leu
            100                 105                 110
Gly Ile Arg Pro Pro Gly Leu Gly Arg Pro Gly Ser Met Ala Glu Tyr
        115                 120                 125
Leu Leu Ile Asp Asp Pro Arg His Leu Val Pro Leu Asp Gly Leu Asp
    130                 135                 140
Pro Val Ala Ala Val Pro Leu Thr Asp Ala Gly Leu Thr Pro Tyr His
145                 150                 155                 160
Ala Ile Lys Arg Ser Leu Pro Lys Leu Val Pro Gly Ser Thr Ala Val
                165                 170                 175
Val Ile Gly Thr Gly Gly Leu Gly His Val Ala Ile Gln Leu Leu Arg
            180                 185                 190
Ala Leu Thr Ser Ala Arg Val Val Ala Leu Asp Val Ser Glu Glu Lys
        195                 200                 205
Leu Arg Leu Ala Arg Ala Val Gly Ala His Glu Ala Val Leu Ser Asp
    210                 215                 220
Ala Lys Ala Ala Asp Ala Val Arg Glu Ile Thr Gly Gly Leu Gly Ala
225                 230                 235                 240
Glu Ala Val Phe Asp Phe Val Gly Val Ala Pro Thr Val Gln Thr Ala
                245                 250                 255
Gly Ala Val Ala Ala Val Glu Gly Asp Val Thr Leu Val Gly Ile Gly
            260                 265                 270
Gly Gly Ser Leu Pro Val Gly Phe Gly Met Leu Pro Phe Glu Val Ser
        275                 280                 285
Val Asn Ala Pro Tyr Trp Gly Ser Arg Ser Glu Leu Thr Glu Val Leu
    290                 295                 300
Asn Leu Ala Arg Ser Gly Ala Val Ser Val His Thr Glu Thr Tyr Ser
305                 310                 315                 320
Leu Asp Asp Ala Pro Leu Ala Tyr Glu Arg Leu His Glu Gly Arg Val
                325                 330                 335
Asn Gly Arg Ala Val Ile Leu Pro His Gly
            340                 345
```

<210> 93
<211> 1032
<212> ADN
<213> Secuencia artificial

<220>
<223> ADH de *Sporobolomyces salmonicolor*

<400> 93

```
atggcaaaga tcgacaacgc agttctgccg gagggttctc tggtgctggt caccggcgcg 60
aacggctttg tcgctagcca tgtggtcgaa caactgctgg aacacggcta taaggtgcgc 120
ggcactgctc gctctgcctc caaactggcg aacctgcaga aacgttggga cgccaaatac 180
cctggtcgtt tcgagactgc cgttgttgaa gacatgctga agcagggtgc atatgatgaa 240
gttattaaag gcgcggcagg tgtcgcccac atcgcgtccg tggtcagctt ttctaacaaa 300
tatgatgagg tggtaactcc tgcgatcggt ggcacgctga atgccctgcg tgccgcagct 360
gctacgcctt ccgtgaaacg ttttgtgctg accagcagca ctgtttctgc actgattcca 420
aaacctaacg tcgaaggtat ttatctggat gagaagagct ggaacctgga aagcattgat 480
aaggctaaaa ccctgcctga atctgatccg cagaaaagcc tgtgggtcta cgccgcaagc 540

aaaacggaag cggaactggc tgcctggaaa ttcatggacg aaaacaaacc gcactttact 600
ctgaatgccg ttctgccaaa ctacactatc ggtaccattt ttgacccaga aacccaatcc 660
ggttccactt ccggctggat gatgtctctg ttcaatggcg aagtatctcc ggcactggcg 720
ctgatgccgc cgcagtacta tgtctctgca gttgatatcg gtctgctgca cctgggttgt 780
ctggttctgc cgcaaatcga acgccgtcgt gtttacggca ccgcaggcac ctttgattgg 840
aacaccgttc tggcgacctt ccgtaaactg tatccgtcca agacgttccc ggctgacttt 900
ccggatcagg gccaggatct gtccaaattt gataccgccc cgagcctgga gattctgaaa 960
tccctgggcc gccctggctg gcgtagcatc gaggaatcta tcaaagatct ggtgggttcc 1020
gagaccgcct aa                                                     1032
```

<210> 94
<211> 343
<212> PRT
<213> Secuencia artificial

<220>
<223> ADH de *Sporobolomyces salmonicolor*

<400> 94

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

<210> 95
<211> 1050
<212> ADN
<213> Secuencia artificial

<220>
<223> ADH de *Aspergillus flavus*

<400> 95

```
atgtctattc cggaaatgca atgggctcaa gtagcggagc agaaaggcgg tccgctgatt 60
tacaaacaga ttccggttcc gaaaccgggc ccggacgaaa tcctggtaaa agtgcgttac 120
tctggtgtgt gccacacgga tctgcacgcg ctgaaaggtg actggccact gcctgtgaaa 180
atgccactgg tgggtggcca tgagggcgcg ggtgttgtcg tcgctcgtgg cgatctggtt 240
actgagttcg aaattggtga ccatgctggc ctgaaatggc tgaacggtag ctgcctggca 300
tgcgagttct gtaaacaggc tgacgaaccg ctgtgcccta acgcgagcct gtctggttat 360
actgtagatg gtactttcca gcagtatgcc attggcaaag ccacccatgc gagcaagctg 420
ccgaaaaacg tgcctctgga tgccgtggca ccggttctgt gcgcgggcat taccgtatac 480
aagggcctga aagaaagcgg tgttcgtcct ggccaaaccg ttgcgatcgt aggtgcgggt 540
ggtggtctgg gctccctggc gctgcagtac gcgaaagcta tgggtattcg cgtggtggcg 600
attgatggcg gtgaagagaa acaagcgatg tgtgaacagc tgggcgcaga agcctacgtt 660
gactttacca aaacgcaaga tctggtagca gatgttaagg ccgccacccc tgaaggtctg 720
ggcgcgcatg cggtaattct gctggcggtc gccgaaaaac catttcagca ggcggccgaa 780
tatgtcagcc gtggtacggt ggttgccatt ggtctgccgg cgggcgcatt cctgcgcgcg 840
ccggtgttta acactgtggt tcgtatgatt aacattaagg gtagctacgt tggcaaccgc 900
caggacggcg ttgaggcagt ggacttcttc gcgcgcggcc tgatcaaagc gccgttcaaa 960
accgctcctc tgcaagatct gccgaaaatc ttcgaactga tggaacaagg taagattgca 1020
ggtcgctacg ttctggaaat cccggaatga                                  1050
```

<210> 96
<211> 349
<212> PRT
<213> Secuencia artificial

<220>
<223> ADH de *Aspergillus flavus*

<400> 96

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


Gly Ala His Ala Val Ile Leu Leu Ala Val Ala Glu Lys Pro Phe Gln
                245                 250                 255
Gln Ala Ala Glu Tyr Val Ser Arg Gly Thr Val Val Ala Ile Gly Leu
            260                 265                 270
Pro Ala Gly Ala Phe Leu Arg Ala Pro Val Phe Asn Thr Val Val Arg
        275                 280                 285
Met Ile Asn Ile Lys Gly Ser Tyr Val Gly Asn Arg Gln Asp Gly Val
    290                 295                 300
Glu Ala Val Asp Phe Phe Ala Arg Gly Leu Ile Lys Ala Pro Phe Lys
305                 310                 315                 320
Thr Ala Pro Leu Gln Asp Leu Pro Lys Ile Phe Glu Leu Met Glu Gln
                325                 330                 335
Gly Lys Ile Ala Gly Arg Tyr Val Leu Glu Ile Pro Glu
            340                 345
```

<210> 97
<211> 729
<212> ADN
<213> Secuencia artificial

<220>
<223> ADH de *Oenococcus oeni*

<400> 97

```
atgtttgatc gcctgaaggg taaggtggcg atcgtcaccg gcggtaatag cggtattggt 60
aaggcaatcg cggcagactt catcgccgag ggcgcgaaag tagtgattac cggtcgcaac 120
caagagaagg gtcgccagac tgcgctggaa atcgccggtg acattctgtt cattcagcag 180
gacgtgagcc aggaggcgga ctggcagaaa gttatctcca aaaccatcga gaagttcggt 240
aagtttgata ttctggttaa caacgctggc gtcggcggcg tgggtaagcc actggcagaa 300
atgtccctgt ctgaattcaa ttggacccaa tccattaatc tgtccggtaa cttcctgggt 360
attcacttcg cgctgaacaa aatgaccgaa cctggtagca ttattgacgt aagcagcgcg 420
gctggcctgc gtggttttcc gggcgcagcc gattatagcg cttccaaagg tggtacgcgt 480
ctgctgacgc gcgccgcggc gctggaagcg ctgcaaatgg gtaaaaagat ccgtgtaaac 540
tctattcacc caggttggat tgataccgat atcgtaccaa aagatatgcg tgagcaggtt 600
atcgcgatga cccctatgca tcacctgggc caaccgaaag atattgccaa gctggcgacg 660
tatctggcat ctgacgagtc cgagtatact acgggctccg aactggcagc cgacggcggt 720
ctgatggca                                                         729
```

<210> 98
<211> 243
<212> PRT
<213> Secuencia artificial

<220>
<223> ADH de *Oenococcus oeni*

<400> 98

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

Gly Phe Pro Gly Ala Ala Asp Tyr Ser Ala Ser Lys Gly Gly Thr Arg
145                 150                 155                 160
Leu Leu Thr Arg Ala Ala Ala Leu Glu Ala Leu Gln Met Gly Lys Lys
                165                 170                 175
Ile Arg Val Asn Ser Ile His Pro Gly Trp Ile Asp Thr Asp Ile Val
            180                 185                 190
Pro Lys Asp Met Arg Glu Gln Val Ile Ala Met Thr Pro Met His His
        195                 200                 205
Leu Gly Gln Pro Lys Asp Ile Ala Lys Leu Ala Thr Tyr Leu Ala Ser
    210                 215                 220
Asp Glu Ser Glu Tyr Thr Thr Gly Ser Glu Leu Ala Ala Asp Gly Gly
225                 230                 235                 240
Leu Met Ala
```

<210> 99
<211> 768
<212> ADN

<213> Secuencia artificial

<220>
<223> ADH de *Ralstonia eutropha*

<400> 99

```
atgcagggca aaattgcact ggttacgggc ggtgcgtctg gtatgggtct ggccttcacc 60
cgtcgcctgg cccaagaagg cgcacgtgtt tacttcaccg acattaacgg cgccgcgggt 120
cacgcgacgg aaagcgagct gcagggtgct ggtctggcgg tcgtattcct gcgtcatgac 180
gtgacccagg aaaacgactg gctgcgcgtc ctggaacata ttggcgctgt tgacggtcgc 240
ctggatgtac tggtgaataa tgccggcatt gcgatttctc acaacattga gacttgtact 300
actgaggatt ttgatcgcac gctgaatgtt aacctgaagt ccgtgtttct gggctgcaaa 360
cacggtctgg ctctgatgaa acagaagggt ggcagcatca ttaacgttag ctctatcacc 420
gcgatttgcg gcgaacctgt ggcgctggcg tatagcgcta gcaaagccgg cgtccgtttt 480
ctgagcaagt ctgttgccct gcattgtgcg gagaagggct acgccatccg tgttaacagc 540
ctgcacccgg gttacattga taccccgctg ctggcaggca gcaacgcggg tggctccctg 600
accgctgacg aggtaaccgc gagccgtgta cgtattggcg gtgaaattcc gctgaaacgt 660
cgcggcaccc cggacgaagt tgcaggcgcg gttctgtacc tggctagcga tgaatccacc 720
tatgttactg gtactgagct ggtaattgat ggcggctacg catgtcat              768
```

<210> 100
<211> 256
<212> PRT
<213> Secuencia artificial

<220>
<223> ADH de *Ralstonia eutropha*

<400> 100

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


Glu Pro Val Ala Leu Ala Tyr Ser Ala Ser Lys Ala Gly Val Arg Phe
145                 150                 155                 160
Leu Ser Lys Ser Val Ala Leu His Cys Ala Glu Lys Gly Tyr Ala Ile
                165                 170                 175
Arg Val Asn Ser Leu His Pro Gly Tyr Ile Asp Thr Pro Leu Leu Ala
            180                 185                 190
Gly Ser Asn Ala Gly Gly Ser Leu Thr Ala Asp Glu Val Thr Ala Ser
        195                 200                 205
Arg Val Arg Ile Gly Gly Glu Ile Pro Leu Lys Arg Arg Gly Thr Pro
    210                 215                 220
Asp Glu Val Ala Gly Ala Val Leu Tyr Leu Ala Ser Asp Glu Ser Thr
225                 230                 235                 240
Tyr Val Thr Gly Thr Glu Leu Val Ile Asp Gly Gly Tyr Ala Cys His
                245                 250                 255
```

<210> 101
<211> 1056
<212> ADN
<213> Secuencia artificial

<220>

<223> ADH de *Termoanaerobium brockii*

<400> 101

```
atgaaaggct tcgccatgct gagcatcggc aaagtgggtt ggattgaaaa agaaaaaccg 60
gcgccaggcc cgttcgatgc aattgtgcgc cctctggcag tagcgccgtg taccagcgat 120
attcatactg tgtttgaagg tgccattggc gagcgtcaca atatgattct gggccatgaa 180
gccgttggtg aagttgttga ggttggcagc gaagtgaagg atttcaaacc gggcgatcgc 240
gttgtcgttc cagcgattac cccggattgg cgcaccagcg aagtccagcg cggctaccat 300
cagcactctg gcggcatgct ggccggctgg aaattcagca atgtaaagga tggtgtgttc 360
ggtgaatttt ttcacgttaa cgacgcagac atgaatctgg cgcacctgcc gaaagaaatc 420
ccgctggaag cagcggttat gattccggat atgatgacca cgggttttca cggcgcagag 480
ctggcggaca ttgaactggg cgctacggta gccgtactgg gcatcggtcc ggtgggcctg 540
atggcagttg caggcgctaa gctgcgcggc gcaggtcgta ttattgccgt tggttctcgc 600
ccggtgtgtg tggacgccgc taagtattat ggtgcaacgg acattgtcaa ttacaaggac 660
ggcccaattg aatctcagat catgaacctg acggaaggta aaggcgttga cgccgcgatt 720
atcgctggcg gcaacgccga catcatggcg accgcagtta aaatcgtcaa gccaggtggt 780
actattgcta acgtcaacta cttcggcgaa ggtgaggtcc tgcctgtccc acgtctggaa 840
tggggttgcg gtatggcaca taaaaccatt aaaggtggcc tgtgcccagg cggccgtctg 900
cgtatggaac gcctgatcga tctggtcttc tacaaacgcg tggatcctag caaactggtg 960
actcacgttt tccgcggctt tgataacatc gaaaaagctt ttatgctgat gaaagataaa 1020
ccgaaagatc tgattaaacc ggttgtcatc ctggct                           1056
```

<210> 102
<211> 352
<212> PRT
<213> Secuencia artificial

<220>
<223> ADH de *Termoanaerobium brockii*

<400> 102

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

```
                100                 105                 110
Ser Asn Val Lys Asp Gly Val Phe Gly Glu Phe Phe His Val Asn Asp
        115                 120                 125
Ala Asp Met Asn Leu Ala His Leu Pro Lys Glu Ile Pro Leu Glu Ala
    130                 135                 140
Ala Val Met Ile Pro Asp Met Met Thr Thr Gly Phe His Gly Ala Glu
145                 150                 155                 160
Leu Ala Asp Ile Glu Leu Gly Ala Thr Val Ala Val Leu Gly Ile Gly
                165                 170                 175
Pro Val Gly Leu Met Ala Val Ala Gly Ala Lys Leu Arg Gly Ala Gly
            180                 185                 190
Arg Ile Ile Ala Val Gly Ser Arg Pro Val Cys Val Asp Ala Ala Lys
        195                 200                 205
Tyr Tyr Gly Ala Thr Asp Ile Val Asn Tyr Lys Asp Gly Pro Ile Glu
    210                 215                 220
Ser Gln Ile Met Asn Leu Thr Glu Gly Lys Gly Val Asp Ala Ala Ile
225                 230                 235                 240
Ile Ala Gly Gly Asn Ala Asp Ile Met Ala Thr Ala Val Lys Ile Val
                245                 250                 255
Lys Pro Gly Gly Thr Ile Ala Asn Val Asn Tyr Phe Gly Glu Gly Glu
            260                 265                 270
Val Leu Pro Val Pro Arg Leu Glu Trp Gly Cys Gly Met Ala His Lys
        275                 280                 285
Thr Ile Lys Gly Gly Leu Cys Pro Gly Gly Arg Leu Arg Met Glu Arg
    290                 295                 300
Leu Ile Asp Leu Val Phe Tyr Lys Arg Val Asp Pro Ser Lys Leu Val
305                 310                 315                 320
Thr His Val Phe Arg Gly Phe Asp Asn Ile Glu Lys Ala Phe Met Leu
                325                 330                 335
Met Lys Asp Lys Pro Lys Asp Leu Ile Lys Pro Val Val Ile Leu Ala
            340                 345                 350
```

<210> 103
<211> 1125
<212> ADN
<213> Secuencia artificial

<220>
<223> ADH de hígado de caballo

<400> 103

```
atgagcacag cagggaaagt aattaaatgc aaagcagctg tgctgtggga ggaaaagaaa 60
ccattttcca ttgaggaggt ggaggtcgca cctccgaagg cccatgaagt ccgtattaag 120
atggtggcca cagggatctg tcgctcagat gaccacgtgg ttagtgggac cctggtcaca 180
cctctgcctg tgatcgcagg ccatgaggca gctggcattg tggagagcat cggggaaggc 240
gtcactacag tacgtccagg tgataaagtc atcccactgt ttactcctca gtgtgggaaa 300
tgccgtgttt gtaaacaccc tgaaggcaac ttctgcttga aaaatgatct gagcatgcct 360
cgtgggacca tgcaggatgg taccagccgt ttcacctgcc gtgggaagcc tatccaccac 420
ttcctgggca ccagcacctt ctcccagtac accgtggtgg acgagatctc agtggccaag 480
atcgatgcgg cctcaccgct ggagaaagtc tgtctgattg gctgtgggtt ttctactggt 540
tatgggtctg cagtcaaggt tgccaaggtc acccagggct ccacctgtgc cgtgtttggc 600
ctgggggggg tgggcctgtc tgttatcatg ggctgtaaag cagccggggc ggcccgtatc 660
attggggtgg acatcaacaa agacaagttt gcaaaggcca aagaagtggg tgccactgag 720
tgtgtcaacc ctcaggacta caagaaacct atccaggagg tgctgacaga aatgagcaat 780
gggggcgtgg atttttcatt tgaagtcatt ggtcgtctgg acaccatggt gactgccttg 840
tcatgctgtc aagaagcata tggtgtgagc gtcatcgtag gggtacctcc tgattcccaa 900
aatctgtcta tgaatcctat gttgctgctg agtgggcgta cctggaaagg ggctattttt 960
ggtggtttta agagtaaaga ttctgtccct aaactggtgg ccgattttat ggctaaaaag 1020
tttgcactgg atcctttaat cacccatgtt ttacctttg aaaaaattaa tgaagggttt 1080
gacctgctgc gctctgggga gagtatccgt accatcctga cgttt                 1125
```

<210> 104
<211> 375
<212> PRT
<213> Secuencia artificial

<220>
<223> ADH de hígado de caballo

<400> 104

```
Met Ser Thr Ala Gly Lys Val Ile Lys Cys Lys Ala Ala Val Leu Trp
1               5                   10                  15
Glu Glu Lys Lys Pro Phe Ser Ile Glu Glu Val Glu Val Ala Pro Pro
            20                  25                  30
Lys Ala His Glu Val Arg Ile Lys Met Val Ala Thr Gly Ile Cys Arg
        35                  40                  45
Ser Asp Asp His Val Val Ser Gly Thr Leu Val Thr Pro Leu Pro Val
    50                  55                  60
Ile Ala Gly His Glu Ala Ala Gly Ile Val Glu Ser Ile Gly Glu Gly
65                  70                  75                  80
Val Thr Thr Val Arg Pro Gly Asp Lys Val Ile Pro Leu Phe Thr Pro
                85                  90                  95
Gln Cys Gly Lys Cys Arg Val Cys Lys His Pro Glu Gly Asn Phe Cys
            100                 105                 110
Leu Lys Asn Asp Leu Ser Met Pro Arg Gly Thr Met Gln Asp Gly Thr
        115                 120                 125
Ser Arg Phe Thr Cys Arg Gly Lys Pro Ile His His Phe Leu Gly Thr
    130                 135                 140
Ser Thr Phe Ser Gln Tyr Thr Val Val Asp Glu Ile Ser Val Ala Lys
145                 150                 155                 160
Ile Asp Ala Ala Ser Pro Leu Glu Lys Val Cys Leu Ile Gly Cys Gly
                165                 170                 175
Phe Ser Thr Gly Tyr Gly Ser Ala Val Lys Val Ala Lys Val Thr Gln
            180                 185                 190
Gly Ser Thr Cys Ala Val Phe Gly Leu Gly Gly Val Gly Leu Ser Val
        195                 200                 205
Ile Met Gly Cys Lys Ala Ala Gly Ala Ala Arg Ile Ile Gly Val Asp
    210                 215                 220
Ile Asn Lys Asp Lys Phe Ala Lys Ala Lys Glu Val Gly Ala Thr Glu
225                 230                 235                 240
Cys Val Asn Pro Gln Asp Tyr Lys Lys Pro Ile Gln Glu Val Leu Thr
                245                 250                 255
Glu Met Ser Asn Gly Gly Val Asp Phe Ser Phe Glu Val Ile Gly Arg
            260                 265                 270
Leu Asp Thr Met Val Thr Ala Leu Ser Cys Cys Gln Glu Ala Tyr Gly
        275                 280                 285
Val Ser Val Ile Val Gly Val Pro Pro Asp Ser Gln Asn Leu Ser Met
    290                 295                 300
Asn Pro Met Leu Leu Leu Ser Gly Arg Thr Trp Lys Gly Ala Ile Phe
305                 310                 315                 320
Gly Gly Phe Lys Ser Lys Asp Ser Val Pro Lys Leu Val Ala Asp Phe
                325                 330                 335
Met Ala Lys Lys Phe Ala Leu Asp Pro Leu Ile Thr His Val Leu Pro
            340                 345                 350
Phe Glu Lys Ile Asn Glu Gly Phe Asp Leu Leu Arg Ser Gly Glu Ser
        355                 360                 365
Ile Arg Thr Ile Leu Thr Phe
    370                 375
```

<210> 105
<211> 759
<212> ADN
<213> *Lactobacillus minor*

<400> 105

```
atgaccgatc ggttgaaggg gaaagtagca attgtaactg gcggtacctt gggaattggc     60
ttggcaatcg ctgataagtt tgttgaagaa ggcgcaaagg ttgttattac cggccgtcac    120
gctgatgtag gtgaaaaagc tgccagatca atcggcggca cagacgttat ccgttttgtc    180
caacacgatg cttctgatga aaccggctgg actaagttgt ttgatacgac tgaagaagca    240
tttggcccag ttaccacggt tgtcaacaat gccggaattg cggtcagcaa gagtgttgaa    300
gataccacaa ctgaagaatg gcgcaagctg ctctcagtta acttggatgg tgtcttcttc    360

ggtacccgtc ttggaatcca acgtatgaag aataaaggac tcggagcatc aatcatcaat    420
atgtcatcta tcgaaggttt tgttggtgat ccagctctgg gtgcatacaa cgcttcaaaa    480
ggtgctgtca gaattatgtc taaatcagct gccttggatt gcgctttgaa ggactacgat    540
gttcgggtta acactgttca tccaggttat atcaagacac cattggttga cgatcttgaa    600
ggggcagaag aaatgatgtc acagcggacc aagacaccaa tgggtcatat cggtgaacct    660
aacgatatcg cttggatctg tgtttacctg gcatctgacg aatctaaatt tgccactggt    720
gcagaattcg ttgtcgacgg agggtacacc gcccaatag                           759
```

<210> 106
<211> 252

<212> PRT
<213> *Lactobacillus minor*

<400> 106

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

<210> 107
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Fórmula de secuencia de cetorreductasa manipulada con esqueleto de *L. kefir*
<223> Constructo sintético

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(7)..(7)
<223> Xaa es un resto no polar o restringido

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(40)..(40)
<223> Xaa es un resto restringido, hidrófilo o básico

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(46)..(46)
<223> Xaa es un resto hidrófilo o básico

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(60)..(60)
<223> Xaa es un resto alifático o no polar

<220>
<221> CARACTERÍSTICA_DIVERSA

<222>(66)..(66)
<223> Xaa es un resto hidrófilo o ácido

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(72)..(72)
<223> Xaa es un resto básico, hidrófilo o polar

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(94)..(94)
<223> Xaa es un resto no polar

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(96)..(96)
<223> Xaa es un resto polar, no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(108)..(108)
<223> Xaa es un resto polar, restringido o hidrófilo

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(134)..(134)
<223> Xaa es un resto no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(141)..(141)
<223> Xaa es un resto no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(144)..(144)
<223> Xaa es un resto no polar, hidrófobo o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(145)..(145)
<223> Xaa es una leucina o un resto aromático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(147)..(147)
<223> Xaa es un resto aromático, no polar, alifático o hidrófobo

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(152)..(152)
<223> Xaa es un resto polar

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(153)..(153)
<223> Xaa es un resto polar, no polar, alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(169)..(169)
<223> Xaa es una cisteína o un resto polar

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(185)..(185)

<223> Xaa es un resto polar

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(186)..(186)
<223> Xaa es un resto no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(190)..(190)
<223> Xaa es una cisteína o un resto restringido

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(195)..(195)
<223> Xaa es un resto no polar, alifático o básico

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(196)..(196)
<223> Xaa es un resto no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(199)..(199)
<223> Xaa es un resto no polar, alifático, aromático o hidrófobo

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(202)..(202)
<223> Xaa es un resto no polar, hidrófobo, alifático o aromático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(206)..(206)
<223> Xaa es un resto no polar, aromático o hidrófobo

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(211)..(211)
<223> Xaa es un resto básico, no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(226)..(226)
<223> Xaa es un resto no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(233)..(233)
<223> Xaa es un resto ácido, no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(245)..(245)
<223> Xaa es un resto de aminoácido no polar o alifático,

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(249)..(249)
<223> Xaa es un resto de aminoácido no polar o aromático

<400> 107

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

<210> 108
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Fórmula de secuencia de cetorreductasa manipulada con esqueleto de *L. brevis*
<223> Constructo sintético

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(7)..(7)
<223> Xaa es un resto no polar o restringido

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(40)..(40)
<223> Xaa es un resto restringido, hidrófilo o básico

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(46)..(46)
<223> Xaa es un resto hidrófilo o básico

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(60)..(60)
<223> Xaa es un resto alifático o no polar

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(66)..(66)
<223> Xaa es un resto hidrófilo o ácido

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(72)..(72)

<223> Xaa es un resto básico, hidrófilo o polar

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(94)..(94)
<223> Xaa es un resto no polar

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(96)..(96)
<223> Xaa es un resto polar, no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(108)..(108)
<223> Xaa es un resto polar, restringido o hidrófilo

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(134)..(134)
<223> Xaa es un resto no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(141)..(141)
<223> Xaa es un resto no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(144)..(144)
<223> Xaa es un resto no polar, hidrófobo o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(145)..(145)
<223> Xaa es una leucina o un resto aromático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(147)..(147)
<223> Xaa es un resto aromático, no polar, alifático o hidrófobo

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(152)..(152)
<223> Xaa es un resto polar

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(153)..(153)
<223> Xaa es un resto polar, no polar, alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(169)..(169)
<223> Xaa es una cisteína o un resto polar

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(185)..(185)
<223> Xaa es un resto polar

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(186)..(186)
<223> Xaa es un resto no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(190)..(190)
<223> Xaa es una cisteína o un resto restringido

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(195)..(195)
<223> Xaa es un resto no polar, alifático o básico

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(196)..(196)
<223> Xaa es un resto no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(199)..(199)
<223> Xaa es un resto no polar, alifático, aromático o hidrófobo

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(202)..(202)
<223> Xaa es un resto polar, hidrófobo, alifático o aromático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(206)..(206)
<223> Xaa es un resto no polar, aromático o hidrófobo

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(211)..(211)
<223> Xaa es un resto básico, no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(226)..(226)
<223> Xaa es un resto no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(233)..(233)
<223> Xaa es un resto ácido, no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(245)..(245)
<223> Xaa es un resto de aminoácido no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(249)..(249)
<223> Xaa es un resto de aminoácido no polar o aromático

<400> 108

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

<210> 109
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Fórmula de secuencia de cetorreductasa manipulada con esqueleto de *L. minor*
<223> Constructo sintético

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(7)..(7)
<223> Xaa es un resto no polar o restringido

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(40)..(40)
<223> Xaa es un resto restringido, hidrófilo o básico

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(46)..(46)
<223> Xaa es un resto hidrófilo o básico

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(60)..(60)
<223> Xaa es un resto alifático o no polar

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(66)..(66)
<223> Xaa es un resto hidrófilo o ácido

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(72)..(72)

<223> Xaa es un resto básico, hidrófilo o polar

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(94)..(94)
<223> Xaa es un resto no polar

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(96)..(96)
<223> Xaa es un resto polar, no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(108)..(108)
<223> Xaa es un resto polar, restringido o hidrófilo

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(134)..(134)
<223> Xaa es un resto no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(141)..(141)
<223> Xaa es un resto no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(144)..(144)
<223> Xaa es un resto no polar, hidrófobo o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(145)..(145)
<223> Xaa es una leucina o un resto aromático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(147)..(147)
<223> Xaa es un resto aromático, no polar, alifático o hidrófobo

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(152)..(152)
<223> Xaa es un resto polar

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(153)..(153)
<223> Xaa es un resto polar, no polar, alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(169)..(169)
<223> Xaa es una cisteína o un resto polar

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(185)..(185)
<223> Xaa es un resto polar

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(186)..(186)
<223> Xaa es un resto no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(190)..(190)
<223> Xaa es una cisteína o un resto restringido

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(195)..(195)
<223> Xaa es un resto no polar, alifático o básico

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(196)..(196)
<223> Xaa es un resto no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(199)..(199)
<223> Xaa es un resto no polar, alifático, aromático o hidrófobo

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(202)..(202)
<223> Xaa es un resto polar, hidrófobo, alifático o aromático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(206)..(206)
<223> Xaa es un resto no polar, aromático o hidrófobo

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(211)..(211)
<223> Xaa es un resto básico, no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(226)..(226)
<223> Xaa es un resto no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(233)..(233)
<223> Xaa es un resto ácido, no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(245)..(245)
<223> Xaa es un resto de aminoácido no polar o alifático

<220>
<221> CARACTERÍSTICA_DIVERSA
<222>(249)..(249)
<223> Xaa es un resto de aminoácido no polar o aromático

<400> 109

```
Met Thr Asp Arg Leu Lys Xaa Lys Val Ala Ile Val Thr Gly Gly Thr
1               5                   10                  15
Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
            20                  25                  30
Lys Val Val Ile Thr Gly Arg Xaa Ala Asp Val Gly Glu Xaa Ala Ala
        35                  40                  45
```

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

<210> 110
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> Variante de *L. kefir*

<400> 110

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

        Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
                    245                 250
```

Reivindicaciones

1. Polipéptido cetorreductasa manipulado capaz de convertir el sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina en el producto (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina a una velocidad que está mejorada en comparación con la de un polipéptido de referencia que tiene la secuencia de aminoácidos de la SEQ ID NO: 2, en el que el polipéptido comprende una secuencia de aminoácidos que es idéntica en al menos un 85% a una secuencia de referencia que tiene la secuencia de aminoácidos de la SEQ ID NO: 2, en el que el polipéptido tiene las siguientes características: (a) el resto correspondiente al resto X94 es una glicina, (b) el resto correspondiente al resto X145 es un aminoácido aromático o leucina; y (c) el resto correspondiente al resto X190 es prolina; y (d) el resto en la posición X153 es treonina o valina, el resto en la posición X195 es metionina, el resto en la posición X206 es fenilalanina, triptófano o tirosina, y/o el resto en la posición X233 es glicina.

2. Polipéptido cetorreductasa según la reivindicación 1, en el que el resto correspondiente a X145 de la SEQ ID NO: 2 está seleccionado del grupo que consiste en fenilalanina y leucina, en el que opcionalmente el resto correspondiente a X145 es fenilalanina.

3. Polipéptido cetorreductasa según la reivindicación 1, en el que la secuencia de aminoácidos comprende adicionalmente una o más de las siguientes características:

el resto correspondiente a X7 es histidina;
el resto correspondiente a X40 es arginina;
el resto correspondiente a X46 es arginina o lisina;
el resto correspondiente a X60 es isoleucina;
el resto correspondiente a X66 es ácido glutámico o ácido aspártico;
el resto correspondiente a X72 es treonina;
el resto correspondiente a X96 es valina;
el resto correspondiente a X108 es histidina;
el resto correspondiente a X134 es metionina;
el resto correspondiente a X141 es valina;
el resto correspondiente a X144 es valina;
el resto correspondiente a X147 es isoleucina, metionina, valina o leucina;
el resto correspondiente a X152 es serina;
el resto correspondiente a X153 es treonina;
el resto correspondiente a X169 es cisteína;
el resto correspondiente a X185 es treonina;
el resto correspondiente a X186 es isoleucina;
el resto correspondiente a X195 es metionina;
el resto correspondiente a X196 es leucina;
el resto correspondiente a X199 es tirosina, triptófano o valina;
el resto correspondiente a X206 es tirosina, triptófano o fenilalanina;
el resto correspondiente a X211 es leucina;
el resto correspondiente a X226 es valina;
el resto correspondiente a X233 es glicina;
el resto correspondiente a X245 es isoleucina;
el resto correspondiente a X249 es triptófano; y
en el que la secuencia de aminoácidos puede tener opcionalmente una o más diferencias de restos en otros restos de aminoácidos en comparación con la secuencia de referencia.

4. Polipéptido cetorreductasa según cualquiera de las reivindicaciones 1-3, en el que el polipéptido es capaz de convertir el sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina en el producto (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina a una velocidad que está mejorada en comparación con la de un polipéptido de referencia que tiene la secuencia de aminoácidos de la SEQ ID NO: 6.

5. Polipéptido cetorreductasa según cualquiera de las reivindicaciones 1-4, en el que el polipéptido comprende una secuencia de aminoácidos seleccionada de entre las SEQ ID NO: 24, 26, 28, 30, 34, 36, 38, 40, 42, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76 y 78.

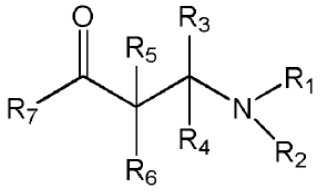
6. Polinucleótido que codifica un polipéptido cetorreductasa manipulado según cualquiera de las reivindicaciones 1-5, en el que la secuencia polinucleotídica está opcionalmente seleccionada de entre las SEQ ID NO: 23, 25, 27, 29, 33, 35, 37, 39, 41, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75 y 77.

7. Vector de expresión que comprende el polinucleótido según la reivindicación 6 unido operativamente a secuencias de control adecuadas para dirigir la expresión en una célula hospedadora, en el que la secuencia de control comprende opcionalmente una señal de secreción.

8. Célula hospedadora que comprende el vector de expresión según la reivindicación 7.

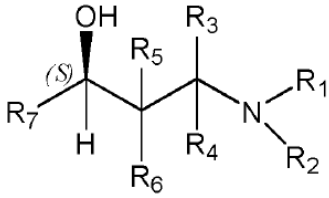
9. Método para producir una *(S)*-3-aril-3-hidroxipropanamina, comprendiendo dicho método:

(a) proporcionar un sustrato 3-aril-3-cetopropanamina con la estructura de fórmula (I):



(I);

(b) poner en contacto el sustrato 3-aril-3-cetopropanamina con el polipéptido cetorreductasa manipulado según cualquiera de las reivindicaciones 1-5 en una mezcla de reacción en condiciones adecuadas para la reducción o conversión del sustrato en un producto *(S)*-3-aril-3-hidroxipropanamina que tiene la estructura de fórmula (II):



(II)

en el que para (I) y (II), cada uno de R_1 y R_2 está seleccionado independientemente del grupo que consiste en hidrógeno, un alquilo inferior opcionalmente sustituido, un cicloalquilo opcionalmente sustituido, un arilo opcionalmente sustituido, o como alternativa, en las que R_1 y R_2 juntos forman un cicloalquilo opcionalmente sustituido o un cicloarilo opcionalmente sustituido que tiene 3-7 átomos de carbono; cada uno de R_3, R_4, R_5 y R_6 está seleccionado independientemente del grupo que consiste en hidrógeno y un alquilo inferior opcionalmente sustituido; y R_7 es un arilo opcionalmente sustituido.

10. Método según la reivindicación 9, en el que uno de R_1 y R_2 es metilo y R_7 es 2-tienilo, y, opcionalmente, en el que R_1 y R_2 son ambos metilo.

11. Método según la reivindicación 10, en el que R_1 y R_2 son ambos metilo y en el que al menos aproximadamente un 95% del sustrato se reduce al producto en menos de aproximadamente 24 horas, en el que la concentración del sustrato es de al menos 100 g/l y la concentración del polipéptido cetorreductasa es inferior a aproximadamente 5 g/l, y en el que el polipéptido cetorreductasa comprende una secuencia de aminoácidos que corresponde a una secuencia de aminoácidos seleccionada del grupo que consiste en la SEQ ID NO: 60, 62, 64, 66, 68, 70, 72, 74, 76 y 78.

12. Método según cualquiera de las reivindicaciones 9-11, en el que el producto se genera en un exceso estereomérico superior a aproximadamente un 99%.

13. Método según la reivindicación 9, en el que la cetorreductasa se aísla y/o purifica y la reacción de reducción se lleva a cabo en presencia de un cofactor para la cetorreductasa y, opcionalmente, un sistema de regeneración para el cofactor, en el que el sistema de regeneración está seleccionado del grupo que consiste en glucosa deshidrogenasa y glucosa, formiato deshidrogenasa y formiato, isopropanol y una alcohol secundario deshidrogenasa, y fosfato y fosfato deshidrogenasa, por ejemplo en el que la alcohol secundario deshidrogenasa es el polipéptido cetorreductasa manipulado.

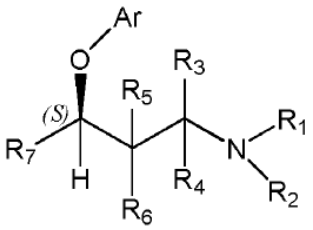
14. Método de producción de una *(S)-N*-metil-3-hidroxi-3-(aril)-propanamina, comprendiendo dicho método el método según la reivindicación 9 y comprendiendo adicionalmente:

(c) desmetilar el producto *(S)*-3-aril-3-hidroxipropanamina de la etapa (b) en una mezcla de reacción en condiciones adecuadas para producir una *(S)-N*-metil-3-hidroxi-3-(aril)-propanamina con la estructura de fórmula (II), en el que uno de R_1 y R_2 es metilo y el otro es hidrógeno, cada uno de R_3, R_4, R_5 y R_6 está seleccionado independientemente del grupo que consiste en hidrógeno y un alquilo inferior opcionalmente sustituido, y R_7 es un arilo opcionalmente sustituido, en el que la etapa de desmetilación comprende opcionalmente poner en

contacto la *(S)*-3-aril-3-hidroxipropanamina con un agente de desmetilación seleccionado del grupo que consiste en un cloroformiato o derivado del mismo o un fosgeno o derivado del mismo.

15. Método de producción de una 3-ariloxi-3-(aril)-propanamina, comprendiendo el método el método según la reivindicación 9 y comprendiendo adicionalmente:

(c) poner en contacto la *(S)*-3-aril-3-hidroxipropanamina con un compuesto de arilo activado en una mezcla de reacción en condiciones suficientes para producir la *(S)*-3-ariloxi-3-arilpropanamina que tiene la estructura de fórmula (VII)



(VII)

en el que para (I), (II) y (VII), cada uno de R_1 y R_2 está seleccionado independientemente del grupo que consiste en hidrógeno, un alquilo inferior opcionalmente sustituido, un cicloalquilo opcionalmente sustituido, un arilo opcionalmente sustituido, o como alternativa, en el que R_1 y R_2 juntos forman un cicloalquilo opcionalmente sustituido o un cicloarilo opcionalmente sustituido que tiene 3-7 átomos de carbono; cada uno de R_3, R_4, R_5 y R_6 está seleccionado independientemente del grupo que consiste en hidrógeno y un alquilo inferior opcionalmente sustituido; y R_7 es un arilo opcionalmente sustituido; y, además, para (VII), Ar es un grupo arilo opcionalmente sustituido, en el que Ar es opcionalmente un arilo seleccionado del grupo que consiste en 1-naftilo, fenilo, 4-trifluorometilfenilo, 2-metilfenilo, 2-metoxifenilo y 2-tiometoxifenilo.

16. Método para la síntesis de *(3S)-N*-metil-3-naftalen-1-iloxi-3-tiofen-2-il-propan-1-amina que incluye como etapa la reducción estereoselectiva del sustrato *N,N*-dimetil-3-ceto-3-(2-tienil)-1-propanamina al producto (*S*)-*N,N*-dimetil-3-hidroxi-3-(2-tienil)-1-propanamina, en el que la etapa comprende poner en contacto el sustrato con el polipéptido cetorreductasa manipulado según cualquiera de las reivindicaciones 1-5 en condiciones de reacción adecuadas para la reducción o conversión del sustrato en el producto.

17. Composición que comprende el compuesto de fórmula estructural (I) y/o de fórmula estructural (II) y cualquiera de los polipéptidos cetorreductasa manipulados según las reivindicaciones 1-5, comprendiendo adicionalmente de manera opcional un sistema de regeneración de cofactor seleccionado del grupo que consiste en glucosa deshidrogenasa y glucosa, formiato deshidrogenasa y formiato, isopropanol y una alcohol secundario deshidrogenasa, y fosfato y fosfato deshidrogenasa.

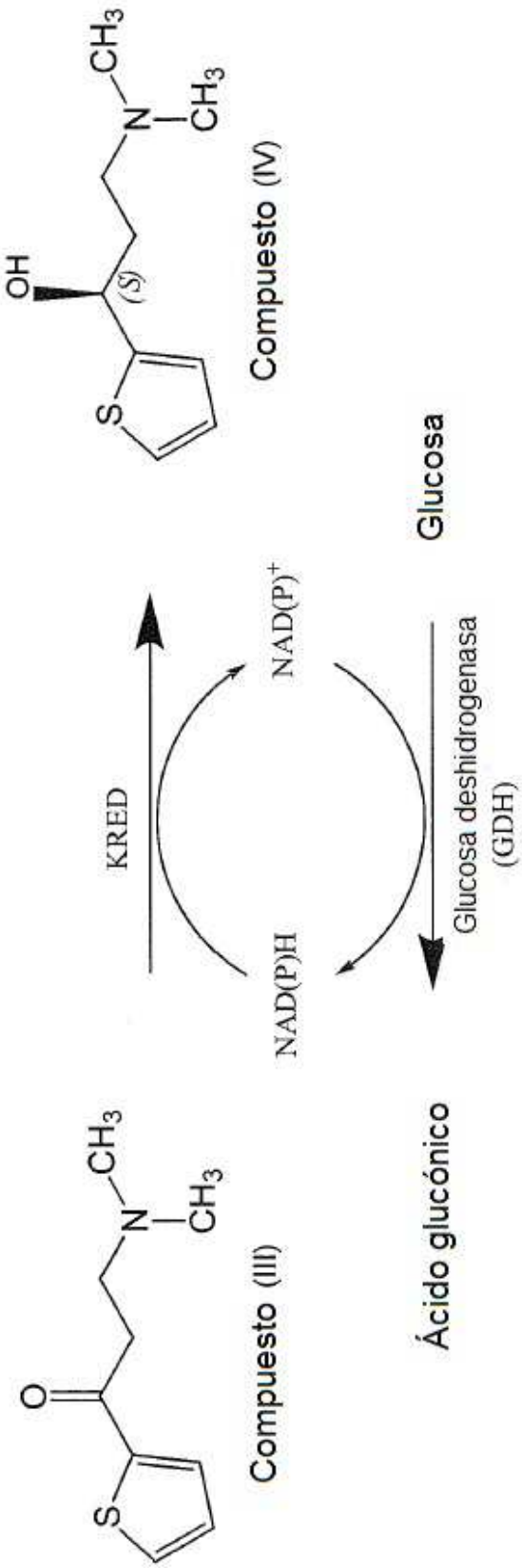

Compuesto (III)
Ácido glucónico
KRED
NAD(P)H
NAD(P)+
Glucosa deshidrogenasa
(GDH)
Glucosa
Compuesto (IV)
(S)
OH
S
N
CH3
CH3

FIG. 1

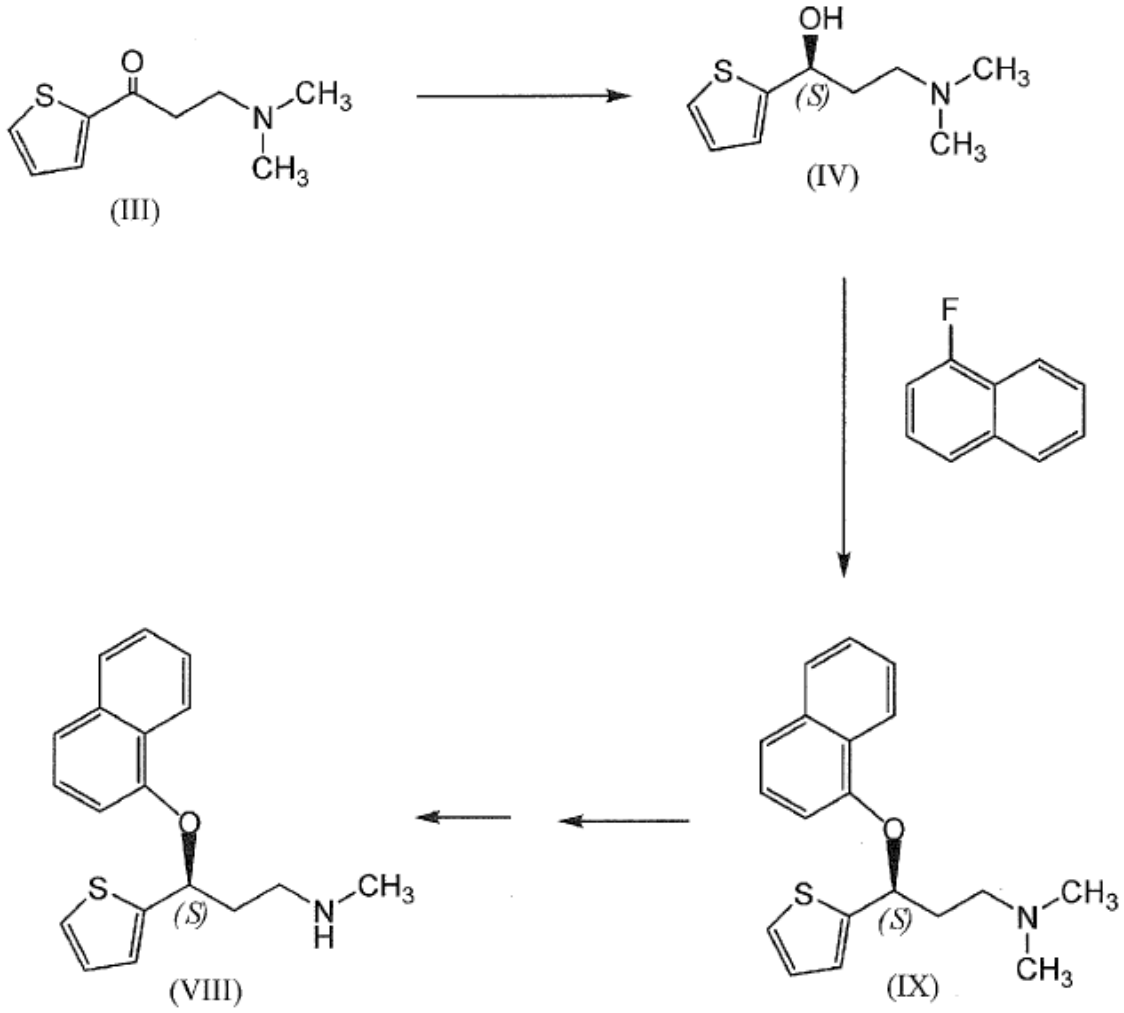

O
S
N
CH3
CH3
(III)
OH
S
(S)
N
CH3
CH3
(IV)
F
O
S
(S)
N
H
CH3
(VIII)
O
S
(S)
N
CH3
CH3
(IX)

FIG. 2