

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 547 528**

51 Int. Cl.:

C12N 9/04 (2006.01)

C12P 13/00 (2006.01)

C12P 17/00 (2006.01)

C12P 17/10 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **01.10.2008 E 08836133 (2)**

97 Fecha y número de publicación de la concesión europea: **24.06.2015 EP 2205727**

54 Título: **Polipéptidos de cetorreductasa para la producción de acetidinona**

30 Prioridad:

01.10.2007 US 976555 P

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

07.10.2015

73 Titular/es:

**CODEXIS, INC. (100.0%)
200 PENOBSCOT DRIVE
REDWOOD CITY, CA 94063, US**

72 Inventor/es:

**CAMPAPIANO, ONORATO;
MUNDORFF, EMILY;
BORUP, BIRTHE y
VOLADRI, RAMA**

74 Agente/Representante:

IZQUIERDO BLANCO, María Alicia

ES 2 547 528 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 de cetorreductasa para la producción de acetidinona

Descripción

5 ANTECEDENTES

Las enzimas que pertenecen a la clase cetorreductasa (KRED) o carbonilo reductasa (EC1.1.1.184) son útiles para la síntesis de alcoholes ópticamente activos del sustrato correspondiente, cetona proestereoisomérica o los sustratos correspondientes aldehído racémicos. Los KREDs comúnmente convierten las cetonas y los sustratos aldehídos a sus correspondientes productos alcohólicos, pero también pueden catalizar la reacción reversa, la oxidación de un sustrato alcohólico al producto correspondiente de cetona/aldehído. La reducción de cetonas y aldehídos y la oxidación de alcoholes por enzimas tales como KRED requiere un cofactor, más comúnmente nicotinamida adenina dinucleótida reducida (NADH) o nicotidamina adenina dinucleótida fosfato reducida (NADPH), y nicotinamida adenina dinucleótida (NAD) o nicotidamina adenina dinucleótida fosfato (NADP) para la reacción de oxidación. NADH y NADPH sirve como donantes de electrones, mientras que NAD y NADP sirven como receptores de electrones. Se ha observado frecuentemente que las cetorreductasas y las deshidrogenasas de alcohol aceptan el cofactor no fosforilado (en su estado oxidado y reducido).

Las enzimas KRED pueden encontrarse en una variedad amplia de bacterias y levaduras (para mediciones: Kraus y Waldman, *Enzyme catalysis in organic synthesis* (Catálisis enzimática en síntesis orgánica) Vols. L y 2.VCH Weinheim 1995; Faber, K., *Biotransformations in organic chemistry* (Transformaciones biológicas en química orgánica), 4ta Ed. Springer, Berlin Heidelberg New York, 2000; Hummel y Kula, 1989, *Eur. J. Biochem.* (Bioquímica) 184:1-13). Varios genes KRED y secuencias enzimáticas han sido reportadas, por ejemplo, *Candida magnoliae* (Cuenta del Banco Genético Númeo. JC7338; GI:11360538) *Candida parapsilosis* (Cuenta del Banco Genético Númeo BAA24528.1; GI:2815409), *Sporobolomyces salmonicolor* (Cuenta del Banco Genético Númeo AF160799; GI:6539734).

Para esquivar cualquier procedimiento sintético químico para la producción de compuestos clave, las cetorreductasas están siendo utilizadas más y más para la conversión enzimática de diferentes sustratos ceto y aldehídos para productos alcohólicos quirales. Estas aplicaciones pueden utilizar células enteras que expresan la cetorreductasa para reducciones cetonas biocatalíticas, o enzimas purificadas en aquellas instancias donde la presencia de varias cetorreductasas en células enteras que afectarían adversamente la estereo - pureza y produciría el producto deseado. Para aplicaciones in - vitro, una enzima regeneradora y cofactor (NADH o NADPH) tal como la glucosa deshidrogenasa (GDH), formiato deshidrogenasa, etc. es utilizado en conjunto con la cetorreductasa. Ejemplos de usos de cetorreductasas para generar compuestos químicos útiles incluyen la reducción asimétrica de ésteres 4-cloroacetoacetatos (Zhou, J. *Ant. Chem. Soc.* 1983 105:5925-5926; Santaniello, J. *Chem. Res.* (S) 1984:132-133; Patente de Estados Unidos número. 5,559,030; Patente de Estados Unidos número. 5,700,670 y patente de Estados Unidos número 5,891,685), reducción de ácidos dicarboxílicos (por ejemplo, patente de Estados Unidos número. 6,399,339), reducción de terc - butil (S) cloro-5-hidroxi-3-oxohexanoato (por ejemplo, patente de Estados Unidos número 6,645,746 y WO 01/40450), reducción de compuestos que se basan en la pirrolotriacina (por ejemplo, la aplicación de Estados Unidos número 2006/0286646); la reducción de acetofenonas sustituidas (por ejemplo, la patente de Estados Unidos número 6,800,477); y la reducción cetotiolanos (WO 2005/054491).

Es deseable identificar otras enzimas cetorreductasas que pueden ser utilizadas para ejecutar conversiones de varios sustratos cetos a sus productos correspondientes alcohólicos quirales. Weckbecker et al. 2006 y Dau ßmann et al. 2006 presenta enzimas cetorreductasas de lactobacilo. WO2005/017135 presenta métodos de mutagénesis.

RESUMEN

Esta presentación suministra polipéptidos cetorreductasas que tienen la habilidad de reducir una mezcla racémica de metil-2-benzamidometil-3-oxobutirato ("el sustrato") a 2S,3R-metil-2-benzamidometil-3-hidroxitirato ("el producto"), la codificación de polinucleótidos tales como polipéptidos, y métodos para usar los polipéptidos. El compuesto 2S, 3R-metil-2-benzamidometil-3-hidroxitirato es un intermediador de la síntesis de acetato de (2R, 3R)-3-((R)-1-(terc-butildimetilsililoxi)etil)-4-oxoacetidin-2-ilo ("acetidinona; acetioxiacetidinona"; registro CAS 76855-69-1), que es un intermediador (penúltimo intermediador) utilizado en la fabricación de varios antibióticos carbapenem. Los antibióticos carbapenem que pueden ser sintetizados de 2S,3R-metil-2-benzamidometil-3-hidroxitirato incluyen, pero no se limitan a imipenem, meropenem, doripenem, ertapenem, biopenem, panipenem, y otros compuestos similares a la tienamicina. Los polipéptidos de cetorreductasa de esta presentación tienen una propiedad mejorada para reducir o convertir el sustrato especificado del producto alcohólico quiral correspondiente en comparación a las enzimas de cetorreductasa de tipo silvestre que ocurren naturalmente obtenidas del Lactobacilo kéfir ("L. kéfir"; IDENTIFICACIÓN SECUENCIAL NÚMERO: 4), Lactobacilo brevis ("L. brevis"; IDENTIFICACIÓN SECUENCIAL NÚMERO: 2), o Lactobacilo minor ("L. minor"; IDENTIFICACIÓN SECUENCIAL NÚMERO: 86). En algunas secciones, los polipéptidos de cetorreductasa diseñados tienen una propiedad mejorada en comparación con otros polipéptidos de cetorreductasa diseñados, tales como IDENTIFICACIÓN SECUENCIAL NÚMERO: 48.

65

El invento suministra un polipéptido de cetorreductasa capaz de convertir el sustrato metil-2-benzamidometil-3-oxobutirato, al producto, 2S, 3R-metil-2-benzamidometil-3-hidroxitirato, con un porcentaje de exceso estereomérico de por lo menos un 60%, que incluye una secuencia de aminoácidos que es por lo menos un 85% idéntica a una secuencia de referencia que se basa en la IDENTIFICACIÓN SECUENCIAL NÚMERO: 2, 4 u 86 que tiene las siguientes características: el residuo correspondiente a X94 es treonina; el residuo correspondiente a X199 es histidina y el residuo correspondiente a X202 es valina o leucina; con la condición que el polipéptido de cetorreductasa tenga una secuencia de aminoácidos en la cual el residuo correspondiente a X 94 es alanina o treonina; el residuo correspondiente a X 199 es alanina, histidina o asparagina; y el residuo correspondiente a X 202 es valina o leucina.

Este invento también suministra un polinucleótido que codifica a un polipéptido del invento.

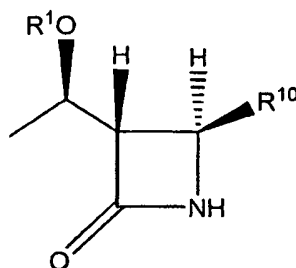
El invento también suministra un vector de expresión que incluye el polinucleótido del invento vinculado en forma operacional a secuencias de control adecuadas para dirigir la expresión en una célula anfitriona.

El invento también suministra una célula anfitriona que incluye el vector de expresión del invento.

El invento suministra un compuesto que incluye una cetorreductasa del invento y el compuesto metil-2-benzamidometil-3-oxobutirato de la fórmula (I) o el compuesto 2S,3R-metil-2-benzamidometil-3-hidroxitirato de la fórmula (II).

El invento también suministra un método para reducir el sustrato de la fórmula (I), metil-2-benzamidometil-3-oxobutirato, al producto de la fórmula (II), 2S,3R-metil-2-benzamidometil-3-hidroxitirato, que se comprende de contactar o incubar el sustrato con un polipéptido de cetorreductasa del invento bajo condiciones de reacción adecuadas para reducir el sustrato del producto de la fórmula (II), y en el cual el producto está opcionalmente presente con un exceso estereomérico mayor que un 99%.

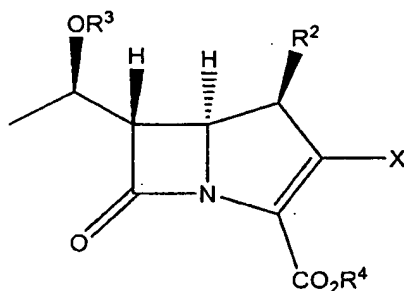
El invento también suministra un método para la síntesis del intermediador de la fórmula (IVa),



(IVa)

Donde R¹ es H o un grupo protector del hidroxilo, y R¹⁰ es un halógeno, u -Oac, donde Ac es acetato, Donde un paso en el método incluye contactar o reaccionar el sustrato de la fórmula (I) con las cetorreductasas del invento bajo condiciones de reacción adecuadas para reducir o convertir el sustrato del producto de la fórmula (II).

El invento también suministra un método para la síntesis del intermediador de la fórmula estructural (IX),

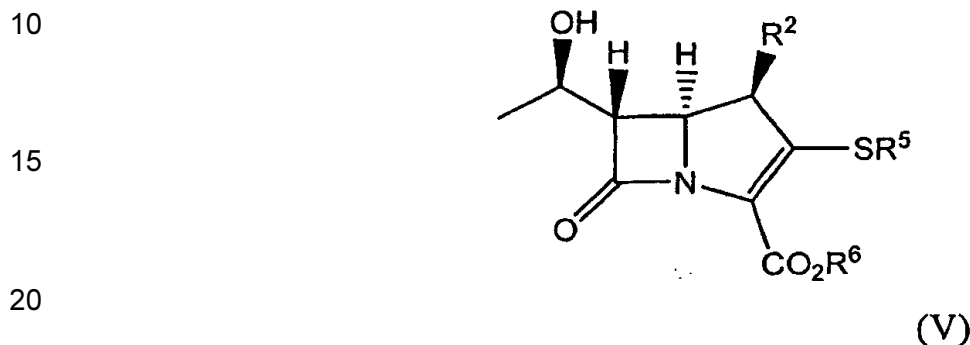


(IX)

Donde R^2 es H o un alquilo C1-C4 (por ejemplo, $-\text{CH}_3$); R^3 es H, o un grupo protector del hidroxilo; R^4 es H, un grupo protector del carboxi, un grupo de amoniaco, un metal alcalino, o un metal alcalinotérreo; y X es OH o un grupo que abandona,

5 Donde un paso en el método incluye el contactar o reaccionar el sustrato de la fórmula (I) con las cetorreductasas del invento bajo condiciones de reacción adecuadas para reducir o convertir el sustrato al producto de la fórmula (II).

El invento también suministra un método para la síntesis de un carbapenem de la fórmula estructural (V):



25 O sus solvatos, hidratos, sales y pro - medicamentos, donde R^2 es H o $-\text{CH}_3$; R^5 es seleccionado de un alquilo sustituido o no sustituido, un arilo sustituido o no sustituido, un heteroalquilo sustituido o no sustituido, un heterocicloalquilo sustituido o no sustituido, y un heteroarilquilo sustituido o no sustituido; y R^6 es H o un pro-grupo, Donde un paso en el método incluye el contactar o reaccionar el sustrato de la fórmula (I) con las cetorreductasas de cualquiera de los inventos bajo las condiciones de reacción adecuadas para reducir o convertir el sustrato del producto de la fórmula (II).

30 En algunas secciones, los polipéptidos de cetorreductasa tienen, en referencia a las secuencias KRED del tipo silvestre *L. kéfir*, *L. brevis* o *L. minor* de las IDENTIFICACIONES SECUENCIALES NÚMERO: 4, 2 y 86, por lo menos la siguiente característica: el residuo 202 es valina o leucina. En algunas secciones, las cetorreductasas de la presentación tienen, en referencia a las secuencias de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 4, 2, u 86, por lo menos dos de las siguientes características: (1) el residuo correspondiente de la posición 94 (es decir, X94) es un residuo alifático o polar, (2) el residuo correspondiente a la posición 199 (es decir, X199) es un residuo alifático, polar o restringido, y (3) el residuo correspondiente a la posición 202 (es decir, X202) es valina o leucina. En algunas secciones, los polipéptidos tienen, en referencia a las secuencias de las IDENTIFICACIONES SECUENCIALES NÚMERO: 4, 2, y 86, por lo menos las siguientes características: (1) el residuo correspondiente a la posición X94 es un residuo polar, (2) el residuo correspondiente a X199 es un residuo alifático, restringido o polar, y (1) el residuo correspondiente a X202 es valina o leucina.

45 Adicionalmente a las características descritas anteriormente, las cetorreductasas pueden tener una o más diferencias de residuos en otras posiciones residuales cuando se las compara a las secuencias de las IDENTIFICACIONES SECUENCIALES NÚMERO: 2, 4, u 86. En algunas secciones, los polipéptidos de cetorreductasa mencionados en este documento incluyen una secuencia de aminoácidos que es por lo menos alrededor de un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o más idénticos a una secuencia referencial que se basa en las IDENTIFICACIONES SECUENCIALES NÚMERO: 2, 4 u 86 con las siguientes características: el residuo correspondiente a X 94 es un residuo alifático o polar, particularmente alanina o treonina; el residuo correspondiente a X 199 es un residuo alifático, restringido o polar, particularmente alanina, histidina, o asparagina; y el residuo correspondiente a X 202 es valina o leucina; con la condición de que la secuencia de aminoácido de cetorreductasa tenga por lo menos las siguientes características, es decir, el residuo correspondiente a X 94 es un residuo alifático o polar; el residuo correspondiente a X 199 es un residuo alifático, restringido o polar; y el residuo correspondiente a X 202 es valina o leucina.

55 En algunas secciones, tal como en donde la propiedad mejorada viene de una sola diferencia residual o una combinación específica de diferencia residuales, las cetorreductasas diseñadas podrían incluir opcionalmente una o más diferencias residuales en otras posiciones en el polipéptido cuando se las compara con la secuencia referencial. En algunas secciones, la diferencia residual incluye mutaciones conservadoras. En algunas secciones, las diferencias residuales adicionales en otras posiciones residuales pueden incorporarse para producir más mejoras en las propiedades enzimáticas. Estas mejoras pueden ser incrementos adicionales en actividades enzimáticas para el sustrato definido, pero también pueden incluir incrementos en estereoselectividad, estabilidad térmica, estabilidad disolvente, y / o inhibición reducida del producto. Varias diferencias residuales que suelen resultar en una o más propiedades enzimáticas mejoradas se suministran en la descripción detallada. En algunas secciones, un polipéptido mejorado de cetorreductasa incluye una secuencia de aminoácidos que corresponde a las fórmulas secuenciales como se muestran en la IDENTIFICACIÓN SECUENCIAL NÚMERO: 83, la IDENTIFICACIÓN SECUENCIAL

NÚMERO: 84 o la IDENTIFICACIÓN SECUENCIAL NÚMERO: 87 (o una de sus regiones, tal como los residuos 90 - 211). La IDENTIFICACIÓN SECUENCIAL NÚMERO: 84 se basa en la secuencia de aminoácidos de tipo silvestre de la cetorreductasa del lactobacilo kéfir (IDENTIFICACIÓN SECUENCIAL NÚMERO: 4), La IDENTIFICACIÓN SECUENCIAL NÚMERO: 83 se basa en la secuencia de aminoácido de tipo silvestre de la cetorreductasa del lactobacilo brevis (IDENTIFICACIÓN SECUENCIAL NÚMERO: 2); y la IDENTIFICACIÓN SECUENCIAL NÚMERO: 87 se basa en la secuencia de aminoácidos de tipo silvestre de la cetorreductasa de tipo silvestre del lactobacilo minor (IDENTIFICACIÓN SECUENCIAL NÚMERO: 86). Las fórmulas secuenciales de las IDENTIFICACIONES SECUENCIALES NÚMEROS: 83, 84 y 87 especifican que el residuo correspondiente a X 94 es un residuo alifático o polar; el residuo correspondiente a X 199 es un residuo alifático, restringido o polar; y el residuo correspondiente a X 202 es valina o leucina. Las fórmulas secuenciales especifican además características de otras posiciones residuales, tal como se menciona en la descripción detallada.

En algunas secciones, el polipéptido diseñado de cetorreductasa puede tener una actividad enzimática incrementada en comparación a la enzima de tipo silvestre de cetorreductasa para reducir el sustrato al producto. Las mejoras en actividades enzimáticas pueden ser medidas al comparar la actividad específica de los polipéptidos de cetorreductasa con aquella de la enzima de tipo silvestre de cetorreductasa usando ensayos enzimáticos estándar. El monto de la mejora puede variar desde 1.5 veces (o el doble) la actividad enzimática del tipo silvestre correspondiente o la enzima de cetorreductasa de referencia, un tanto como 2 veces, 5 veces, 10 veces, 20 veces, 25 veces, 50 veces, 75 veces, 100 veces, o más. En secciones específicas, la enzima diseñada de cetorreductasa muestra actividades enzimáticas mejoradas que son por lo menos 1.5 veces, 2 veces, 3 veces, 4 veces, 5 veces, 10 veces, 25 veces, 50 veces, 100 veces, 500 veces, o 1000 veces mayores que aquellas del tipo silvestre o que la enzima de cetorreductasa referencial. Las mejoras en la actividad enzimática pueden incluir incrementos de estereoselectividad, estereoespecificidad, estabilidad térmica, estabilidad disolvente o inhibición reducida del producto.

En algunas secciones, los polipéptidos de cetorreductasa son mejorados en comparación a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 4, la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48 y / o la IDENTIFICACIÓN SECUENCIAL NÚMERO: 66 con respecto a su tasa de actividad enzimática, es decir, su tasa para convertir el sustrato al producto. En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato del producto a una tasa que es por lo menos 5 veces, 10 veces, 25 veces, 50 veces, 75 veces, 100 veces, 150 veces, 200 veces, 250 veces, 300 veces, 500 veces, o 1000 veces sobre la tasa de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 4 o la IDENTIFICACIÓN SECUENCIAL NÚMERO: 90.

En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, con un exceso porcentual estereomérico de por lo menos alrededor del 85% o con un exceso porcentual estereomérico que es mayor que aquel de KRED de L. kéfir de tipo silvestre (IDENTIFICACIÓN SECUENCIAL NÚMERO: 4). Ejemplo de polipéptidos con esta propiedad incluyen, pero no se limitan a, polipéptidos que tienen secuencias de aminoácidos correspondientes a las IDENTIFICACIONES SECUENCIALES NÚMEROS: 6, 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, y 62.

En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato, metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, con un exceso porcentual estereomérico de por lo menos alrededor del 60 - 89% y a una tasa que es por lo menos alrededor de 1 - 15 veces mayor que la tasa capaz de los polipéptidos de la secuencia de aminoácidos de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48. Ejemplos de polipéptidos con esta propiedad incluyen, pero no se limitan a, polipéptidos que contienen secuencias de aminoácidos correspondientes a las IDENTIFICACIONES SECUENCIALES NÚMEROS: 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 46, 50, 52, 54, 56, 58, 60, y 62. Puesto que el polipéptido referencial que tiene la secuencia de aminoácidos de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48 es capaz de convertir el sustrato al producto a una tasa (por ejemplo, 100% de conversión en 20 horas o 1g / L de sustrato con alrededor de 10 g / l de KRED, en 50% de IPA a un pH 8) y con una estereoselectividad que es mejorada a la de tipo silvestre (IDENTIFICACIÓN SECUENCIAL NÚMERO: 4), los polipéptidos aquí mencionados que están mejorados sobre la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48 también están mejorados en comparación del tipo silvestre.

En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato, metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, con un exceso porcentual estereomérico de por lo menos alrededor del 90 - 94%. Ejemplos de polipéptidos con esta propiedad incluyen, pero no se limitan a, los polipéptidos que tienen secuencias de aminoácidos correspondientes a las IDENTIFICACIONES SECUENCIALES NÚMEROS: 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 40, 42, 50, 52, 56, 58, 60, y 62.

En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato, metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, con un exceso porcentual estereomérico de por lo menos alrededor de 95 - 99%. Ejemplos de polipéptidos con esta propiedad incluyen, pero no se limitan a, los polipéptidos que tienen las secuencias de aminoácidos correspondientes a las

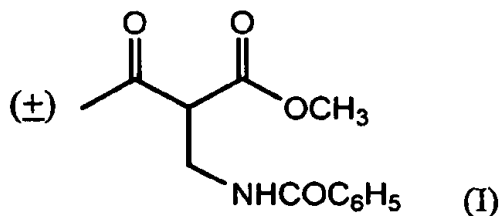
IDENTIFICACIONES SECUENCIALES NÚMEROS: 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 42, 50, 52, 56, 58, 60, y 62.

- 5 En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato, metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, con un exceso porcentual estereomérico de por lo menos alrededor del 99%. Ejemplos de polipéptidos con esta propiedad incluyen, pero no se limitan a, polipéptidos que tienen secuencias de aminoácidos que corresponden a las IDENTIFICACIONES SECUENCIALES NÚMEROS: 6, 8, 10, 12, 14, 20, 22, 24, 30, 32, 34, 60, y 62.
- 10 En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato, metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, a una tasa que es por lo menos alrededor de 15 - 30 veces mayor que la tasa de capacidad del polipéptido que tiene la secuencia de aminoácidos de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48. Ejemplos de polipéptidos con esta propiedad incluyen, pero no se limitan a los polipéptidos que tienen las secuencias de aminoácidos correspondientes a las
- 15 IDENTIFICACIONES SECUENCIALES NÚMEROS: 6, 8, 10, 12, 14, 20, 22, 24, 26, 28, 30, 32, 34, 50, 60, y 62.
- En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato, metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, a una tasa que es por lo menos alrededor de 30 - 40 veces mayor que la tasa de capacidad del polipéptido que tiene la secuencia de
- 20 aminoácidos de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48. Ejemplos de polipéptidos con esta propiedad incluyen, pero no se limitan a, los polipéptidos que contienen las secuencias de aminoácidos correspondientes a las IDENTIFICACIONES SECUENCIALES NÚMEROS: 6, 8, 10, 12, 14, 20, 22, 24, 26, 30, 34, 60, y 62.
- En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato, metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, a una tasa que es por lo menos alrededor de 40 - 50 veces mayor a la tasa de capacidad del polipéptido que tiene la secuencia de
- 25 aminoácidos de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48. Ejemplos de polipéptidos ejemplares con esta propiedad incluyen, pero no se limitan a, los polipéptidos que contienen las secuencias de aminoácidos correspondientes a las IDENTIFICACIONES SECUENCIALES NÚMEROS: 6, 8, 10, 12, 14, 22, y 60.
- 30 En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato, metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, a una tasa que es por lo menos alrededor de 50 veces mayor a la tasa de capacidad del polipéptido que tiene la secuencia de aminoácidos de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48. Ejemplos de polipéptidos con esta propiedad incluyen, pero no se limitan a, los polipéptidos que contienen las secuencias de aminoácidos correspondientes a las identificaciones
- 35 secuenciales números: 6, 8, 10, y 12.
- En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato, metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, a una tasa que es por lo menos alrededor de 50 veces mayor a la tasa de capacidad del polipéptido que tiene la secuencia de aminoácidos de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48 y con un exceso estereomérico de por lo menos el 99%. Ejemplos de polipéptidos con esta propiedad incluyen, pero no se limitan a, los polipéptidos que contienen las
- 40 secuencias de aminoácidos correspondientes a las identificaciones secuenciales números: 6, 8, 10 y 12.
- En algunas secciones, los polipéptidos de cetorreductasa son capaces de retener su capacidad para convertir el sustrato, metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, después de un tratamiento de calor a 40 °C durante 21 horas. Ejemplos de polipéptidos con esta propiedad incluyen, pero no se limitan a, los polipéptidos que tienen las secuencias de aminoácidos correspondientes a las IDENTIFICACIONES
- 45 SECUENCIALES NÚMEROS: 6, 10 y 44.
- 50 En otro aspecto, esta presentación suministra los polinucleótidos que codifican a las cetorreductasas diseñadas aquí descritas o los polinucleótidos que hibridan a aquellos polinucleótidos bajo condiciones altamente rigurosas. El polinucleótido puede incluir elementos regulatorios promotores y de otro tipo útiles para la expresión de la cetorreductasa codificada diseñada, y puede utilizar codones optimizados para sistemas específicos de expresiones
- 55 deseadas. Ejemplos de polinucleótidos incluyen, pero no se limitan a las IDENTIFICACIONES SECUENCIALES NÚMEROS: 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, y 81. Los ejemplos de polinucleótidos también incluyen polinucleótidos que codifican los polipéptidos que corresponden a las fórmulas secuenciales de las IDENTIFICACIONES
- 60 SECUENCIALES NÚMEROS: 83, 84 y 87.
- En otro aspecto, esta presentación suministra células anfitrionas que tienen los polinucleótidos y / o los vectores de expresión aquí descritos. Las células anfitrionas pueden ser *L. kékír*, o *L. brevis* o *L. minor*, o pueden ser un organismo diferente. Las células anfitrionas pueden ser utilizadas para la expresión y aislamiento de las enzimas
- 65 diseñadas de cetorreductasa aquí descritas, o, alternamente, pueden ser utilizadas directamente para la conversión del sustrato al producto estereoisomérico.

Ya sea, ejecutando el método con células completas, extractos celulares o enzimas purificadas de cetorreductasa, una sola enzima de cetorreductasa puede ser utilizada o, alternativamente, mezclas de dos o más enzimas de cetorreductasa pueden ser utilizadas.

5 Tal como se mencionó anteriormente, en algunas secciones, las enzimas de cetorreductasa aquí descritas son capaces de catalizar la reacción de reducción del grupo ceto en el compuesto de la fórmula estructural (I), metil-2-benzamidometil-3-oxobutirato ("el sustrato"):

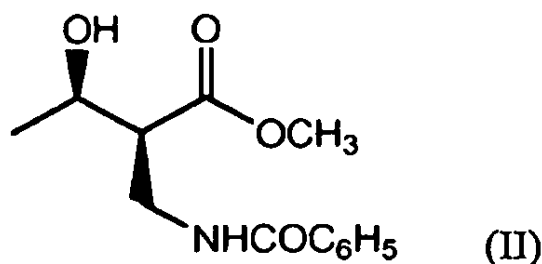
10



15

20 al producto correspondiente de alcohol estereoisomérico de la fórmula estructural (II), 2S,3R-metil-2-benzamidometil-3-hidroxitbutirato ("el producto"):

25



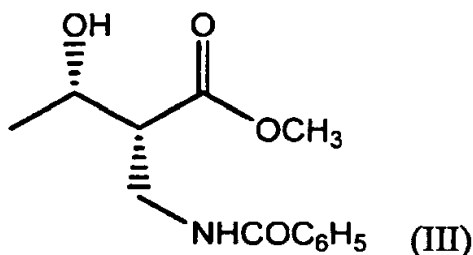
30

35

40 En algunas secciones, las enzimas de cetorreductasa aquí descritas son capaces de catalizar la reacción de reducción del grupo ceto en el compuesto de la fórmula estructural (I) del producto correspondiente de alcohol estereoisomérico de la fórmula estructural (III), 2R,3R-metil-2-benzamidometil-3-hidroxitbutirato, (el cual también puede ser referido como un "producto"):

40

45



50

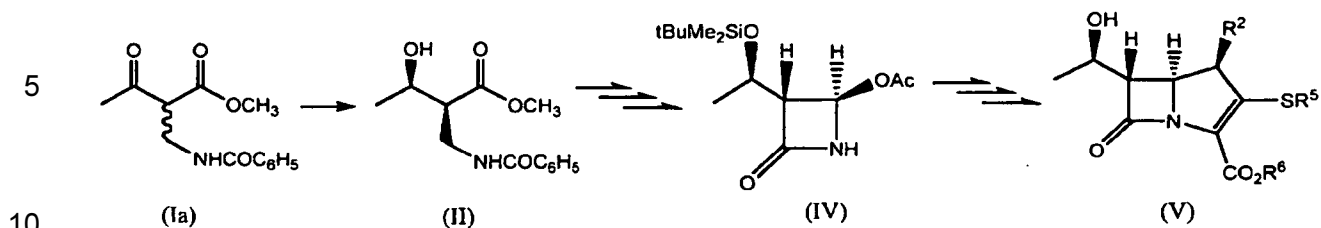
55

Asimismo, en algunas secciones aquí indicadas, existen métodos para reducir el sustrato de la fórmula estructural (I) al producto alcohólico de la fórmula estructural (II) o de la fórmula estructural (III), método que incluye el contactar o incubar el sustrato con un polipéptido de cetorreductasa de la presentación bajo condiciones de reacción adecuadas para la reducción o conversión del sustrato al producto de la fórmula estructural (II) o de la fórmula estructural (III).

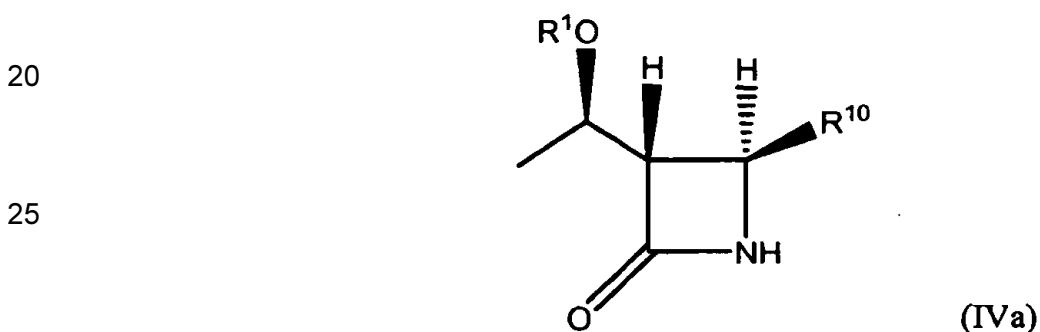
60

En algunas secciones, el producto 2S,3R-metil-2-benzamidometil-3-hidroxitbutirato de la fórmula estructural (II) puede utilizarse para sintetizar compuestos intermedios y cabapenems, tal como se ilustra en el Esquema 3:

65

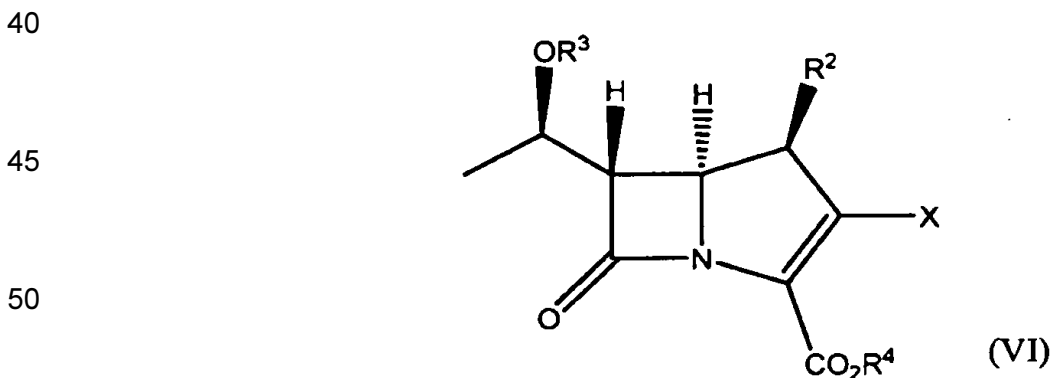


15 Asimismo, en algunas secciones, las cetorreductasas de la presentación pueden ser utilizadas en un método para la síntesis del intermediador de la fórmula estructural (IVa),



35 Dónde R^1 es H o un grupo protector de hidroxilos, y R^{10} es un halógeno (por ejemplo, Cl), o -OAc (Ac es acetato). Asimismo, en un método para la síntesis del intermediador de la fórmula estructural (IVa), un paso en el método incluye el contactar o reaccionar el sustrato de la fórmula (I) con una cetorreductasa de la presentación bajo condiciones de reacción adecuadas para la reducción o conversión del sustrato al producto de la fórmula (II).

40 En algunas secciones, las cetorreductasas de la presentación pueden ser usadas en la síntesis del intermediador de la fórmula estructural (VI):

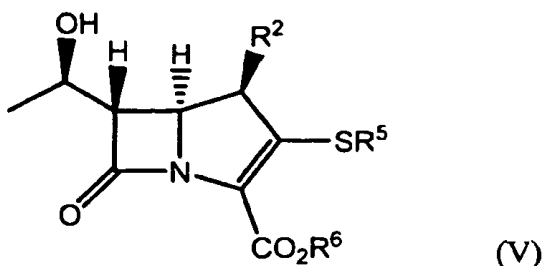


60 Donde R^2 es H o un alquilo C1-C4 (por ejemplo, $-CH_3$); R^3 es H, o un grupo protector de hidroxilo; R^4 es H, o un grupo protector de carboxis, un grupo de amoniaco, un metal alcalino o un metal alcalinotérreo; y X es OH, o un grupo que abandona. Ejemplos de grupos que abandonan incluyen, pero no se limitan a, $-OP(O)(OR')$ o $OS(O_2)R''$, donde R' y R'' pueden ser alquilo C1-C6, alquiloarilo C1-C6, arilo, alquilo perfluorado C1-C6. Asimismo, en algunas secciones, en un método para la síntesis del intermediador de la fórmula (VI), un paso en el método incluye el contactar o reaccionar el sustrato de la fórmula (I) con una cetorreductasa de la presentación bajo condiciones de reacción adecuadas para reducir o convertir el sustrato del producto de la fórmula (II).

65 En algunas secciones, las cetorreductasas de la presentación pueden ser utilizadas en el proceso para la síntesis de carbapenems que se basan en el compuesto terapéutico de la fórmula estructural (V):

5

10



15

20

O sus solvatos, hidratos, sales y pro - medicamentos, donde R² es H o -CH₃; R⁵ puede ser varios sustituyentes, incluyendo, pero sin limitarse a, alquilo sustituido o no sustituido, arilo sustituido o no sustituido, heteroalquilo sustituido o no sustituido; y R⁶ es H, o un progrupo, tal como un grupo de ésteres hidrolizables. Asimismo, en el método para la síntesis del compuesto de la fórmula estructural (V), un paso en el método puede comprender el contactar o reaccionar el sustrato de la fórmula (I) con las cetorreductasas de la presentación bajo condiciones de reacción adecuadas para la reducción o conversión del sustrato del producto de la fórmula (II). Ejemplos de carbapenems de la fórmula estructural (V) incluyen, pero no se limitan a, imipenem, meropenem, Doripenem, ertapenem, Biopenem y panipenem.

25

En algunas secciones, la presentación también suministra el uso de las cetorreductasas en métodos para sintetizar los compuestos sulopenems y en los intermediadores usados en la síntesis de sulopenems.

DESCRIPCIÓN BREVE DE LAS FIGURAS

30

La figura uno ilustra el rol de las cetorreductasas (KRED – ketoreductases) en la conversión del compuesto sustrato de la fórmula (I) al producto correspondiente de la fórmula (II). La reducción utiliza una KRED del invento y un cofactor tal como el NADPH. Alcohol isopropilo (IPA – isopropyl alcohol) es utilizado para convertir / reciclar el NADP* a NADPH.

DESCRIPCIÓN DETALLADA

135

Definiciones

Tal como se usa aquí, los siguientes términos tienen el propósito de tener los siguientes significados.

40

“Cetorreductasa” y “KRED” se utilizan intercambiamente en este documento para referirse a un polipéptido que tiene una capacidad enzimática para reducir un grupo carboxilo a su alcohol correspondiente. Más específicamente, los polipéptidos de cetorreductasa del invento son capaces de reducir en una forma estereoselectiva el compuesto de la fórmula (I), mencionada anteriormente, al producto correspondiente de la fórmula (II), mencionada anteriormente. El polipéptido utiliza comúnmente un cofactor de nicotinamida adenina dinucleótida reducida (NADH) o nicotinamida – adenina – dinucleótido - fosfato reducido (NADPH – nicotinamide adenine dinucleotide phosphate) como el agente reductor. Las cetorreductasas como se utilizan aquí incluyen cetorreductasas que ocurren naturalmente (tipo silvestre) así como los polipéptidos que no ocurren naturalmente y que son diseñados y generados por la manipulación artificial.

50

“Secuencia de codificación” se refiere a la opción del ácido nucleico (por ejemplo, un gen) que codifica la secuencia de aminoácidos de una proteína.

55

“Que ocurren naturalmente” o “de tipo silvestre” se refiere a la forma que se encuentra en la naturaleza. Por ejemplo, una secuencia de un polipéptido o polinucleótido que ocurren naturalmente o de tipo silvestre es una secuencia presente en un organismo que puede ser aislado de una fuente en la naturaleza y que no ha sido modificado intencionalmente por cualquier manipulación humana.

60

“Recombinante” cuando se lo utiliza en referencia a, por ejemplo, una célula, un ácido nucleico o un polipéptido, se refiere a un material, o un material proveniente de la forma natural o silvestre del material, que ha sido modificado en una forma por la cual de no ser así no existiría en la naturaleza, o es idéntica pero no se produce o se deriva de materiales sintéticos y / o por medio de manipulación usando técnicas recombinantes. Ejemplos no limitantes incluyen, entre otros, genes que expresan células recombinantes que no se encuentran dentro de la forma natural (no recombinante) de la célula o genes que se expresan naturalmente que de otra manera se expresan en un nivel diferente.

65

“Porcentaje de la identidad secuencial” y “homología porcentual” se utilizan intercambiamente en este documento para referirse a las comparaciones entre polinucleótidos y polipéptidos, y se determinan al comparar dos secuencias

alineadas óptimamente sobre una ventana de comparación, donde la porción de la secuencia del polinucleótido o polipéptido en la ventana de comparación puede incluir adiciones o sustracciones (por ejemplo, vacíos) cuando se los compara con la secuencia referencial (que no contiene adiciones o sustracciones) para la alineación óptima de las dos secuencias. El porcentaje puede ser calculado al determinar el número de posiciones en las cuales la base idéntica de ácido nucleico o residuo de aminoácidos ocurre en ambas secuencias para producir el número de posiciones emparejadas, dividiendo el número de posiciones emparejadas para el número total de posiciones en la ventana de comparación y multiplicando el resultado por 100 para obtener el porcentaje de identidad secuencial. Alternativamente, el porcentaje puede ser calculado al determinar el número de posiciones en el cual cualquiera de la base idéntica de ácido nucleico o de residuo de aminoácidos ocurre en ambas secuencias o una base de ácido nucleico o de residuo de aminoácidos está alineada con un vacío para obtener el número de posiciones emparejadas, dividiendo el número de posiciones emparejadas por el número total de posiciones en la ventana de comparación y multiplicando el resultado por 100 para obtener el porcentaje de la identidad secuencial. Aquellos con experiencia en la industria podrán apreciar que existen muchos algoritmos establecidos disponibles para alinear dos secuencias. Se pueden realizar alineamientos óptimos de las secuencias para su comparación, por ejemplo, por medio del algoritmo de homología local de Smith y Waterman, 1981, *Adv. Appl. Math.* 2:482, por medio del algoritmo de alineación de homología de Needleman y Wunsch, 1970, *J. Mol. Biol.* 48:443, por medio del método de búsqueda de similitudes de Pearson y Lipman, 1988, *Proc. Natl. Acad. Sci. EE.UU.* 85:2444, por medio de implementaciones informáticas de estos algoritmos (GAP, BESTFIT, FASTA, y TFASTA en el paquete de software GCG Wisconsin), o por medio de una inspección visual (refiérase generalmente, a *Current Protocols in Molecular Biology - Protocolos Actuales en Biología Molecular*), F. M. Ausubel et al., eds., *Current Protocols (Protocolos Actuales)*, un proyecto conjunto entre Greene Publishing Associates, Inc. y John Wiley & Sons, Inc., (suplemento de 1995) (Ausubel)). Ejemplos de algoritmos que son adecuados para determinar la identidad secuencial porcentual y la similitud secuencial 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. (Ácidos nucleicos Res.)* 3389-3402, respectivamente. El software para realizar análisis BLAST está disponible únicamente por medio de la página web del National Center for Biotechnology Information (Centro Nacional de Información Biotecnológica). Este algoritmo incluye el identificar primero parejas secuenciales de alto puntaje (HSPs – high scoring sequence pairs) al identificar palabras cortas de longitud W en la secuencia consultada, que cuadren o satisfagan algún puntaje T de límites de valores positivos cuando se alinea con una palabra del mismo largo en una secuencia de la base de datos. T es referido como, el límite de puntaje de la palabra del vecindario (Altschuk et al, mencionado anteriormente). Estas coincidencias iniciales de palabras del vecindario actúan como semillas para iniciar búsquedas para encontrar HSPs más largos que las contenga. Las coincidencias de palabras se extienden en ambas direcciones a lo largo de la secuencia mientras el puntaje acumulado de alineación pueda incrementarse. Los puntajes acumulados son calculados utilizando, para secuencias nucleótidas, los parámetros M (puntaje de premio para una pareja de residuos que coinciden; siempre >0) y N (puntaje de penalización para residuos que no coincida; siempre <0). Para secuencias de aminoácidos, se utiliza una matriz de puntaje para calcular el puntaje acumulado. La extensión de coincidencias de palabras en cada dirección es interrumpida cuando: el puntaje acumulado de alineación cae por debajo de la cantidad x de su valor de alcance máximo; el puntaje acumulado baja a cero o más abajo, debido a la acumulación de una o más alineaciones de residuos con puntaje negativo; o el final de cualquiera de las secuencias es alcanzado. Los parámetros del algoritmo BLAST W, T y X determinan la sensibilidad y velocidad de la alineación. El programa BLASTN (para secuencias nucleótidas) utiliza como predeterminados una longitud de palabra (W) de 11, una expectativa (E) de 10, M = 5, N = -4, y una comparación de ambas hebras. Para secuencias de aminoácidos, el programa BLASTP usa como predeterminados una longitud de palabra (W) de 3, con una expectativa (E) de 10, y la matriz de puntaje BLOSUM62 (refiérase a Henikoff y Henikoff, 1989, *Proc Natl Acad Sci EE.UU.* 89:10915). Un ejemplo de determinación de alineación secuencial e identidad secuencial porcentual puede emplear los programas BESTFIT o GAP en el paquete de software GCG Wisconsin (Accelrys, Madison, WI), usando los parámetros predeterminados que fueron suministrados.

“Secuencia referencial” se refiere a una secuencia definida utilizada como una base para una comparación secuencial. Una secuencia referencial puede ser el subconjunto de una secuencia más larga, por ejemplo, segmento de una secuencia completa genética o de polipéptidos. Generalmente, una secuencia referencial tiene un largo de por lo menos 20 residuos nucleótidos o de aminoácidos, por lo menos 25 residuos de largo, por lo menos 50 residuos de largo, o el largo completo del ácido nucleico o el polipéptido. Puesto que dos polinucleótidos o polipéptidos pueden cada uno (1) tener una secuencia (es decir, una porción de la secuencia, letra) que es similar entre las dos secuencias, y (2) puede tener además una secuencia que es divergente entre las dos secuencias, las comparaciones secuenciales entre las dos (o más) polinucleótidos o polipéptido son realizados comúnmente al comparar las secuencias de los dos polinucleótidos en una “ventana de comparación” para identificar y comparar las regiones locales de similitud secuencial.

En algunas secciones, una “secuencia referencial” puede basarse en una secuencia de aminoácidos primaria, donde la secuencia referencial es una secuencia que tiene uno o más cambios en la secuencia primaria. Por ejemplo, una “secuencia referencial que se basa en la IDENTIFICACIÓN SECUENCIAL NÚMERO: 4 que tiene en el residuo correspondiente a X202 una leucina o valina” se refiere a una secuencia referencial en la cual el residuo correspondiente en X 202 en la IDENTIFICACIÓN SECUENCIAL NÚMERO: 4 ha sido cambiada a leucina o valina.

65

“Ventana de comparación” se refiere a un segmento conceptual de por lo menos 20 posiciones nucleotidasas o residuos de aminoácidos contiguos donde una secuencia puede ser comparada a una secuencia referencial de por lo menos 20 nucleótidos o aminoácidos contiguos y donde la porción de la secuencia en la ventana de comparación podría fluir adiciones o sustracciones (es decir, vacíos) del 20% o menos cuando se compara con la secuencia referencial (la cual no incluye adiciones o sustracciones) para la alineación óptima de las dos secuencias. La ventana de comparación puede ser más larga que 20 residuos contiguos, e incluye, opcionalmente, 30, 40, 50, 100 o ventanas más largas.

“Identidad sustancial” se refiere a una secuencia de polinucleótidos o polipéptidos que tiene por lo menos un 80% de identidad secuencial, por lo menos un 85% de identidad porcentual y del 89 al 95% de identidad secuencial, más comúnmente por lo menos un 99% de identidad secuencial cuando se compara a una secuencia referencial en una ventana de comparación de por lo menos 20 posiciones residuales, frecuentemente en una ventana de por lo menos 30 - 50 residuos, donde el porcentaje de identidad secuencial se calcula al comparar la secuencia referencial a una secuencia que incluye sustracciones o adiciones que totalizan el 20% o menos de la secuencia referencial en la ventana de comparación. En secciones específicas aplicadas a los polipéptidos, el término “identidad sustancial” se refiere a que dos secuencias de polipéptidos, cuando están alineados óptimamente, como lo hacen los programas GAP o BESTFIT utilizando los pesos predeterminados de vacíos, comparten por lo menos un 80% de la identidad secuencial, preferiblemente por lo menos un 89% de identidad secuencial, por lo menos un 95% de identidad porcentual o más (por ejemplo, 99% de identidad secuencial). Preferiblemente, las posiciones residuales que no son idénticas difieren debido a sustituciones conservadoras de aminoácidos.

“Correspondiente a”, “en referencia a” o “en relación a” cuando se utilizan en el contexto de la numeración de una secuencia de aminoácidos o polinucleótidos se refiere a la numeración de los residuos de una secuencia referencial especificada cuando la secuencia de aminoácidos o polinucleótidos en cuestión es comparada a la secuencia referencial. En otras palabras, el número residual o posición residual de un polímero específico es designado en relación a la secuencia referencial en vez de en relación a la posición numérica actual del residuo dentro de la secuencia de aminoácidos o polinucleótidos en cuestión. Por ejemplo, una secuencia específica de aminoácidos, en la cual una cetorreductasa diseñada, puede alinearse a una secuencia referencial al introducir vacíos para optimizar la coincidencia residual entre las dos secuencias. En estos casos, aunque los vacíos están presentes, la numeración de los residuos en la secuencia de aminoácidos o poli nucleótidos en cuestión es hecha en relación a la secuencia referencial a la cual se ha alineado.

“Estereoselectividad” se refiere a la formación preferencial en una reacción química o enzimática de un estereoisómero por sobre otro. La estereoselectividad puede ser parcial, cuando la formación de un estereoisómero es preferida sobre otro, o puede ser completa cuando sólo se forma estereoisómero. Cuando los estereoisómeros son enantiómeros, la estereoselectividad es referida como enantioselectividad, la fracción (comúnmente reportada como un porcentaje) de un enantiómero es la suma de ambos. Comúnmente en la industria se reporta alternamente (típicamente como un porcentaje) como un exceso enantiomérico (e.e.) Calculados consecuentemente de acuerdo a la fórmula $[\text{enantiómero mayor} - \text{enantiómero menor}] / [\text{enantiómero mayor} + \text{enantiómero menor}]$. En los casos en que los estereoisómeros son diastereoisómeros, la estereoselectividad es referida como de diaestereoselectividad, la fracción (reportada comúnmente como un porcentaje) de un diaestereómero en una mezcla de dos diaestereómeros, comúnmente reportada alternamente como el exceso diaestereomérico (d.e. – diastomeric excess). El exceso enantiomérico y el exceso diaestereomérico son tipos de excesos estereoméricos.

“Altamente estéreselectivo” se refiere a un polipéptido de cetorreductasa que es capaz de convertir o reducir el sustrato al producto correspondiente que tiene la fórmula química (II) o (III) con por lo menos alrededor de un 85% de exceso estereomérico.

“Estereoespecificidad” se refiere a la conversión preferencial en una reacción química o enzimática de un estereoisómero sobre otro. La estereoespecificidad puede ser parcial, cuando la conversión de un estereoisómero es preferida por sobre otro, o puede ser completa cuando solamente un estereoisómero es convertido.

“Quimioselectividad” se refiere a la formación preferencial en una reacción química o enzimática de un producto sobre otro.

“Propiedad enzimática mejorada” se refiere a un polipéptido de cetorreductasa que muestra una mejora en cualquier propiedad enzimática en comparación a una cetorreductasa referencial. Para los polipéptidos diseñados de cetorreductasa aquí descritos, la comparación es hecha generalmente con la enzima de tipo silvestre de cetorreductasa, aunque en algunas secciones, la cetorreductasa referencial podría ser otra cetorreductasa diseñada mejorada. Las propiedades enzimáticas en las cuales una mejora es deseable incluyen, pero no se limitan a, la actividad enzimática (que puede ser expresada en términos de conversión porcentual del sustrato), estabilidad térmica, estabilidad disolvente, perfil de actividad pH, requerimientos de los cofactores, refractariedad a los inhibidores (por ejemplo, inhibición del producto), estereoespecificidad y estéreselectividad (incluyendo la enantioselectividad).

5 “Actividad enzimática incrementada” se refiere a una propiedad mejorada de los polipéptidos diseñados de cetorreductasa, que pueden presentarse por medio de un incremento en actividades específicas (por ejemplo, producto producido / tiempo / peso proteína) o un incremento en la conversión porcentual del sustrato del producto (por ejemplo, la conversión porcentual del monto de inicio del sustrato al producto en un período de tiempo especificado usando un monto especificado de KRED) en comparación con la enzima referencial de cetorreductasa. Ejemplos de métodos para determinar la actividad enzimática se suministran en los Ejemplos. Cualquier propiedad que se relaciona a la actividad enzimática puede ser afectada, incluyendo las propiedades enzimáticas clásicas de Km, Vmax o kcat, cuyos cambios pueden conllevar a incrementar la actividad enzimática. Las mejoras en actividad enzimática podrían ser de alrededor de 1.5 veces la actividad enzimática de la enzima correspondiente de cetorreductasa de tipo silvestre, hasta 2 veces, 5 veces, 10 veces, 20 veces, 25 veces, 50 veces, 75 veces, 100 veces o más actividad enzimática que la cetorreductasa que ocurren naturalmente u otra cetorreductasa diseñada de la cual se derivaron los polipéptidos de cetorreductasa. Se entiende por una persona con experiencia que la actividad de cualquier enzima es limitada por la difusión de tal forma que la tasa de producción catalítica no puede exceder la tasa de difusión del sustrato, incluyendo cualquier cofactor requerido. El máximo teórico del límite de difusión, o Kcat / Km, es generalmente alrededor de 10^8 a 10^9 ($M^{-1} s^{-1}$). Por lo tanto, cualquier mejora en la actividad enzimática de la cetorreductasa tendrá un límite superior relacionado a la tasa de difusión de los sustratos que son influenciados por la enzima cetorreductasa. La actividad cetorreductasa puede ser medida por medio de cualquier ensayo estándar utilizado para medir la cetorreductasa, tal como la disminución en la absorción de fluorescencia de NADPH debido a su oxidación con la reducción concomitante de una cetona a un alcohol, o por medio del producto producido en un ensayo de emparejamiento. Las comparaciones de las actividades enzimáticas son hechas utilizando una preparación definida de enzimas, un ensayo definido bajo una condición establecida, y uno o más sustratos definidos, tal como se describe en mayor detalle en este documento. Generalmente, cuando se comparan los lisados, los números de células y el monto de proteínas en los ensayos son determinados así como el uso de sistemas idénticos de expresión y células anfitrionas idénticas para minimizar variaciones en el monto enzimático producido por las células anfitrionas y presentes en los lisados.

15 “Conversión” se refiere a la reducción enzimática del sustrato al producto correspondiente. “Conversión porcentual” se refiere al porcentaje del sustrato que es reducido al producto dentro de un período de tiempo bajo las condiciones especificadas. Por lo tanto, la “actividad enzimática” o la “actividad” de un polipéptido de cetorreductasa puede expresarse como una “conversión porcentual” del sustrato del producto.

20 “Térmicamente estable” se refiere al polipéptido de cetorreductasa que mantiene una actividad similar (más que el 60% al 80%, por ejemplo) después de su exposición a altas temperaturas (por ejemplo, 40 - 80 °C) por un período de tiempo (por ejemplo, 0.5 - 24 horas) en comparación con la enzima que no ha sido tratada.

25 “Estabilidad de los disolventes” se refiere a un polipéptido de cetorreductasa que mantiene una actividad similar (más que, un ejemplo, 60% a 80%) después de su exposición a concentraciones variables (por ejemplo, 5 - 99%) de solventes (isopropilalcohol, tetrahidrofurano, 2-metilhidrofurano, acetona, tolueno, butilacetato, terc-butiletter metilo, etc.) durante un período de tiempo (por ejemplo, 0.5 - 24 horas) en comparación con la enzima que no ha sido tratada.

30 “Con pH estable” se refiere a un polipéptido de cetorreductasa que mantiene una actividad similar (más que, por ejemplo, el 60% al 80%) después de su exposición a un pH alto o bajo (por ejemplo, 4.5 - 6 u 8 a 12) durante un período de tiempo (por ejemplo, 0.5 - 24 horas) en comparación a la enzima que ha sido tratada.

35 “Con estabilidad térmica y disolvente” se refiere a un polipéptido de cetorreductasa que es estable térmicamente y estable en lo que se refiere a disolventes.

40 “Derivado de” como se utiliza aquí en el contexto de enzimas de cetorreductasa diseñadas, identifica la enzima de donde se origina la cetorreductasa, y / o la codificación genética de aquella enzima cetorreductasa, de donde se basó el diseño. Por ejemplo, la enzima de cetorreductasa diseñada de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 60 fue obtenida al evolucionar artificialmente por varias generaciones el gen que codifica a la enzima cetorreductasa del lactobacilo kéfir de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 4. Por lo tanto, esta enzima cetorreductasa es “derivada de” la cetorreductasa de tipo silvestre de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 4.

45 “Aminoácido o residuo hidrofílico” se refiere a un aminoácido o residuo que tiene una cadena lateral que muestra una hidrofobicidad menor que cero de acuerdo al consenso estandarizado de la escala de hidrofobicidad de Eisenberg et al., 1984, J. Mol. Biol. (Biología Molecular) 179:125-142. Los aminoácidos hidrofílicos 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).

50 “Aminoácido o residuo ácido” se refiere a un aminoácido o residuo hidrofílico que tiene una cadena lateral que muestra un valor pK menor que aproximadamente 6 cuando el aminoácido es incluido en un péptido o polipéptido. Los aminoácidos ácidos tienen típicamente cadenas laterales cargadas negativamente al pH fisiológico debido a la pérdida de un ion de hidrógeno. Los aminoácidos a cívicos codificados genéticamente incluyen L-Glu (E) y L-Asp (D).

- 5 “Aminoácido o residuo básico” se refiere a un aminoácido o residuo hidrofílico que tiene una cadena lateral que incluye un valor pK mayor que aproximadamente 6 cuando el aminoácido es incluido en un péptido o polipéptido. Los aminoácidos básicos típicamente tienen cadenas laterales con carga positiva con un pH fisiológico debido a la asociación con un ion de hidronio. Aminoácidos básicos codificados genéticamente incluyen L-Arg (R) y L-Lys (K).
- 10 “Aminoácido o residuo polar” se refiere a un aminoácido o residuo hidrofílico que tiene una cadena lateral que no tiene carga con un pH fisiológico, pero tiene por lo menos un enlace en el cual la pareja de electrones compartidos en común por dos átomos es sostenido más de cerca por uno de los átomos. Los aminoácidos polares codificados genéticamente incluyen L-Asn (N), L-Gln (Q), L-Ser (S) y L-Thr (T).
- 15 “Aminoácido o residuo hidrofóbico” se refiere a un aminoácido o residuo que tiene una cadena lateral que contiene una hidrofobicidad > 0 de acuerdo a la escala de hidrofóbico oxidar de consenso estandarizado de Eisenberg et al., 1984, J. Mol Biol. 179:125-142. Aminoácidos hidrofóbicos codificados genéticamente incluyen a 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).
- 20 “Aminoácido o residuo aromático” se refiere a un aminoácido o residuo hidrofílico o hidrofóbico que tiene una cadena lateral que incluye por lo menos un anillo aromático o heteroaromático. Los aminoácidos aromáticos codificados genéticamente incluyen L-Phe (F), L-Tyr (Y) y L-Trp (W). Aunque le debe al pKa su átomo de nitrógeno entero aromático, L-His (H) es clasificado algunas veces como un residuo básico, o como un residuo aromático puesto que su cadena lateral incluye un anillo entero aromático, en este documento la histidina es clasificada como un residuo hidrofílico o como un “residuo restringido” (refiérase a continuación).
- 25 “Aminoácido o residuo restringido” se refiere a un aminoácido o residuo que tiene una geometría restringida. En este documento, los residuos restringidos incluyen a L-pro (P) y L-his (H). La histidina tiene una geometría restringida porque tiene un anillo imidazol relativamente pequeño. Prolina tiene una geometría restringida porque además tiene un anillo de cinco miembros.
- 30 “Aminoácido o residuo no polar” se refiere a un aminoácido o residuo hidrofóbico que tiene una cadena lateral que no está cargada un pH fisiológico que tiene enlaces en los cuales la pareja de electrones compartidos en común por los dos átomos es igualmente mantenido en general por cada uno de los dos átomos (es decir, la cadena lateral no es polar). Los aminoácidos no polares codificados genéticamente incluyen a L-Gly (G), L-Leu (L), L-Val (V), L-Ile (I), L-Met (M) y L-Ala (A).
- 35 “Aminoácido o residuo alifático” se refiere a un aminoácido o residuo hidrofóbico que tiene una cadena lateral de hidrocarbóno alifático. Los aminoácidos alifáticos su codificados genéticamente incluyen a L-Ala (A), L-Val (V), L-Leu (L) y L-Ile (I).
- 40 “Cisteína (Cysteine)” o L-Cys (C) es inusual puesto que puede formar puentes de disulfuro con otros aminoácidos L-Cys (C) u otros aminoácidos que contienen sulfanilo- o sulfhidrilo-. Los “recibos similares a la cisteína” incluyen cisteína y otros aminoácidos que contienen fracciones de sulfhidrilo que están disponibles para la formación de puentes de disulfuro. La capacidad del L-Cys (C) (y otros aminoácidos que tienen cadenas laterales que contienen -SH) para existir en un péptido ya sea en una forma reducida o libre de -SH o componentes de disulfuro oxidados afecta si es que el L-Cys (C) contribuye con una característica hidrofóbica o hidrofílica neta a un péptido. Mientras un L-Cys (C) muestra una hidrofobicidad de 0.29 de acuerdo a la escala de consenso estandarizado de Eisenberg (Eisenberg et al., 1984, mencionado anteriormente), debe quedar claro que para propósitos de esta presentación L-Cys (C) es clasificado en su propio grupo único punto
- 45
- 50 “Aminoácido o residuo pequeño” se refiere a un aminoácido o residuo que tiene una cadena lateral que está compuesta de un total de tres o menos carbonos y / o heteroátomos (excluyendo el carbono - α y los hidrógenos). Los aminoácidos o residuos pequeños pueden ser categorizados además en aminoácidos o residuos pequeños alifáticos, no polares, polares o ácidos, de acuerdo a las definiciones mencionadas anteriormente. Los aminoácidos pequeños codificados genéticamente incluyen a L-Ala (A), L-Val (V), L-Cys (C), L-Asn (N), L-Ser (S), L-Thr (T) y L-Asp (D).
- 55
- 60 “Aminoácido o residuo que contiene hidroxilo” se refiere a un aminoácido que contiene una fracción de hidroxilo (-OH). Los aminoácidos que contienen hidroxilo y que son codificados genéticamente incluyen L-Ser (S) L-Thr (T) y L-Tyr (Y).
- 65 Sustituciones o mutaciones “conservadoras” de aminoácidos se refieren a la intercambiabilidad de residuos que tienen cadenas laterales similares, y por lo tanto involucran comúnmente la sustitución del aminoácido en el polipéptido con aminoácidos dentro de la misma o una clase definida similar de aminoácidos. Sin embargo, tal como se utiliza aquí, en algunas secciones, las mutaciones conservadoras no incluyen sustituciones de un residuo hidrofílico a hidrofílico, hidrofóbico a hidrofóbico, que contienen hidroxilo a otros que contienen hidroxilo, o de pequeños a pequeños, si la mutación conservadora puede en vez de eso ser una sustitución desde un residuo alifático a un alifático, un residuo no polar a un no polar, un residuo polar a un polar, un residuo ácido a un ácido, un

residuo básico a un básico, un residuo aromático a un aromático o un residuo restringido a un restringido. Además, tal como se utiliza aquí, A, V, L o I pueden ser mutados conservadoramente a otro residuo alifático o a otro residuo no polar. La tabla a continuación muestra ejemplos de sustituciones conservadoras.
[0076] Tabla 1

5

10

15

20

Residuo	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
Y, W, F	Otro aromático (Y, W, F)
C	Ninguno

25

“Sustitución no conservadora” se refiere a la sustitución o mutación de un aminoácido en el polipéptido con un aminoácido con propiedades de la cadena lateral que difieren significativamente. La sustituciones no conservadoras podrían utilizar aminoácidos entre, en desde dentro, de los grupos definidos que se listaron anteriormente. En una sección, una mutación no conservadora afecta (a) la estructura del esqueleto del péptido en el área de sustitución (por ejemplo, prolina en vez de glicina) (b) la carga o la hidrofobicidad, o (c) una parte importante de la cadena lateral.

30

“Eliminación” se refiere a la modificación del polipéptido al remover uno o más aminoácidos del polipéptido referencial. Las eliminaciones incluyen la remoción de 1 o más aminoácidos, 2 o más aminoácidos, 5 o más aminoácidos, 10 o más aminoácidos, 15 o más aminoácidos, o 20 o más aminoácidos, hasta el 10% del número total de aminoácidos, hasta el 20% del número total de aminoácidos, o hasta el 30% del número total de aminoácidos que componen al polipéptido reteniendo la actividad enzimática y / o reteniendo las propiedades mejoradas de una enzima cetorreductasa diseñada. Las eliminaciones pueden ser dirigidas a las porciones internas y / o terminales del polipéptido. En varias secciones, la eliminación puede incluir un segmento o pueden ser no continuas.

35

40

“Inserción” se refiere a la modificación del polipéptido al añadir uno o más aminoácidos del polipéptido referencial. En algunas secciones, las enzimas mejoradas diseñadas de cetorreductasa contienen inserciones de uno o más aminoácidos al polipéptido de cetorreductasa que ocurren naturalmente así como inserciones de uno o más aminoácidos a otros polipéptidos mejorados de cetorreductasa. Las inserciones pueden ser en las porciones internas del polipéptido, o al término del carboxilo o del amino. Las inserciones, como se utilizan aquí, incluyen proteínas de fusión como se conoce en la industria. La inserción puede ser un segmento contiguo de aminoácidos o segmentos separados por uno o más aminoácidos en el polipéptido que ocurre naturalmente.

45

“Fragmento” tal como se utiliza en este documento se refiere a un polipéptido que tiene una eliminación amino - terminal y / o carboxilo - terminal, pero donde la secuencia remanente de aminoácidos es idéntica a la posición correspondiente en la secuencia. Los fragmentos pueden ser de por lo menos 14 aminoácidos de largo, por lo menos 20 aminoácidos de largo, por lo menos 50 aminoácidos de largo o más largos, y hasta el 70%, 80%, 90%, 95%, 98% y 99% del largo total del polipéptido de cetorreductasa, por ejemplo el polipéptido de las IDENTIFICACIONES SECUENCIALES NÚMERO: 2, 4 u 86.

50

55

“Polipéptido aislado” se refiere a un polipéptido que está separado sustancialmente de otros contaminantes que lo acompañan naturalmente, por ejemplo, proteínas, lípidos y polinucleótidos. El término abarca polipéptidos que han sido removidos o purificados de sus en tornos o sistema de expresión que ocurren naturalmente (por ejemplo, célula anfitriona o síntesis in vitro). Las enzimas mejoradas de cetorreductasa pueden estar presentes dentro de una célula, presentes en el medio celular, o preparadas en varias formas, tales como preparaciones rizadas o aisladas. Como tales, en algunas secciones, la enzima mejorada de cetorreductasa puede ser un polipéptido aislado.

60

“Polipéptido sustancialmente puro” se refiere a una composición en la cual la especie de polipéptido es la especie predominante presente (es decir, en una base molar o de masa es más abundante que cualquier otra especie molecular individual en la composición), y es generalmente una composición sustancialmente purificada cuando la especie en cuestión incluye por lo menos alrededor del 50%, de las especies macromoleculares presentes por mol o masa porcentual. Generalmente, una composición sustancialmente pura de cetorreductasa incluirá alrededor del 60% o más, alrededor del 70% o más, alrededor del 80% o más, alrededor del 90% o más, alrededor del 95% o más, y alrededor del 98% o más de todas las especies macro moleculares por mol o por masa porcentual presentes en la composición. En algunas secciones, la especie en cuestión es purificada a una homogeneidad esencial (es decir, los

65

contaminantes no pueden ser detectados en la composición por medio de métodos de detección convencionales) donde la composición consiste esencialmente de una sola especie macromolecular. Las especies solventes, moléculas pequeñas (menos de 500 Daltons), y especies iónicas elementales no son consideradas especies macro moleculares. En algunas secciones, los polipéptidos aislados mejorados de cetorreductasas es una composición de polipéptidos sustancialmente pura.

“Hibridación rigurosa” como se la utiliza en este documento se refiere a las condiciones bajo las cuales los híbridos ácido nucleicos son estables. Tal como se conoce en la industria, la estabilidad de los híbridos se refleja en la temperatura de derretimiento (T_m – melting temperature) de los híbridos. En general, la estabilidad de un híbrido es una función de la fortaleza iónica, temperatura, contenido G / C, y la presencia de agentes caotrópicos. Los valores T_m para polinucleótidos pueden calcularse usando métodos conocidos para predecir las temperaturas de derretimiento (referirse, por ejemplo, a Baldino et al., Methods Enzymology (Métodos Enzimológicos) 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; Kierzck et al., Biochemistry (Bioquímica) 25:7840-7846; Rychlik et al., 1990, Nucleic Acids Res (Ácidos Nucleicos Res) 18:6409-6412 (erratum, 1991, Nucleic Acids Res (Res Ácidos Nucleicos) 19:698); Sambrook et al., mencionado anteriormente); Suggs et al., 1981, In Developmental Biology Using Purified Genes (En Biología de Desarrollo Usando Genes Purificados) (Brown et al., eds.), pg. 683-693, Academic Press; y Wetmur, 1991, Crit Rev Biochem Mol Biol 26:227-259. En algunas secciones, los polinucleótidos codifica el polipéptido aquí presentado e hibridan bajo condiciones definidas, tales como moderadamente rigurosas o altamente rigurosas, al complemento de una secuencia que codifica una enzima diseñada de cetorreductasa de esta presentación.

“Rigurosidad de la hibridación” se relaciona a las condiciones de hibridación, tales como las condiciones de lavado, en la hibridación de ácidos nucleicos. Generalmente, las reacciones de hibridación son realizadas bajo condiciones de rigurosidad baja, seguidas por lavados con varios niveles de rigurosidad pero más altos. El término “hibridación moderadamente rigurosa” se refiere a las condiciones que permiten que un ADN - objetivo enlace un ácido nucleico complementario que tiene alrededor de un 60% de identidad, preferiblemente alrededor de un 75% de identidad, alrededor del 85% identidad que el del ADN objetivo; con más de alrededor de un 90% de identidad que el del polinucleótido objetivo. Ejemplos de condiciones moderadamente rigurosas son equivalentes para la hibridación en un 50% de formamida, 5x de la solución de Denhart, 5xSSPE, 0.2% SDS a 42 °C, seguido de lavados en 0.2 x SSPE, 0.2% de SDS a 42 °C. “Hibridación de alta rigurosidad” se refiere generalmente a las condiciones que están a alrededor de 10 °C o menos de la temperatura de derretimiento térmico T_m tal como se determina bajo la condición de la solución para una secuencia polinucleótida definida. En algunas secciones, una condición altamente rigurosa se refiere a las condiciones que permiten la hibridación de solamente aquellas secuencias de ácidos nucleicos que forman híbridos estables en 0.018M NaCl a 65 °C. (Es decir, si un híbrido no es estable en 0.018M NaCl a 65 °C, no será estable bajo condiciones de alta rigurosidad, tal como se lo contempla en este documento). Las condiciones de alta rigurosidad pueden ser suministradas, por ejemplo, por la hibridación en condiciones equivalentes a un 50% de formamida, 5x de solución de Denhart, 5 x SSPE, 0.2% de SDS a 42 °C, seguido por lavados en 0.1 x SSPE y 0.1% de SDS a 65 °C. Otra condición de alta rigurosidad es la hibridación en condiciones equivalentes a la hibridación en 5X SSC que contiene 0.1 por ciento (masa: volumen), SDS a 65 °C y lavados en 0.1 x SSC que contiene 0.1 por ciento de SDS a 65 °C. Otras condiciones de hibridación de alta rigurosidad, así como condiciones de rigurosidad moderada, están descritas en las referencias ya mencionadas.

Polinucleótido “Heterólogo” se refiere a cualquier polinucleótido que es introducido a una célula anfitriona por medio de técnicas de laboratorio, e incluye polinucleótidos que son removidos de una célula anfitriona, sujetos a manipulación de laboratorio, y entonces reintroducidos a una célula anfitriona.

“Optimización de codones” se refiere a cambios en los codones del polinucleótido que codifica una proteína los cuales son usados preferencialmente en un organismo en particular de tal forma que la proteína codificada es expresada eficientemente en el organismo objetivo. Aunque el código genético se degenera en el sentido que la mayoría de aminoácidos se representan por medio de varios codones, a los que se les llama codones “sinónimos”, es muy conocido en la industria que el uso de codones por organismos específicos no es aleatorio y se inclina bastante hacia tripletas específicas de codones. Esta inclinación de uso de codones podría ser más alta en referencia a un gen específico, genes de una función común o de origen ancestral, proteínas altamente expresadas versus proteínas con un bajo número de copias, y las regiones acumuladas de codificaciones proteínicas del genoma de un organismo. En algunas secciones, los polinucleótidos que codifican a las enzimas de cetorreductasa podrían ser optimizados en lo que se refiere a los codones para una producción óptima del organismo anfitrión seleccionado para la expresión.

“Uso preferido, óptimo, alto de codones con inclinación hacia los codones” se refiere intercambiamente a codones que son utilizados a una frecuencia más alta en las regiones de codificación proteínicas que otros codones que codifican el mismo aminoácido. Los codones preferidos pueden ser determinados en relación a la utilización de codones en un solo gen, un conjunto de genes de función u origen común, genes altamente expresados, la frecuencia de codones en las regiones de codificación proteínicas agregadas de todo el organismo, la frecuencia de codones en las regiones de codificación proteínica agregada de organismos relacionados, o sus combinaciones. Los codones que incrementan su frecuencia con los niveles de expresión genética son comúnmente codones óptimos

para su expresión. Una variedad de métodos son conocidos para determinar la presencia de codones (por ejemplo, uso de codones, uso relativo de codones sinónimos) y la preferencia de codones en organismos específicos, incluyendo análisis multivariante, por ejemplo, usando algoritmos de agrupamiento o análisis de correspondencias, y el número efectivo de codones utilizados en un gen (referirse a GCG CodonPreference (Preferencia de Codones), Genetics Computer Group Wisconsin (Grupo de Genética y Computadoras de Wisconsin) paquete; CodonW, John Peden, University of Nottingham; McInerney, J. O, 1998, Bioinformatics (Bioinformática) 14:372-73; Stenico et al., 1994, Nucleic Acids Res. (Ácidos Nucleicos Res.) 22:2437-46; Wright, F., 1990, Gene 87:23-29). Tablas del uso de codones están disponibles para una lista creciente de organismos (referirse, por ejemplo, a Wada et al., 1992, Nucleic Acids Res. (Res. Ácidos Nucleicos.) 20:2111-2118; Nakamura et al., 2000, Nucl. Acids Res. (Res. Ácidos Nucleicos.) 28:292; Duret, et al., mencionado anteriormente; Henaut y Danchin, "Escherichia coli y Salmonella," 1996, Neidhardt, et al. Eds., ASM Press, Washington D.C., pg. 2047-2066. La fuente de datos para obtener el uso de codones puede basarse en cualquier secuencia nucleótida disponible capaz de codificar una proteína. Estos conjuntos de datos incluyen secuencias de ácidos nucleicos ya conocidos para codificar proteínas expresadas (por ejemplo, secuencias completas de codificación proteínica – CDS – coding sequences), etiquetas de secuencias expresadas (ESTS – expressed sequence tags), o regiones de codificación previstas (refiérase por ejemplo, a Mount, D., Bioinformatics: Sequence and Genome Analysis (Bioinformática: Análisis de Secuencias y de Genomas), capítulo 8, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 2001; Uberbacher, E. C., 1996, Methods Enzymol. (Enzimol. Métodos) 266:259-281; Tiwari et al., 1997, Comput. Appl. Biosci. 13:263-270).

“Secuencia de control” se define aquí para incluir todos los componentes, que son necesarios o ventajosos para la expresión de un polipéptido de esta presentación. Cada secuencia de control puede ser nativa o foránea a la secuencia de ácido nucleico que codifica al polipéptido. Tales secuencias de control incluyen, pero no se limitan a, un líder, una secuencia de poliadenilación, secuencia de propéptidos, promotores, secuencias de señales de péptidos, y terminadores de transcripción.

“Enlazados operacionalmente” se define en este documento como una configuración en la cual una secuencia de control es colocada apropiadamente (es decir, en una relación funcional) en una posición relativa a un polinucleótido de interés de tal forma que la secuencia de control dirija o regule la expresión del oligonucleótido y / o polipéptido de interés.

“Secuencia promotora” es una secuencia de ácidos nucleicos que es reconocida por una célula anfitriona para la expresión de un polinucleótido de interés, tal como una secuencia de codificación. La secuencia de control puede incluir una secuencia promotora apropiada. La secuencia promotora contiene secuencias de control transcripcional, que regulan la expresión de un polinucleótido de interés. La promotora puede ser cualquier secuencia de ácido nucleico que muestra actividad transcripcional en la célula anfitriona escogida incluyendo actividad mutante, truncada, y promotores híbridos, y puede ser obtenida de genes que codifican polipéptidos extracelulares o intracelulares ya sean homólogos o heterólogos en relación a la célula anfitriona.

1.2. Enzimas Cetorreductasas

En un aspecto, esta presentación suministra enzimas diseñadas de cetorreductasa (“KRED”) que son capaces de reducir estereoselectivamente o convertir la mezcla racémica de metil-2-benzamidometil-3-oxobutirato (“el sustrato”) a 2S,3R-metil-2-benzamidometil-3-hidroxibutirato (“el producto 2S,3R”). Estos polipéptidos de cetorreductasa (también descritos en este documento como “cetorreductasas selectivas 2S,3R”) tienen una propiedad mejorada para reducir o convertir metil-2-benzamidometil-3-oxobutirato a 2S,3R-metil-2-benzamidometil-3-hidroxibutirato cuando se los compara con la enzima KRED que ocurre naturalmente, de tipo silvestre obtenida de *L. kéfir* (IDENTIFICACIÓN SECUENCIAL NÚMERO: 4), *L. brevis* (IDENTIFICACIÓN SECUENCIAL NÚMERO: 2), o *L. minor* (IDENTIFICACIÓN SECUENCIAL NÚMERO: 86) o cuando se compara con otra enzima de cetorreductasa diseñada.

En algunas secciones, la propiedad mejorada es comparada con un polipéptido de tipo silvestre u otro diseñado, tal como la IDENTIFICACIÓN SECUENCIAL NÚMERO: 64, en referencia a un incremento en la estereoselectividad para reducir o convertir el sustrato metil-2-benzamidometil-3-oxobutirato a 2S,3R-metil-2-benzamidometil-3-hidroxibutirato, es decir, en este documento, un incremento en el exceso estereomérico del producto. En algunas secciones, la propiedad mejorada del polipéptido de cetorreductasa en referencia a un incremento es su habilidad de convertir o reducir un mayor porcentaje del sustrato al producto. En algunas secciones, la propiedad mejorada del polipéptido de cetorreductasa se referencia a un incremento en la tasa de conversión del sustrato el producto, que puede manifestarse por la habilidad de usar menos del polipéptido mejorado en comparación con la secuencia de referencia de tipo silvestre o diseñada para reducir o convertir el mismo monto de producto. En algunas secciones, la propiedad mejorada del polipéptido de cetorreductasa se refiere a su estabilidad y estabilidad térmica. En algunas secciones, el polipéptido de cetorreductasa tiene más de una propiedad mejorada, tal como el incremento de estereoselectividad y una actividad enzimática mejorada.

Tal como se describirá en más detalle a continuación, los polipéptidos de cetorreductasa capaces de convertir el sustrato al producto 2S,3R incluye una secuencia de aminoácidos en la cual el residuo correspondiente a X 202 de

las IDENTIFICACIONES SECUENCIALES NÚMEROS: 2, 4 u 86 es valina o leucina. En algunas secciones, los polipéptidos de cetorreductasa incluyen una secuencia de aminoácidos en la cual el residuo correspondiente a X 94 de las IDENTIFICACIONES SECUENCIALES NÚMEROS: 2, 4 u 86 es un residuo alifático o polar, específicamente alanina o treonina; y el residuo correspondiente a X 202 de las IDENTIFICACIONES SECUENCIALES NÚMEROS: 2, 4 u 86 es valina o leucina. En algunas secciones, los polipéptidos de cetorreductasa selectivos 2S,3R incluyen una secuencia de aminoácidos en la cual el residuo correspondiente a X 94 de las IDENTIFICACIONES SECUENCIALES NÚMEROS: 2, 4 u 86 es un residuo alifático o polar, específicamente alanina o treonina; el residuo correspondiente a X 199 de las IDENTIFICACIONES SECUENCIALES NÚMERO: 2, 4 u 86 es un residuo restringido, polar o alifático, específicamente histidina, asparagina o alanina; y el residuo correspondiente a X 202 de las IDENTIFICACIONES SECUENCIALES NÚMEROS: 2, 4 u 86 es valina o leucina.

Tal como se mencionó anteriormente, las cetorreductasas de la presentación pueden ser descritas en referencia a la secuencia de aminoácidos de una cetorreductasa que ocurren naturalmente de *L. kéfir*, *L. brevis* o *L. minor* (también referidas como "ADH" o "alcohol deshidrogenasa") u otra cetorreductasa diseñada. Como tal, la posición residual de aminoácidos es determinada en las cetorreductasas empezando por el residuo iniciador de metionina (M) (es decir, M representa la posición residual 1), aunque se entenderá por una persona con experiencia que este residuo iniciador de metionina puede ser removido por medio de maquinaria de procesamiento biológico, tal como una célula anfitriona o un sistema de conversión in vitro, para generar una proteína madura que no tenga el residuo iniciador de metionina. La posición residual de aminoácidos en la cual un aminoácido o un cambio de aminoácidos en específico está presente en una secuencia de aminoácidos es a veces referido en este documento en los términos "Xn", o "posición n", donde n se refiere a la posición residual. En los casos en los que los residuos de aminoácidos en la misma posición residual difieran entre cetorreductasas, los residuos diferentes pueden ser denotados por medio de una "r" con el arreglo siendo "residuo kéfir / residuo brevis / minor". Una mutación por sustitución, que es el reemplazo de un residuo de aminoácidos en una secuencia referencial, por ejemplo las cetorreductasas de tipo silvestre de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 2 y la IDENTIFICACIÓN SECUENCIAL NÚMERO: 4 y la IDENTIFICACIÓN SECUENCIAL NÚMERO: 86, con un residuo de aminoácidos diferente podría denotarse por medio del símbolo "-". En este documento, en algunas secciones, las mutaciones son descritas a veces como una mutación "a un" tipo de aminoácidos. Por ejemplo, el residuo 199 de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 4 puede ser mutada "a un" residuo polar. Pero el uso de la frase "a un" no excluye mutaciones de un aminoácido de una clase a otro aminoácido de la misma clase. Por ejemplo, el residuo 199 de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 4 es un residuo alifático, leucina, pero puede ser mutado a un residuo alifático diferente, por ejemplo, la mutación puede ser a un "L199A", una mutación (199->A). La secuencia de aminoácidos de la cetorreductasa que ocurre naturalmente (también referida como "ADH" o "alcohol deshidrogenasa") de *L. kéfir*, *L. brevis* o de *L. minor*, puede obtenerse del poli nucleótido conocido para codificar la actividad cetorreductasa (por ejemplo, Acceso del Banco Genético Número AAP94029 GI:33112056 o IDENTIFICACIÓN SECUENCIAL NÚMERO: 3 para *L. kéfir*; Acceso del Banco Genético Número. CAD66648 GI:28400789 o IDENTIFICACIÓN SECUENCIAL NÚMERO: 1 para *L. brevis*; e IDENTIFICACIÓN SECUENCIAL NÚMERO: 86 para *L. minor*).

En algunas secciones los polipéptidos de cetorreductasa pueden tener una variedad de modificaciones a la secuencia referencial (por ejemplo, un polipéptido que ocurren naturalmente o un polipéptido diseñado) para resultar en la propiedad mejorada de cetorreductasa. Tal como se utiliza en este documento, las "modificaciones" incluyen sustituciones, eliminaciones e inserciones de aminoácidos. Cualquier, o una combinación de modificaciones, pueden ser introducidas al polipéptido que ocurren naturalmente o que es diseñado para generar encimas diseñadas. En aquellas secciones, el número de modificaciones a la secuencia de aminoácidos puede incluir uno o más aminoácidos, dos o más aminoácidos, tres o más aminoácidos, cuatro o más aminoácidos, cinco o más aminoácidos, seis o más aminoácidos, ocho o más aminoácidos, 10 o más aminoácidos, 15 o más aminoácidos, 20 o más aminoácidos, hasta el 10% lo del número total de aminoácidos, hasta el 10% del número total de aminoácidos, hasta el 15% del número total de aminoácidos, hasta el 20% to del número total de aminoácidos, o hasta el 30% del número total de aminoácidos de la secuencia referencial de polipéptidos. En algunas secciones, el número de modificaciones al polipéptido que ocurren naturalmente o un polipéptido diseñado que produce una propiedad de cetorreductasa mejorada puede incluir desde alrededor de 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 alrededor de 1-40 modificaciones de la secuencia referencial. En algunas secciones, el número de modificaciones pueden ser 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35 o alrededor de 40 residuos de aminoácidos. Las modificaciones deben incluir inserciones, eliminaciones, sustituciones o sus combinaciones.

En algunas secciones, las modificaciones incluyen sustituciones de aminoácidos a la secuencia referencial. Las sustituciones que pueden producir una propiedad mejorada de la cetorreductasa podrían ubicarse en uno o más aminoácidos, dos o más aminoácidos, tres o más aminoácidos, cuatro o más aminoácidos, cinco o más aminoácidos, seis o más aminoácidos, ocho o más aminoácidos, 10 o más aminoácidos, 15 o más aminoácidos, o 20 o más aminoácidos, hasta el 10% del número total de aminoácidos, hasta el 10% del número total de aminoácidos, hasta el 20% del número total de aminoácidos, o hasta el 30% del número total de aminoácidos de la secuencia enzimática referencial. En algunas secciones, el número de sustituciones al polipéptido que ocurren naturalmente o a un polipéptido diseñado que produce una propiedad mejorada de la cetorreductasa puede incluir desde alrededor de 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 alrededor de 1-40 sustituciones de aminoácidos de la secuencia referencial. En algunas secciones

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 alrededor de 40 residuos de aminoácidos.

5 En algunas secciones, el polipéptido mejorado de cetorreductasa incluye una secuencia de aminoácidos que es por lo menos alrededor del 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% o más idéntica a la secuencia referencial que se basa en la IDENTIFICACIÓN SECUENCIAL NÚMERO: 2, 4 u 86 que tiene en el residuo correspondiente a X 202 a una leucina o valina, con la condición de que el polipéptido de cetorreductasa tenga una secuencia de aminoácidos en la cual el residuo que corresponde a X 202 es leucina o valina. En algunas secciones, el residuo correspondiente a X 202 es leucina. En algunas secciones, estos polipéptidos de cetorreductasa pueden tener una o más diferencias en otras posiciones residuales en comparación de la secuencia referencial de aminoácidos. Las diferencias incluyen varias modificaciones, tales como sustituciones, eliminaciones e inserciones. Las sustituciones pueden ser no conservadoras, conservadoras o una combinación entre sustituciones no conservadoras y conservadoras. En algunas secciones, estos polipéptidos de cetorreductasa pueden tener opcionalmente, alrededor de 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 alrededor de 1-40 diferencias residuales en otros residuos de aminoácidos en comparación a la secuencia referencial. En algunas secciones, 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 alrededor de 40 diferencias residuales en otros residuos de aminoácidos en comparación con la secuencia referencial. En algunas secciones, la secuencia referencial es la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48.

20 En algunas secciones, el polipéptido mejorado de cetorreductasa incluye una secuencia de aminoácidos que es por lo menos alrededor de un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% o más idéntica a la secuencia referencial que se basa en las IDENTIFICACIONES SECUENCIALES NÚMEROS: 2, 4 u 86 que tienen las siguientes características: el residuo correspondiente a X 94 es un residuo alifático o polar, específicamente de alanina o treonina; y el residuo correspondiente a X 202 es valina o leucina; con la condición que el polipéptido de cetorreductasa tenga una secuencia de aminoácidos con por lo menos las características que se acaba de mencionar (es decir, el residuo correspondiente a X 94 es un residuo alifático o polar, y el residuo correspondiente a X 202 es valina o leucina). En algunas secciones, la cetorreductasa tiene una secuencia de aminoácidos en la cual el residuo correspondiente a X 94 es un residuo polar, y el residuo correspondiente a X 202 es valina o leucina. En algunas secciones, la cetorreductasa tiene una secuencia de aminoácidos en la cual el residuo correspondiente a X 94 es treonina y el residuo correspondiente a X 202 es valina o leucina. En algunas secciones, estos polipéptidos de cetorreductasa pueden tener opcionalmente desde alrededor de 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 alrededor de 1-40 diferencias residuales en otras posiciones residuales de aminoácidos en comparación con la secuencia referencial. En algunas secciones, 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 alrededor de 40 diferencias residuales en otros residuos de aminoácidos en comparación con la secuencia referencial. En algunas secciones, la secuencia referencial es la IDENTIFICACIÓN SECUENCIAL NÚMERO: 26 o 28.

40 En algunas secciones, el polipéptido mejorado de cetorreductasa incluye una secuencia de aminoácidos que es por lo menos alrededor de un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% o más idéntica a una secuencia referencial que se basa en las IDENTIFICACIONES SECUENCIALES NÚMEROS: 2, 4 u 86 que tienen las siguientes características: el residuo correspondiente a X 94 es un residuo alifático o polar, específicamente alanina o treonina; el residuo correspondiente de a X 199 es un residuo alifático, restringido o polar, específicamente alanina, asparagina, o histidina; y el residuo correspondiente a X 202 es valina o leucina; con la condición que el polipéptido de cetorreductasa tiene una secuencia de aminoácidos que tiene por lo menos las características que se acabaron de mencionar (es decir, el residuo correspondiente a X 94 es un residuo alifático o polar; el residuo correspondiente a X 199 es un residuo alifático, restringido o polar; y el residuo correspondiente a X 202 es valina o leucina). En algunas secciones, el residuo correspondiente a X 94 es un residuo polar; el residuo correspondiente a X 199 es un residuo alifático, restringido o polar; y el residuo correspondiente a X 202 es valina o leucina. En algunas secciones, la cetorreductasa tiene una secuencia de aminoácidos en la cual el residuo correspondiente a X 94 es treonina; el residuo correspondiente a X 199 es alanina, asparaagina, o histidina; y el residuo correspondiente a X 202 es valina o leucina. En algunas secciones, estos polipéptidos de cetorreductasa pueden tener una o más diferencias residuales en otras posiciones residuales en comparación con la secuencia referencial de aminoácidos. Las diferencias incluyen varias modificaciones, tales como sustituciones, eliminaciones e inserciones. Las sustituciones pueden ser no conservadoras, conservadoras o una combinación de sustituciones no conservadoras y conservadoras. En algunas secciones, estos polipéptidos de cetorreductasa pueden tener opcionalmente desde alrededor de 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 alrededor de 1-40 diferencias residuales en otras posiciones residuales de aminoácidos en comparación con la secuencia referencial. En algunas secciones, 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 alrededor de 40 diferencias residuales en otros residuos de aminoácidos en comparación con la secuencia referencial. En algunas secciones, la secuencia referencial es la IDENTIFICACIÓN SECUENCIAL NÚMERO: 22, 24 o 30.

65 En vista de lo anterior, las cetorreductasas selectivas 2S,3R pueden ser descritas en referencia a las características de las varias combinaciones de residuos correspondientes a X 94, X 199, y X 202. Por ejemplo, un polipéptido

selectivo 2S,3R caracterizado por las propiedades en los residuos correspondientes a X 94 y X 202 se refiere a las descripciones aquí provistas para la combinación de las posiciones residuales especificadas. Similarmente, un polipéptido selectivo 2S,3R caracterizado por las propiedades en los residuos correspondientes a X 94, X 199 y X 202 se refiere a las descripciones aquí provistas para la combinación de los residuos especificados. Tal como se describe en mayor detalle a continuación, las cetorreductasas pueden tener una o más características adicionales en la secuencia de aminoácidos en comparación a la secuencia referencial.

En algunas secciones, un polipéptido de cetorreductasa selectivo 2S,3R tiene una secuencia de aminoácidos que se basa en las fórmulas secuenciales descritas en la IDENTIFICACIÓN SECUENCIAL NÚMERO: 83, la IDENTIFICACIÓN SECUENCIAL NÚMERO: 84 o la IDENTIFICACIÓN SECUENCIAL NÚMERO: 87 (o una de sus regiones, con los residuos 90 - 211). La IDENTIFICACIÓN SECUENCIAL NÚMERO: 84 se basa en la secuencia de aminoácidos de tipo silvestre de la cetorreductasa L. kéfir (IDENTIFICACIÓN SECUENCIAL NÚMERO: 4), la IDENTIFICACIÓN SECUENCIAL NÚMERO: 83 se basa en la secuencia de aminoácidos de tipo silvestre de la cetorreductasa L. brevis (IDENTIFICACIÓN SECUENCIAL NÚMERO: 2), y la IDENTIFICACIÓN SECUENCIAL NÚMERO: 87 se basa en la secuencia de aminoácidos de tipo silvestre de la cetorreductasa L. minor (IDENTIFICACIÓN SECUENCIAL NÚMERO: 86). La IDENTIFICACIÓN SECUENCIAL NÚMERO: 83, 84 u 87 especifican que el residuo correspondiente a X 94 es un residuo alifático o polar; el residuo correspondiente a X 199 es un residuo alifático, restringido o polar; y el residuo correspondiente a X 202 es valina o leucina. La fórmula secuencial especifica las características para otras porciones residuales, tal como se describe a continuación.

La Tabla 2 a continuación suministra ejemplos de cetorreductasas selectivas 2S,3R, IDENTIFICACIONES SECUENCIALES NÚMEROS: 1 - 62, con sus actividades asociadas. Las secuencias a continuación se derivan de las secuencias de cetorreductasa de tipo silvestre L. kéfir (IDENTIFICACIONES SECUENCIALES NÚMEROS: 3 y 4) a menos que se especifique de otra forma. En la Tabla 2 a continuación, cada fila lista los NÚMEROS DE IDENTIFICACIONES SECUENCIALES, donde el número impar se refiere a la secuencia de nucleótidos que codifica la secuencia de aminoácidos provista en el número par. La columna que lista el número de mutaciones referencial al número de sustituciones de aminoácidos en comparación a la secuencia de aminoácidos KRED L. kéfir de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 4, y las sustituciones específicas están listadas en la columna "mutaciones desde kéfir". En la columna de actividad, un "+" corresponde a una mejora de 1 - 15 veces en comparación de la habilidad del polipéptido que tiene la secuencia de aminoácidos de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48 para convertir el sustrato del producto de la fórmula (II). Dos señales de suma "++" indican que el polipéptido es alrededor de 15 a 30 veces mejorado en comparación de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48. Tres señales de suma "+++" indican que el polipéptido es alrededor de 30 a 40 veces mejorado en comparación de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48. Cuatro señales de suma "++++" indican que el polipéptido es alrededor de 40 a 50 veces mejorado en comparación de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48, y cinco señales de suma "+++++" indican que el polipéptido es mayor que 50 veces mejorado en comparación de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48. Una señal "+" abajo de la columna de estabilidad indica que el polipéptido es capaz de retener la actividad enzimática para convertir el sustrato al producto de la fórmula (II) después de 21 horas de tratamiento de calor a 40 °C. Para la columna de selectividad, una sola señal de suma "+" indica que el polipéptido es capaz de convertir el sustrato al producto de la fórmula (II) con alrededor de un 60 - 89% de exceso estereomérico; dos señales de suma "++" indican que el polipéptido es capaz de convertir el sustrato al producto de la fórmula (II) con alrededor de un 90 - 94% de exceso estereomérico; tres señales de suma "+++" indican que el polipéptido es capaz de convertir el sustrato al producto de la fórmula (II) con alrededor de un 95 - 99% de exceso estereomérico; y cuatro señales de suma "++++" indican que el polipéptido es capaz de convertir el sustrato al producto de la fórmula (II) con un exceso estereomérico que sobrepasa el 99%. Asimismo en algunas secciones, las cetorreductasas selectivas 2S,3R pueden incluir una secuencia correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 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, o 62.

Tabla 2

Tabla 2: Lista de secuencias y mejoras correspondientes de actividades					
NÚMERO DE IDENTIFICACIÓN SECUENCIAL	Diferencias residuales de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 2	# de mutaciones desde kéfir	Actividad	Estabilidad	Selectividad
1/2	L. brevis				
3/4	L. kéfir				
47/48	A202V	1	Mejorado por sobre wt		+
37/38	A94T; E105G; L153A; L199A; A202L; M206F	6	+		+

Tabla 2: Lista de secuencias y mejoras correspondientes de actividades						
	NÚMERO DE IDENTIFICACIÓN SECUENCIAL	Diferencias residuales de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 2	# de mutaciones desde kéfir	Actividad	Estabilidad	Selectividad
5	15/16	A94T; S96F; A202V	3	+		+++
10	55/56	L153A; L199A; A202L	3	+		+++
	57/58	T86I; L199N; A202L	3	+		+++
15	51/52	L153A; A202L	2	+		+++
	53/54	L153A; A202V	2	+		+
	31/32	A94T; L199A; A202V	3	++		++++
20	33/34	A94T; L153A; L199I4; A202L	4	+++		++++
	49/50	L153A; L199H; A202L	3	++		+++
	19/20	A94T; L199N; A202V	3	+++		++++
25	45/46	L153S; A202L	2	+		+
	35/36	A94T; L153A; L199A; A202V	4	+		++
	25/26	A94T; A202L	2	+++		+++
	27/28	A94T; A202V	2	++		+++
30	29/30	A94T; L199A; A202L	3	+++		++++
	21/22	A94T; L199H; A202L	3	++++		++++
35	23/24	A94T; L199H; A202V	3	+++		++++
	41/42	L153A; L199N; A202L	3	+		+++
40	39/40	A94T; S96F; M129T; A202V; M206F	5	+		++
	17/18	A80T; L153A; A202V;	3	+		+++
	43/44	F147M; A202V	2	+	+	
45	9/10	H4OR; A94T; F147L; L199H; A202L	5	+++++	+	++++
	11/12	H4OR; A94T; L199H; A202L	4	+++++		++++
50	5/6	A94T; F147L; L199H; A202L	4	+++++	+	++++
	7/8	A94T; L199H; A202L	3	+++++		++++
55	59/60	111F; H4OR; A94F; S96V; F147M; L195V; V196L; L199W; I226V; G248K; Y249W	11	++++		++++
60	61/62	T2A; R4C; H4OR; A94G; S96V; F147M; V196L; L199W; I226V; G248K; Y249W	11	+++		++++
65	13/14	H4OR; A94F; S96V; F147M; L195V; V196L; L199W; I226V; Y249W	9	++++		++++

- 5 En algunas secciones, una cetorreductasa selectiva mejorada 2S,3R incluye una secuencia de aminoácidos que es por lo menos alrededor de un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% idéntica a la secuencia referencial que se basa en la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 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, o 62 con la condición que la secuencia de aminoácidos de cetorreductasa tenga por lo menos las siguientes características: el residuo correspondiente a X 94 es un residuo alifático o polar, específicamente alanina o treonina; el residuo correspondiente a X 199 es un residuo alifático, restringido o polar, particularmente alanina, histidina o asparagina; y el residuo correspondiente a X 202 es valina o leucina. En algunas secciones, los polipéptidos de cetorreductasa tienen 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 alrededor de 1-40 diferencias residuales en otras posiciones residuales de aminoácidos en comparación con la secuencia referencial. En algunas secciones, 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 alrededor de 40 diferencias residuales en otros residuos de aminoácidos. En algunas secciones, las diferencias incluyen mutaciones conservadoras.
- 10
- 15 En algunas secciones, una cetorreductasa selectiva 2S,3R incluye una secuencia de aminoácidos que es por lo menos alrededor de un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% idéntica a la secuencia referencial que se basa en la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 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, o 62, con la condición de que la secuencia de aminoácidos de cetorreductasa incluya cualquier conjunto de mutaciones contenido en cualquiera de las secuencias de polipéptidos listadas en la Tabla 2 en comparación a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 2, 4 u 86. En algunas secciones, los polipéptidos de cetorreductasas 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 alrededor de 1-40 diferencias residuales en otras posiciones residuales de aminoácidos en comparación a la secuencia referencial. En algunas secciones, 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 alrededor de 40 diferencias residuales en otros residuos de aminoácidos. En algunas secciones, las diferencias incluyen mutaciones conservadoras.
- 20
- 25
- 30 En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato, metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, con un exceso porcentual estereomérico de por lo menos alrededor de un 85% o con un exceso porcentual estereomérico que sea mayor al de la KRED de tipo silvestre L. kéfir (IDENTIFICACIÓN SECUENCIAL NÚMERO: 4). Ejemplos de polipéptidos que tienen esta capacidad incluyen, pero no se limitan a, polipéptidos con una secuencia de aminoácidos correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 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, o 62.
- 35
- 40 En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato, metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, con un exceso porcentual estereomérico de por lo menos de alrededor de 60 - 89% y a una tasa que es por lo menos alrededor de 1 - 15 veces mayor que la tasa de capacidad del polipéptido que tiene la secuencia de aminoácidos de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48. Ejemplos de polipéptidos que tienen esta capacidad incluyen, pero no se limitan a, los polipéptidos que tienen una secuencia de aminoácidos correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 46, 50, 52, 54, 56, 58, 60, o 62. Puesto que el polipéptido referencial que tiene la secuencia de aminoácidos de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48 es capaz de convertir el sustrato al producto a una tasa (por ejemplo, 100% de conversión en 20 horas de 1 g / litro de sustrato con alrededor de 10 g / litro de la KRED, en un 50% de IPA a pH 8) y con una estereoselectividad que es mejorada sobre el tipo silvestre (IDENTIFICACIÓN SECUENCIAL NÚMERO: 4), los polipéptidos aquí descritos que son mejorados sobre la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48 también están mejorados sobre el tipo silvestre.
- 45
- 50 En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato, metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, con un exceso porcentual estereomérico de por lo menos alrededor de 90 - 94%. Ejemplos de polipéptidos que son capaces incluyen, pero no se limitan a, polipéptidos que tienen una secuencia de aminoácidos correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 40, 42, 50, 52, 56, 58, 60, o 62.
- 55
- 60 En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, con un exceso porcentual estereomérico de por lo menos alrededor de 95 - 99%. Ejemplos de polipéptidos que son capaces incluyen, pero no se limitan a, polipéptidos que tienen una secuencia de aminoácidos correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 42, 50, 52, 56, 58, 60, o 62.
- 65
- En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, con un exceso porcentual estereomérico de por lo menos alrededor de 99%. Ejemplos de polipéptidos que son capaces incluyen,

pero no se limitan a, polipéptidos que tienen una secuencia de aminoácidos correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 8, 10, 12, 14, 20, 22, 24, 30, 32, 34, 60, o 62.

5 En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, a una tasa que es por lo menos alrededor de 15 - 30 veces mayor que la tasa de capacidad del polipéptido que tiene la secuencia de aminoácidos de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48. Ejemplos de polipéptidos que tienen esta capacidad incluyen, pero no se limitan a, polipéptidos que tienen una secuencia de aminoácidos correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 8, 10, 12, 14, 20, 22, 24, 26, 28, 30, 32, 34, 50, 60, o 62.

10 En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, a una tasa que es por lo menos alrededor de 30 - 40 veces mayor que la tasa de capacidad del polipéptido que tiene la secuencia de aminoácidos de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48. Ejemplos de polipéptidos que tienen esta capacidad incluyen, pero no se limitan a, polipéptidos que tienen una secuencia de aminoácidos correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 8, 10, 12, 14, 20, 22, 24, 26, 30, 34, 60, o 62.

15 En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, a una tasa que es por lo menos alrededor de 40 - 50 veces mayor que la tasa de capacidad del polipéptido que tiene la secuencia de aminoácidos de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48. Ejemplos de polipéptidos que tienen esta capacidad incluyen, pero no se limitan a, polipéptidos que tienen una secuencia de aminoácidos correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 8, 10, 12, 14, 22, o 60.

20 En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, a una tasa que es por lo menos alrededor de 50 veces mayor que la tasa de capacidad del polipéptido que tiene la secuencia de aminoácidos de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48. Ejemplos de polipéptidos que tienen esta capacidad incluyen, pero no se limitan a, polipéptidos que tienen una secuencia de aminoácidos correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 8, 10, o 12.

25 En algunas secciones, los polipéptidos de cetorreductasa son capaces de convertir el sustrato metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, a una tasa que es por lo menos alrededor de 50 veces mayor que la tasa de capacidad del polipéptido que tiene la secuencia de aminoácidos de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48 y con un exceso estereomérico de por lo menos el 99%. Ejemplos de polipéptidos que tienen esta capacidad incluyen, pero no se limitan a, polipéptidos que tienen una secuencia de aminoácidos correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 8, 10, y 12.

30 En algunas secciones, los polipéptidos de cetorreductasa son capaces de retener su habilidad para convertir el sustrato, metil-2-benzamidometil-3-oxobutirato, al producto, 2S,3R-metil-2-benzamidometil-3-hidroxitirato, después de recibir un tratamiento de calor a 40 °C durante 21 horas. Ejemplos de polipéptidos que tienen esta capacidad incluyen, pero no se limitan a, polipéptidos que tienen una secuencia de aminoácidos correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 10 o 44.

35 En algunas secciones, los polipéptidos de cetorreductasa selectivos 2S,3R de la presentación pueden incluir una secuencia de aminoácidos que es por lo menos alrededor de un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% idéntico a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 2 o a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 4 o a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 86 (o una de sus regiones o dominios, tal como los residuos 90 - 211) con la condición que los residuos correspondientes al residuo 202 de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 2 o la IDENTIFICACIÓN SECUENCIAL NÚMERO: 4 o la IDENTIFICACIÓN SECUENCIAL NÚMERO: 86 es valina o leucina, el residuo correspondiente al residuo 94 de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 2, 4 u 86 es treonina, y el residuo correspondiente de al residuo 199 de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 2, 4 u 86 es histidina, y adicionalmente tiene una o más de las siguientes sustituciones que permiten que el polipéptido sea mejorado aún más (en referencia a estereoselectividad, actividad enzimática y / o estabilidad térmica) sobre la cetorreductasa de tipo silvestre kéfir u otra cetorreductasa diseñada (tal como la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48): 2->A; 4->C; 11->F; 40->H; 80->T; 86->I; 96->F, V; 105->G; 129->T; 147->M, L; 153->A, S; 195->V; 196->L; 206->F; 226->V; 248->K; y 249->W.

40 En algunas secciones, los polipéptidos de cetorreductasa 2S,3R aquí descritos pueden incluir una secuencia de aminoácidos que es por lo menos alrededor de un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% idéntica a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 2, 4 u 86 (o una de sus regiones o dominios, tal como los residuos 90 - 211) con la condición de que los residuos correspondientes al residuo 202 de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 2, 4 u 86 sea valina o leucina, el residuo correspondiente al residuo 94 de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 2, 4 u 86 sea treonina, y el residuo correspondiente al residuo 199 de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 2, 4 u 86 sea histidina, y adicionalmente tenga una o más de las siguientes funciones de tal forma que el polipéptido sea mejorado aún más (en referencia a la estereoselectividad,

actividad enzimática y / o estabilidad térmica) por sobre la cetorreductasa de tipo silvestre kéfir u otra cetorreductasa diseñada (tal como la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48): 40->H; y 147->L, M.

5 Cómo puede ser apreciado por aquellas personas con conocimiento en la industria, algunas de las categorías que se definieron anteriormente de residuos de aminoácidos, a menos que se especifique de otra forma, no son mutuamente exclusivas. Por lo tanto, los aminoácidos que tiene cadenas laterales que exhiben dos o más propiedades físico - químicas pueden ser incluidas en varias categorías. La clasificación apropiada de cualquier aminoácido o residuo será aparente para aquellas personas con conocimiento en la industria, especialmente en luz de la presentación detallada suministrada aquí.

10 En algunas secciones, las enzimas de cetorreductasa diseñadas y mejoradas incluyen eliminaciones de polipéptidos de cetorreductasa que ocurren naturalmente o eliminaciones de otros polipéptidos de cetorreductasa diseñados. En algunas secciones, cada una de las enzimas de cetorreductasa diseñadas y mejoradas aquí descritas pueden incluir eliminaciones de los polipéptido aquí descritos. Por lo tanto, para cada una de las secciones de los polipéptidos de cetorreductasa de la presentación, las eliminaciones pueden incluir uno o más aminoácidos, dos o más aminoácidos, tres o más aminoácidos, cuatro o más aminoácidos, cinco o más aminoácidos, seis o más aminoácidos, ocho o más aminoácidos, 10 o más aminoácidos, 15 o más aminoácidos, o 20 o más aminoácidos, hasta el 10% del número total de los aminoácidos, hasta el 20% del número total de los aminoácidos, o hasta el 30% del número total de los aminoácidos de los polipéptidos de cetorreductasa, mientras la actividad funcional de la cetorreductasa se mantenga. En algunas secciones, las eliminaciones pueden incluir, 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 alrededor de 1-40 residuos de aminoácidos. En algunas secciones, el número de eliminaciones pueden ser 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35 o alrededor de 40 aminoácidos. En algunas secciones, las eliminaciones pueden incluir eliminaciones de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, o 20 residuos de aminoácidos.

25 Tal como se describe aquí, los polipéptidos de cetorreductasa de la presentación pueden estar en la forma de polipéptidos de fusión en los cuales los polipéptidos de cetorreductasa están fusionados con otros polipéptidos, tales como, por ejemplo y no como limitación, etiquetas de anti cuerpos (por ejemplo, epitopo de myc), secuencias de purificaciones (por ejemplo, etiquetas de His), y señales de localización celular (por ejemplo, señales de secreción) por lo tanto, los polipéptidos de cetorreductasa pueden ser utilizados con o sin fusiones a otros polipéptidos.

35 Los polipéptidos aquí descritos no son restringidos a los aminoácidos codificados genéticamente. Adicionalmente a los aminoácidos codificados genéticamente, los polipéptidos aquí descritos pueden incluir, ya sea parcialmente o completamente, aminoácidos no codificados que ocurren naturalmente y / o sintéticos. Ciertos aminoácidos no codificados que se encuentran comúnmente que se cubren en este documento incluyen, pero no se limitan a: los D-estereómeros de los aminoácidos codificados genéticamente; 2,3- ácido diaminopropiónico (Dpr); ácido α-aminoisobutírico (Aib); ácido ε-aminoheptanoico (Aha); ácido S-aminovalérico (Ava); N-metilglicina o sarcosina (MeGly o Sar); ornitina (Orn); citrulina (Cit); t-butilalanina (Bua); t-butilglicina (Bug); N-metilisoleucina (Melle); fenilglicina (Phg); ciclohexilalanina (Cha); norleucina (Nle); naftalalanina (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 (Mpfc); 2,4-difluorofenilalanina (Opff); 3,4-difluorofenilalanina (Mpf); pirid-2-ilalanina (2pAla); pirid-3-ilalanina (3pAla); pirid-4-ilalanina (4pAla); naft- 1-ilalanina (1nAla); naft-2-ilalanina (2nAla); tiasolilalanina (taAla); benzotienilalanina (bAla); tienilalanina (tAla); furilalanina (fAla); homofenilalanina (hPhe); homotirosina (hTyr); homotriptofano (hTrp); pentafluorofenilalanina (5ff); stirilalanina (sAla); autrilalanina (aAla); 3,3-difenilalanina (Dfa); ácido 3-amino-5-fenipentanoico (Afp); penicilamina (Pen); 1,2,3,4-ácido tetrahidroisoquinolin-3-carboxílico (Tic); β-3-2-tienilalanina (Thi); sulfóxido de metionina (Mso Methionine sulfoxide); N(w)-nitroarginina (nArg); homolisina (hLys); fosfometilfenilalanina (pmPhe); fosfoserina (pSer - phosphoserine); fosfotreonina (pThr - phosphothreonine); ácido homoaspártico (hAsp); ácido homoglutánico (hGlu); ácido 1-aminociclopent-(2 o 3)-ene-4-carboxílico; ácido pipercolico (PA - pipercolic acid), ácido acetidin-3-carboxílico (ACA); 1-ácido aminociclopentan-3-carboxílico; alilglicina (aOly); propargilglicina (pgGly); homoalanina (hAla); norvalina (nVal); homoleucina (hLeu); homovalina (hVal); homoisoleucina (hIle); homoarginina (hArg); lisina de N-acetilo (AcLys); ácido 2,4-diaminobutírico (Dbu); ácido 2,3-diaminobutírico (Dab); N-metilvalina (MeVal); homocisteína (hCys); homoserina (hSer); hidroxiprolina (Hyp) y homoprolina (hPro). Aminoácidos no codificados adicionales de los cuales los polipéptidos aquí descritos pueden estar presentes serán aparentes a aquellas personas con conocimiento en la industria (referirse, origen, a los varios aminoácidos suministrados en Fasman, 1989, CRC Practical Handbook of Biochemistry and Molecular Biology (Manual Práctico de Bioquímica y Biología Molecular), CRC Press, Boca Raton, FL, at pg. 3-70 y las referencias aquí citadas. Estos aminoácidos podrían tener la configuración L- o D-.

65 Aquellas personas con conocimiento en la industria reconocerán que los aminoácidos o residuos que tienen grupos protectores de la cadena lateral también pueden tener los polipéptidos aquí descritos. Ejemplos no limitantes de

aquellos aminoácidos protegidos, que en este caso pertenecen a la categoría aromática, incluyen (los grupos protectores están listados dentro de los paréntesis), pero no se limitan a: Arg(tos), Cys(metilbencilo), Cys (nitropiridinesulfenilo), Glu(S-éster bencilo), Gln(xantilo), Asn(N-6-xantilo), His(bom), His(bencilo), His(tos), Lys(fmoc), Lys(tos), Ser(O-bencilo), Thr (O-bencilo) y Tyr(O-bencilo).

Los aminoácidos que no codifican están restringidos en su conformación de los cuales los polipéptido se aquí descritos podrían ser compuestos incluyen, pero no se limitan a, aminoácidos N-metilos (configuración-L); ácido 1-aminociclopent-(2 o 3)-ene-4- carboxílico; ácido piperídico; ácido acetidin-3- carboxílico; homoprolina (hPro); y ácido aminociclopentan-3- carboxílico.

Tal como se describió anteriormente las varias modificaciones introducidas al polipéptido que ocurren naturalmente para generar una enzima diseñada de cetorreductasa puede ser enfocada a una propiedad específica de la enzima.

1.3. Polinucleótidos que codifican a las cetorreductasas diseñadas

En otro aspecto, esta presentación suministra polinucleótidos que codifican las enzimas diseñadas de cetorreductasa. Los oligonucleótidos pueden ser enlazados operativamente a uno o más secuencias heterólogas regulatorias que controlan la expresión genética para crear un poli nucleótido recombinante capaz de expresar el polipéptido. Estructuras de expresión que contienen polinucleótidos heterólogos que codifican a la cetorreductasa diseñada pueden introducirse en células anfitrionas apropiadas para expresar el polipéptido de cetorreductasa correspondiente.

Debido al conocimiento de los codones correspondientes a los varios aminoácidos, la disponibilidad de una secuencia proteínica suministra una descripción de todos los polinucleótidos que tiene la capacidad de codificar el sujeto. La degeneración del código genético, en los casos en los que los mismos aminoácidos han sido codificados por medio de cordones alternos o sinónimos permite que un gran número de ácidos nucleicos sean elaborados, todos los cuales codifican las enzimas de cetorreductasa mejoradas que se presentan en este documento. Por lo tanto, al haberse identificado una secuencia de aminoácidos específica, aquellas personas con conocimiento en la industria podrían elaborar cualquier número de ácidos nucleicos diferentes al modificar la secuencia de uno o más codones en una forma en la cual no se cambia la secuencia de aminoácidos de la proteína. En este aspecto, esta presentación contempla específicamente cada una y todas las variaciones de polinucleótidos que pudiesen elaborarse al seleccionar combinaciones que se basan en las posibles opciones de codones, y todas aquellas variaciones deben ser consideradas específicamente como presentadas para cualquier polipéptido de esta presentación, incluyendo las secuencias de aminoácidos presentadas en la Tabla 1.

En varias secciones, los codones son seleccionados preferiblemente para adaptarse a la célula anfitriona en la cual la proteína está siendo producida. Por ejemplo, los codones preferidos usados en bacterias son utilizados para expresar el gen en las bacterias; los codones preferidos utilizados en la levadura son utilizados para la expresión en la levadura; y los codones preferidos utilizados en mamíferos son utilizados para la expresión en células mamíferas. Como ejemplo, el polinucleótido de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 1 ha sido optimizado en lo que se refiere a codones para la expresión en *E. coli*, pero de otra manera codifican la cetorreductasa que ocurre naturalmente del *Lactobacilo kéfir*.

En ciertas secciones, todos los codones no necesitan ser reemplazados para optimizar el uso de cetorreductasas por parte de los codones puesto que la secuencia natural incluirá a los codones preferidos y puesto que el uso de los codones preferidos podrían no ser requeridos para todos los residuos de aminoácidos. Consecuentemente, los polinucleótidos en los cuales se han optimizado los codones que codifican las enzimas de cetorreductasa podrían contener codones preferidos alrededor de un 40%, 50%, 60%, 70%, 80% o más de un 90% de las posiciones de codones de toda la región de codificación.

En algunas secciones, el polinucleótido incluye una secuencia de nucleótidos que codifica un polipéptido de cetorreductasa con una secuencia de aminoácidos que tienen por lo menos alrededor de 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% o más identidad secuencial a cualquiera de los polipéptidos de cetorreductasa diseñados referenciales aquí descritos. Asimismo, en algunas secciones, el polinucleótido codifica una secuencia de aminoácidos que es por lo menos alrededor de un 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más idéntico a una secuencia referencial que se basa en la IDENTIFICACIÓN SECUENCIAL NÚMERO: 2, 4 u 86 teniendo las siguientes características: el residuo correspondiente a la posición X 94 es un residuo alifático o polar, específicamente alanina o treonina; el residuo correspondiente a X 199 es un residuo alifático, restringido o polar, específicamente alanina, histidina o asparagina; y el residuo correspondiente a X 202 es valina o leucina, con la condición que el polipéptido de cetorreductasa codificado tenga una secuencia de amino con las características que se acabaron de mencionar, es decir, el residuo correspondiente a la posición X 94 es un residuo alifático o polar; el residuo correspondiente a X 199 es un residuo alifático, restringido o polar; y el residuo correspondiente a X 202 es valina o leucina. En algunas secciones, el polinucleótido codifica una cetorreductasa que tiene una secuencia de aminoácidos en la cual el residuo correspondiente a X 94 es treonina; el residuo correspondiente a X 199 es alanina, histidina o asparagina; el residuo correspondiente a X 202 es valina o leucina. En algunas secciones, el polinucleótido codifica una secuencia

de aminoácidos correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 8, 10, 12, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, o 58.

- 5 En algunas secciones, el polinucleótido incluye una secuencia de nucleótidos que codifica un polipéptido de cetorreductasa con una secuencia de aminoácidos que tienen por lo menos alrededor de un 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% o más identidad secuencial al polipéptido que contiene un aminoácido correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 14, 60 o 62.
- 10 En algunas secciones, el polinucleótido incluye una secuencia de nucleótidos que codifica un polipéptido de cetorreductasa con una secuencia de aminoácidos que tienen por lo menos alrededor de un 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% o más identidad secuencial al polipéptido que tiene una secuencia de aminoácidos correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 64, 66, 68, 70, 72, 74, 76, 78, 80 u 82.
- 15 En algunas secciones, los polinucleótidos que codifican a la cetorreductasa son seleccionados de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, u 81. En algunas secciones, los polinucleótidos son capaces de hibridar bajo condiciones altamente rigurosas a un polinucleótido que incluye la IDENTIFICACIÓN SECUENCIAL NÚMERO: 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53,
- 20 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, u 81, donde los polinucleótidos que se utilizan bajo condiciones altamente rigurosas codifican a una cetorreductasa capaz de reducir o convertir estereoselectivamente el sustrato de la fórmula (I) al producto de la fórmula (II), o reducir o convertir estereoselectivamente el sustrato de la fórmula (I) al producto de la fórmula (III).
- 25 En algunas secciones, los polinucleótidos codifican a los polipéptido aquí descritos pero tienen alrededor de un 80% o más de identidad secuencial, alrededor de un 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% o más identidad secuencial a nivel de los nucleótidos en comparación de un polinucleótido referencial que codifica la cetorreductasa diseñada. En algunas secciones, el polinucleótido referencial es seleccionado de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33,
- 30 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, u 81.
- Un polinucleótido aislado que codifica a un polipéptido mejorado de cetorreductasa podría ser manipulado en una variedad de formas para facilitar la expresión del polipéptido. La manipulación del polinucleótido aislado antes de su inserción a un vector podría ser deseable o necesario dependiendo del vector de expresión. Las técnicas para
- 35 modificar a las secuencias de polinucleótidos y de ácidos nucleicos que utilizan métodos de ADN recombinante son bien conocidos en la industria. Se puede obtener orientación en Sambrook et al., 2001, *Molecular Cloning: A Laboratory Manual* (Clonación Molecular: Un Manual de Laboratorio), 3ra Ed., Cold Spring Harbor Laboratory Press; *Current Protocols in Molecular Biology* (Protocolos Actuales en Biología Molecular), Ausubel, F. ed., Greene Pub. Associates, 1998, actualizaciones hasta 2006.
- 40 Para células anfitrionas bacterianas, promotores adecuados para dirigir la transcripción de las estructuras de ácidos nucleicos de la presentación, incluyen los promotores obtenidos del gen *E. coli lac operon*, *E. coli trp operon*, bacteriofago λ *Streptomyces coelicolor agarase* (*dagA*), el gen *Bacilo subtilis levansucrase* (*sacB*), el gen *Bacilo licheniformis alfa - amilasa* (*amyL*), el gen *Bacilo stearotermofilo maltogénico amilasa* (*amyM*), el gen *Bacilo amiloliquefaciens alfa-amilasa* (*amyQ*), el gen *Bacilo licheniformis penicillinasa* (*penP*), los genes *Bacilo subtilis xilA* y *xylB*, y el gen procarriótico beta - lactamasa (Villa-Kamaroff et al., 1978, *Proc. Natl Acad. Sci. EE.UU.* 75: 3727-3731), así como el promotor *tac* (DeBoer et al., 1983, *Proc. Natl Acad. Sci. EE.UU.* 80: 21-25). Más promotores se describen en "Useful proteins from recombinant bacteria" ("Proteínas útiles provenientes de bacterias recombinantes") in *Scientific American* (en Americano Científico), 1980, 242:74-94; y en Sambrook et al.,
- 45 mencionado anteriormente.
- 50 Para células anfitrionas de hongos filamentosos, promotores adecuados para dirigir la transcripción de las este curas de ácidos nucleicos de esta presentación incluyen los promotores obtenidos de los genes amilasa *Aspergillus oryzae TAKA*, proteínasa aspártica *Rhizomucor miehei*, amilasa - alfa neutral *Aspergillus niger*, amilasa - alfa de ácido estable *Aspergillus niger*, glucoamilasa *Aspergillus niger* o *Aspergillus awamori* (*glaA*), Lipasa *Rhizomucor miehei*, proteasa alcalina *Aspergillus oryzae*, isomerasa de fosfato *Aspergillus oryzae* triose, acetadimasa *Aspergillus nidulans*, y proteasa tipo tripsina *Fusarium oxysporum* (WO 96/00787), así como el promotor NA2-tpi (un híbrido de los genes amilasa - alfa *Aspergillus niger* e isomerasa de fosfato *Aspergillus oryzae* triose), y sus promotores mutantes, truncados e híbridos.
- 55 En una levadura anfitriona, promotores útiles pueden provenir de los genes para enolasa *Saccharomyces cerevisiae* (ENO-1), galactoquinasa *Saccharomyces cerevisiae* (GAL1), *Saccharomyces cerevisiae* deshidrogenasa / glicer aldehído-3-fosfato dehidrogenasa alcohólica (ADH2/GAP), y 3-fosfoglicerato quinasa *Saccharomyces cerevisiae*. Otros promotores útiles para las células anfitrionas de levadura se describen por Romanos et al., 1992,
- 60 Yeast (Levadura) 8:423-488.
- 65

- Las secuencias de control también podría ser una secuencia terminadora de transcripción, una secuencia reconocida por una célula anfitriona para terminar la transcripción. La secuencia de terminación está vinculada operacionalmente al término 3' de la secuencia de ácido nucleico que codifica al polipéptido. Cualquier terminador que es funcional en la célula anfitriona escogida puede ser utilizado en este invento.
- 5 Ejemplos de terminadores de transcripciones para células anfitrionas de hongos filamentosos pueden obtenerse de los genes para amilasa *Aspergillus oryzae* TAIGA, glucoamilasa *Aspergillus niger*, antranilato de sintasa *Aspergillus nidulans*, glucosidasa – alfa *Aspergillus niger* y proteasa tipo tripsina *Fusarium oxysporum*.
- 10 Ejemplos de terminadores para células anfitrionas de levadura pueden ser obtenidos de los genes para enolasa *Saccharomyces cerevisiae*, citocromo C *Saccharomyces cerevisiae* (CYC1), y dehidrogenasa gliceraldehido-3-fosfato *Saccharomyces cerevisiae*. Otros terminadores útiles para células anfitrionas de levadura son descritos por Romanos et al., 1992, mencionado anteriormente.
- 15 Las secuencias de control también puede ser una secuencia líder adecuada, una región no interpretada de un ARNm que es importante para la interpretación de la célula anfitriona. La secuencia líder es vinculada operacionalmente al terminal 5' de la secuencia de ácido nucleico que codifica el polipéptido. Cualquier secuencia líder que es funcional en la célula anfitriona escogida puede ser utilizada. Ejemplos de líderes para células anfitrionas de hongos filamentosos se obtienen de los genes para amilasa *Aspergillus oryzae* TAKA e isomerasa de fosfato *Aspergillus nidulans* triose. Líderes adecuados para células anfitrionas de levadura se obtienen de los genes para enolasa *Saccharomyces cerevisiae* (ENO-1), quinasa de 3-fosfoglicerato *Saccharomyces cerevisiae*, factor alfa de *Saccharomyces cerevisiae*, deshidrogenasa de alcohol / deshidrogenasa de gliceraldehido-3-fosfato *Saccharomyces cerevisiae* (ADH2/GAP).
- 20 Las secuencias de control también pueden ser una secuencia de poliadenilación, una secuencia vinculada operacionalmente a la terminal 3' de la secuencia de ácidos nucleicos en donde, cuando transcrita, es reconocida por la célula anfitriona como una señal para añadir residuos de poliadenosina al ARNm transcrito. Cualquier secuencia de poliadenilación que es funcional en la célula anfitriona escogida puede ser utilizada en este invento. Ejemplos de secuencias de poliadenilación para células anfitrionas de hongos filamentosos pueden provenir de los genes para amilasa *Aspergillus oryzae* TAKA, glucoamilasa *Aspergillus niger*, sintasa de antranilato *Aspergillus nidulans*, proteasa tipo tripsina *Fusarium oxysporum*, y glucosidasa – alfa *Aspergillus niger*. Secuencias de poliadenilación útiles para células anfitrionas de levadura se describen por Guo y Sherman, 1995, *Mol Cell Bio* 15:5983-5990.
- 25 La secuencia de control también puede ser una región codificante del péptido de señalización que a su vez codifica para una secuencia de aminoácidos vinculada al terminal amino de un polipéptido y dirige al polipéptido codificado en el sendero de secreción de la célula. El terminal 5' de la secuencia de codificación de la secuencia de ácido nucleico puede contener inherentemente una región codificante del péptido de señalización vinculada naturalmente en el marco de lectura de interpretación con el segmento de la región de codificación que codifica al polipéptido secretado. Alternamente, el terminal 5' de la secuencia de codificación podría contener una región codificante del péptido de señalización que es foránea a la secuencia de codificación. La región foránea codificante del retiro de señalización podría ser requerida cuando la secuencia de codificación no contiene naturalmente de una región de codificación del péptido de señalización.
- 30 Alternamente, la región foránea codificante del péptido de señalización podría reemplazar simplemente de a la región natural codificante del péptido de señalización para mejorar la secreción del polipéptido. Sin embargo, cualquier región codificante del péptido de señalización que dirige al polipéptido expresado en el sendero de secreción de la célula anfitriona escogida podría utilizarse en este invento.
- 35 Regiones efectivas codificantes del péptido de señalización para células anfitrionas bacterianas son las regiones codificantes del péptido de señalización obtenidos de los genes para amilasa maltogénica del Bacilo NCIB 11837, amilasa alfa del Bacilo *stearothermophilus*, subtilisina del Bacilo *licheniformis*, beta-lactamasa del Bacilo *licheniformis*, proteasa neutral del Bacilo *stearothermophilus* (nprT, nprS, nprM), y Bacilo *subtilis* prsA. Más péptidos de señalización se describen por Simonen y Palva, 1993, *Microbiol Rev* 57: 109-137.
- 40 Regiones efectivas codificantes de péptidos de señalización para células anfitrionas de hongos filamentosos pueden ser las regiones codificantes de péptidos de señalización obtenidos de los genes para amilasa de *Aspergillus oryzae* TAKA, amilasa neutral de *Aspergillus niger*, glucomilasa de *Aspergillus niger*, proteinasa aspártica de *Rhizomucor miehei*, celulasa de *Humicola insolens*, y lipasa de *Humicola lanuginosa*.
- 45 Péptidos de señalización útiles para células anfitrionas de levadura pueden provenir de los genes para factor alfa de *Saccharomyces cerevisiae* e invertasa de *Saccharomyces cerevisiae*. Otras regiones codificantes de péptidos de señalización útiles se describen por Romanos et al., 1992, mencionado anteriormente.
- 50 Las secuencias de control también podrían ser una región codificante de péptidos que a su vez codifica para una secuencia de aminoácidos posicionada en el terminal amino de un polipéptido. El polipéptido resultante es conocido
- 55
- 60
- 65

como una proenzima o un propolipéptido (o un zimogeno en algunos casos). Un polipéptido es generalmente inactivo y puede ser convertido a un polipéptido activo maduro por medio de una escisión catalítica o autocatalítica del polipéptido que proviene del propolipéptido. La región codificante del polipéptido puede ser obtenida de los genes para la proteasa alcalina del *Bacilo subtilis* (*aprE*), la proteasa neutral del *Bacilo subtilis* (*nprT*), el factor alfa de *Saccharomyces cerevisiae*, la proteinasa aspártica de *Rhizomucor miehei*, y la lactasa de la *Myceliophthora thermophila* (WO 95/33836).

En los casos en los cuales están presentes péptidos de señalización y regiones de propéptidos en el terminal amino de un polipéptido, la región del polipéptido es posicionada junto al terminal amino de un polipéptido y la región del péptido de señalización está posicionada junto al terminal amino de la región del polipéptido.

También podría ser deseable el agregar secuencias regulatorias, que permitirán la regulación de la expresión del polipéptido en relación al crecimiento de la célula anfitriona. Ejemplos de sistemas regulatorios son aquellos que causan que la expresión del gen se active o se desactive como respuesta a un estímulo químico o físico, incluyendo la presencia de un compuesto regulatorio. En las células anfitrionas procariotas, secuencias regulatorias adecuadas incluyen los sistemas de operadores *lac*, *tac* y *trp*. En las células anfitrionas de levadura, sistemas regulatorios adecuados incluyen, por ejemplo, el sistema *ADH2* o el sistema *GALI*. En hongos filamentosos, las secuencias regulatorias adecuadas incluyen al promotor amilasa – alfa *TAKA*, el promotor glucoamilasa de *Aspergillus*, y el promotor glucoamilasa de *Aspergillus oryzae*.

Otros ejemplos de secuencias regulatorias son aquellas que permiten la amplificación genética. En sistemas eucarióticos, estas incluyen al gen de reductasa de dihidrofolato, que es amplificado en la presencia de metotrexato, y los genes de metalotioneína, que son amplificados por metales pesados. En estos casos, la secuencia de ácidos nucleicos que codifican al polipéptido *KRED* de este invento estarán enlazados operablemente con la secuencia regulatoria.

Por lo tanto, en otra sección, esta presentación también está dirigida a un vector de una expresión recombinante que contiene un polinucleótido que codifica a un polipéptido diseñado de cetorreductasa o a una de sus variantes, y una o más regiones que regulan expresiones tales como un promotor y un terminador, y origen de replicación, etcétera, dependiendo del tipo de anfitrión al cual van a ser introducidos. Las varias secuencias de ácidos nucleicos y de control aquí descritas pueden ser unidas para producir un vector de expresión recombinante que podría incluir uno o más lugares convenientes de restricción para permitir la inserción o sustitución de la secuencia de ácido nucleico que codifica al polipéptido en aquellos sitios. Alternamente, la secuencia de ácido nucleico de esta presentación puede expresarse al insertar la secuencia de ácido nucleico o una estructura de ácido nucleico que contenga la secuencia en un vector apropiado para su expresión. Al crear el vector de expresión, la secuencia de codificación es ubicada en el vector para que la secuencia de codificación sea vinculada operacionalmente con las secuencias de control apropiadas para su expresión.

El vector de expresión recombinante puede ser cualquier vector (por ejemplo, un plásmido o un virus), que puede estar sujeto convenientemente a procedimientos de ADN recombinante y puede provocar la expresión de la secuencia de polinucleótidos. La selección del vector comúnmente dependerá en la compatibilidad del vector con la célula anfitriona en la que el vector será introducido. Los vectores podrían ser plásmidos, circulares, lineares o cerrados.

El vector de expresión puede ser un vector autónomamente replicante, es decir, un vector que existe como una entidad extracromosómica, cuya replicación es independiente de una replicación cromosómica, por ejemplo, un plásmido, un elemento extra cromosómico, un mini cromosoma, o un cromosoma artificial. El vector podría contener algún sistema para asegurar su auto replicación. Alternamente, el vector podría ser uno que, cuando se introduce en la célula anfitriona, es integrado al genoma y replicado junto con el o los cromosomas en los cuales ha sido integrado. Además, se podría usar un solo vector o plásmido o de dos o más vectores plásmidos que contengan juntos la totalidad de ADN a ser introducida en el genoma de la célula anfitriona, o transposón.

El vector de expresión de este invento contiene preferiblemente uno o más marcadores seleccionables, que permiten una selección fácil de las células transformadas. Un marcador seleccionable es un producto genético que suministra una resistencia biocida o viral, resistencia a los metales pesados, prototofía a auxótrofos, y similares. Ejemplos de marcadores seleccionables bacterianos son los genes del *Bacilo subtilis* o *Bacilo lichniformis*, o marcadores, que otorgan resistencia antibiótica tales como la ampicilina, kanamicina, cloranfenicol (ejemplo 1) o resistencia a la tetraciclina. Marcadores adecuados para células anfitrionas de levadura son *ADE2*, *HIS3*, *LEU2*, *LYS2*, *MET3*, *TRP1*, y *URA3*.

Marcadores seleccionables para su uso en células anfitrionas de hongos filamentosos incluyen, pero no se limitan a, *amdS* (acetamidasa), *argB* (ornitina carbamoiltransferasa), *bar* (fosfotricin acetiltransferasa), *hph* (higromicina fosfotransferasa), *niaD* (nitrato reductasa), *pyrG* (orotidin-5'-fosfato decarboxilasa), *sC* (sulfato adeniltransferasa), y *trpC* (antranilato sintasa), así como sus equivalentes. Las secciones para el uso en una célula *Aspergillus* incluyen los genes *amdS* y *pyrG* de *Aspergillus nidulans* o *Aspergillus oryzae* y la barra genética de *Streptomyces hygroscopicus*.

5 Los vectores de expresión de este invento contienen preferiblemente un elemento o varios elementos que permiten la integración del vector al genoma de la célula anfitriona o una replicación autónoma del vector en la célula independiente del genoma. Para integración adentro del genoma de la célula anfitriona, el vector puede basarse en la secuencia de ácido nucleico que codifica al polipéptido o cualquier otro elemento del vector para su integración al genoma por medio de recombinaciones homólogas y no homólogas.

10 Alternamente, el vector de expresión puede contener secuencias adicionales de ácidos nucleicos para dirigir la integración por medio de recombinaciones homólogas al genoma de la célula anfitriona. Las secuencias adicionales de los ácidos nucleicos permiten al vector integrarse al genoma de la célula anfitriona en una ubicación o ubicaciones específicas en el cromosoma o cromosomas. Para incrementar la posibilidad de integración en una ubicación precisa, los elementos integradores deberán contener preferiblemente un número suficiente de ácidos nucleicos, tal como 100 a 10,000 parejas base, preferiblemente de 400 a 10,000 parejas base, y más preferiblemente de 800 a 10,000 parejas base, que sean altamente homólogas en referencia a la secuencia objetiva correspondiente para mejorar la probabilidad de una recombinación homóloga. Los elementos integradores podrían ser cualquier secuencia que sea homóloga con la secuencia objetiva en el genoma de la célula anfitriona. Es más, los elementos integradores podría ser secuencias de ácidos nucleicos no codificantes o codificantes. Por otro lado, el vector puede estar integrado al genoma de la célula anfitriona por medio de recombinación no homóloga.

20 Para una replicación autónoma, el vector podría tener adicionalmente un origen de replicación que permita al vector replicarse autónomamente en la célula anfitriona en cuestión. Ejemplos de orígenes bacterianos de replicación son P15S ori (como se muestra del plásmido de la figura 5) o los orígenes de replicación de plásmidos pBR322, pUC19, pACYC177 (plásmidos que tienen PISA ori), o pACYC184 que permite la replicación en *E. coli*, y pUB110, pE194, pTA1060, o pAM131 que permite la replicación en bacilos. Ejemplos de orígenes de replicación para su utilización en células anfitrionas de levadura son los 2 orígenes de micrones de replicación, 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 hace que su funcionamiento sea sensible a la temperatura en la célula anfitriona (refiérase, por ejemplo, a Ehrlich, 1978, Proc Natl Acad Sci. EE.UU. 75:1433).

30 Más de una copia de una secuencia de ácido nucleico de este invento podría insertarse en una célula anfitriona para incrementar la producción del producto genético. Un incremento en el número de copias de la secuencia de ácido nucleico puede obtenerse al integrar por lo menos una copia adicional de la secuencia en el genoma de la célula anfitriona o por medio de la inclusión de un gen marcador seleccionable amplificable con la secuencia de ácido nucleico donde las células que contienen copias amplificadas del gen marcador seleccionables, y por lo tanto copias adicionales de la secuencia de ácido nucleico, pueden seleccionarse por medio de cultivos de las células en la presencia del agentes seleccionables apropiados.

40 Muchos de los vectores de expresión que se utilizan en este invento no están disponibles comercialmente. Los vectores apropiados de expresión que son comerciales incluyen p3xFLAGTM™, vectores de expresión de Sigma-Aldrich Chemicals, St. Louis MO., que incluyen un promotor CMV y un lugar de poliadenilación hGH para la expresión en células anfitrionas de mamíferos y un origen pBR322 de replicación y marcadores de resistencia de ampicilina para la amplificación en *E. coli*. Otros vectores de expresión apropiados son pBluescriptII SK(-) y pBK-CMV, que son comercialmente disponibles de Stratagene, LaJolla CA, y plásmidos que se derivan de pBR322 (Gibco BRL), pUC (Gibco BRL), pREP4, pCEP4 (Invitrogen) o pPoly (Lathe et al., 1987, Gene (Gen) 57:193-201).

45 1.4. Células Anfitrionas para la Expresión de Polipéptidos de Cetorreductasa

50 En otro aspecto, esta presentación suministra una célula anfitriona que contiene un polinucleótido que codifica a un polipéptido mejorado de cetorreductasa de esta presentación, polinucleótido que está enlazado operativamente a una o más secuencias de control para la expresión de la enzima de cetorreductasa en la célula anfitriona. Las células anfitrionas que son útiles para la expresión de los polipéptidos KRED codificados por los vectores de expresión de este invento son bien conocidas en la industria e incluyen pero no se limitan a, células bacterianas, tales como células de *E. coli*, lactobacilo kefir, lactobacilo brevis, *Streptomyces* y *Salmonella typhimurium*; células de hongos, tales como las células de levadura (por ejemplo, *Saccharomyces cerevisiae* o *Pichia pastoris* (número de acceso ATCC. 201178)); células de insecto o como las células de *Drosophila S2* y *Spodoptera Sf9*; células animales tales como las células de CHO, COS, BHK, 293, y melanoma Bowes; células vegetales. Medios de cultivos y condiciones de crecimiento apropiados para las células que se acaban de mencionar son bien conocidos en la industria.

60 Polinucleótidos para la expresión de la cetorreductasa pueden ser introducidos en células por varios métodos conocidos en la industria. Las técnicas incluyen, entre otras, electroporación, bombardeo de partículas biolísticas, transfección mediada por liposomas, transfección de cloruro de calcio, y la fusión de protoplastos. Varios métodos para introducir polinucleótidos en células serán evidentes para el experto en la materia.

65 Un ejemplo de célula anfitriona es *Escherichia coli* W3110. El vector de expresión fue creado al vincular operativamente por el nucleótido que codifica una cetorreductasa mejorada al plásmido pCK110900 vinculado

operativamente al promotor lac bajo control del represor lacI. El vector de expresión también contuvo el origen P15a de replicación y el gen de resistencia cloranfenicol. Las células que contienen el polinucleótido sujeto en la Escherichia coli W3110 fueron aislados al exponer a las células a la selección del cloranfenicol.

5 1.5. Métodos para Generar Polipéptidos Diseñados de Cetorreductasa

En algunas secciones, para hacer polinucleótidos y polipéptido mejorados KRED de esta presentación, las enzimas de cetorreductasa que ocurren naturalmente y que catalizan la reacción de reducción son obtenidas (o derivadas) del lactobacilo kéfir, el lactobacilo brevis o el lactobacilo minor. En algunas secciones, la secuencia de polinucleótida paterna es optimizada en lo que se refiere a los codones para mejorar la expresión de la cetorreductasa en una célula anfitriona específica. Como una ilustración, la secuencia de polinucleótidos paterna que codifica al polipéptido KRED de tipo silvestre del lactobacilo kéfir se construyó de oligonucleótidos preparados que se basaron en la secuencia conocida de polipéptidos de la secuencia de KRED del lactobacilo kéfir disponible en la base de datos del Banco Genético (número de acceso del banco genético AAP94029 GI:33112056). La secuencia de polinucleótidos paternal, diseñada como la IDENTIFICACIÓN SECUENCIAL NÚMERO: 1, fue optimizada en lo que se refiere a codones para la expresión en E. coli y el polinucleótido optimizado en lo que se refiere a codones fue clonado en un vector de expresión, colocando la expresión del gen de cetorreductasa bajo el control del promotor lac y el gen represor lacI. Los clones que expresan la cetorreductasa activa en el e. coli fueron identificados y los genes fueron secuenciados para confirmar su identidad. La secuencia designada (IDENTIFICACIÓN SECUENCIAL NÚMERO: 1) fue la secuencia paterna utilizada como punto de inicio para la mayoría de experimentos y la construcción de la biblioteca de las cetorreductasas diseñadas que evolucionaron a partir de la cetorreductasa del lactobacilo kéfir.

Las cetorreductasas diseñadas pueden ser obtenidas del polinucleótido que codifica a la cetorreductasa que ocurre naturalmente por medio de mutagénesis y / o métodos dirigidos de evolución, tal como se mencionó anteriormente. Un ejemplo de una técnica dirigida de evolución es la mutagénesis y / o el arrastramiento de ADN tal como se describió en Stemmer, 1994, Proc Natl Acad Sci EE.UU. 91:10747-10751; WO 95/22625; WO 97/0078; WO 97/35966; WO 98/27230; WO 00/42651; WO 01/75767 y la patente de Estados Unidos 6,537,746. Otros procedimientos de evolución dirigida que pueden ser utilizados incluyen, entre otros, el proceso escalonado de extensión (StEP - staggered extension process), recombinación in vitro (Zhao et al., 1998, Nat. Biotechnol. 16:258-261), PCR mutagénico (Caldwell et al., 1994, PCR Methods Appl. (Appl. Métodos) 3:S136-S140), y mutagénesis de cinta (Black et al., 1996, Proc Natl Acad Sci EE.UU. 93:3525-3529).

Los clones obtenidos siguiendo el tratamiento de mutagénesis son examinados para encontrar las cetorreductasas diseñadas que tengan una propiedad enzimática mejorada y deseada. El medir la actividad enzimática de las bibliotecas de expresión puede realizarse usando la técnica bioquímica estándar para monitorear la tasa de disminución (por medio de una disminución en la absorción o fluorescencia) de la concentración NADH o NADPH, mientras se convierte en NAD⁺ o NADP⁺. (Para un ejemplo, refiérase al ejemplo 7.) En esta reacción, la NADH o NADPH es consumida (oxidada) por la cetorreductasa mientras la cetorreductasa reduce un sustrato de cetona al correspondiente grupo hidroxilo. La tasa de disminución de la concentración de NADH o NADPH, tal como se mide por la disminución de absorción o fluorescencia, por unidad de tiempo indica la actividad relativa (enzimática) del polipéptido KRED en un monto fijo del lisado (o una de sus formas convertidas en polvo liofilizado). Donde la propiedad enzimática mejorada deseada es estabilidad térmica, la actividad enzimática puede ser medida después de exponer a las preparaciones enzimáticas a una temperatura definida y medir el monto de actividad enzimática que queda después de los tratamientos de calor. Los clones que contienen un polinucleótido que codifica a una cetorreductasa son aislados, y secuenciados para identificar los cambios secuenciales de los nucleótidos (si existiesen), y utilizados para expresar la enzima en una célula anfitriona.

En los casos en que la secuencia del polipéptido mejorado es conocida, los polinucleótidos que codifican a la enzima pueden prepararse por medio de métodos estándar en su fase sólida, de acuerdo a los métodos sintéticos conocidos. En algunas secciones, fragmentos de hasta 100 bases pueden ser sintetizados individualmente, y después juntados (por ejemplo, por medio de métodos de litigación enzimática o química, o métodos mediados por polimerasas) para formar cualquier secuencia continua deseada. Por ejemplo, los polinucleótidos y los oligonucleótidos del invento pueden prepararse por medio de síntesis química usando, por ejemplo, el método clásico de fosforamidita descrito por Beaucage et al., 1981, Tet Lett 22:1859-69, o el método descrito por Matthes et al., 1984, EMBO J. 3:801-05, por ejemplo, como es comúnmente practicado en los métodos sintéticos automatizados. De acuerdo al método de fosforamidita, los oligonucleótidos son sintetizados, por ejemplo, en un sintetizador automático de ADN, purificados, reconocidos, ligados y clonados en vectores apropiados. Adicionalmente, esencialmente cualquier ácido nucleico puede ser obtenido de cualquier variedad de 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 de cetorreductasa diseñadas en una célula anfitriona pueden ser recuperadas de las células y o del medio de cultivo usando una o más técnicas bien conocidas para la purificación proteínica, incluyendo, entre

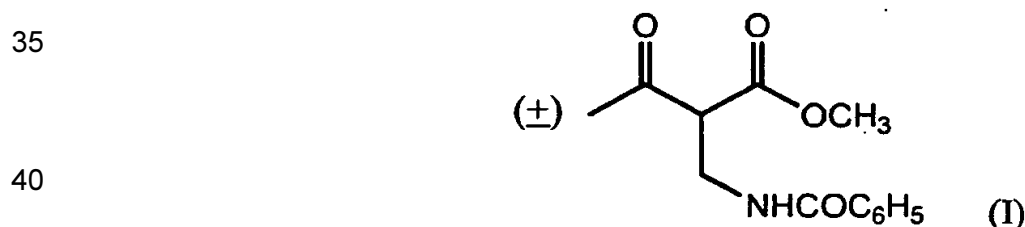
otras el tratamiento de lisozima, sonicación, filtración, precipitación salina, ultracentrifugación y cromatografía. Soluciones adecuadas para lisar y para la extracción altamente eficiente de proteínas de bacterias, tales como el *E. coli*, están comercialmente disponibles bajo el nombre de marca de CelLytic B™ de Sigma - Aldrich de St. Louis MO.

5 Las técnicas cromatográficas para el aislamiento de los polipéptidos de cetorreductasa incluyen, entre otros, cromatografía líquida de alto rendimiento para cromatografía por fases en reversa, cromatografía por intercambio de iones, electroforesis por medio de gel, y cromatografía de afinidad. Las condiciones para purificar una enzima específica dependerán, en parte, de factores tales como la carga neta, la hidrofobicidad, la hidrofili-
10 cidad, la masa molecular, la forma molecular, etcétera, y será aparente para aquellos que tienen conocimiento en la industria.

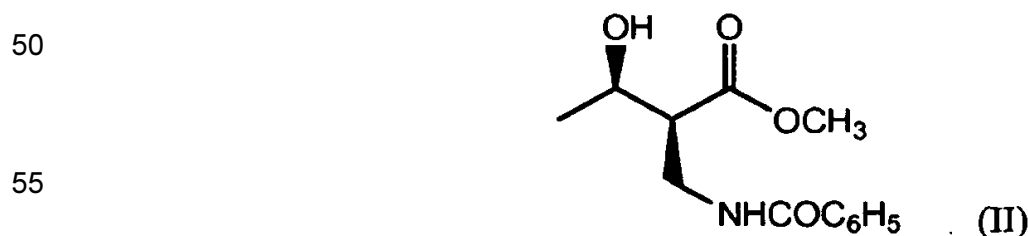
En algunas secciones, se podrá utilizar técnicas de afinidad para aislar las enzimas de cetorreductasa mejoradas. Para purificación por cromatografía de afinidad, cualquier anticuerpo que vincula específicamente al polipéptido de cetorreductasa puede ser utilizado. Para la producción de anticuerpos, varios animales anfitriones, incluyendo, pero sin limitarse a, conejos, ratones, ratas, etcétera, pueden ser inmunizados por medio de inyecciones con un polipéptido diseñado. El polipéptido puede estar adherido a un portador adecuado, tal como BSA, por medio de un grupo funcional de cadena lateral o vinculadores adheridos a un grupo funcional de la cadena lateral. Varios adyuvantes pueden ser utilizados para incrementar la respuesta inmunológica, dependiendo de las especies anfitrionas, incluyendo, pero sin limitarse a, los de Freund (completos e incompletos), geles minerales tales como hidróxido de aluminio, sustancias activas superficiales tales como la lisolecitina, polioles plurónicos, polianiones, péptidos, emulsiones de aceite, hemocianina de lapa de ojo de cerradura, dinitrofenol, y adyuvantes humanos potencialmente útiles tales como BCG (bacilos Calmette Guerin) y *Corynebacterium parvum*.

25 1.6. Métodos para Usar las Enzimas de Cetorreductasa Diseñadas y Sus Compuestos Preparados a partir de estas

En algunas secciones, las enzimas de cetorreductasa aquí descritas son capaces de catalizar la reacción de reducción del grupo ceto en el compuesto de la fórmula estructural (I), metil-2-benzamidometil-3-oxobutirato ("el sustrato"):



45 al correspondiente producto alcohólico estereoisomérico de la fórmula estructural (II), 2S,3R-metil-2-benzamidometil-3-hidroxi-3-butirato ("el producto"):

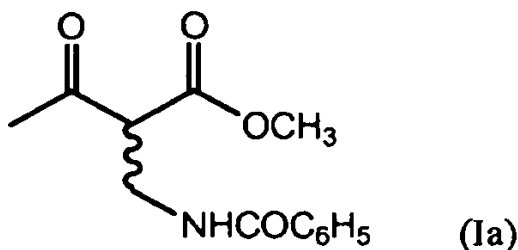


60 En algunas secciones, el sustrato de la fórmula (I) es una mezcla racémica, tal como se muestra en la fórmula (Ia).

65

5

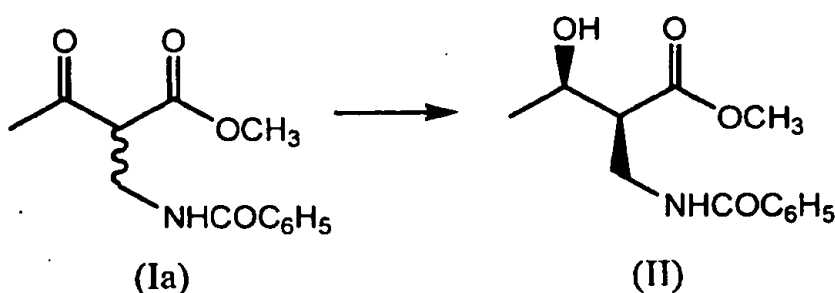
10



15 Y las cetorreductasas selectivas 2S,3R pueden ser utilizadas para reducir o convertir el sustrato racémico en la reacción que se muestra en el Esquema 1 a continuación para preparar el producto de la fórmula (II):

20

25



30

Esquema 1

35

40

45

50

Asimismo, en algunas secciones, las cetorreductasas de la presentación pueden ser utilizadas en un método para reducir estereoselectivamente el sustrato de la fórmula (I) al producto correspondiente de la fórmula estructural (II) con por lo menos alrededor de un 60% de exceso estereomérico, método que comprende el contactar o incubar el sustrato de la fórmula (I) o la fórmula (Ia) con un polipéptido de cetorreductasa selectivo 2S,3R de la presentación bajo condiciones de reacción apropiadas para la reducción o conversión de sustrato del producto de la fórmula (II). En algunas secciones de este método el producto de la fórmula (II) es producido con un mínimo de alrededor del 85% de exceso estereomérico. En algunas secciones de este método, los polipéptidos de cetorreductasa selectivos 2S,3R tienen, en referencia a las secuencias KRED de tipo silvestre de *L. kékfir*, *I. brevis* o *L. minor* de las IDENTIFICACIONES SECUENCIALES NÚMEROS: 4, 2 y 86 por lo menos las siguientes características: el residuo 202 es valina o leucina. En algunas secciones, los polipéptidos de cetorreductasa selectivos 2S,3R tienen, en referencia a las secuencias KRED de tipo silvestre de *L. kékfir*, *I. brevis* o *L. minor* de las IDENTIFICACIONES SECUENCIALES NÚMEROS: 4, 2 y 86 por lo menos las siguientes características: (1) el residuo correspondiente a X 94 es un residuo alifático o polar; (2) el residuo correspondiente a X 199 es un residuo alifático, restringido o polar; y (3) el residuo correspondiente a X 202 es valina o leucina. En algunas secciones, los polipéptidos de cetorreductasa selectivos 2S,3R tienen, en referencia a las secuencias KRED de tipo silvestre de *L. kékfir*, *L. brevis* o *L. minor* de las IDENTIFICACIONES SECUENCIALES NÚMEROS: 4, 2 y 86 por lo menos las siguientes características: (1) el residuo correspondiente a X 94 es un residuo polar; (2) el residuo correspondiente a X 199 es un residuo restringido, y (3) el residuo correspondiente a X 202 es valina o leucina.

55

60

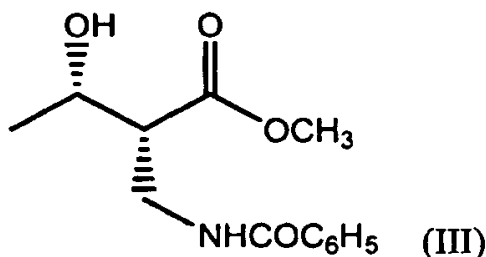
En algunas secciones del método para convertir el sustrato al producto de la fórmula estructural (II), el sustrato es convertido al producto con un exceso estereomérico porcentual de por lo menos alrededor de 60 - 89% y una tasa que es de por lo menos alrededor de 1 - 15 veces más grande que la tasa de capacidad del polipéptido que tiene la secuencia de aminoácidos de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48. Ejemplos de polipéptidos que pueden ser usados en este método incluyen, pero no se limitan a, polipéptidos que tienen las secuencias de aminoácidos correspondientes a las IDENTIFICACIONES SECUENCIALES NÚMEROS: 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 46, 50, 52, 54, 56, 58, 60, y 62.

65

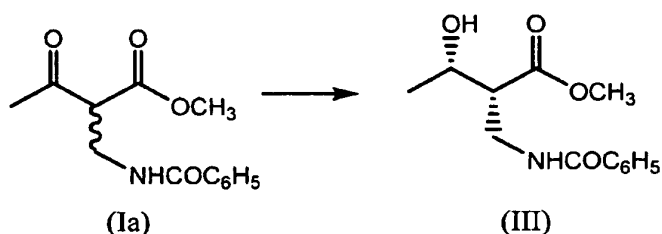
En algunas secciones del método para convertir el sustrato del producto de la fórmula estructural (II), el sustrato es convertido al producto con un exceso estereomérico porcentual de por lo menos alrededor del 90 - 94%. Ejemplos de polipéptidos que pueden ser usados en este método incluyen, pero no se limitan a, polipéptidos que tienen las secuencias de aminoácidos correspondientes a las IDENTIFICACIONES SECUENCIALES NÚMEROS: 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 40, 42, 50, 52, 56, 58, 60, y 62.

En algunas secciones del método para convertir el sustrato del producto de la fórmula estructural (II), el sustrato es convertido a un producto con un exceso estereomérico porcentual de por lo menos alrededor del 95 - 99%. Ejemplos de polipéptidos que pueden ser usados en este método incluyen, pero no se limitan a, polipéptidos que tienen las secuencias de aminoácidos correspondientes a las IDENTIFICACIONES SECUENCIALES NÚMEROS: 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 42, 50, 52, 56, 58, 60, y 62.

En algunas secciones, las enzimas de cetorreductasa aquí descritas son capaces de catalizar la reacción de reducción del grupo ceto en el compuesto de la fórmula estructural (I) al producto alcohólico estereoisomérico correspondiente de la fórmula estructural (III), 2R,3R-metil-2-benzamidometil-3-hidroxiacetato, (lo cual también puede ser referido como un "producto"):



En algunas secciones, el sustrato de la fórmula (I) es una mezcla racémica como se muestra en la fórmula (Ia), y las cetorreductasas selectivas 2R,3R pueden ser utilizadas para reducir o convertir el sustrato racémico en la reacción conocida del Esquema 3 más adelante para preparar el producto de la fórmula (III):

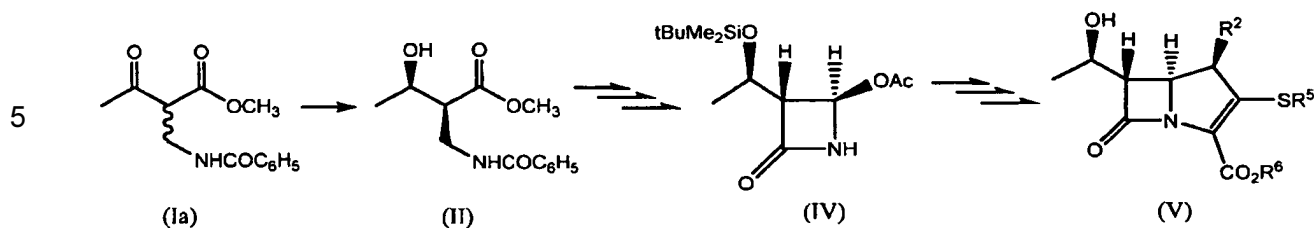


Esquema 2

En algunas secciones del método para convertir el sustrato al producto de la fórmula estructural (III), el sustrato es convertido a un producto con un exceso estereomérico porcentual de por lo menos alrededor del 85%. Ejemplos de polipéptidos que pueden ser utilizados en este método incluyen, pero no se limitan a, polipéptidos que tienen la secuencia de aminoácidos correspondientes a las IDENTIFICACIONES SECUENCIALES NÚMEROS: 68, 72, 74, 76, 78, y 82.

En algunas secciones del método para convertir el sustrato al producto de la fórmula estructural (III), el sustrato es convertido a un producto con una tasa que es por lo menos alrededor de una vez mayor que la tasa de capacidad del polipéptido que tiene la secuencia de aminoácidos de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 66. Ejemplos de polipéptidos que pueden ser utilizados en este método incluyen, pero no se limitan a, polipéptidos que tienen las secuencias de aminoácidos correspondientes a las IDENTIFICACIONES SECUENCIALES NÚMERO: 64, 68, 70, 72, 74, 76, 78, 80, y 82.

En algunas secciones, las cetorreductasas aquí descritas pueden ser utilizadas en un método para sintetizar un carbapenem, el método comprende el proceso general que se muestra continuación (Esquema 3),



10

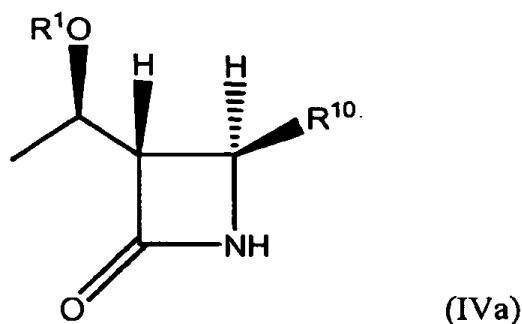
Esquema 3

15 Donde la KRED puede ser cualquiera de los polipéptidos de cetorreductasa aquí descritos. El carbapenem de la fórmula general (V) está descrito en mayor detalle a continuación. Para la síntesis de varias terapias que se basan en carbapenems, un intermediador importante es el compuesto de la fórmula estructural (IVa),

20

25

30



35 Dónde R¹ es H o un grupo protector de hidroxilo, y R¹⁰ es un halógeno (por ejemplo, Cl), u -OAc (Ac es acetato). Los varios grupos protectores del hidroxilo son conocidos en la industria, e incluyen, en forma de ejemplo y no de limitación, ésteres sililos (por ejemplo, dimetilo de sililo de 1-butilo) y ésteres bencilos sustituidos (refiérase a, Green y Wuts, Protective Groups in Organic Synthesis (Grupos Protectores en Síntesis Orgánica), Wiley-Interscience, Nueva York, 1999). Asimismo, en un método para la síntesis del intermediador de la fórmula estructural (IVa), un paso en el método comprende el contactar o reaccionar el sustrato de la fórmula (I) con la cetorreductasa de la presentación bajo condiciones adecuadas de reacción para reducir o convertir el sustrato al producto de la fórmula (II). La síntesis del intermediador de la fórmula (IVa) en el cual R¹⁰ es -OAc se describe en Tetrahedron Lett. 23:2293 (1982); Tetrahedron Lett. 39:2399 (1983); Tetrahedron Lett. 39:2505 (1983); EP0290385; y EP0369691. En algunas secciones, el intermediador de la fórmula (IVa) en donde R¹⁰ es -OAc puede sintetizarse tal como se ilustra en el Esquema 4:

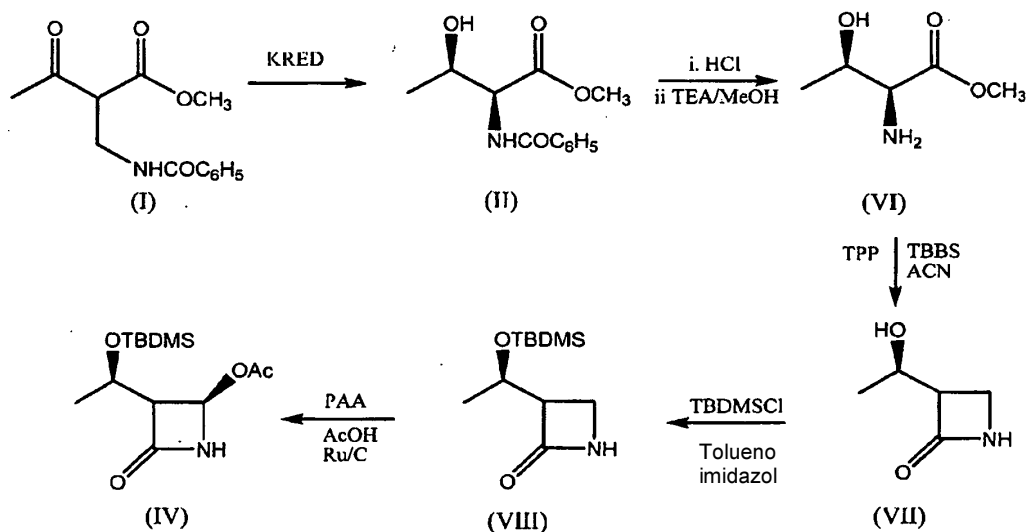
45

50

55

60

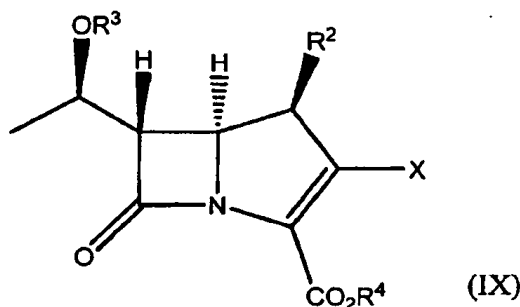
65



Esquema 4

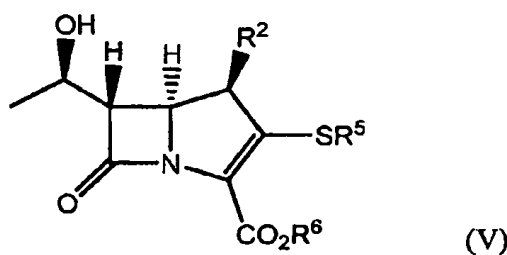
[0245] En el esquema cuatro, el método incluye: (a) reducir el sustrato de la fórmula (I) al producto de la fórmula (II) al contactar o reaccionar el sustrato con una cetorreductasas selectiva 2S,3R de esta presentación bajo condiciones de reacción apropiadas para reducir el sustrato al producto de la fórmula (II); (b) remover el grupo $-C(O)C_6H_5$ para formar el compuesto (VI); (c) convertir el compuesto de la fórmula (VI) a la β -lactama de la fórmula (VII) al reaccionar con trifenilfosfina y N-terc-butil-2-benzotiazolilsulfenamida (refiérase, por ejemplo, a Tetrahedron Lett. 1995, 36(21):3703); (d) reaccionar con terc-butil dimetil clorosilano (TBDMS-Cl) para formar el compuesto de la fórmula (VIII); y (e) acetoxilatar el compuesto de la fórmula (VIII) en presencia de un ácido peracético y rutenio para formar el intermediador de la fórmula (IV) (refiérase, por ejemplo, a Murahashi et al., 1990, J. Am. Chem. Soc. 112(21):7820-7822).

[0246] En algunas secciones, otro intermediador en la síntesis de tratamientos que se basan en carbapenems es el intermediador de la fórmula estructural (IX):



Donde R^2 es H o un alquilo C1-C4 (por ejemplo, $-CH_3$); R^3 es H, o un grupo protector del hidroxilo; R^4 es H, un grupo protector de carboxi, el grupo de amoniaco, metal alcalino, o un metal alcalinotérreo; y X es OH, o un grupo que abandona. Ejemplos de grupos que abandonan incluyen, pero no se limitan a, $-OP(O)(OR')$ u $OS(O_2)R''$, donde R' and R'' pueden ser alquilo C1-C6, alcarilo C1-C6, arilo, alquilo perfluorado C1 -C6. Grupos protectores para R^4 pueden ser, pero no se limitan a, bencilo, p-nitrobencilo, y metoximetilo. Descripciones del intermediador de la fórmula estructural (IX) y su síntesis son suministradas en varias referencias, por ejemplo, la patente de Estados Unidos número 5,317,016; la patente de Estados Unidos número 4,933,333; y WO 0236594. Asimismo, en algunas secciones, en un método para la síntesis del intermediador de la fórmula (IX), un paso en el método comprende el contactar o reaccionar el sustrato de la fórmula (I) con las cetorreductasas de la presentación bajo condiciones de reacción adecuadas para reducir o convertir el sustrato al producto de la fórmula (II).

En algunas secciones, las cetorreductasas de la presentación pueden ser utilizadas en un proceso para la síntesis de compuestos terapéuticos que se basan en carbapenem de la fórmula estructural (V):



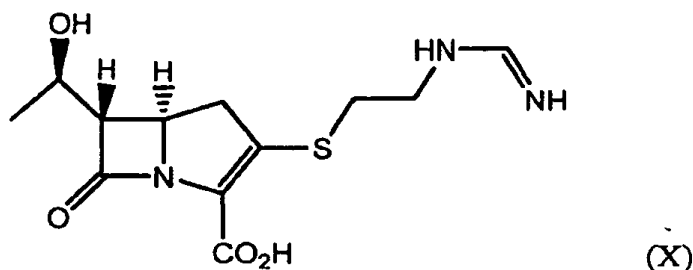
o sus solvatos, hidratos, sales y pro - medicamentos, donde R^2 es H o $-CH_3$; R^1 pueden ser varios sustituyentes, incluyendo, pero sin limitarse a, alquilo sustituido o no sustituido, arilo sustituido o no sustituido, heteroalquilo sustituido o no sustituido, heterocicloalquilo sustituido o no sustituido y heteroariloalquilo sustituido o no sustituido; y R^6 es H, o un progrupo, tal como un grupo de ésteres hidrolizables. Tal como se utiliza aquí, progrupo se refiere a un tipo de grupo protector que, cuando se usa para enmascarar un grupo funcional dentro de un medicamento activo para formar una pro-fracción, convierte al medicamento en un pro-medicamento. Los pro-grupos están adheridos comúnmente al grupo funcional del medicamento por medio de enlaces que son disociables bajo condiciones específicas de uso. Por lo tanto, un pro-grupo es aquella porción de una pro-fracción que se disocia para liberar al grupo funcional bajo condiciones específicas de uso. Un grupo carboxilo puede enmascararse como fracciones de ésteres (incluyendo ésteres y tioésteres de sililos), amidas o hidracidas, que podrían ser hidrolizadas in vivo para suministrar el grupo carboxilo. Otros ejemplos específicos de pro-grupos adecuados y sus respectivas pro-fracciones serán aparentes para aquellas personas con conocimiento en la industria. Asimismo, en el método para la síntesis del compuesto de la fórmula estructural (V), un paso en el método puede comprender el contactar o reaccionar el sustrato de la fórmula (I) con la cetorreductasa de la presentación bajo condiciones de reacción adecuadas para reducir o convertir el sustrato

del producto de la fórmula (II). La síntesis de varios compuestos terapéuticos que se basan en carapebenem son descritos en, entre otras, las patentes de Estados Unidos números: 4,925,836; 5,317,016; WO 02/036594; WO2004/067532; y WO 2007/104221144.

5 En algunas secciones, las cetorreductasas de la presentación pueden ser utilizadas para la síntesis de Imipenem de la siguiente estructura (X):

10

15



20

25

o sus solvatos, hidratos, pro - medicamentos y sales. Imipenem tiene un amplio espectro de actividad en contra de bacterias gram positivas y gram negativas aeróbicas y anaeróbicas. Es particularmente efectivo en contra de *Pseudomonas aeruginosa* y las especies de enterococcus. Asimismo, en un método para la síntesis de Imipenem de la fórmula estructural (X), un paso en el método comprende el contactar o reaccionar el sustrato de la fórmula (I) con las cetorreductasas de la presentación bajo condiciones adecuadas de reacción para reducir o convertir el sustrato del producto de la fórmula (II). El proceso para preparar Imipenem a partir del intermediador de la fórmula estructural (VI) se describe en WO 0236594.

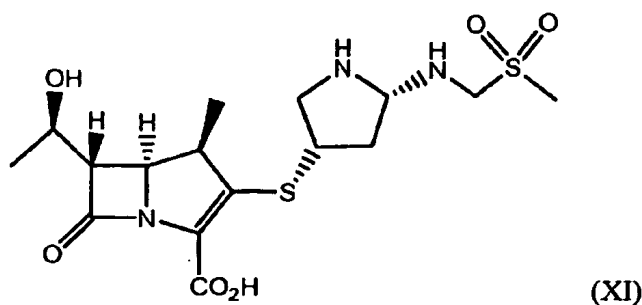
30

En algunas secciones, las cetorreductasas de la presentación pueden ser utilizadas en la síntesis de Doripenem de la fórmula estructural (XI):

35

40

45



50

o sus solvatos, hidratos o sales. Doripenem tiene un espectro y potencia en contra del cocci gram-positivo similar al Imipenem o Ertapenem, y una actividad Gram- negativa similar al Meropenem. Asimismo en un método para la síntesis de Doripenem de la fórmula estructural (XI), un paso en el método comprende contactar o reaccionar el sustrato de la fórmula (I) con una cetorreductasa de este invento bajo condiciones de reacción apropiadas para reducir o convertir el sustrato al producto de la fórmula (II). El proceso para preparar Doripenem del intermediador de la fórmula (VI) está descrito en la patente de Estados Unidos número 5,317,016.

55

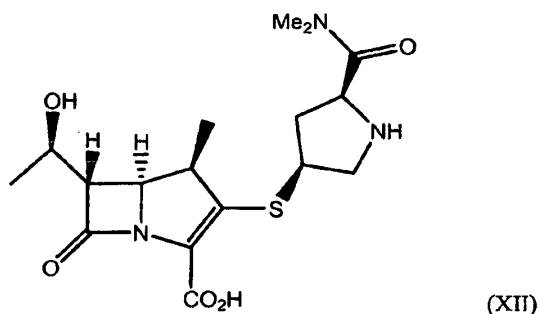
En algunas secciones, las cetorreductasas de la presentación pueden ser utilizadas en la síntesis de Meropenem de la fórmula estructural (XII):

60

65

5

10



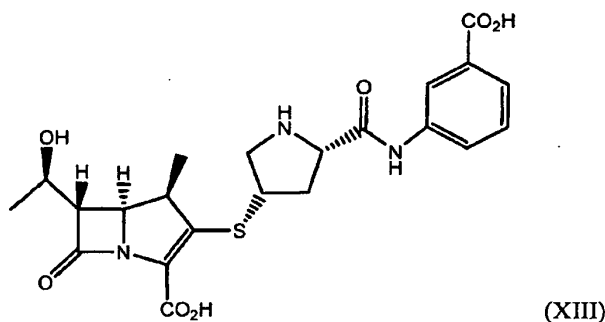
15 o sus solvatos, hidratos, pro - medicamentos y sales. Meropenem es indicado para el tratamiento de infecciones intra - abdominales, por ejemplo, apendicitis y peritonitis; el tratamiento de la meningitis bacteriana; y el tratamiento de la piel o de infecciones estructurales de la piel causadas por organismos susceptibles. Asimismo, en un método para la síntesis de Meropenem de la fórmula (XII), un paso en el método comprende contactar o reaccionar el sustrato de la fórmula (I) con una cetorreductasa de esta presentación bajo condiciones apropiadas de reacción para reducir o convertir el sustrato al producto de la fórmula (II). Un proceso para preparar Meropenem está descrito en WO 2007/104221.

20 En algunas secciones, las cetorreductasas de la presentación pueden ser utilizadas para la síntesis de Ertapenem de la fórmula estructural (XIII):

25

30

35



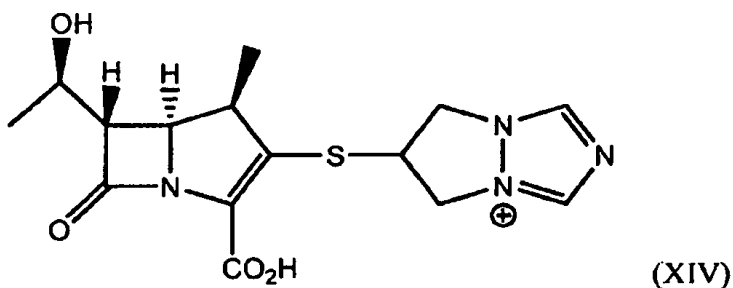
40 o su solvatos, hidratos, pro - medicamentos y sales. Ertapenem es efectivo en contra de las bacterias Gram negativas, pero no es activo en contra de MRSA, el enterococci resistente a la histidina, Pseudomonas aeruginosa o las especies Acinetobacter. Ertapenem también tiene una actividad clínicamente útil en contra de las bacterias anaeróbicas. Ertapenem es utilizado principalmente en contra del amplio espectro de bacterias que producen (ESBL) de beta-lactamasa y Gram- negativas que producen AmpC de alto nivel. Asimismo, en el método para la síntesis de Ertapenem de la fórmula (XIII), un paso del método comprende el contactar o reaccionar el sustrato de la fórmula (I) con las cetorreductasas de la presentación bajo condiciones apropiadas de reacción para reducir o convertir el sustrato al producto de la fórmula (II). Los procesos para preparar Ertapenem están descritos en la patente de Estados Unidos número 5,478,820 y la patente de Estados Unidos número 7,342,005.

50

55 En algunas secciones, las cetorreductasas de la presentación pueden ser utilizadas para la síntesis de Biapenem de la fórmula estructural (XIV):

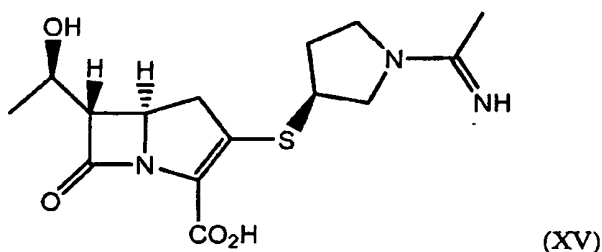
60

65



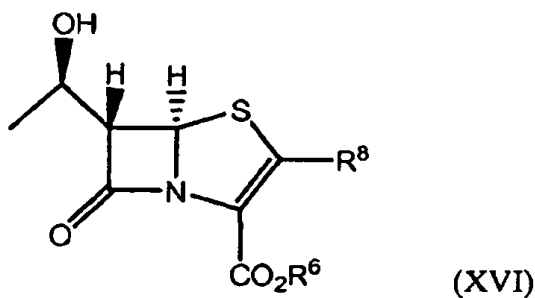
o sus solvatos, hidratos, pro - medicamentos y sales. Biapenem es un carbapenem parenteral que posee actividades anti bacterianas en contra de una amplia gama de bacterias Gram- positivas y -negativas, y es estable para la deshidropeptidasa-I (DHP-I) renal humana. También muestra una actividad in vitro en contra de bacterias anaeróbicas. Asimismo, en un método para la síntesis de Biapenem de la fórmula (XIV), un paso comprende contactar o reaccionar el sustrato de la fórmula (I) con las cetorreductasas de la presentación bajo condiciones apropiadas de reacción para reducir y convertir el sustrato al producto de la fórmula (II). El proceso para preparar Piapenem a partir del intermediador (IVa) se describe en la patente de Estados Unidos número 4,925,836.

En algunas secciones, las cetorreductasas de la presentación pueden ser utilizadas en la síntesis de Panipenem de la fórmula estructural (XII):



o sus solvatos, hidratos, pro - medicamentos y sales. Asimismo, en un método para la síntesis de Panipenem de la fórmula (XV), un paso en el método comprende el contactar o reaccionar el sustrato de la fórmula (I) con las cetorreductasas de la presentación bajo condiciones apropiadas de reacción para reducir o convertir el sustrato al producto de la fórmula (II). Panipenem a partir del intermediador de la fórmula (II) está descrito en Miyadera et al., 1983, J Antibiotic (J Antibiótico) (Tokio) 36(8): 1 034-1039; y Hirai et al., 1999, J Synth Org. Chem. 57(5):382-386.

En algunas secciones, las cetorreductasas de la presentación pueden ser utilizadas en un método para la síntesis de componentes de sulopenem de la fórmula (XVI):

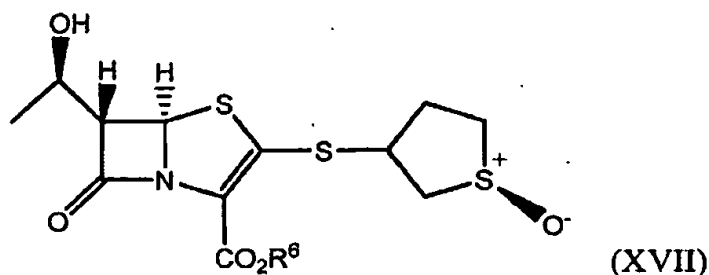


Donde R⁶ es H o un pro - grupo, y R⁸ puede ser un alquilo sustituido o no sustituido, un arilo sustituido o no sustituido; un heteroalquilo sustituido o no sustituido, un heterocicloalquilo sustituido o no sustituido y un heteroariloalquilo sustituido o no sustituido, o -SR⁹, donde R⁹ puede ser uno de los sustituyentes ya descritos para R⁸. Otros sustituyentes para R⁸ que están en compuestos terapéuticos bioactivos de sulopenem son conocidos en la industria, y se describen algunos de ellos a continuación. Las descripciones de los sulopenems pueden encontrarse en, entre otras fuentes, la patente de Estados Unidos número 3,951,954; la patente de Estados Unidos número 4,234,579, la patente de Estados Unidos número 4,287,181; la patente de Estados Unidos número 4,452,796; la patente de Estados Unidos número 4,343,693; la patente de Estados Unidos número 4,348,264; la patente de Estados Unidos número 4,416,891; la patente de Estados Unidos número 4,457,924; y la patente de Estados Unidos número 5,013,729; la patente de Estados Unidos número 5,506,225; Volkman et al., 1992, J. Org. Chem. 57:4352-4361; y Volkman y O'Neill, 1991, Strategies and Tactics in Organic Synthesis (Estrategias y Tácticas en Síntesis Orgánica), Vol 13, pg 485-534. Asimismo, en un método para la síntesis de sulopenem de la fórmula estructural (XVI), un paso comprende contactar o reaccionar el sustrato de la fórmula (I) con las cetorreductasas de la presentación bajo condiciones de reacción adecuadas para reducir o convertir el sustrato al producto de la fórmula (II).

En algunas secciones, las cetorreductasas de la presentación pueden ser utilizadas en un método para la síntesis de un sulopenem de la fórmula (XVII):

5

10



15

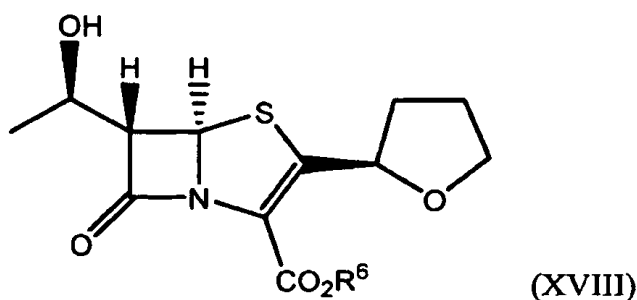
20

o sus solvatos, hidratos pro - medicamentos y sales, donde R^6 es H o un pro - grupo. Asimismo en un método para la síntesis del compuesto de la fórmula (XVII), un paso en el método comprende el contactar o reaccionar el sustrato de la fórmula (I) con las cetorreductasas de esta presentación bajo condiciones adecuadas de reacción para reducir o convertir el sustrato al producto de la fórmula (II). Un proceso para la síntesis del sulopenem de la fórmula (XVII) está descrito en la patente de Estados Unidos número 5,013,729. Varios pro - grupos para R^6 se describen en la publicación de la aplicación de Estados Unidos 2008/0009474, la publicación de la aplicación de Estados Unidos 2008/0125408; y Wujcik et al., 2008, Rapid Commun. Mass Spectrom (Espectro Rápido de Masa Commun.). 22:3195-3206.

25

30

35



40

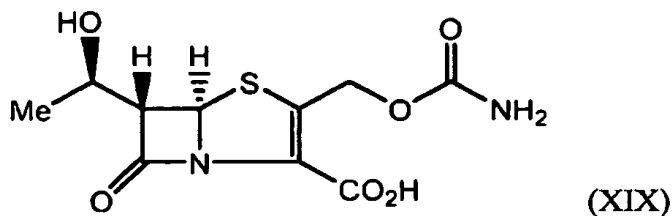
45

o sus solvatos, hidratos pro - medicamentos y sales, donde R^6 es H o un pro - grupo. Asimismo, en un método para la síntesis de Panipenem de la fórmula (XVIII), un paso en el método comprende contactar o reaccionar el sustrato de la fórmula (I) con las cetorreductasas de la presentación bajo condiciones de reacción apropiadas para reducir o convertir el sustrato al producto de la fórmula (II). El proceso para la síntesis del sulopenem de la fórmula (XVIII) está descrito en la patente de Estados Unidos número 5,506,225 y WO 2007/039885.

[0257] Otro compuesto de sulopenem que puede ser sintetizado usando las cetorreductasas aquí presentadas es el sulopenem de la fórmula (XIX):

50

55



60

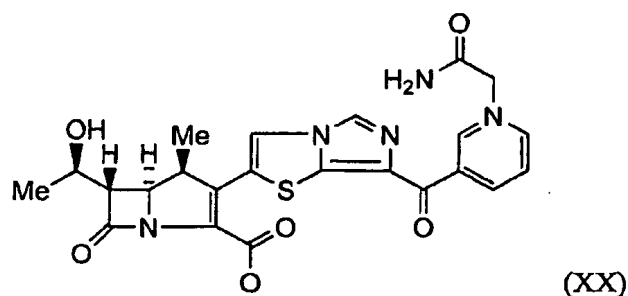
Asimismo, en un método para la síntesis del compuesto de la fórmula (XIX) o sus solvatos, hidratos, pro - medicamentos, y sales, un paso comprende el contactar o reaccionar el sustrato de la fórmula (I) con las cetorreductasas de la presentación bajo condiciones adecuadas de reacción para reducir o convertir el sustrato al producto de la fórmula (II). Ritipenem de la fórmula estructural (XIX) se describe en la patente de Estados Unidos número 4,482,565.

65

En algunas secciones, las cetorreductasas de la presentación pueden ser utilizadas en un método para la síntesis de un sulopenem de la fórmula estructural (XX):

5

10



(XX)

15

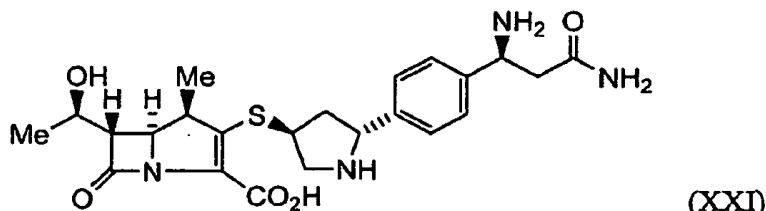
o sus solvatos, hidratos, pro - medicamentos y sales. Asimismo, en un método para la síntesis del compuesto de la fórmula (XX), un paso comprende contactar o reaccionar el sustrato de la fórmula (I) con las cetorreductasas de la presentación bajo condiciones de reacción adecuadas para reducir o convertir el sustrato al producto de la fórmula (II). El compuesto de la fórmula estructural (XX) se describe en EP1336612.

20

En algunas secciones, las cetorreductasas de la presentación pueden ser utilizadas en un método para la síntesis del compuesto de la fórmula estructural (XXI):

25

30



(XXI)

35

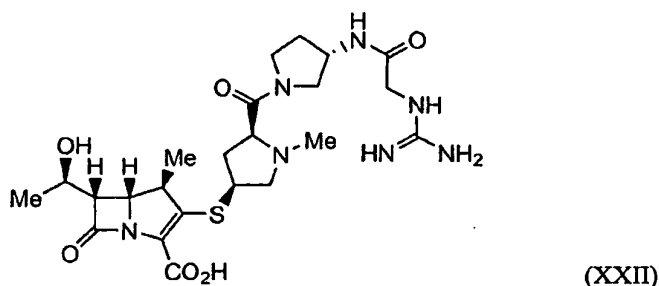
o sus solvatos, hidratos, pro - medicamentos y sales. Asimismo, en un método para la síntesis del compuesto de la fórmula (XXI), un paso en el método comprende contactar o reaccionar el sustrato de la fórmula (I) con las cetorreductasas de la presentación bajo condiciones de reacción apropiadas para la reducción o conversión del sustrato del producto de la fórmula (II). El compuesto de la fórmula estructural (XXI) se describe en JP2002114788.

40

En algunas secciones, las cetorreductasas de la presentación pueden ser utilizadas en un método para la síntesis de sulopenem de la fórmula estructural (XXII):

45

50



(XXII)

55

60

o sus solvatos, hidratos, pro - medicamentos y sales. Asimismo, en un método para la síntesis del compuesto de la fórmula (XXIII), un paso comprende el contactar o reaccionar el sustrato de la fórmula (I) con las cetorreductasas de la presentación bajo condiciones apropiadas de reacción para la reducción o conversión del sustrato al producto de la fórmula (II). El compuesto de la fórmula estructural (XXII) es descrito en la patente de Estados Unidos número 6,924,279 y la patente de Estados Unidos número 7,034,150.

65

En algunas secciones, las cetorreductasas de la presentación pueden estar presentes como composiciones, por ejemplo composiciones de reacción, con los sustratos activados por las cetorreductasas, y / o con los productos generados por la reacción de las cetorreductasas. Asimismo, en algunas secciones, una composición puede comprender una cetorreductasa selectiva 2S,3R de esta presentación, y un sustrato de la fórmula estructural (I). En algunas secciones, una composición puede comprender una cetorreductasa selectiva 2S,3R de esta

presentación, y un producto de la fórmula estructural (II). En algunas secciones, la composición puede comprender una cetorreductasa 2S,3R de la presentación, el sustrato de la fórmula (I) y el producto de la fórmula (II).

5 En algunas secciones, una composición puede comprender una cetorreductasa selectiva 2R,3R de esta presentación, y un sustrato de la fórmula estructural (I). En algunas secciones, una composición puede comprender una cetorreductasa selectiva 2R,3R de esta presentación, y un producto de la fórmula estructural (III). En algunas secciones, la composición puede comprender una cetorreductasa 2R,3R de la presentación, el sustrato de la fórmula (I) y el producto de la fórmula (III).

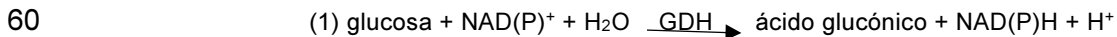
10 Puesto que las reacciones de cetorreductasa pueden ser ejecutadas en la presencia de un sistema regenerador de cofactores (NADH o NADPH), las condiciones de reacción pueden incluir además elementos de un sistema regenerador de cofactores, que se describe en mayor detalle a continuación. Asimismo, en algunas secciones, las composiciones mencionadas anteriormente de cetorreductasas pueden incluir además un sistema regenerador de cofactores que comprende deshidrogenasa de glucosa y glucosa; deshidrogenasa de formato y formato; o isopropanol y un alcohol secundario de deshidrogenasa. En algunas secciones, el alcohol secundario deshidrogenasa es una cetorreductasa diseñada. Otras enzimas y sustratos que pueden utilizarse para el reciclamiento de los cofactores serán bien conocidos por una persona con experiencia.

15 Tal como es conocido por personas con experiencia en la industria, las reacciones de reducción catalizadas de cetorreductasa comúnmente requieren un cofactor. Reacciones de reducción que son catalizadas por las enzimas de cetorreductasa diseñadas aquí descritas también requieren comúnmente un cofactor, aunque muchas secciones de las cetorreductasas diseñadas requieren mucho menos cofactores que las reacciones catalizadas con enzimas de cetorreductasas de tipo silvestre. Tal como se utiliza aquí, el término "cofactor" se refiere a un compuesto no proteínico que opera en combinación con una enzima cetorreductasa. Los cofactores adecuados para su uso con las enzimas de cetorreductasas diseñadas aquí descritas incluyen, pero no se limitan a, NADP⁺ (Nicotinamida adenina dinucleótido fosfato - Nicotinamide adenine dinucleotide phosphate), NADPH (la forma reducida de NADP⁺), NAD⁺ (Nicotinamida adenina dinucleótido) y NADH (la forma reducida de NAD⁺). Generalmente, la forma reducida del cofactor es agregada a la mezcla de reacción. La forma NAD(P)H reducida puede ser regenerada opcionalmente a partir de la forma NAD(P)⁺ oxidada usando un sistema de regeneración de cofactores.

20 El término "sistema de regeneración de cofactores" se refiere a un conjunto de reactivos que participan en una reacción que reduce la forma oxidada del cofactor (por ejemplo, NADP⁺ a NADPH). Los cofactores oxidados por medio de la reducción catalizada de la cetorreductasa del sustrato ceto se regeneran en una forma reducida por el sistema de regeneración de cofactores. Los sistemas de regeneración de cofactores comprenden un reductor estequiométrico que es una fuente para reducir los equivalentes del hidrógeno y es capaz de reducir la forma oxidada del cofactor. El sistema de regeneración de cofactores puede comprender además un catalizador, por ejemplo, un catalizador enzimático, que cataliza la reducción de la forma oxidada del cofactor por medio de un reductor. Los sistemas de regeneración de cofactores que sirven para regenerar NADH o NADPH a partir de NAD⁺ o NADP⁺, respectivamente, son conocidos en la industria y pueden ser utilizados en los métodos aquí descritos.

25 Ejemplos adecuados de sistemas de regeneración de cofactores que pueden ser utilizados incluyen, pero no se limitan a, glucosa y deshidrogenasa de glucosa, formato y deshidrogenasa de formato, glucos-6-fosfato y deshidrogenasa de glucos-6-fosfato, un alcohol secundario (por ejemplo, isopropanol) y deshidrogenasa de un alcohol secundario, fosfito y deshidrogenasa de fosfito, hidrógeno molecular e hidrogenase, y similares. Éstos sistemas pueden ser utilizados en combinación con cualquiera de NADP⁺ / NADPH o NAD⁺ / NADH como cofactores. La regeneración electroquímica utilizando hidrogenase también podría utilizarse como un sistema de regeneración de cofactores. Referirse, por ejemplo, a las patentes de Estados Unidos números 5,538,867 y 6,495,023. Los sistemas químicos de regeneración de cofactores que comprenden un catalizador de metal y un agente reductor (por ejemplo, hidrógeno molecular o formato) también son adecuados. Ver, por ejemplo, la publicación PCT WO 2000/053731.

30 Los términos "deshidrogenasa de glucosa" y "GDH" son utilizados intercambiamente en este documento para referirse a enzimas dependientes de NAD⁺ o NADP⁺ que catalizan la conversión de D- glucosa y NAD⁺ o NADP⁺ a ácido glucónico y NADH o NADPH, respectivamente. La ecuación (1) abajo, describe la reducción catalizada de deshidrogenasa de glucosa de NAD⁺ o NADP⁺ por la glucosa.



40 Las deshidrogenasas de glucosa que son adecuadas para su uso en la práctica de los métodos aquí descritos incluyen deshidrogenasas las de glucosa que ocurren naturalmente, así como las deshidrogenasas de glucosa que no ocurren naturalmente. Las deshidrogenasas de glucosa que ocurren naturalmente que codifican genes han sido reportadas en la literatura. Por ejemplo, el gen del bacilo subtilis 61297 GDH fue expresado en E. coli y se reportó que mostraba las mismas propiedades fisicoquímicas que las enzimas producidas en su huésped

nativo (Vasanth et al., 1983, Proc. Natl. Acad. Sci. EE.UU. 80:785). La secuencia genética del gen GDH de *B. subtilis*, que corresponde a la cuenta del Banco Genético número MI2276, fue reportada por Lampel et al., 1986, J. Bacteriol. 166:238-243, y en forma corregida por Yamane et al., 1996, Microbiology (Microbiología) 142:3047-3056 con la cuenta del Banco Genético número D50453. Genes GDH que ocurren naturalmente también incluyen aquellos que codifican el GDH a partir de *B. cereus* ATCC 14579 (Nature (Naturaleza), 2003, 423:87-91; cuenta del Banco Genético número AE017013) y *B. megaterium* (Eur. J. Biochem., 1988, 174:485-490, cuenta del Banco Genético número X12370; J. Ferment, Bioeng., 1990, 70:363-369, cuenta del Banco Genético número GI216270). Deshidrogenasa de glucosa del bacilo sp. son provistas en la publicación PCT WO 2005/018579 con las IDENTIFICACIONES SECUENCIALES NÚMEROS: 10 y 12 (codificadas por las secuencias de polinucleótidos correspondientes a las IDENTIFICACIONES SECUENCIALES NÚMEROS: 9 y 11, respectivamente, de la publicación PCT).

Las deshidrogenasas de glucosa que no ocurren naturalmente podrían ser generadas utilizando métodos conocidos, como por ejemplo, mutagénesis, evolución dirigida, y similares. Las enzimas GDH que tienen actividad apropiada, ya sea naturalmente o no naturalmente pueden ser identificadas fácilmente usando el ensayo descrito en el ejemplo 4 de la publicación PCT WO 2005/018579. Ejemplos de deshidrogenasa de glucosa que no ocurre naturalmente están indicadas en la publicación PCT WO 2005/018579 con las IDENTIFICACIONES SECUENCIALES NÚMEROS: 62, 64, 66, 68, 122, 124, y 126. La secuencia de polinucleótidos que las codifican están indicadas en la publicación PCT WO 2005/018579 con las IDENTIFICACIONES SECUENCIALES NÚMEROS: 61, 63, 65, 67, 121, 123, y 125, respectivamente. Deshidrogenasas de glucosas adicionales que no ocurren naturalmente que son adecuadas para el uso en las reacciones de reducción de cetorreductasas catalizadas aquí presentadas son suministradas en las publicaciones de aplicación de Estados Unidos números 2005/00952619 y 2005/0153417.

Las deshidrogenasas de glucosa utilizadas en las reacciones de reducción catalizada por cetorreductasa aquí descritas podrían exhibir una actividad de por lo menos 10 μmol / minutos / miligramos y a veces por lo menos alrededor de 10² micromoles / minutos / miligramos o alrededor de 10³ micromoles / minutos / miligramos, y hasta alrededor de 10⁴ micromoles / minutos / miligramos o más en el ensayo descrito en el ejemplo 4 de la publicación PCT WO 2005/018579.

Las reacciones de reducción catalizadas por las cetorreductasas aquí descritas son ejecutadas generalmente en un solvente. Solventes adecuados incluyen agua, solventes orgánicos (por ejemplo, acetato etílico, acetato butílico, 1-octanol, heptano, octano, éter t-butilo de metilo (MTBE), tolueno, y similares), líquidos iónicos (por ejemplo, 1 -etil 4-metilimidazolio tetrafluoroborato, 1-butil-3-metilimidazolio tetrafluoroborato, 1-butil-3-metilimidazolio hexafluorofosfato, y similares). En algunas secciones, solventes acuosos, incluyendo agua y sistemas co - solventes acuosos son utilizados.

Ejemplos de sistemas co-solventes acuosos tienen agua y uno o más solventes orgánicos. En general, un componente solvente orgánico de un sistema co-solvente acuoso es seleccionado de tal forma que no desactiva completamente a la enzima cetorreductasa. Sistemas co-solventes apropiados pueden ser identificados fácilmente al medir la actividad enzimática de la enzima diseñada específica de cetorreductasa con un sustrato definido de interés en el sistema solvente candidato, utilizando un ensayo de actividad enzimática, tal como aquellos descritos aquí.

El componente de solvente orgánico de un sistema co-solvente acuoso puede ser escalable con el componente acuoso, suministrando una sola fase líquida, o puede ser parcialmente escalable o inmiscible con el componente acuoso, suministrando dos fases líquidas. Generalmente, cuando un sistema co-solvente acuoso es utilizado, es seleccionado para ser bifásico, con agua dispersada en un solvente orgánico, o viceversa. Generalmente, cuando un sistema co-solvente acuoso es utilizado, es deseable seleccionar un solvente orgánico que pueda ser separado fácilmente de la fase acuosa. En general, la tasa de agua en comparación del solvente orgánico en el sistema co-solvente está típicamente en el rango de alrededor de 90:10 alrededor de 10:90 (volumen / volumen) de solvente orgánico a agua, y entre 80:20 y 20:80 (volumen / volumen) de solvente orgánico a agua. El sistema co-solvente puede ser formado previamente antes de su adición a la mezcla reactiva, o puede ser formado in situ en el portador de la reacción.

El solvente acuoso (agua o el sistema co-solvente acuoso) puede ser amortiguado o no amortiguado en lo que se refiere al pH. Generalmente, la reducción puede ser ejecutada a un pH de alrededor de 10 o menos, usualmente en el rango de desde alrededor de 5 a 10. En algunas secciones, las reducciones ejecutada con un pH de alrededor de 9 o menos, usualmente en el rango de alrededor de 5 a 9. En algunas secciones, la reducciones son ejecutadas con un pH de alrededor de 8 o menos, a menudo en el rango de alrededor de 5 a de 8, y usualmente en el rango de alrededor de 6 a 8. La reducción también puede ser ejecutada con un pH de alrededor de 7.8 o más bajo, o 7.5 o más bajo. Alternamente, la reducción puede ser ejecutada con un pH neutral, es decir, alrededor de 7.

5 Durante el transcurso de las reacciones de reducción, el pH de la mezcla de reducción puede cambiar. El pH de la mezcla de reacción puede mantenerse en el pH deseado o dentro del rango deseado de pH por medio de la adición de un ácido o una base durante el transcurso de la reacción. Alternamente, el pH puede ser controlado al usar un solvente acuoso que contiene un amortiguador. Amortiguadores adecuados para mantener los rangos de pH deseados son conocidos en la industria e incluyen, por ejemplo, tampón fosfato, amortiguador de trietanolamina, y similares. Combinaciones de amortiguadores, ácidos o bases también pueden usarse.

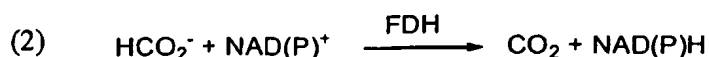
10 Cuando el sistema de regeneración de cofactores de glucosa / deshidrogenasa de glucosa se utiliza, la co-producción de ácido glucónico (pKa=3.6), tal como se presenta o en la ecuación (3) causa que el pH de la mezcla de reacción baje si el ácido glucónico acuoso resultante no es neutralizado de otra forma. El pH de la mezcla de reacción puede mantenerse al nivel deseado por medio de técnicas de amortiguamiento estándar, donde el amortiguador neutraliza el ácido glucónico hasta la capacidad de amortiguamiento suministrada, o por medio de la adición de una base concurrente en el transcurso de la conversión. Combinaciones de adición de amortiguadores y de bases también pueden utilizarse. Amortiguadores apropiados para mantener los rangos deseados de pH ya fueron descritos. Las bases adecuadas para la neutralización del ácido glucónico son bases orgánicas, por ejemplo aminos, alcóxidos, y similares, y bases inorgánicas, por ejemplo, sales de hidróxidos (por ejemplo, NaOH), sales carbonatadas (por ejemplo, NaHCO₃), sales de bicarbonatos (por ejemplo, K₂CO₃), sales básicas de fosfatos (por ejemplo, K₂HPO₄, Na₃PO₄), y similares. La adición de una base concurrente durante el transcurso de la conversión podría ser hecha manualmente mientras se monitorea el pH de la mezcla de reacción o, más convenientemente, usando un medidor automático como un medidor de pH. Una combinación de la capacidad de amortiguamiento parcial y la adición de bases también se puede usar para el control del proceso.

25 Cuando se utiliza la adición de bases para neutralizar al ácido glucónico durante la reacción de reducción catalizada por las cetorreductasas, el progreso de la conversión puede ser monitoreado por el monto de base añadida para mantener el pH. Típicamente, las bases añadidas a mezclas de reacción no amortiguada o parcialmente amortiguada durante el transcurso de la reducción son añadidas en las soluciones acuosas.

30 En algunas secciones, el sistema de regeneración de cofactores puede comprender una deshidrogenasa de formato. Los términos "deshidrogenasa de formato" y "FDH" son utilizados intercambiablemente que en este documento para referirse a una enzima dependiente de NAD⁺ o NADP⁺ que cataliza la conversión del formato y el NAD⁺ o NADP⁺ a dióxido de carbono y NADH o NADPH, respectivamente. Las deshidrogenasas de formato que son adecuadas para su uso como sistemas regeneradores de cofactores en las reacciones de reducción catalizadas por medio de cetorreductasas aquí descritas incluyen deshidrogenasas de formato que ocurren naturalmente, así como deshidrogenasas de formato que no ocurren naturalmente. Las deshidrogenasa de formato incluyen aquellas correspondientes a las IDENTIFICACIONES SECUENCIALES NÚMEROS: 70 (*Pseudomonas* sp) y 72 (*Candida boidini*) de la publicación PCT WO 2005/018579, que están codificadas por las secuencias de polinucleótidos correspondientes a las IDENTIFICACIONES SECUENCIALES NÚMEROS: 69 y 71, respectivamente, de la publicación PCT 2005/018579. Las deshidrogenasa de formato utilizadas en los métodos aquí descritos, ya sean de ocurrencia natural o de ocurrencia no natural, podrían exhibir una actividad de por lo menos alrededor de 1 μmol / minutos / miligramos, a veces por lo menos alrededor de 10 μmol / minutos / miligramos, o por lo menos alrededor de 10² micromoles / minutos / miligramos, hasta alrededor de 10³ micromoles / minutos / miligramos o más altos, y pueden ser examinados fácilmente en lo que se refiere a actividades en el ensayo descrito en el ejemplo 4 de la publicación PCT WO 2005/018579.

45 Tal como se utiliza aquí, el término "formato" se refiere a una anión de formato (HCO₂⁻), ácido fórmico (HCO₂H), y sus mezclas. El formato puede ser suministrado en la forma de una sal, típicamente una sal alcali o de amoníaco (por ejemplo, HCO₂Na, KHCO₂NH₄, y similares), en la forma de ácido fórmico, típicamente ácido fórmico acuoso, o sus mezclas. El ácido fórmico es un ácido moderado. En soluciones acuosas con varias unidades de pH de su pKa (pKa = 3.7 en agua), el formato está presente como HCO₂⁻ y HCO₂H en concentraciones en equilibrio. Los valores de pH superiores a alrededor de pH 4, el formato está presente en una forma predominante como HCO₂⁻. Cuando el formato es suministrado como un ácido fórmico, la mezcla de reacción es amortiguada comúnmente o se la hace menos ácida al añadir una base para suministrar el pH deseado, que comúnmente es alrededor de pH 5 o más. Bases adecuadas para la neutralización del ácido fórmico incluyen, pero no se limitan a, bases orgánicas, por ejemplo aminos, alcóxidos y similares, y bases inorgánicas, por ejemplo, sales de hidróxido (por ejemplo, NaOH), sales de carbonato (por ejemplo, NaHCO₃), sales de bicarbonato (por ejemplo, K₂CO₃), sales básicas de fosfato (por ejemplo, K₂HPO₄, Na₃PO₄), y similares.

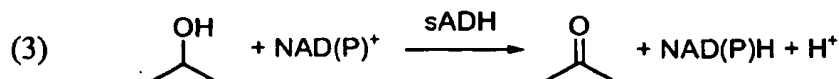
60 Para los valores pH superiores a pH 5, al cual el formato está presente de una forma predominante como HCO₂⁻, la ecuación (2) a continuación, describe la reducción catalizada por deshidrogenasa de formato de NAD⁺ o NADP⁺ por el formato.



5 Cuando se utilizan formatos o deshidrogenasa de formatos como el sistema de regeneración de cofactores, el pH de la mezcla de reacción puede ser mantenido en los niveles deseados por medio de técnicas de amortiguamiento estándar, donde el amortiguamiento libera protones hasta su capacidad de amortiguamiento suministrada, o por medio de adición de un ácido concurrente durante el transcurso de la conversión. Ácidos adecuados para ser añadidos durante el transcurso 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 hidrohálidos (tales como el ácido hidrocórico), ácido sulfúrico, ácido fosfórico, y similares, sales ácidas, por ejemplo, sales de dedihidrogenofosfato (por ejemplo, KH_2PO_4), sales de bisulfato (por ejemplo, NaHSO_4) y similares. Algunas secciones utilizan ácido fórmico, con el cual la concentración de formato y el pH de la solución son mantenidos.

10 Cuando se utiliza una adición de ácido para mantener el pH durante la reacción de reducción utilizando el sistema de regeneración de cofactores de formato / deshidrogenasa de formato, el progreso de la conversión puede monitorearse por el monto de ácido añadido para mantener el pH. Típicamente, ácidos añadidos a mezclas de reacción no amortiguadas o parcialmente amortiguadas durante el transcurso de la conversión son añadidos en forma de soluciones acuosas.

15 Los términos “deshidrogenasa de alcohol secundario” y “sADH” (secondary alcohol dehydrogenasa) son utilizados intercambiamente en este 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. La ecuación (3), a continuación, describe la reducción de NAD^+ o NADP^+ por un alcohol secundario, ilustrado por isopropanol.



30 Las deshidrogenasas de alcohol secundario que son apropiadas para su uso como sistemas regeneradores de cofactores en las reacciones de reducción catalizadas por la cetorreductasa aquí indicadas incluyen deshidrogenasas de alcohol secundario que ocurren naturalmente, así como deshidrogenasas de alcohol secundario que no ocurren naturalmente. Las deshidrogenasas de alcohol secundario que ocurren naturalmente incluyen deshidrogenasas de alcohol conocidas que provienen de, *Thermoanaerobium brockii*, *Rhodococcus erythropolis*, *Lactobacillus kefir*, y *Lactobacillus brevis*, y deshidrogenasas de alcohol secundario que no ocurren naturalmente incluyen a sus deshidrogenasas de alcohol diseñadas derivadas. Deshidrogenasas de alcohol secundario utilizadas en los métodos aquí descritos, ya sean de ocurrencia natural o no natural, podrían exhibir una actividad de por lo menos 1 μmol / minuto / miligramos, a veces por lo menos alrededor de 10 μmol / minutos / miligramos, o por lo menos alrededor de 10^2 μmol / minutos / miligramos, o hasta 10^3 μmol / minutos / miligramos o niveles más altos.

40 Alcoholes secundarios apropiados incluyen alcanoles y aril-alquil carbinoles secundarios inferiores. 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 una sección, el alcohol secundario es isopropanol, aril-alquil carbinoles incluyen 1-ariletanoles sustituidos y no sustituidos.

45 Cuando se utiliza un alcohol secundario o una deshidrogenasa de alcohol secundario como el sistema de regeneración de cofactores, el NAD^+ o NADP^+ resultante es reducido por medio de la oxidación emparejada del alcohol secundario a cetona por parte de la deshidrogenasa de alcohol secundario. Algunas cetorreductasas también tienen la actividad de deshidrogenar un reductor de alcohol secundario. En algunas secciones el usar un alcohol secundario como reductor, la cetorreductasa diseñada y la deshidrogenasa de alcohol son la misma enzima.

50 Al ejecutar secciones de las reacciones de reducción catalizadas por la cetorreductasas aquí descritas utilizando un sistema de regeneración de cofactores, podría suministrarse inicialmente la forma oxidada o reducida del cofactor. Tal como se describió anteriormente, el sistema de regeneración de cofactores convierte al cofactor oxidado a su forma reducida, la cual se utiliza entonces en la reducción del sustrato de cetorreductasa.

55 En algunas secciones, los sistemas de regeneración de cofactores no son utilizados. Para reacciones de reducción a ser ejecutadas sin el uso de un sistema de regeneración de cofactores, el cofactor es añadido a la mezcla de la reacción en una forma reducida.

60 En algunas secciones, cuando el proceso es ejecutado usando células completas del organismo anfitrión, la célula completa podría por naturaleza suministrar el cofactor. Alternamente o en combinación, la célula podría suministrar por naturaleza o recombinantemente la deshidrogenasa de glucosa.

65 Al ejecutar las reacciones de reducción estereoselectivas aquí descritas, la enzima de cetorreductasa diseñada, y cualquier enzima que conforme el sistema opcional de regeneración de cofactores, podría añadirse a la mezcla de la reacción en la forma de enzimas purificadas, células completas transformadas con genes que codifican las enzimas,

5 y / o extractos celulares y / o lisados de aquellas células. Los genes que codifican la enzima cetorreductasa diseñada y las enzimas opcionales de regeneración de los cofactores pueden ser transformadas a células anfitrionas por separado o juntas en la misma célula anfitriona. Por ejemplo, en algunas secciones un conjunto de células anfitrionas puede transformarse con genes que codifican a la enzima cetorreductasa diseñada y otro conjunto puede ser transformado con genes que codifican las enzimas de regeneración de cofactores. Ambos conjuntos de células transformadas pueden ser utilizados juntos en la mezcla de la reacción en la forma de células completas, o en la forma de lisados o extractos derivados de estas células. En otras secciones, las células anfitrionas pueden transformarse con genes que codifican las enzimas de cetorreductasa diseñadas y las enzimas regeneradoras de cofactores.

10 Células completas transformadas con genes que codifican las enzimas de cetorreductasa diseñadas y / o las enzimas opcionales regeneradoras de cofactores, o sus extractos celulares y / o lisados, pueden utilizarse en una variedad de formas diferentes, incluyendo sólidas (un ejemplo, liofilizadas, secas por medio de spray, y similares) o semisólidas (por ejemplo, una pasta cruda).

15 Los extractos celulares o lisados celulares podrían ser parcialmente purificados por medio de precipitación (sulfato de amoníaco, polietileimina, tratamiento del corazón o similares, seguido por un procedimiento de desalación previo a la liofilización (por ejemplo, ultrafiltración, diálisis y similares). Cualquiera de las reparaciones celulares puede ser estabilizada al vincular transversalmente usando agentes de enlace transversal, tales como, por ejemplo, glutaraldehído o la inmovilización a una fase sólida (por ejemplo, Euoergit C y similares).

20 Los reactivos sólidos (por ejemplo, enzimas, sales, etcétera) pueden ser suministrados a la reacción en varias formas, incluyendo polvos (por ejemplo, liofilizados, secados por spray, y similares), soluciones, emulsiones, suspensiones y similares. Los reactivos pueden ser liofilizados o secarse con spray fácilmente utilizando métodos y equipos que son conocidos para aquellas personas con experiencia en la industria. Por ejemplo, la solución proteínica puede ser congelada a -80 °C en pequeñas partes, luego agregada a una cámara de liofilización previamente enfriada, seguido por la aplicación de un vacío. Después de la remoción de agua de las muestras, la temperatura es elevada típicamente a 4 °C durante dos horas antes de liberarse al vacío y recuperarse las muestras liofilizadas.

25 Las cantidades de reacciones utilizadas en la reacción de reducción variarán generalmente dependiendo en las cantidades del producto deseado, y concurrentemente el monto de sustrato de cetorreductasa utilizado. Las siguientes guías pueden ser utilizadas para determinar los montos de cetorreductasa, cofactores y sistemas opcionales de regeneración de cofactores a ser utilizados. Generalmente, los sustratos ceto pueden ser utilizados con una concentración de alrededor de 20 a 300 g / litro usando desde alrededor de 50 mg a 5 g de cetorreductasa y alrededor de 10 mg a 150 mg de cofactores. Aquellos que tienen conocimiento en la industria entenderán fácilmente como variar estas cantidades para personalizarlas al nivel deseado de productividad y escala de producción. Las cantidades apropiadas del sistema opcional de regeneración de cofactores puede determinarse fácilmente por medio de experimentación rutinaria que se basa en el monto de cofactores y / o cetorreductasa utilizados. En general, el reductor (por ejemplo, glucosa, formato, isopropanol) es utilizado a niveles que superan el nivel equimolar de sustrato de cetorreductasa para lograr esencialmente una conversión completa o casi completa del sustrato de cetorreductasa.

30 La orden de adición de los reactivos no es crítica. Los reactivos pueden ser agregados juntos al mismo tiempo a un solvente (por ejemplo, un solvente monofásico, un sistema co - solvente de acuoso bifásico, y similares), o alternamente, algunos de los reactivos pueden ser agregados por separado, y algunos en diferentes puntos de tiempo. Por ejemplo, el sistema de regeneración de cofactores, cofactor, cetorreductasa y sustrato de cetorreductasa pueden ser agregados primero al solvente.

35 Para una eficiencia mejorada de mezcla cuando un sistema co-solvente acuoso es utilizado, el sistema de regeneración de cofactores, la cetorreductasa, y el cofactor pueden ser agregados y mezclados primeros en la fase acuosa. La fase orgánica puede ser agregada y mezclada, seguido por una adición del sustrato de cetorreductasa. Alternamente, el sustrato de cetorreductasa puede ser mezclado previamente en la fase orgánica, antes de su adición a la fase acuosa.

40 Condiciones apropiadas para ejecutar las reacciones de reducción catalizadas por la cetorreductasa aquí descritas incluyen una amplia variedad de condiciones que suelen ser optimizadas fácilmente por experimentación rutinaria que incluye, pero no se limita a, contactar la enzima de cetorreductasa diseñada y el sustrato en un pH, temperatura y producto detector experimental, por ejemplo, usando los métodos aquí descritos.

45 La reducción catalizada por cetorreductasa es ejecutada típicamente a una temperatura en el rango de alrededor de 15 °C a 75 °C. Para algunas secciones, la reacción es ejecutada a una temperatura en el rango de alrededor de 20 °C a 55 °C. En otras secciones, es ejecutada a una temperatura en el rango de alrededor de 20 °C a 45 °C. La reacción también puede ser ejecutada a las condiciones ambientales.

La reacción de reducción se le permite proceder generalmente hasta que está esencialmente completa, o cerca de estar completa, y la reducción del sustrato ha sido obtenida. La reducción del sustrato al producto puede monitorearse usando métodos conocidos al detectar el sustrato y / o el producto. Métodos apropiados incluyen cromatografía de gases, HPLC y similares. Las producciones de conversión del producto por medio de reducción de alcohol generado en la mezcla de reacción son generalmente superiores que alrededor del 50%, pueden ser superiores que alrededor del 60%, también pueden ser superiores que alrededor de 70%, pueden ser superiores que alrededor del 80%, pueden ser superiores que alrededor de 90% y a menudo son superiores que alrededor de 97%.

EJEMPLOS

[0301] Varias características y secciones de la presentación son ilustradas en los siguientes ejemplos representativos, que tienen la intención de ser ilustrativos y no limitantes.

Ejemplo 1: La producción de polvos de cetorreductasa; procedimiento de agitación de matraces

Una sola colonia microbiana de *E. coli* que contiene un plásmido con el gen de cetorreductasa de interés fue inoculada dentro de 50 ml de caldo triptico (12 g / litro becto-triptona, 24 g / litros de extracto de levadura, 4 ml / litros de glicerol, 65mM de fosfato de potasio, pH 7.0) que contiene 30 µg / mililitros de cloranfenicol y 1% de glucosa en un matraz de 250 ml Erlenmeyer. Las células fueron cultivadas durante la noche (por lo menos 16 horas) en una incubadora a 30 °C agitándose a 250 revoluciones por minuto. El cultivo fue diluido en 250 ml de caldo Terrific que contiene 1mM MgSO₄, 30 µg / mililitros de cloranfenicol en un matraz de 1 l a una densidad óptica de 600 nm (OD600) de 0.2 y se le permite crecer a 30 °C. La expresión del gen de cetorreductasa fue inducida con 1mM IPTG cuando el OD600 del cultivo es 0.6 a 0.8 e incubado durante la noche (por lo menos 16 horas). Las células fueron cultivadas por medio de centrifugación (500 revoluciones por minuto, 15 minutos, 4 °C) y el material flotante fue descartado. El pellet celular fue re - suspendido con un volumen igual frío (4 °C) 100mM de amortiguador de trietanolamina (cloruro), pH 7.0 (incluyendo 1mM MgSO₄ en el caso de ADH-LK y ADH-LB y sus cetorreductasa diseñadas co - cultivadas por medio de centrifugación como se mencionó anteriormente. Las células lavadas fueron re - suspendidas en 12 ml del amortiguador de trietanolamina fría (cloruro) y pasadas a través de una prensa francesa dos veces a 120,000 psi mientras se mantiene a 4 °C. Los vestigios celulares fueron removidos por medio de centrifugación (9000 revoluciones por minuto, 45 minutos, 4 °C). Los vestigios despejados del lisado fueron recaudados y almacenados a -20 °C. La liofilización del lisado despejado congelado suministrado un polvo seco de enzimas crudas de cetorreductasa.

Ejemplo 2: Producción de Cetorreductasas; Procedimiento de Fermentación

En un material celular agitado 15 l de fermentador, 6.0 litros de medio de crecimiento con 0.88 gramos / litros de sulfato de amonio, 0.98 gramos / litros de citrato de sodio; 12.5 gramos / litros de trihidrato de fosfato de hidrógeno de dipotasio, 6.25 gramos / litro de fosfato de dihidrógeno de potasio, 6.2 gramos / litro de Tastona de extracto 154 de levadura, 0.083 gramos / litros de citrato de amonio férrico, y 8.3 mililitros / litro de una solución de elemento de rastreo que contiene 2 g / litro de dihidrato de cloruro de calcio, 2.2 gramos / litro de septahidrato de sulfato de zinc, 0.5 gramos / litro de monohidrato de sulfato de manganeso, 1 g / litro de heptahidrato de sulfato cuproso, 0.1 g / litro de tetrahidrato molibdato de amonio y 0.02 gramos / litro de decahidrato de tetraborato de sodio a una temperatura de 30 °C. El fermentador fue inoculado con el cultivo exponencial reciente de *E. coli* W3110 que contenía un plásmido con el gen de cetorreductasa de interés, cultivado en un frasco de agitación tal como se describió en el ejemplo 3 a un inicio de OD600 de 0.5 a 2.0. El fermentador fue agitado a 500 - 1500 revoluciones por minuto y se suplementó aire al contenedor de la fermentación a 1.0 - 15.0 litros / minuto para mantener disuelto el nivel de oxígeno a un 30% de saturación o más. El pH del cultivo fue controlado a 7.0 por medio de la adición de un 20% volumen / volumen de hidróxido de amonio. El crecimiento del cultivo fue mantenido por medio de la adición de una solución de alimentación que contenía 500 g / litro de cerelosa, 12 g / litro de cloruro de amonio y 10.4 gramos / litro de heptahidrato de sulfato de magnesio. Después de que el cultivo alcanzó un OD600 de 50, la expresión de cetorreductasa fue inducida por medio de la adición de isopropil-β-D-tiogalactósido (IPTG) a una concentración final de 1mM. Se continuó con el cultivo durante 14 horas. El cultivo fue enfriado a 4 °C y mantenido a 4 °C hasta ser cosechado. Las células cosechadas por medio de centrifugación a 5000G durante 40 minutos en un centrifugador Sval RC12BP a 4 °C. Las células cosechadas fueron usadas inmediatamente en el siguiente proceso de recuperación o fueron almacenadas a 4 °C hasta su utilización.

El pellet celular fue re - suspendido en dos volúmenes de 100 mM de amortiguador trietanolamina (cloruro), pH 6.8, a 4 °C a cada volumen de pasta celular húmeda. La cetorreductasa intracelular fue liberada de las células al pasar la suspensión a través de un homogeneizador acoplado con un ensamblaje de válvulas homogeneizadoras de dos etapas usando una presión de 12,000 psig. La homogenización celular fue enfriada a 4 °C inmediatamente después de la interrupción. Una solución de 10% masa / volumen de polietileneimina, pH 7.2, fue agregada a lisado a una concentración final de 0.5 por ciento masa / volumen y agitado durante 30 minutos. La suspensión resultante fue clarificada por medio de centrifugación a 5000G en una centrifugación de laboratorio estándar durante 30 minutos. El material flotante fue decantado y concentrado 10 veces usando una membrana de ultrafiltración celulosa con un corte de masa molecular de 30 Kd. La concentración final fue dispensada en contenedores hondos, congelados a -20 °C y liofilizados en polvos. Los polvos de cetorreductasa fueron almacenados a -80 °C.

Ejemplo 3: métodos analíticos para la detección de estereoisómeros de metil-2-benzamidometil-3-hidroxitirato

Para análisis de rutina, los dos estereoisómeros de metil-2-benzamidometil-3-oxobutirato y 4 estereoisómeros de metil-2-benzamidometil-3-hidroxitirato fueron separados por medio de HPLC quiral de fase normal en una columna Chiralpak IA (4.6 x 150 mm (Chiral technologies, número de catálogo 80324); isocrática (92% de heptano, 8% de etanol); 40 °C; 1.5 mililitros / minutos de caudal; volumen de la muestra: 10 µl; detección: absorción UV 254 nm). Tiempos de retención: 2S,3R- estereoisómero: 6.8 minutos, 2R,3S-diastereómero: 8.1 min, 2R,3R-estereoisómero. 9.3 minutos, 2S,3S-diastereómero: 10.1 minutos, sustrato de estereoisómeros: 11.6 y 13.5 minutos.

Ejemplo 4: Examinación previa de fluorescencia de NADPH de alto caudal para identificar variaciones activas en el alcohol isopropílico (IPA – Isopropyl alcohol)

Las bibliotecas de plásmidos obtenidas por medio de evolución dirigida y que contenían genes evolucionados de cetorreductasa fueron transformadas a E. coli y puestas en platos en un caldo Luria-Bertani (LB) que contenía 1% de glucosa y 30 µg / mililitros de cloranfenicol (CAM). Después de la incubación durante 16 horas a 30 °C, las colonias fueron levantadas usando un recogedor de colonias robótico Q-bot® (GEnetix USA, Inc., Beaverton, OR) en platos de microtitulación de pozos profundos de 96 pozos que contienen 180 µl de Luria-Bertani (LB), 1% de glucosa, 30 µg / mililitros de cloranfenicol (CAM), y 2mM de MgSO₄. Las células fueron cultivadas durante la noche a 30 °C agitándolas a 200 revoluciones por minuto. 20 µl de este cultivo fueron transferidos a platos de 96 pozos profundos que contenían 380 µl de caldo Terrific (TB), 2mM de MgSO₄ y 30 µg / mililitros de CAM. Después de la incubación de los platos en pozos profundos a 30 °C agitándolos a 250 revoluciones por minuto durante 2.5 a 3 horas (OD⁶⁰⁰ 0.6 - 0.8), se indujo la expresión genética recombinante de los cultivos celulares por medio de tiogalactósido de isopropilo (IPTG) a una concentración final de 1mM. Los platos fueron incubados a 30 °C y agitado a 250 revoluciones por minuto durante 15 - 23 horas.

Las células fueron granuladas por medio de centrifugación, re - suspendidas en amortiguadores de lisis de 400 µl y lisadas agitándolas a la temperatura del cuarto durante una hora. El amortiguador de lisis contenía 100mM de amortiguador de trietanolamina (sulfato), pH 7.0 - 7.2, 1 mg / mililitro de lisozima y 200 µg / mililitro de sulfato de poliximina B y 2mM de MgSO₄. Los platos fueron girados en la centrifugación a 4000 revoluciones por minuto durante 10 minutos a 4 °C y el material flotante que se despejó (el material flotante despejado) fue utilizado en el ensayo fluorescente.

En platos de microtitulación negros de 96 pozos se agregó 20 µl de lisado (una dilución 10 veces en 100 mM de amortiguador de trietanolamina (sulfato) pH 7.0) a 180 µl de una mezcla de un ensayo que consistía de 100 mM de amortiguador de trietanolamina (sulfato) pH 7.0, 1mM MgSO₄, 1 g / L de NADP, y 50% de isopropilalcohol y el progreso de la reacción fue medido siguiendo el incremento en fluorescencia de NADPH a 445 nm después de su proceso de excitación a 330 nm en Flexstation (Molecular Devices, USA).

Ejemplo 5: ensayo de HPLC de alto caudal para actividades de la cetorreductasa en metil-2-benzamidometil-3-oxobutirato usando isopropilalcohol para el reciclamiento de cofactores.

Los lisados fueron preparados tal como se describió en el Ejemplo 4. La actividad de cetorreductasa fue medida al transferir cantidades medidas de los lisados celulares a los pozos de los platos con pozos profundos de Costar (número de catálogo 3961). A 50 µl de 3 mg / mililitros de Na-NADP en 100 mM de amortiguador de trietanolamina (sulfato) pH 7.0, se agregaron 150 µl de 3.3 miligramos / mililitros de metil-2-benzamidometil-3-oxobutirato en IPA, 100 µl de lisado despejado. Los platos fueron sellados y calentados con cinta adhesiva laminada de sellado al calor de aluminio / polipropileno (Velocity 1) (Menlo Park, CA), número de catálogo 06643-001), incubado a la temperatura del cuarto durante 24 horas agitándose. Al final de la reacción, se agregaron 990 µl de MTBE a cada pozo, los platos fueron sellados nuevamente y agitados durante 20 minutos. La fase orgánica fue separada de la fase acuosa por medio de centrifugación (4000 revoluciones por minuto, 5 minutos, 4 °C) y 200 µl de la capa orgánica fue transferida a un nuevo plato de 96 pozos hondos para ser analizada utilizando los métodos descritos en el Ejemplo 3.

Ejemplo 6: Ensayo HPLC de alto caudal para la actividad de la cetorreductasa en metil-2-benzamidometil-3-oxobutirato usando isopropilalcohol para reciclamiento de cofactores a una concentración más alta del sustrato.

Pozos de un plato de microtitulación de 96 pozos se agregó 10 µl de 0.6 miligramos / mililitros de Na-NADP en 100 mM de amortiguador de trietanolamina (sulfato) pH 7.0, 150 µl de 100 mg / mililitros de metil-2-benzamidometil-3-oxobutirato en IPA, 10 µl de lisado despejado y 130 µl de 100mM TEA. Los platos fueron sellados, calentados e incubados a la temperatura del cuarto durante 24 horas agitándose. Al final de la reacción, se agregó a cada pozo 990 µl de MTBE, los platos fueron calentados nuevamente y agitados durante 20 minutos. La fase orgánica fue separada de la fase acuosa por medio de centrifugación (4000 revoluciones por minuto, 5 minutos, 4 °C) y 200 µl de la capa orgánica fueron transferidos a un nuevo plato de 96 pozos hondos usando los métodos descritos en el Ejemplo 3.

Este ejemplo describe el método que fue utilizado para identificar variantes KRED para la reducción estereoselectiva del metil-2-benzamidometil-3-oxobutirato.

Ejemplo 7: Evaluación de variantes ADH-LK para la reducción metil-2-benzamidometil-3-oxobutirato

Algunas variantes ADH-LK fueron evaluadas para la reducción estereoselectiva de metil-2-benzamidometil-3-oxobutirato tal como se describió en el ejemplo 5. Las muestras fueron analizadas tal como se describió en el ejemplo 3..

La tabla 4 lista los NÚMEROS DE IDENTIFICACIONES SECUENCIALES correspondientes a las cetorreductasas, el número de mutaciones de aminoácidos provenientes de ADH-LK, la conversión porcentual y el porcentaje del (2S,3R)- estereoisómero deseado.

TABLA CUATRO:

TABLA 4			
NÚMERO DE IDENTIFICACIÓN SECUENCIAL	Número de mutaciones provenientes de ADH-LK	Actividad ^A	stereoselectividad ^B
90	1	++	+
92	4	++	+
96	1	+	+
98	1	++	+

A: +: >30% de conversión, ++: >60% de conversión; B +: >95% de isómero (2S,3R).

La variante ADH-LK con la IDENTIFICACIÓN SECUENCIAL NÚMERO: 90 fue utilizada para obtener una mayor evolución en un KRED eficiente para la producción de (2S,3R)-metil-2-benzamidometil-3-hidroxitirato. Similarmente, una variante ADH-LK con IDENTIFICACIÓN SECUENCIAL NÚMERO: 94 que contenía 10 mutaciones en comparación a ADH-LK que dio principalmente el estereoisómero (2R,3R), fue utilizado para una mayor evolución de un KRED eficiente para la producción de (2R,3R)-metil-2-benzamidometil-3-hidroxitirato.

Este ejemplo muestra que las variantes de ADH-LK que pueden ser útiles para la reducción estereoselectiva del metil-2-benzamidometil-3-oxobutirato. Nuevas variantes de las IDENTIFICACIONES SECUENCIALES NÚMEROS: 90 y 94 fueron generadas al removerse un sitio interno Bgll. Las nuevas secuencias correspondientes son las IDENTIFICACIONES SECUENCIALES NÚMEROS: 48 y 66 respectivamente

Ejemplo 8: La reducción del metil-2-benzamidometil-3-oxobutirato a (2S,3R)-metil-2-benzamidometil-3-hidroxitirato por medio de cetorreductasas diseñadas derivadas de ADH-LK.

Las cetorreductasas mejoradas derivadas de la variante de ADH-LK con la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48 fueron evaluadas en referencia a su actividad incrementada y estereoselectividad (2S,3R) por medio del método descrito en el Ejemplo 6.

5

Tabla 5: Actividad de variantes ADH-LK para la reducción de Metil-2-benzamidometil-3-oxobutirato a (2S,3R)-metil-2-benzamidometil-3-hidroxitirato

10	Número de identificación secuencial	Mutaciones a partir de kéfir	# Mutaciones provenientes de ADH LK (IDENTIFICACIÓN SECUENCIAL: 4)	Actividad ^a	Estabilidad ^b	Selectividad ^c
15	2	brevis				
	4	kéfir				
20	48	A202V;	1	+		+
	38	A94T; E105G; L153A; L199A; A202L; M206F;	6	+		+
25	16	A94T; S96F; A202V;	3	+		+++
	56	L153A; L199A; A202L;	3	+		+++
30	58	T861; L199N; A202L;	3	+		+++
	52	L153A; A202L;	2	+		+++
35	54	L153A; A202V;	2	+		+
	32	A94T; L199A; A202V;	3	++		++++
40	34	A94T; L153A; L199H; A202L;	4	+++		++++
	50	L153A; L199H; A202L;	3	++		+++
45	20	A94T; L199N; A202V;	3	+++		++++
	46	L153S; A202L;	2	+		+
50	36	A94T; L153A; L199A; A202V;	4	+		++
	26	A94T; A202L;	2	+++		+++
55	28	A94T; A202V;	2	++		+++
	30	A94T; L199A; A202L;	3	+++		++++
60	22	1 A94T; L199H; A202L;	3	++++		++++

65

5
10
15
20
25
30
35
40
45
50
55
60
65

Tabla 5: Actividad de variantes ADH-LK para la reducción de Metil-2-benzamidometil-3-oxobutirato a (2S,3R)-metil-2-benzamidometil-3-hidroxitirato					
Número de identificación secuencial	Mutaciones a partir de kéfir	# Mutaciones provenientes de ADH LK (IDENTIFICACIÓN SECUENCIAL: 4)	Actividad ^a	Estabilidad ^b	Selectividad ^c
24	A94T; L199H; A202V;	3	+++		++++
42	L153A; L199N; A202L;	3	+		+++
40	A94T; S96F; M129T; A202V; M206F;	5	+		++
18	A80T; L153A; A202V;	3	+		+++
44	F147M; A202V;	2	+	+	
10	H4OR; A94T; F147L; L199H; A202L;	5	+++++	+	++++
12	H4OR; A94T; L199H; A202L;	4	+++++		++++
6	A94T; F147L; L199H; A202L;	4	+++++	+	++++
8	A94T; L199H; A202L;	3	+++++		++++
60	I I I F; H4OR; A94F; S96V; F147M; L195V; V196L; L199W; 1226V; G248K; Y249W;	11	++++		++++
62	T2A; R4C; H4OR; A94G; S96V; F147M; V196L; L199W; 1226V; G248K; Y249W;	11	+++		++++
14	H4OR ; A94F; S96V; F147M; L195V; V196L; L199W; 1226V; Y249W;	9	++++		++++

+: 1 - 15 veces más actividad que KRED con la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48; ++: 15 - 30 veces más activo que KRED con la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48; +++: 30 - 40 veces más activo que KRED con la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48; ++++: 40 - 50 veces más activo que KRED con la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48; +++++: >50 veces más activo que KRED con la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48,

^b +: retiene la actividad después de 21 horas de pre - incubación a 40 °C,

^c +: 60 - 89 % del producto (2S,3R); ++: 90 - 94% del producto (2S,3R); +++: 95 - 99% del producto (2S,3R); +++++: >99% del producto (2S,3R).

Ejemplo 9: Reducción de Metil-2-benzamidometil-3-oxobutirato a (2R, 3R)-metil-2-benzamidometil-3-hidroxitirato por Derivados de Cetorreductasas Diseñadas de ADH-LK.

- 5 Las cetorreductasas mejoradas que provienen de la variante ADH-LK con la IDENTIFICACIÓN SECUENCIAL NÚMERO: 66 fueron evaluadas en lo referente a una actividad incrementada y a la estereoselectividad (2R,3R) por medio del método descrito en el Ejemplo 6.

10

Tabla 6: Actividad de variantes ADH-LK para la reducción de Metil-2-benzamidometil-3-oxobutirato a (2R,3R)-metil-2-benzamidometil-3-hidroxitirato

15

20

25

30

35

40

45

50

55

60

65

Número de identificación secuencial	Secuencia - mutaciones de codificación	Número de mutaciones a partir de kéfir	Actividad ^a	Selectividad ^b
66	H4OR; A94G; S96V; E145F; F147M; Y190P; V196L; L199W; I225V; Y249W	10	+	+
74	II IF; H4OR; A94E; S96V; E145F; F147M; Y190P; L195V; V196L; L199W; I226V; Y249H;	12	++	++
82	D3V; AI OT; H4OR; A94G; S96V; F147M; Y190P; V196L; L199W; I226V; G248K; Y249H;	12	+	++
68	H4OR; A94F; S96V; E145F; F147M; Y190P; L195V; V196L; L199W; I226V; G248R; Y249W;	12	++	++
76	I1 1L; H4OR; A94E; S96V; F147M; Y190H; V196L; I226V; G248K; Y249H;	10	+++	++
72	H4OR; T54A; A94F; S96V; E105K; E145D; F147M; V196L; L199W; I226V; Y249W;	11	+++	++
78	11 IF; H4OR; A94G; S96V; E145F; F147M; Y190H; L195V; V196L; L199W; A202V; I226V; Y249H; A251T;	14	+++	++
70	H4OR; E78D; A94E; S96V; F147M; Y190H; L195V; V196L; I226V; Y249H; T250Y;	11	+++	+
80	K8N; V9G; 111F; H4OR; A94G; S96V; E145F; F147M; Y190P; V196L; I226V; G248K; Y249R;	13	+++	+

Tabla 6: Actividad de variantes ADH-LK para la reducción de Metil-2-benzamidometil-3-oxobutirato a (2R,3R)-metil-2-benzamidometil-3-hidroxitirato

Número de identificación secuencial	Secuencia - mutaciones de codificación	Número de mutaciones a partir de kéfir	Actividad ^a	Selectividad ^b
64	V121; H4OR; A94E; S96V; F147M; Y190P; L195V; V196L; L199W; I226V; G248R; Y249W;	12	+++	+
<p>a +: 1 vez más activa que KRED con la IDENTIFICACIÓN SECUENCIAL NÚMERO: 66; ++: 1 - 2 veces más activa que KRED con la IDENTIFICACIÓN SECUENCIAL NÚMERO: 66; +++: >2 veces más activa que KRED con la IDENTIFICACIÓN SECUENCIAL NÚMERO: 66;</p> <p>b +: <85% de producto (2R,3R); >85% de producto (2R,3R).</p>				

Ejemplo 10: Producción preparativa a escala de (2S,3R)-metil-2-benzamidometil-3-hidroxitirato.

Un matraz de 250 ml de cuello 3 con un agitador superior fue cargado con 2-benzamidometil-3-cetobutirato de metilo (25 g), isopropilalcohol (37.5ml) y 0.1 M de trietanolamina (cloruro) / 0.04 M de amortiguador MgSO₄ con pH 7.2 (30 ml). La mezcla de reacción fue agitada y la temperatura fue adecuada a 37 °C usando un baño de aceite. La reacción fue iniciada con la adición de 0.5 mililitros 19 g / litros de NADP-Na seguida por 2.5 mililitros 30 g / litros de KRED de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 10; mientras las soluciones en 0.1 M de trietanolamina (cloruro) / 0.04 M de amortiguador MgSO₄ con un pH 7.2. El progreso de la reacción fue seguido al tomar porciones de 5 µl durante el transcurso de la reacción que fueron diluidos con 1 mililitro de acetonitrilo, filtrado a través de un cernidor tipo jeringa de 0.25 micrómetros y analizado tal como se describió en el ejemplo 3. Cuando la conversión excedió el 96%, se agregó cloruro de sodio acuoso saturado (12.5 mililitros) a la mezcla de reacción, seguido por 40 ml de acetato etílico. La mezcla de la reacción fue agitada durante 15 minutos más, luego filtrada a través de una almohadilla de celita (5 g en un filtro de vidrio poroso) al vacío. El pastel del filtro fue lavado con 20 ml de acetato etílico y las dos fases del filtro se les permitió separarse. La capa orgánica fue lavada tres veces con 20 ml de agua y fue concentrada en un evaporado rotatorio a un aceite de peso constante. Después de la remoción del acetato etílico, se agregó tolueno (10 ml) y se destiló al vacío. Se obtuvo (2S, 3R)-metil-benzamidometil-3-hidroxitirato (26.47 gramos) en forma de un aceite que contenía ~10% de tolueno tal como se determinó por ¹H-NMR- CDCl₃.

LISTA DE SECUENCIAS

<110> CAMPOPIANO, Onorato
MUNDORFF, Emily
BORUP, Birthe
VOLADRI, Rama

<120> POLIPÉPTIDOS DE CETORREDUCTASA PARA LA PRODUCCIÓN DE ACETITINONA

<130> 376247-022

<141> 2008-10-01

<150> 60/976,555

<151> 2007-10-01

<160> 87

<210> 1

<211> 762

<212> ADN

<213> Secuencia Artificial

<220>

<223> Optimizado en lo que se refiere a Codones Secuencia L. Brevis

<400> 1

```

atgtctaacc gtctggatgg caaagtagcc atcattaccg gcgggactct gggatatcgg 60
ttggcaatcg ccacgaaatt tgtagaggag ggtgcgaaag taatgattac tggtcgtcac 120
5 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
ggcaccctgc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
10 atgagcagta ttgaggggtt cgtaggcgat ccgagcctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcaactgaa ggactacgat 540
gtgctgtca acacagtaca tccgggctat atcaagacct cgctggtcga tgatctgccg 600
ggtgctgagg aagcgatgtc acagcgtacg aaaaccctta tgggccacat tggcgaaccg 660
15 aatgacatcg catatatctg tgtgtacctg gcactcaatg aatcgaatt tgcgacgggt 720
tccgaatttg tggtcgacgg cgggtatacc gcacagtaat ga 762

```

```

20 <210> 2
    <211> 252
    <212> PRT
    <213> Secuencia Artificial
    <220>
    <223> Traducción de la secuencia L. Brevis con Codones Optimizados
25 <400> 2

```

```

Met Ser Asn Arg Leu Asp Gly Lys Val Ala Ile Ile Thr Gly Gly Thr
1      5      10      15
30 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 Lys Ser Val Gly Thr Pro Asp Gln Ile Gln Phe Phe Gln His Asp Ser
    35      40      45      50
Ser Asp Glu Asp Gly Trp Thr Lys Leu Phe Asp Ala Thr Glu Lys Ala
65 Phe Gly Pro Val Ser Thr Leu Val Asn Asn Ala Gly Ile Ala Val Asn
40      55      60      65      70      75      80
Lys Ser Val Glu Glu Thr Thr Thr Ala Glu Trp Arg Lys Leu Leu Ala
45      85      90      95      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
50 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
55 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
60 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
65 Ser Glu Phe Val Val Asp Gly Gly Tyr Thr Ala Gln
245      250

```

ES 2 547 528 T3

<210> 3
 <211> 759
 <212> ADN
 <213> Secuencia Artificial

5

<220>
 <223> Secuencia de L. Kéfir Codones Optimizados

<400> 3

10

```

atgaccgatc gtctgaaggg caaagtagcc atcgtaaccc gccggactct gggatatcgg 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
15 ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cagtttccaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcaccgctc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggtt cgtagggcat ccgacgctgg gggcatacaa cgcttccaag 480
20 ggggcggtac gtatcatgtc gaaaagcgcg gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggctat atcaagaccg cgctggtcga tgatctggaa 600
gggtgctgagg aaatgatgtc acagcgtacg aaaaccctta tgggccacat tggcgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcactctgac aatcgaaatt tgcgacgggt 720
25 gcagaatttg tggtcgacgg cgggtatacc gcacagtga
    
```

25

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

30

<220>
 <223> Traducción de la secuencia L. Kéfir con Codones Optimizados

35

<400> 4

40

45

50

55

60

65

5 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
 10 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
 15 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
 20 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
 25 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
 30 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
 35 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

40 <210> 5
 <211> 759
 <212> ADN
 <213> Secuencia Artificial

45 <220>
 <223> variante de L. kéfir

<400>5
 50 atgaccgatc gtctgaaggc caaagtagcc atcgtaaccg gcgggacact gggatcgggt 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggcgtcac 120
 gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatggtat tcgctttgtc 180
 cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 ttcggcccgg ttacgaccgt cgtgaacaat gcagggatta ccgtagcaa aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgta atctggatgg tgttttttc 360
 55 ggcacccgctc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
 atgagcagta ttgaggggct ggtagggcat ccgacgctgg gggcatacaa cgcttccaag 480
 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggctat atcaagacc cgctggtcga tgatcatgaa 600
 ggtctggagg aatgatgtc acagcgtacg aaaacccta tgggccacat tgggtaaccg 660
 60 aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

65 <210> 6
 <211> 252
 <212> PRT
 <213> Secuencia Artificial

<220>

<223> variante de L. kéfir

5 <400> 6

	Met	Thr	Asp	Arg	Leu	Lys	Gly	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr
	1				5					10					15	
10	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			
15	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
20	Phe	Gly	Pro	Val	Thr	Thr	Val	Val	Asn	Asn	Ala	Gly	Ile	Thr	Val	Ser
					85						90				95	
	Lys	Ser	Val	Glu	Asp	Thr	Thr	Thr	Glu	Glu	Trp	Arg	Lys	Leu	Leu	Ser
				100					105					110		
25	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				
30	Glu	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	
35	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	His	Glu	Gly	Leu	Glu	Glu	Met	Met	Ser	Gln
			195					200					205			
40	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
45	Ala	Glu	Phe	Val	Val	Asp	Gly	Gly	Tyr	Thr	Ala	Gln				
					245					250						

<210> 7

<211> 759

<212> ADN

<213> Secuencia Artificial

50

<220>

<223> variante de L. Kefir

55

<400> 7

60

65

atgaccgatc gtctgaaggc caaagtagcc atcgtaaccg gcgggacact gggatcgggt 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
 5 gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
 cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 ttcggcccgg ttacgaccgt cgtgaacaat gcagggatta ccgtagcaa aagcgttgaa 300
 gacactacca cggaggaatg gcgtaactg ctgtccgtta atctggatgg tgttttttc 360
 ggcaccgctc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
 10 atgagcagta ttgaggggtt cgtaggcgtat ccgacgctgg gggcatacaa cgcttccaag 480
 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
 gtgctgtca acacagtaca tccgggctat atcaagaccg cgctggtcga tgatcatgaa 600
 ggtctggagg aatgatgtc acagcgtacg aaaaccctta tgggccacat tgggtgaaccg 660
 15 aatgacatcg catggatctg tgtgtacctg gcattctgacg aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210> 8

<211> 252

<212> PRT

20 <213> Secuencia Artificial

<220>

<223> variante de L. kéfir

25 <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
30				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
35		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	Val	Val	Asn	Asn	Ala	Gly	Ile	Thr	Val	Ser	
				85					90					95		
40	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			
45	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
50				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	His	Glu	Gly	Leu	Glu	Glu	Met	Met	Ser	Gln
				195				200					205			
55	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
60	Ala	Glu	Phe	Val	Val	Asp	Gly	Gly	Tyr	Thr	Ala	Gln				
				245						250						

<210> 9

<211> 759

65 <212> ADN

<213> Secuencia Artificial

<220>

<223> variante de L. kéfir

5 <400> 9

atgaccgatc gtctgaaggg caaagtagcc atcgtaacccg gcgggacact gggatcgggt 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcgc 120
 gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatggtat tgcgtttgtc 180
 10 cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 ttcggcccgg ttacgaccgt cgtgaacaat gcagggatta ccgtagcaa aageggtgaa 300
 gacactacca cggaggaatg gcgtaaaactg ctgtccggtta atctggatgg tgtttttttc 360
 ggcacccgctc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
 atgagcagta ttgaggggct ggtaggcgtat ccgacgctgg gggcatacaa cgcttccaag 480
 15 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggctat atcaagacc cgtgtgtcga tgatcatgaa 600
 ggtctggagg aatgatgtc acagcgtacg aaaacccta tgggccacat tggatgaaccg 660
 aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
 20 gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210> 10

<211> 252

<212> PRT

25 <213> Secuencia Artificial

<220>

<223> variante de L. kéfir

30 <400> 10

	Met	Thr	Asp	Arg	Leu	Lys	Gly	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr
	1				5					10					15	
35	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			
40	Lys	Ser	Ile	Gly	Gly	Thr	Asp	Val	Ile	Arg	Phe	Val	Gln	His	Asp	Ala
	50						55					60				
45	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	Thr	Val	Ser
					85					90					95	
50	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			
55	Met	Lys	Asn	Lys	Gly	Leu	Gly	Ala	Ser	Ile	Ile	Asn	Met	Ser	Ser	Ile
	130						135					140				
	Glu	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
60					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	His	Glu	Gly	Leu	Glu	Glu	Met	Met	Ser	Gln
			195					200					205			
65	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> 11
 <211> 759
 <212> ADN
 <213> Secuencia Artificial

5

<220>
 <223> variante de L. kéfir

<400> 11

10 atgaccgatc gtctgaaggc caaagtagcc atcgtaaccg gcgggacact gggatcgggt 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggcgcgcgc 120
 gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
 15 cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 ttcggccccg ttacgaccgt cgtgaacaat gcagggatta ccgtagcaa aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgcta atctggatgg tgtttttttc 360
 ggcaccctgc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
 atgagcagta ttgaggggtt cgtagggcat ccgacgctgg gggcatacaa cgcttccaag 480
 20 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggctat atcaagacce cgctggctga tgatcatgaa 600
 ggtctggagg aatgatgtc acagcgtacg aaaaccctta tgggccacat tggatgaaccg 660
 aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaat tgcgacgggt 720
 25 gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

25

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

30

<220>
 <223> variante de L. kéfir

<400> 12

35 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
 40 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
 45 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Thr Val Ser
 85 90 95
 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
 100 105 110
 50 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
 55 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
 60 Thr Pro Leu Val Asp Asp His Glu Gly Leu 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
 65 Ala Glu Phe Val Val Asp Gly Gly Tyr Thr Ala Gln
 245 250

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

5

<220>
 <223> variante de L. kéfir

<400> 13

10

```

atgaccgatc gtctgaaggc caaagtagcc atcgtaaccg gcgggacact gggatatcgg 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 gcagggattt ttgttgtaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcaccgctc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaagggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgactgaa ggactacgat 540
gtgctgtca acacagtaca tccgggctat atcaagacc cggtgctcga tgattgggaa 600
ggtgctgagg aatgatgtc acagcgtacg aaaacccta tgggccacat tgggaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgatg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga 759
    
```

25

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

30

<220>
 <223> variante de L. kéfir

35

<400> 14

40

45

50

55

60

65

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

40 <210> 15
 <211> 759
 <212> ADN
 <213> Secuencia Artificial

45 <220>
 <223> variante de L. kéfir

<400> 15

50 atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggactct gggtatcggg 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
 gcggatgtag gcgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
 cagcagatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 55 ttcggcccgg ttacgaccgt cgtgaacaat gcagggatta cggtttttaa aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgta atctggatgg tgttttttc 360
 ggcaccctgc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
 atgagcagta ttgaggggtt cgtagggcat ccgacgctgg gggcatacaa cgcttccaag 480
 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
 60 gtgctgtca acacagtaca tccgggctat atcaagacc cgctggtcga tgatctggaa 600
 ggtgtggagg aatgatgtc acagcgtagc aaaacccta tgggccacat tgggtgaaccg 660
 aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

65

<210> 16
 <211> 252
 <212> PRT
 <213> Secuencia Artificial
 5
 <220>
 <223> variante de L. kéfir
 <400> 16
 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
 15 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 Thr Val Phe
 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 Val 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> 17
 <211> 759
 <212> ADN
 <213> Secuencia Artificial
 50
 <220>
 <223> variante de L. kéfir
 <400> 17
 55
 60
 65

5
10
15

```

atgaccgatc gtctgaaggc caaagtagcc atcgtaaccc gcgggactct gggatcggg 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggagaca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg ccgtagcaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcaccogtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggtt cgtaggcgat ccgacggcag gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgctgtca acacagtaca tccgggctat atcaagacc cgctggtcga tgatctggaa 600
ggtgtggagg aatgatgtc acagcgtacg aaaacccta tgggccacat tgggtaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759
    
```

20

```

<210> 18
<211> 252
<212> PRT
<213> Secuencia Artificial
    
```

25

```

<220>
<223> variante de L. kéfir
    
```

30

```

<400> 18
    
```

30
35
40
45
50
55
60
65

```

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 Thr
 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 Ala 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 Val 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> 19
 <211> 759
 <212> ADN
 <213> Secuencia Artificial

5

<220>
 <223> variante de L. kéfir

<400> 19

10

```

atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggactct gggatatcgg 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 gcagggatta ccgtagcaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcaccctgc tgggcattca gcgcatgaaa aataaaggot tgggcgctag catcatcaat 420
atgagcagta ttgaggggtt cgtaggcgtat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgcg gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggctat atcaagacct cgctggtcga tgataatgaa 600
ggtgtggagg aaatgatgtc acagcgtacg aaaacccta tgggccacat tggngaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759
    
```

25

<210> 20
 <211> 252
 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> variante de L. kéfir
 <400> 20

35

40

45

50

55

60

65

5 Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10
 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
 10 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 Thr Val Ser
 85 90 95
 15 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
 20 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
 25 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
 30 Thr Pro Leu Val Asp Asp Asn Glu Gly Val 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
 35 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

40 <210> 21
 <211> 759
 <212> ADN
 <213> Secuencia Artificial

45 <220>
 <223> variante de L. kéfir

<400> 21

50 atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg ggggactct gggatcgggt 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgctcac 120
 gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
 cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 55 ttcggcccgg ttacgaccgt cgtgaacaat gcagggatta ccgtttagcaa aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgтта atctggatgg tgtttttttc 360
 ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
 atgagcagta ttgaggggtt cgtaggcgat ccgacgctgg gggcatacaa cgctccaag 480
 ggggcgggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
 60 gtgcgtgtca acacagtaca tccgggctat atcaagacc cgctggtcga tgatcatgaa 600
 ggtctggagg aatgatgtc acagcgtacg aaaacccta tgggccacat tgggtgaaccg 660
 aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

65

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

5

<220>
 <223> variante de L. kéfir

10

<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
15				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
20		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	Thr	Val	Ser
25				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			
30	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
35	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		
40	Thr	Pro	Leu	Val	Asp	Asp	His	Glu	Gly	Leu	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				
45	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						

50

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

55

<220>
 <223> variante de L. kéfir

60

<400> 23

65

5 atgaccgatc gtctgaaggg caaagtagcc atcgtaaccc gcgggactct gggatccggt 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
 gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tgcctttgtc 180
 cagcacgatg catccgatga agcaggctgg acgaaactgt tgcacaccac cgaggaggca 240
 ttcggcccgg ttacgaccgt cgtgaacaat gcagggatta ccgtagcaa aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
 ggcacccgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
 10 atgagcagta ttgagggggt cgtaggcgtat ccgacgctgg gggcatacaa cgcttccaag 480
 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcaactgaa ggactacgat 540
 gtgcggtgtca acacagtaca tccgggctat atcaagaccc cgctggctcga tgatcatgaa 600
 ggtgtggagg aatgatgtc acagcgtacg aaaaccoccta tgggccacat tggatgaaccg 660
 15 aatgacatcg catggatctg tgtgtacctg gcacactgac aatcgaaatt tgcgacgggt 720
 gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

20 <210> 24
 <211> 252
 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> variante de L. kéfir
 25 <400> 24

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 His Ala Asp Val Gly Glu Lys Ala Ala
 35 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
 40 65 70 75 80
 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Thr Val Ser
 85 90 95
 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
 100 105 110
 45 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
 50 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
 55 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 His Glu Gly Val Glu Glu Met Met Ser Gln
 195 200 205
 60 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
 65 Ala Glu Phe Val Val Asp Gly Gly Tyr Thr Ala Gln
 245 250

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

5

<220>
 <223> variante de L. kéfir

10 <400> 25

```

atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg ggggactct gggtatcggg 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
15 cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggatta ccgtagcaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcaccctgc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
20 atgagcagta ttgaggggtt cgtagggcat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgcg gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggctat atcaagacc cgctggtcga tgatctggaa 600
ggtctggagg aatgatgtc acagcgtacg aaaaccctta tgggccacat tgggtgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
25 gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759
    
```

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

30

<220>
 <223> variante de L. kéfir

35 <400> 26

```

Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
1 5 10 15
40 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
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
50 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Thr Val Ser
85 90 95
Lys Ser Val Glu Asp Thr Thr Thr Glu Trp Arg Lys Leu Leu Ser
100 105 110
55 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
60 Glu Gly Phe Val Gly Asp Pro Thr Leu Gly Ala Tyr Asn Ala Ser Lys
145 150 155 160
    
```

65

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
 5 180 185 190
Thr Pro Leu Val Asp Asp Leu Glu Gly Leu 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
 10 210 215 220
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
 15 245 250

<210> 27
 <211> 759
 <212> ADN
 20 <213> Secuencia Artificial

<220>
 <223> variante de L. kéfir

25 <400> 27

atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggactct gggatatcggg 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
 30 **gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180**
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggccccg ttacgaccgt cgtgaacaat gcagggatta ccgtagcaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgta atctggatgg tgttttttc 360
ggcaccctgc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
 35 **atgagcagta ttgaggggtt cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480**
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggctat atcaagacc cgctggctga tgatctggaa 600
ggtgtggagg aaatgatgtc acagcgtacg aaaaccctc tgggccacat tggatgaaccg 660
 40 **aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720**
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210> 28
 <211> 252
 <212> PRT
 45 <213> Secuencia Artificial

<220>
 <223> variante de L. kéfir

50 <400> 28

55

60

65

	Met	Thr	Asp	Arg	Leu	Lys	Gly	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr
	1				5					10					15	
5	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			
10	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
15	Phe	Gly	Pro	Val	Thr	Thr	Val	Val	Asn	Asn	Ala	Gly	Ile	Thr	Val	Ser
					85					90					95	
	Lys	Ser	Val	Glu	Asp	Thr	Thr	Thr	Glu	Glu	Trp	Arg	Lys	Leu	Leu	Ser
				100					105					110		
20	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				
25	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	
30	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	Val	Glu	Glu	Met	Met	Ser	Gln
			195				200						205			
35	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
40	Ala	Glu	Phe	Val	Val	Asp	Gly	Gly	Tyr	Thr	Ala	Gln				
					245					250						

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

<220>
 <223> variante de L. kéfir

<400> 29

50	atgaccgatc	gtctgaaggg	caaagtagcc	atcgtaaccg	gcgggactct	gggtatcggg	60
	ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttattac	tggtcgtcac	120
	gcggatgtag	gtgaaaaggc	cgccaaatca	atcggcggca	ctgatgttat	tcgctttgtc	180
	cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
55	ttcggcccgg	ttacgaccgt	cgtgaacaat	gcagggatta	ccgttagcaa	aagcgttgaa	300
	gacactacca	cggaggaatg	gcgtaaactg	ctgtccggtta	atctggatgg	tgttttttc	360
	ggcacccgtc	tgggcattca	gcgcatgaaa	aataaaggct	tgggcgctag	catcatcaat	420
	atgagcagta	ttgagggggt	cgtaggcgat	ccgacgctgg	gggcatacaa	cgcttccaag	480
60	ggggcggtag	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcactgaa	ggactacgat	540
	gtgcgtgtca	acacagtaca	tccgggctat	atcaagaccc	cgctggtcga	tgatgcagaa	600
	ggtctggagg	aaatgatgtc	acagcgtacg	aaaacccta	tgggccacat	tggtgaaccg	660
	aatgacatcg	catggatctg	tgtgtacctg	gcactctgacg	aatcgaaatt	tgcgacgggt	720
65	gcagaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

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

5

<220>
 <223> variante de L. kéfir

10

<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
 15 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 Thr 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 Ala Glu Gly Leu 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

50

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

55

<220>
 <223> variante de L. kéfir

60

<400> 31

65

atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggactct gggatcgggt 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgctcac 120
 gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatggtat tcgctttgtc 180
 5 cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 ttcggcccgg ttacgaccgt cgtgaacaat gcagggatta ccgttagcaa aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
 ggcaccctgc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
 10 atgagcagta ttgaggggtt cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
 ggggcggtac gtatcatgtc gaaaagcgcg gcgctggatt gcgcactgaa ggactacgat 540
 gtgctgtca acacagtaca tccgggctat atcaagacc cgctggtcga tgatgcagaa 600
 ggtgtggagg aatgatgtc acagcgtacg aaaaccctta tgggccacat tgggtgaaccg 660
 aatgacatcg catggatctg tgtgtacctg gcactctgacg aatcgaaatt tgcgacgggt 720
 15 gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210> 32
 <211> 252
 <212> PRT
 20 <213> Secuencia Artificial

<220>
 <223> variante de L. kéfir

25 <400> 32

Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 30 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 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
 40 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Thr Val Ser
 85 90 95
 Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
 100 105 110
 45 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
 50 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
 55 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 Ala Glu Gly Val Glu Glu Met Met Ser Gln
 195 200 205
 60 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> 33
 <211> 759
 <212> ADN
 <213> Secuencia Artificial

5

<220>
 <223> variante de L. kéfir

<400> 33

10

```

atgaccgatac gtctgaaggg caaagtagcc atcgtaaccg gcgggactct gggatatcgg 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgctcac 120
gaggatgatg gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
15 ttcggcccgg ttacgaccgt cgtgaacaat gcagggatta ccgtagcaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaaactg ctgtccgta atctggatgg tgttttttc 360
ggcaccctgc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggtt cgtaggcgat ccgacggcag gggcatacaa cgcttccaag 480
20 ggggpcggtac gtatcatgtc gaaaagcgcg gcgctggatt gcgcactgaa ggactacgat 540
gtgctgttca acacagtaca tccgggctat atcaagaccc cgctggtcga tgatcatgaa 600
ggctctggagg aatgatgtc acagcgtacg aaaaccctta tgggccacat tggatgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759
    
```

25

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

30

<220>
 <223> variante de L. kéfir

<400> 34

35

```

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 Thr 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 Ala 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 His Glu Gly Leu 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
    
```

65

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

5
 <220>
 <223> variante de L. kéfir

10
 <400> 35

```

atgaccgatc gtctgaaggc caaagtagcc atcgtaaccg gcgggactct gggatcgggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tgcctttgtc 180
cagcacgatg catccgatga agcaggctgg acgaaactgt tgcacaccac cgaggaggca 240
15 ttcggcccgg ttacgaccgt cgtgaacaat gcagggatta ccgtagcaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgta atctggatgg tgttttttc 360
ggcaccctgc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggtt cgtaggcgat ccgacggcag gggcatacaa cgcttccaag 480
20 ggggocggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcaactgaa ggactacgat 540
gtgocgtgtca acacagtaca tccgggctat atcaagacc cgctggtcga tgatgcagaa 600
ggtgtagagg aatgatgctc acagcgtacg aaaacccta tgggccacat tggatgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759
    
```

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

30
 <220>
 <223> variante de L. kéfir

35
 <400> 36

```

Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
1      5      10
Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val 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 Thr 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 Ala Gly Ala Tyr Asn Ala Ser Lys
145    150    155    160
55 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 Ala Glu Gly Val Glu Glu Met Met Ser Gln
195    200    205
60 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
65 Ala Glu Phe Val Val Asp Gly Gly Tyr Thr Ala Gln
245    250
    
```

ES 2 547 528 T3

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

5

<220>
 <223> variante de L. kéfir

<400> 37

10 **atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggactct gggatcgggt 60**
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tgcctttgtc 180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 15 **ttcggcccgg ttacgaccgt cgtgaacaat gcagggatta ccgtagcaa aagcgttgaa 300**
gacactacca cgggggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcaccctgc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggtt cgtagggcat ccgacggcag gggcatacaa cgcttccaag 480
 20 **ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540**
gtgctgtca acacagtaca tccgggctat atcaagacc cgctgggtcga tgatgcagaa 600
ggtctggagg aaatgttttc acagcgtacg aaaaccctta tgggccacat tggatgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcactctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

25

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

30

<220>
 <223> variante de L. kéfir

<400> 38

35

40

45

50

55

60

65

5 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
 10 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
 15 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Thr Val Ser
 85 90 95
 Lys Ser Val Glu Asp Thr Thr Thr Gly 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
 20 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 Ala Gly Ala Tyr Asn Ala Ser Lys
 145 150 155 160
 25 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
 30 Thr Pro Leu Val Asp Asp Ala Glu Gly Leu 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
 35 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> 39
 <211> 759
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> variante de L. kéfir

<400> 39

50 atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggactct gggatcgggt 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
 gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
 55 cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 ttcggcccgg ttacgaccgt cgtgaacaat gcagggatta ccgtttttaa aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
 ggcacccgtc tgggcattca gcgcacgaaa aataaaggct tgggcgctag catcatcaat 420
 atgagcagta ttgagggggt cgtagcgcat ccgacgctgg gggcatacaa cgcttccaag 480
 60 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggctat atcaagacc cgctggtcga tgatctggaa 600
 ggtgtggagg aatgttttc acagcgtacg aaaacccta tgggccacat tgggtgaaccg 660
 aatgacatcg catggatctg tgtgtacctg gcactctgacg aatcgaaatt tgcgacgggt 720
 65 gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210> 40
 <211> 252
 <212> PRT
 <213> Secuencia Artificial
 5
 <220>
 <223> variante de L. kéfir
 <400>40
 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 Thr Val Phe
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
Thr 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 Val 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 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
 50
 <210> 41
 <211> 759
 <212> ADN
 <213> Secuencia Artificial
 55
 <220>
 <223> variante de L. kéfir
 <400> 41
 60
 65

atgaccgatc gtctgaaggc caaagtagcc atcgtaaccg gcgggactct gggatcgggt 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
 gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tgcgtttgtc 180
 5 cagcacgatg catccgatga agcaggctgg acgaaactgt tgcacaccac cgaggaggca 240
 ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cagtttccaa aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
 ggcaccgctc tgggcatcca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
 atgagcagta ttgaggggtt cgtagggcat ccgacggcag gggcatacaa cgcttccaag 480
 10 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
 gtgctgtca acacagtaca tccgggctat atcaagacc cgctggtcga tgataatgaa 600
 ggtctggagg aatgatgtc acagcgtacg aaaacccta tgggccacat tggatgaaccg 660
 aatgacatcg catggatctg tgtgtacctg gcactctgacg aatcgaaatt tgcgacgggt 720
 15 gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210> 42

<211> 252

<212> PRT

20 <213> Secuencia Artificial

<220>

<223> variante de L. kéfir

25 <400> 42

	Met	Thr	Asp	Arg	Leu	Lys	Gly	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr
	1				5					10					15	
30	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				
35	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
40	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		
45	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				
50	Glu	Gly	Phe	Val	Gly	Asp	Pro	Thr	Ala	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		
55	Thr	Pro	Leu	Val	Asp	Asp	Asn	Glu	Gly	Leu	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				
60	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						

65

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

5

<220>
 <223> variante de L. kéfir

<400> 43

10 **atgaccgatac gtctgaaggc caaagtagcc atcgtaaccg gcgggactct gggatcggg 60**
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggcgcgcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 15 **ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cagtttcaa aagcgttgaa 300**
gacactacca cggaggaatg gcgtaaacctg ctgtccgtta atctggatgg tgttttttc 360
ggcaccgctc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaagggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
 20 **ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540**
gtgcgtgtca acacagtaca tccgggctat atcaagacc cgctggtcga tgatctggaa 600
ggtgtagagg aatgatgtc acagcgtacg aaaacccta tgggccacat tggatgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcacatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

25

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

30

<220>
 <223> variante de L. kéfir

<400> 44

35

40

45

50

55

60

65

	Met	Thr	Asp	Arg	Leu	Lys	Gly	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr
	1				5					10					15	
5	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			
10	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
15	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		
20	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				
25	Glu	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	
30	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	Val	Glu	Glu	Met	Met	Ser	Gln
			195					200					205			
35	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
40	Ala	Glu	Phe	Val	Val	Asp	Gly	Gly	Tyr	Thr	Ala	Gln				
					245					250						

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

<220>
 <223> variante de L. kéfir

<400> 45

50	atgaccgatac	gtctgaagg	caaagtagcc	atcgtaaccg	gcgggactct	gggtatcgg	60
	ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttattac	tggtcgtcac	120
	gcggatgtag	gtgaaaaggc	cgccaaatca	atcggcggca	ctgatgttat	tcgctttgtc	180
	cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
55	ttcggccccg	ttacgaccgt	cgtgaacaat	gcagggattg	cagtttccaa	aagcgttgaa	300
	gacactacca	cggaggaatg	gcgtaaaactg	ctgtccgтта	atctggatgg	tgtttttttc	360
	ggcaccgctc	tgggcattca	gcgcatgaaa	aataaaggct	tgggcgctag	catcatcaat	420
	atgagcagta	ttgagggggt	cgtaggcgat	ccgacgagcg	gggcatacaa	cgcttccaag	480
60	ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcactgaa	ggactacgat	540
	gtgctgtca	acacagtaca	tccgggctat	atcaagacc	cgctggtcga	tgatctggaa	600
	ggtctggagg	aaatgatgtc	acagcgtacg	aaaacccta	tgggccacat	tggtgaaccg	660
	aatgacatcg	catggatctg	tgtgtacctg	gcactctgacg	aatcgaaatt	tgcgacgggt	720
65	gcagaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

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

5

<220>
 <223> variante de L. kéfir

10

<400> 46

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

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

<220>
 <223> variante de L. kéfir

<400> 47

60

65

atgaccgatc gtctgaaggc caaagtagcc atcgtaacccg gcgggactct gggatatcggg 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
 gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
 5 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
 10 atgagcagta ttgagggggt cgtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
 ggggcggtac gtatcatgtc gaaaagcgcg gcgctggatt gcgactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggctat atcaagacc cgctgggtcga tgatctggaa 600
 ggtgtagagg aatgatgtc acagcgtacg aaaaccecta tgggccacat tgggtaaccg 660
 aatgacatcg catggatctg tgtgtacctg gcactctgacg aatcgaaatt tgcgacgggt 720
 15 gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

<210> 48
 <211> 252
 <212> PRT
 20 <213> Secuencia Artificial

<220>
 <223> variante de L. kéfir

25 <400> 48

Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
 1 5 10 15
 30 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 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
 40 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
 45 115 120 125
 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
 50 145 150 155 160
 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
 55 Thr Pro Leu Val Asp Asp Leu Glu Gly Val 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
 60 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

65

ES 2 547 528 T3

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

5

<220>
 <223> variante de L. kéfir

10

<400> 49

atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggactct gggatcgggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcggtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tgcgtttgtc 180
 15 **cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240**
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cagtttccaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcaccgctc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
 20 **atgagcagta ttgaggggtt cgtaggcgat ccgacggcag gggcatacaa cgcttccaag 480**
ggggcggtac gtatcatgtc gaaaagcgcg gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tcggggctat atcaagacct cgctggtcga tgatcatgaa 600
ggtctggagg aatgatgtc acagcgtacg aaaacccta tgggccacat tgggtaaccg 660
aatgacatcg catggatctg tgtgtacctg gcactctgacg aatcgaaatt tgcgacgggt 720
 25 **gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759**

30

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

<220>
 <223> variante de L. kéfir

35

<400> 50

40

45

50

55

60

65

	Met	Thr	Asp	Arg	Leu	Lys	Gly	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr
	1				5					10					15	
5	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			
10	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	
15	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			
20	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	Ala	Gly	Ala	Tyr	Asn	Ala	Ser	Lys
	145					150						155				160
25	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		
30	Thr	Pro	Leu	Val	Asp	Asp	His	Glu	Gly	Leu	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
35	Ala	Glu	Phe	Val	Val	Asp	Gly	Gly	Tyr	Thr	Ala	Gln				
					245						250					

40 <210> 51
 <211> 759
 <212> ADN
 <213> Secuencia Artificial

45 <220>
 <223> variante de L. kéfir

<400> 51

	atgaccgatc	gtctgaaggg	caaagtagcc	atcgtaaccg	gcgggactct	gggtatcggg	60
50	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
55	ggcaccgctc	tgggcattca	gcgcatgaaa	aataaaggct	tgggcgctag	catcatcaat	420
	atgaccagta	ttgaggggtt	cgtagggcgt	ccgactggag	gggcatacaa	cgcttccaag	480
	ggggcggtag	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcactgaa	ggactacgat	540
	gtgctgtca	acacagtaca	tccgggctat	atcaagacc	cgctggtcga	tgatctggaa	600
	ggtctggagg	aaatgatgtc	acagcgtacg	aaaaccctta	tgggccacat	tggtgaaccg	660
60	aatgacatcg	catggatctg	tgtgtacctg	gcattctgacg	aatcgaaatt	tgcgacgggt	720
	gcagaatttg	tggtcgacgg	cggtataacc	gcacagtga			759

65 <210> 52
 <211> 252
 <212> PRT

<213> Secuencia Artificial

<220>

<223> variante de L. kéfir

5

<400> 52

	Met	Thr	Asp	Arg	Leu	Lys	Gly	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr
	1				5					10					15	
10	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			
15	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
20	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		
25	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				
30	Glu	Gly	Phe	Val	Gly	Asp	Pro	Thr	Ala	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	
35	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	Leu	Glu	Glu	Met	Met	Ser	Gln
			195					200					205			
40	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
45	Ala	Glu	Phe	Val	Val	Asp	Gly	Gly	Tyr	Thr	Ala	Gln				

245

250

50

<210> 53

<211> 759

<212> ADN

<213> Secuencia Artificial

55

<220>

<223> variante de L. kéfir

<400> 53

60

65

```

atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggactct gggatcgggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
5 cagcaccgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggccccg ttacgaccgt cgtgaacaat gcagggattg cagtttccaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcaccocgtc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggtt cgtagggcat ccgacggcag gggcatacaa cgcttccaag 480
10 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcggtgtca acacagtaca tccgggctat atcaagaccg cgctggtcga tgatctggaa 600
ggtgtagagg aatgatgtc acagcgtacg aaaacccta tgggccacat tggatgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcactctgacg aatcgaaatt tgcgacgggt 720
15 gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759

```

```

<210> 54
<211> 252
<212> PRT
20 <213> Secuencia Artificial

```

```

<220>
<223> variante de L. kéfir

```

```

25 <400> 54

```

```

Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Val Thr Gly Gly Thr
1 5 10 15
30 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 35 40 45
Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
50 50 55 60
Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
65 65 70 75 80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Ala Val Ser
40 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
45 115 120 125
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 Ala Gly Ala Tyr Asn Ala Ser Lys
50 145 150 155 160
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
55 180 185 190
Thr Pro Leu Val Asp Asp Leu Glu Gly Val 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
60 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
65 245 250

```

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

5

<220>
 <223> variante de L. kéfir

<400> 55

10

```

atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggactct gggatcgggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcggtcac 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg catccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg cagtttccaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaaactg ctgtccggtta atctggatgg tgtttttttc 360
ggcaccgctc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaggggtt cgtagggcat cgcacggcag gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgcg gcgctggatt gcgcactgaa ggactacgat 540
gtgctgttca acacagtaca tccgggctat atcaagaccc cgctggtcga tgatgcagaa 600
ggtctggagg aatgatgtc acagcgtacg aaaacccta tgggccacat tggatgaaccg 660
aatgacatcg catggatctg tgtgtacctg gcatctgacg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtatacc gcacagtga 759
    
```

25

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

30

<220>
 <223> variante de L. kéfir

<400> 56

35

```

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 Ala 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
    
```

65

			180					185					190			
	Thr	Pro	Leu	Val	Asp	Asp	Ala	Glu	Gly	Leu	Glu	Glu	Met	Met	Ser	Gln
5			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
10						230					235					240
	Ala	Glu	Phe	Val	Val	Asp	Gly	Gly	Tyr	Thr	Ala	Gln				
					245					250						

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

<220>
 <223> variante de L. kéfir
 <400> 57

	atgaccgatc	gtctgaaggg	caaagtagcc	atcgtaaccg	gcgggactct	gggtatcggg	60
	ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttattac	tggtcgtcac	120
25	gcggatgtag	gtgaaaaggc	cgccaaatca	atcggcggca	ctgatgttat	tcgctttgtc	180
	cagcacgatg	catccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
	ttcggcccgg	ttacgatcgt	cgtgaacaat	gcagggattg	cagtttccaa	aagcgttgaa	300
	gacactacca	cggaggaatg	gcgtaaactg	ctgtccgtta	atctggatgg	tgtttttttc	360
30	ggcaccctgc	tgggcattca	gcgcatgaaa	aataaaggct	tgggcgctag	catcatcaat	420
	atgagcagta	ttgaggggtt	cgtaggcgat	ccgacgctgg	gggcatacaa	cgcttccaag	480
	ggggcggtac	gtatcatgtc	gaaaagcgcg	gcgctggatt	gcgcactgaa	ggactacgat	540
	gtgctgtca	acacagtaca	tccgggctat	atcaagacct	cgctggtcga	tgataatgaa	600
	ggtctggagg	aatgatgtc	acagcgtacg	aaaaccctca	tgggccacat	tggtgaaccg	660
35	aatgacatcg	catggatctg	tgtgtacctg	gcacatgacg	aatcgaaatt	tgcgacgggt	720
	gcagaatttg	tggtcgacgg	cgggtatacc	gcacagtga			759

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

<220>
 <223> variante de L. kéfir
 <400> 58

50
 55
 60
 65

	Met	Thr	Asp	Arg	Leu	Lys	Gly	Lys	Val	Ala	Ile	Val	Thr	Gly	Gly	Thr
	1				5					10					15	
5	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			
10	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
15	Phe	Gly	Pro	Val	Thr	Ile	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		
20	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				
25	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	
30	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	Asn	Glu	Gly	Leu	Glu	Glu	Met	Met	Ser	Gln
			195				200						205			
35	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
40	Ala	Glu	Phe	Val	Val	Asp	Gly	Gly	Tyr	Thr	Ala	Gln				
				245						250						

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

<220>
 <223> variante de L. kéfir

50	<400> 59															
	atgaccgatc	gtctgaaggg	caaagtagcc	ttcgtaacccg	gcgggacact	gggtatcggt	60									
	ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttattac	tggtcgtcgt	120									
	gcggatgtag	gtgaaaaggc	cgccaaatca	atcggcggca	ctgatgttat	tcgctttgtc	180									
55	cagcacgatg	cgtccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240									
	ttcggcccgg	ttacgaccgt	cgtgaacaat	gcagggattt	ttgttgtaa	aagcgttgaa	300									
	gacactacca	cggaggaatg	gcgtaaactg	ctgtccgтта	atctggatgg	tgtttttttc	360									
	ggcaccgctc	tgggcattca	gcgcatgaaa	aataaaggct	tgggcgctag	catcatcaat	420									
	atgagcagta	ttgaagggat	ggtagggcat	ccgacgctgg	gggcatacaa	cgcttccaag	480									
60	ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcactgaa	ggactacgat	540									
	gtgctgtca	acacagtaca	tccgggctat	atcaagacc	cggtgctcga	tgattgggaa	600									
	ggtgctgagg	aaatgatgtc	acagcgtacg	aaaacccta	tgggccacat	tggtgaaccg	660									
	aatgacatcg	catgggtctg	tgtgtacctg	gcatctgatg	aatcgaaatt	tgcgacgggt	720									
65	gcagaatttg	tggtcgacgg	caagtggacc	gcacagtga			759									

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

5

<220>
 <223> variante de L. kéfir

<400> 60

10

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

15

Leu Gly Ile Gly Leu Ala Ile Ala Asp Lys Phe Val Glu Glu Gly Ala
20 25 30

20

Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
35 40 45

25

Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
50 55 60

30

Ser Asp Glu Ala Gly Trp Thr Lys Leu Phe Asp Thr Thr Glu Glu Ala
65 70 75 80

35

Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Phe Val Val
85 90 95

40

Lys Ser Val Glu Asp Thr Thr Thr Glu Glu Trp Arg Lys Leu Leu Ser
100 105 110

45

Val Asn Leu Asp Gly Val Phe Phe Gly Thr Arg Leu Gly Ile Gln Arg
115 120 125

50

Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
130 135 140

55

Glu Gly Met Val Gly Asp Pro Thr Leu Gly Ala Tyr Asn Ala Ser Lys
145 150 155 160

60

Gly Ala Val Arg Ile Met Ser Lys Ser Ala Ala Leu Asp Cys Ala Leu
165 170 175

65

Lys Asp Tyr Asp Val Arg Val Asn Thr Val His Pro Gly Tyr Ile Lys
180 185 190

70

Thr Pro Val Leu Asp Asp Trp Glu Gly Ala Glu Glu Met Met Ser Gln
195 200 205

75

Arg Thr Lys Thr Pro Met Gly His Ile Gly Glu Pro Asn Asp Ile Ala
210 215 220

80

Trp Val Cys Val Tyr Leu Ala Ser Asp Glu Ser Lys Phe Ala Thr Gly
225 230 235 240

85

Ala Glu Phe Val Val Asp Gly Lys Trp Thr Ala Gln
245 250

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

90

<220>
 <223> variante de L. kéfir

95

<400> 61

100

atggccgatt gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggatatcgg 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgctcgt 120
 gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tgcgtttgtc 180
 5 cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
 ggcacccgctc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
 atgagcagta ttgaagggat ggtaggcgtat cgcacgctgg gggcatacaa cgcttccaag 480
 10 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
 gtgcgtgtca acacagtaca tccgggctat atcaagacc cgtctgctcga tgattgggaa 600
 ggtgctgagg aatgatgtc acagcgtacg aaaacccta tgggccacat tgggtgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg gcactctgat aatcgaaatt tgcgacgggt 720
 15 gcagaatttg tggtcgacgg caagtggacc gcacagtga 759

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

20 <220>
 <223> variante de L. kéfir

25 <400> 62

 Met Ala Asp Cys 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
 30 20 25 30
 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
 35 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
 45 115 120 125
 Met Lys Asn Lys Gly Leu Gly Ala Ser Ile Ile Asn Met Ser Ser Ile
 130 135 140
 Glu Gly Met Val Gly Asp Pro Thr Leu Gly Ala Tyr Asn Ala Ser Lys
 50 145 150 155 160
 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
 55 180 185 190
 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
 60 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 Lys Trp Thr Ala Gln
 245 250
 65

ES 2 547 528 T3

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

5

<220>
 <223> variante de L. kéfir

<400> 63

10

```

atgaccgatc gtctgaaggg caaagtagcc atcataaccg gcgggacact gggatatcgg 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgtcg 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
15 ttcggccccg ttacgaccgt cgtgaacaat gcagggattg aagttggtta aagcgttgaa 300
gacactacca cggaggaatg gcgtaaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcaccgctc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaagggat ggtagggcat ccgacgctgg gggcatacaa cgcttccaag 480
20 ggggcggtac gtatcatgtc gaaaagcgcg gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccc atcaagacc cggtgctcga tgattgggaa 600
ggtgctgagg aatgatgtc acagcgtacg aaaaccctta tgggccacat tggatgaccg 660
aatgacatcg catgggtctg tgtgtacctg gcactctgat aatcgaaatt tgcgacgggt 720
25 gcagaatttg tggtcgacgg caggtggacc gcacagtga 759
    
```

25

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

30

<220>
 <223> variante de L. kéfir

<400> 64

35

```

Met Thr Asp Arg Leu Lys Gly Lys Val Ala Ile Ile Thr Gly Gly Thr
1          5          10          15
40 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
45 Lys Ser Ile Gly Gly Thr Asp Val Ile Arg Phe Val Gln His Asp Ala
    
```

45

50

55

60

65

	50		55		60												
	Ser	Asp	Glu	Ala	Gly	Trp	Thr	Lys	Leu	Phe	Asp	Thr	Thr	Glu	Glu	Ala	
	65					70					75					80	
5	Phe	Gly	Pro	Val	Thr	Thr	Val	Val	Asn	Asn	Ala	Gly	Ile	Glu	Val	Val	
					85					90					95		
	Lys	Ser	Val	Glu	Asp	Thr	Thr	Thr	Glu	Glu	Trp	Arg	Lys	Leu	Leu	Ser	
				100					105					110			
10	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					
15	Glu	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		
20	Lys	Asp	Tyr	Asp	Val	Arg	Val	Asn	Thr	Val	His	Pro	Gly	Pro	Ile	Lys	
				180					185					190			
	Thr	Pro	Val	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					
25	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	Arg	Trp	Thr	Ala	Gln					
				245						250							

30

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

35

<220>
 <223> variante de L. kéfir

40

<400> 65

45

```

atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggatcggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggcgcgct 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg gggttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaaactg ctgtccgtta atctggatgg tgtttttttc 360
ggcaccgctc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtagggcat cgcacgctgg gggcatacaa cgcttccaag 480
ggggcgggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgctgtca acacagtaca tccgggcccc atcaagacc cgctgctcga tgattgggaa 600
ggtgctgagg aatgatgtc acagcgtacg aaaaccctta tgggccacat tggatgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgatg aatcgaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga 759
    
```

55

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

60

<220>
 <223> variante de L. kéfir

65

<400> 66

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

<210> 67

<211> 759

<212> ADN

<213> Secuencia Artificial

<220>

<223> variante de L. kéfir

<400> 67

50 atgaccgatc gtctgaaggc caaagtagcc atcgtaaccg gcgggacact gggatcgggt 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgctcgt 120
 gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
 55 cagcaccgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 ttcggcccgg ttaccgaccgt cgtgaacaat gcagggattt ttggtggtta aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccggtta atctggatgg tgtttttttc 360
 ggcacccgctc tgggcattca gcgcatgaaa aataaaggct tgggocgctag catcatcaat 420
 atgagcagta ttttcgggat ggtagggcat ccgacgctgg gggcatacaa cgcttccaag 480
 60 ggggocggtac gtatcatgct gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
 gtgctgtca acacagtaca tccgggcccg atcaagacc cggtgctcga tgattgggaa 600
 ggtgctgagg aatgatgct acagcgtacg aaaacccta tgggccacat tggatgaaccg 660
 aatgacatcg catgggtctg tgtgtacctg gcatctgatg aatcgaaatt tgcgacgggt 720
 65 gcagaatttg tggtcgacgg caggtggacc gcacagtga 759

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

5

<220>
 <223> variante de L. kéfir

<400> 68

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
 15 **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
 20 **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 Phe Val Val
85 90 95
 25 **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
 30 **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
 35 **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
 40 **Thr Pro Val 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
 45 **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 Arg Trp Thr Ala Gln
245 250

50 <210> 69
 <211> 759
 <212> ADN
 <213> Secuencia Artificial

55 <220>
 <223> variante de L. kéfir

<400> 69

60

65

5
 10
 15

```

atgaccgatc gtctgaaggg caaagtagcc atcgtaaccg gcgggacact gggatcggg 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgctcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgatgaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg aagttggtta aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcaccctgc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgaaggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgctgtca acacagtaca tccgggccat atcaagacc cggtgctcga tgatctggaa 600
ggtgctgagg aatgatgtc acagcgtacg aaaacccta tgggccacat tgggtaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcactctgatg aatcgaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggcattac gcacagtga 759
  
```

20

```

<210> 70
<211> 252
<212> PRT
<213> Secuencia Artificial
  
```

25

```

<220>
<223> variante de L. kéfir
  
```

25

```

<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
30 20 25 30
Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
35 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 Asp Glu Ala
65 70 75 80
Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Glu 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
Glu 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 His Ile Lys
180 185 190
Thr Pro Val 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 His Tyr Ala Gln
245 250
  
```

65

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

5
 <220>
 <223> variante de L. kéfir

10
 <400> 71

```

atgaccgatc gtctgaaggc caaagtagcc atcgtaaccg gcgggacact gggtatcggc 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgctcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggcg ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
15 ttcggcccgg ttacgaccgt cgtgaacaat gcagggattt ttggtgtaa aagcgttgaa 300
gacactacca cgaaggaatg gcgtaaactg ctgtccgcta atctggatgg tgtttttttc 360
ggcaccgctc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttgatgggat ggtagcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcaactgaa ggactacgat 540
20 gtgctgtca acacagtaca tccgggctat atcaagacc cgctgctcga tgattgggaa 600
ggtgctgagg aatgatgtc acagcgtacg aaaacccta tgggccacat tggatgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcactctgat aatcgaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggtggacc gcacagtga 759
    
```

25
 <210> 72
 <211> 252
 <212> PRT
 <213> Secuencia Artificial

30
 <220>
 <223> variante de L. kéfir

<400> 72

```

35 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
40 Lys Val Val Ile Thr Gly Arg Arg Ala Asp Val Gly Glu Lys Ala Ala
    35 40 45
Lys Ser Ile Gly Gly Ala 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
45 Phe Gly Pro Val Thr Thr Val Val Asn Asn Ala Gly Ile Phe Val Val
    85 90 95
Lys Ser Val Glu Asp Thr Thr Thr Lys Glu Trp Arg Lys Leu Leu Ser
    100 105 110
50 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
Asp Gly Met Val Gly Asp Pro Thr Leu Gly Ala Tyr Asn Ala Ser Lys
    145 150 155 160
55 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 Leu Asp Asp Trp Glu Gly Ala Glu Glu Met Met Ser Gln
    195 200 205
60 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
65 Ala Glu Phe Val Val Asp Gly Gly Trp Thr Ala Gln
    245 250
    
```


ES 2 547 528 T3

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

5

<220>
 <223> variante de L. kéfir

<400> 73

10

```

atgaccgatc gtctgaaggg caaagtagcc ttcgtaaccg gcgggacact gggatcgggt 60
ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgctcgt 120
gcggatgtag gtgaaaaggc cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg aagttgttaa aagcgttgaa 300
gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgttttttc 360
ggcaccgctc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
atgagcagta ttttcgggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgcactgaa ggactacgat 540
gtgcgtgtca acacagtaca tccgggcccg atcaagacc cggtgctcga tgattgggaa 600
ggtgctgagg aatgatgtc acagcgtacg aaaacccta tgggccacat tgggtgaaccg 660
aatgacatcg catgggtctg tgtgtacctg gcatctgatg aatcgaaatt tgcgacgggt 720
gcagaatttg tggtcgacgg cgggcatacc gcacagtga 759
    
```

25

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

30

<220>
 <223> variante de L. kéfir

35

<400> 74

40

45

50

55

60

65

	Met	Thr	Asp	Arg	Leu	Lys	Gly	Lys	Val	Ala	Phe	Val	Thr	Gly	Gly	Thr
	1				5					10					15	
5	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			
10	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
15	Phe	Gly	Pro	Val	Thr	Thr	Val	Val	Asn	Asn	Ala	Gly	Ile	Glu	Val	Val
					85					90					95	
	Lys	Ser	Val	Glu	Asp	Thr	Thr	Thr	Glu	Glu	Trp	Arg	Lys	Leu	Leu	Ser
				100					105					110		
20	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				
25	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	
30	Lys	Asp	Tyr	Asp	Val	Arg	Val	Asn	Thr	Val	His	Pro	Gly	Pro	Ile	Lys
				180					185					190		
	Thr	Pro	Val	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				
35	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	His	Thr	Ala	Gln				
					245					250						

40 <210> 75
 <211> 759
 <212> ADN
 <213> Secuencia Artificial

45 <220>
 <223> variante de L. kéfir

<400> 75

50	atgaccgatc	gtctgaaggg	caaagtagcc	ttcgtaaccc	gcgggacact	gggtatcggg	60
	ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttattac	tggtcgtcgt	120
	gcggatgtag	gtgaaaaggc	cgccaaatca	atcggcggca	ctgatgttat	tcgctttgtc	180
	cagcacgatg	cgtccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
55	ttcggcccgg	ttacgaccgt	cgtgaacaat	gcagggattg	gggttggtta	aagcgttgaa	300
	gacactacca	cggaggaatg	gcgtaaactg	ctgtccgtta	atctggatgg	tgtttttttc	360
	ggcaccgctc	tgggcattca	gcgcatgaaa	aataaaggct	tgggcgctag	catcatcaat	420
60	atgagcagta	ttttcgggat	ggtaggcgat	ccgacgctgg	gggcatacaa	cgcttccaag	480
	ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcactgaa	ggactacgat	540
	gtgcgtgtca	acacagtaca	tccgggccat	atcaagacc	cggtgctcga	tgattgggaa	600
	ggtgttgagg	aatgatgtc	acagcgtacg	aaaaccctta	tgggccacat	tggtgaaccg	660
	aatgacatcg	catgggtctg	tgtgtacctg	gcatctgatg	aatcgaaatt	tgcgacgggt	720
65	gcagaatttg	tggtcgacgg	cgggcatacc	acacagtga			759

<210> 76
 <211> 252
 <212> PRT
 <213> Secuencia Artificial
 5
 <220>
 <223> variante de L. kéfir
 <400> 76
 10
Met Thr Asp Arg Leu Lys Gly Lys Val Ala Phe 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 His Ile Lys
180 185 190
Thr Pro Val Leu Asp Asp Trp Glu Gly Val 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 His Thr Thr Gln
245 250

50 <210> 77
 <211> 759
 <212> ADN
 <213> Secuencia Artificial
 <220>
 55 <223> variante de L. kéfir
 <400> 77

60

65

atgaccgatc gtctgaaggg caaagtagcc ttggtaacgg gcgggacact gggatcgggt 60
 ttggcaatcg ccgataaatt tgtagaggag ggtgcgaaag tagttattac tggtcgctcgt 120
 gcggatgtag gtgaaaaggg cgccaaatca atcggcggca ctgatgttat tcgctttgtc 180
 5 cagcacgatg cgtccgatga agcaggctgg acgaaactgt tcgacaccac cgaggaggca 240
 ttcggcccgg ttacgaccgt cgtgaacaat gcagggattg aagttgttaa aagcgttgaa 300
 gacactacca cggaggaatg gcgtaaactg ctgtccgtta atctggatgg tgtttttttc 360
 ggcacccgctc tgggcattca gcgcatgaaa aataaaggct tgggcgctag catcatcaat 420
 10 atgagcagta ttgaagggat ggtaggcgat ccgacgctgg gggcatacaa cgcttccaag 480
 ggggcggtac gtatcatgtc gaaaagcgca gcgctggatt gcgactgaa ggactacgat 540
 gtgctgtca acacagtaca tccgggccat atcaagacc cgtgctcga tgatctggaa 600
 ggtgctgagg aatgatgtc acagcgaacg aaaaccctc tgggccacat tgggtgaaccg 660
 15 aatgacatcg catgggtctg tgtgtacctg gcactctgat aatcgaatt tgcgacgggt 720
 gcagaatttg tggtcgacgg aaagcatacc gcacagtga 759

<210> 78

<211> 252

<212> PRT

20 <213> Secuencia Artificial

<220>

<223> variante de L. kéfir

25 <400> 78

	Met	Thr	Asp	Arg	Leu	Lys	Gly	Lys	Val	Ala	Leu	Val	Thr	Gly	Gly	Thr
	1				5					10					15	
30	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			
35	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
40	Phe	Gly	Pro	Val	Thr	Thr	Val	Val	Asn	Asn	Ala	Gly	Ile	Glu	Val	Val
				85					90					95		
	Lys	Ser	Val	Glu	Asp	Thr	Thr	Thr	Glu	Glu	Trp	Arg	Lys	Leu	Leu	Ser
				100					105					110		
45	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				
50	Glu	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	His	Ile	Lys
			180					185					190			
55	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				
60	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	Lys	His	Thr	Ala	Gln				
				245						250						

65

ES 2 547 528 T3

<210> 79
 <211> 759
 <212> ADN
 <213> Secuencia Artificial

5

<220>
 <223> variante de L. kéfir

<400> 79

10	atgaccgatc	gtctgaaggg	caatggagcc	ttcgtaaccg	gcgggacact	gggtatcggg	60
	ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttattac	tggtcgtcgt	120
	gcggatgtag	gtgaaaaggc	cgccaaatca	atcggcggca	ctgatgttat	tcgctttgtc	180
	cagcacgatg	cgtccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
15	ttcggccccg	ttacgaccgt	cgtgaacaat	gcagggattg	gggttgtaa	aagcgttgaa	300
	gacactacca	cggaggaatg	gcgtaaactg	ctgtccgcta	atctggatgg	tgtttttttc	360
	ggcaccgctc	tgggcattca	gcgcatgaaa	aataaaggct	tgggcgctag	catcatcaat	420
	atgagcagta	ttttcgggat	ggtaggcgat	ccgacgctgg	gggcatacaa	cgcttccaag	480
20	ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcactgaa	ggactacgat	540
	gtgctgtca	acacagtaca	tccgggcccc	atcaagacct	cgctgctcga	tgatctggaa	600
	ggtgctgagg	aaatgatgtc	acagcgtacg	aaaaccctta	tgggccacat	tggtgaaccg	660
	aatgacatcg	catgggtctg	tgtgtacctg	gcatctgatg	aatcgaatt	tgcgacgggt	720
	gcagaatttg	tggtcgacgg	aaagaggacc	gcacagtga			759

25

<210> 80
 <211> 252
 <212> PRT
 <213> Secuencia Artificial

30

<220>
 <223> variante de L. kéfir
 <400> 80

35

40

45

50

55

60

65

	Met	Thr	Asp	Arg	Leu	Lys	Gly	Asn	Gly	Ala	Phe	Val	Thr	Gly	Gly	Thr
	1				5					10					15	
5	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			
10	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
15	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		
20	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				
25	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	
30	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				
35	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	Lys	Arg	Thr	Ala	Gln				
				245						250						

<210> 81
 <211> 759
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> variante de L. kéfir

<400> 81

	atgaccgctc	gtttgaagg	caaagtaacc	atcgtaacgg	gcgggacact	gggtattggt	60
	ttggcaatcg	ccgataaatt	tgtagaggag	ggtgcgaaag	tagttattac	tggtcgtcgt	120
	gcggatgtag	gtgaaaaggc	cgccaaatca	atcggcggca	ctgatgttat	tcgctttgtc	180
55	cagcacgatg	cgtccgatga	agcaggctgg	acgaaactgt	tcgacaccac	cgaggaggca	240
	ttcggcccgg	ttacgaccgt	cgtgaacaat	gcagggattg	gggttgtaa	aagcgttgaa	300
	gacactacca	cggaggaatg	gcgtaaactg	ctgtccgtta	atctggatgg	tgtttttttc	360
	ggcaccgctc	tgggcattca	gcgcatgaaa	aataaaggct	tgggcgctag	catcatcaat	420
	atgagcagta	ttgaagggat	ggtaggcgat	ccgacgctgg	gggcatacaa	cgcttccaag	480
60	ggggcggtac	gtatcatgtc	gaaaagcgca	gcgctggatt	gcgcactgaa	ggactacgat	540
	gtgctgtca	acacagtaca	tccgggcccg	atcaagacc	cgctgctcga	tgattgggaa	600
	ggtgctgagg	aatgatgtc	acagcgtacg	aaaacccta	tgggccacat	tggtgaaccg	660
	aatgacatcg	cgtgggtctg	tgtgtacctg	gcacatctgatg	aatcgaaatt	tgcgacgggt	720
65	gcagaatttg	tggtcgcacgg	caagcatacc	gcacagtga			759

- 5
 <210> 82
 <211> 252
 <212> PRT
 <213> Secuencia Artificial
- 10
 <220>
 <223> variante de L. kéfir
- 15
 <400> 82
- 20
 <210> 83
 <211> 252
 <212> PRT
 <213> Secuencia Artificial
- 25
 <220>
 <223> Fórmula Secuencial de Cetorreductasa Diseñada con base al esqueleto de L. brevis
 <223> Estructura Sintética
- 30
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(2)..(2)
 <223> Xaa es un residuo polar, no polar o alifático
- 35
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(4)..(4)
 <223> Xaa es un residuo básico o cisteína
- 40
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(11)..(11)
 <223> Xaa es un residuo alifático, no polar o aromático
- 45
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(40).. (40)
 <223> Xaa es un residuo restringido o básico
- 50
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(80).. (80)
 <223> Xaa es un residuo alifático, polar o no polar
- 55
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(86)..(86)
 <223> Xaa es un residuo polar, alifático o no polar
- 60
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(94)..(94)
 <223> Xaa es un residuo alifático o polar
- 65
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(96)..(96)
 <223> Xaa es un residuo polar, aromático, alifático o no polar
- 70
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(105)..(105)
 <223> Xaa es un residuo alifático, no polar, básico o ácido
- 75
 <220>

- <221> CARACTERÍSTICA_MISC
 <222>(129)..(129)
 <223> Xaa es un residuo polar o no polar
- 5 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(147)..(147)
 <223> Xaa es un residuo aromático, no polar o alifático
- 10 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(153)..(153)
 <223> Xaa es un residuo alifático, no polar o polar
- 15 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(190)..(190)
 <223> Xaa es un residuo aromático o restringido
- 20 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(195)..(195)
 <223> Xaa es un residuo no polar o alifático
- 25 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(196)..(196)
 <223> Xaa es un residuo no polar o alifático
- 30 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(199)..(199)
 <223> Xaa es un residuo alifático, polar o restringido
- 35 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(202)..(202)
 <223> Xaa es valina o leucina
- 40 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(206)..(206)
 <223> Xaa es un residuo no polar o aromático
- 45 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(226)..(226)
 <223> Xaa es un residuo no polar o alifático
- 50 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(248)..(248)
 <223> Xaa es un residuo no polar o básico
- 55 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(249)..(249)
 <223> Xaa es un residuo aromático
- 60 <400> 83
- 65

	Met	Xaa	Asn	Xaa	Leu	Asp	Gly	Lys	Val	Ala	Xaa	Ile	Thr	Gly	Gly	Thr
	1				5					10					15	
5	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	Lys	Ala	Ala
			35					40					45			
10	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	Xaa
	65					70					75					80
15	Phe	Gly	Pro	Val	Ser	Xaa	Leu	Val	Asp	Asn	Ala	Gly	Ile	Xaa	Val	Xaa
						85				90						95
	Lys	Ser	Val	Glu	Glu	Thr	Thr	Thr	Xaa	Glu	Trp	Arg	Lys	Leu	Leu	Ala
				100					105					110		
20	Val	Asn	Leu	Asp	Gly	Val	Phe	Phe	Gly	Thr	Arg	Leu	Gly	Ile	Gln	Arg
			115					120					125			
	Xaa	Lys	Asn	Lys	Gly	Leu	Gly	Ala	Ser	Ile	Ile	Asn	Met	Ser	Ser	Ile
		130					135					140				
25	Glu	Gly	Xaa	Val	Gly	Asp	Pro	Ser	Xaa	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	
30	Lys	Asp	Tyr	Asp	Val	Arg	Val	Asn	Thr	Val	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			
35	Arg	Thr	Lys	Thr	Pro	Met	Gly	His	Ile	Gly	Glu	Pro	Asn	Asp	Ile	Ala
		210					215					220				
	Tyr	Xaa	Cys	Val	Tyr	Leu	Ala	Ser	Asn	Glu	Ser	Lys	Phe	Ala	Thr	Gly
	225					230					235					240
40	Ser	Glu	Phe	Val	Val	Asp	Gly	Xaa	Xaa	Thr	Ala	Gln				
					245					250						

<210> 84

<211> 252

<212> PRT

45 <213> Secuencia Artificial

<220>

<223> Fórmula secuencial de cetorreductasa diseñada con base en el esqueleto de L. kéfir

50 <223> Estructura Sintética

<220>

<221> CARACTERÍSTICA_MISC

<222>(2)..(2)

55 <223> Xaa es un residuo polar, no polar o alifático

<220>

<221> CARACTERÍSTICA_MISC

<222>(4)..(4)

60 <223> Xaa es un residuo básico o de cisteína

<220>

<221> CARACTERÍSTICA_MISC

<222>(11)..(11)

65 <223> Xaa es un residuo alifático, no polar o aromático

- 5
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(40)..(40)
 <223> Xaa es un residuo restringido o básico
- 10
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(80)..(80)
 <223> Xaa es un residuo alifático, no polar o polar
- 15
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(86)..(86)
 <223> Xaa es un residuo polar, alifático o no polar
- 20
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(94)..(94)
 <223> Xaa es un residuo alifático polar
- 25
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(96)..(96)
 <223> Xaa es un residuo polar, aromático, alifático o no polar
- 30
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(105)..(105)
 <223> Xaa es un residuo alifático, no polar, básico o ácido
- 35
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(129)..(129)
 <223> Xaa es un residuo no polar o polar
- 40
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(147)..(147)
 <223> Xaa es un residuo aromático, no polar o alifático
- 45
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(153)..(153)
 <223> Xaa es un residuo alifático, no polar o polar
- 50
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(190)..(190)
 <223> Xaa es un residuo aromático o restringido
- 55
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(195)..(195)
 <223> Xaa es un residuo no polar o alifático
- 60
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(196)..(196)
 <223> Xaa es un residuo no polar o alifático
- 65
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(199)..(199)
 <223> Xaa es un residuo alifático, polar o restringido
- <220>

<221> CARACTERÍSTICA_MISC
 <222>(202)..(202)
 <223> Xaa es valina o leucina

5 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(206)..(206)
 <223> Xaa es un residuo no polar o aromático

10 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(226)..(226)
 <223> Xaa es un residuo no polar o alifático

15 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(248)..(248)
 <223> Xaa es un residuo no polar o básico

20 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(249)..(249)
 <223> Xaa es un residuo aromático

25 <400> 84

	Met	Xaa	Asp	Xaa	Leu	Lys	Gly	Lys	Val	Ala	Xaa	Val	Thr	Gly	Gly	Thr
	1				5					10					15	
30	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	Lys	Ala	Ala
				35				40					45			
35	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	Xaa
	65					70					75					80
40	Phe	Gly	Pro	Val	Thr	Xaa	Val	Val	Asn	Asn	Ala	Gly	Ile	Xaa	Val	Xaa
					85						90				95	
	Lys	Ser	Val	Glu	Asp	Thr	Thr	Thr	Xaa	Glu	Trp	Arg	Lys	Leu	Leu	Ser
				100					105					110		
45	Val	Asn	Leu	Asp	Gly	Val	Phe	Phe	Gly	Thr	Arg	Leu	Gly	Ile	Gln	Arg
				115					120				125			
	Xaa	Lys	Asn	Lys	Gly	Leu	Gly	Ala	Ser	Ile	Ile	Asn	Met	Ser	Ser	Ile
	130						135					140				
50	Glu	Gly	Xaa	Val	Gly	Asp	Pro	Thr	Xaa	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	
55	Lys	Asp	Tyr	Asp	Val	Arg	Val	Asn	Thr	Val	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			
60	Arg	Thr	Lys	Thr	Pro	Met	Gly	His	Ile	Gly	Glu	Pro	Asn	Asp	Ile	Ala
	210						215					220				
	Trp	Xaa	Cys	Val	Tyr	Leu	Ala	Ser	Asp	Glu	Ser	Lys	Phe	Ala	Thr	Gly
	225					230					235					240
65	Ala	Glu	Phe	Val	Val	Asp	Gly	Xaa	Xaa	Thr	Ala	Gln				
				245						250						

<210> 85
 <211> 759
 <212> ADN
 <213> Lacto bacilo Minor

5
 <400> 85

```

atgaccgatac ggttgaaggg gaaagtagca attgtaactg gcggtacctt ggaattggc      60
ttggcaatcgc ctgataagtt tgttgaagaa ggcgcaaagg ttggtattac 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
ggtaccgctc ttggaatcca acgtatgaag aataaaggac tcggagcatc aatcatcaat      420
atgtcatcta tcgaagggtt tgttggatg ccagctctgg gtgcatacaa cgcttcaaaa      480
ggtgctgtca gaattatgct taaatcagct gccttggatt gcgctttgaa ggactacgat      540
gttcgggtta aactgtttca tccaggttat atcaagacac cattggttga cgatctttaa      600
ggggcagaag aatgatgctc acagcggacc aagacaccaa tgggtcatat cggtgaacct      660
aacgatatcg cttggatctg tgtttacctg gcatctgacg aatctaaatt tgccactggt      720
gcagaattcg ttgtcgacgg agggtagacc gcccaatag      759
    
```

<210> 86
 <211> 252
 <212> PRT
 <213> Lacto bacilo Minor

25
 <400> 86

```

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
35 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
40 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
45 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
50 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
55 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
60 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
65 Ala Glu Phe Val Val Asp Gly Gly Tyr Thr Ala Gln
    245 250
    
```

- <210> 87
 <210> 87
 <211> 252
 <212> PRT
 5 <213> Secuencia Artificial
- <220>
 <223> Fórmula Secuencial de Cetorreductasa Diseñada en base al esqueleto de L. menor
- 10 <223> Estructura Sintética
 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(2)..(2)
 <223> Xaa es un residuo polar, no polar o alifático
- 15 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(4)..(4)
 <223> Xaa es un residuo básico o de cisteína
- 20 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(11)..(11)
 <223> Xaa es un residuo alifático, no polar o aromático
- 25 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(40)..(40)
 <223> Xaa es un residuo restringido o básico
- 30 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(80)..(80)
 <223> Xaa es un residuo alifático, no polar o polar
- 35 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(86)..(86)
 <223> Xaa es un residuo polar, alifático o no polar
- 40 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(94)..(94)
 <223> Xaa es un residuo alifático o polar
- 45 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(96)..(96)
 <223> Xaa es un residuo polar, aromático, alifático o no polar
- 50 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(105)..(105)
 <223> Xaa es un residuo alifático, no polar, básico o ácido
- 55 <220>
 <221> CARACTERÍSTICA_MISC
 <222> (129) . . (129)
 <223> Xaa es un residuo no polar o polar
- 60 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(147)..(147)
 <223> Xaa es un residuo aromático, no polar o alifático
- 65 <220>

- <221> CARACTERÍSTICA_MISC
 <222>(153)..(153)
 <223> Xaa es un residuo alifático, no polar o polar
- 5 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(190)..(190)
 <223> Xaa es un residuo aromático o restringido
- 10 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(195)..(195)
 <223> Xaa es un residuo no polar o alifático
- 15 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(196)..(196)
 <223> Xaa es un residuo no polar o alifático
- 20 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(199)..(199)
 <223> Xaa es un residuo alifático, polar o restringido
- 25 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(202)..(202)
 <223> Xaa es valina o leucina
- 30 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(206)..(206)
 .<223> Xaa es un residuo no polar o aromático
- 35 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(226)..(226)
 <223> Xaa es un residuo no polar o alifático
- 40 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(248)..(248)
 <223> Xaa es un residuo no polar o básico
- 45 <220>
 <221> CARACTERÍSTICA_MISC
 <222>(249)..(249)
 <223> Xaa es un residuo aromático
- 50 <400> 87
- 55
- 60
- 65

	Met	Xaa	Asp	Xaa	Leu	Lys	Gly	Lys	Val	Ala	Xaa	Val	Thr	Gly	Gly	Thr
	1				5					10					15	
5	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	Lys	Ala	Ala
			35					40					45			
10	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	Xaa
	65					70					75					80
15	Phe	Gly	Pro	Val	Thr	Xaa	Val	Val	Asn	Asn	Ala	Gly	Ile	Xaa	Val	Xaa
					85					90					95	
	Lys	Ser	Val	Glu	Asp	Thr	Thr	Thr	Xaa	Glu	Trp	Arg	Lys	Leu	Leu	Ser
				100					105					110		
20	Val	Asn	Leu	Asp	Gly	Val	Phe	Phe	Gly	Thr	Arg	Leu	Gly	Ile	Gln	Arg
			115					120					125			
	Xaa	Lys	Asn	Lys	Gly	Leu	Gly	Ala	Ser	Ile	Ile	Asn	Met	Ser	Ser	Ile
	130						135					140				
25	Glu	Gly	Xaa	Val	Gly	Asp	Pro	Ala	Xaa	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
30	Lys	Asp	Tyr	Asp	Val	Arg	Val	Asn	Thr	Val	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			
35	Arg	Thr	Lys	Thr	Pro	Met	Gly	His	Ile	Gly	Glu	Pro	Asn	Asp	Ile	Ala
	210						215					220				
	Trp	Xaa	Cys	Val	Tyr	Leu	Ala	Ser	Asp	Glu	Ser	Lys	Phe	Ala	Thr	Gly
40	225					230					235					240
	Ala	Glu	Phe	Val	Val	Asp	Gly	Xaa	Xaa	Thr	Ala	Gln				
					245					250						

Reivindicaciones

- 5 1. Un polipéptido de cetorreductasa capaz de convertir el sustrato, metil-2-benzamidometil-3-oxobutirato, al producto, 2S, 3R-metil-2-benzamidometil-3-hidroxibutirato, con un exceso estereomérico porcentual de por lo menos 60%, lo que comprende una secuencia de aminoácidos:
- 10 (i) que es por lo menos 85% idéntica a la secuencia referencial que se basa en la IDENTIFICACIÓN SECUENCIAL NÚMERO: 2, 4 o 86 teniendo las siguientes características: el residuo correspondiente a X 94 es treonina; el residuo correspondiente a X 199 es histidina y el residuo correspondiente a X 202 es valina o leucina; y
- (ii) en la cual el residuo correspondiente a X 94 es alanina o treonina; el residuo correspondiente a X 199 es alanina, histidina o asparagina; y el residuo correspondiente a X 202 es valina o leucina.
- 15 2. El polipéptido de la declaración 1 en el cual el residuo correspondiente a X 94 es treonina.
3. El polipéptido de la declaración 1 o 2 en el cual la secuencia de aminoácidos de cetorreductasas tiene adicionalmente una o más de las siguientes características:
- 20 el residuo correspondiente a X2 es alanina;
 el residuo correspondiente a X4 es cisteína;
 el residuo correspondiente a X 11 es fenilalanina;
 el residuo correspondiente a X 40 es arginina;
 el residuo correspondiente a X 80 es treonina;
 el residuo correspondiente a X 86 es isoleucina;
- 25 el residuo correspondiente a X 96 es valina o fenilalanina;
 el residuo correspondiente a X 105 es glicina;
 el residuo correspondiente a X 129 es treonina;
 el residuo correspondiente a X 147 es metionina o leucina;
 el residuo correspondiente a X 153 es alanina o serina;
- 30 el residuo correspondiente a X 190 es histidina o prolina;
 el residuo correspondiente a X 195 es valina;
 el residuo correspondiente a X 196 es leucina;
 el residuo correspondiente a X 206 es fenilalanina;
 el residuo correspondiente a X 226 es valina;
- 35 el residuo correspondiente a X 248 es lisina o arginina;
 el residuo correspondiente a X 249 es triptófano;
 donde opcionalmente la secuencia de aminoácidos tiene uno o más diferencias de residuos en otras posiciones residuales de aminoácidos en comparación de la secuencia referencial.
- 40 4. El polipéptido de la declaración 1, que comprende una secuencia de aminoácidos correspondiente de las IDENTIFICACIONES SECUENCIALES NÚMEROS: 6, 8, 10, 12, 20, 22, 24, 30, 32, 34, 36, 38, 40, 42, 50, 56, y 58.
5. El polipéptido de la declaración 1, que es:
- 45 a) capaz de convertir el sustrato al producto con un exceso estereomérico porcentual de por lo menos 90%, y que opcionalmente comprende una secuencia de aminoácidos correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 8, 10, 12, 20, 22, 24, 30, 32, 34, 36, 40, 42, 50, 56, o 58;
- 50 b) Capaz de convertir el sustrato al producto con un exceso estereomérico porcentual de por lo menos 95%, y que opcionalmente comprende una secuencia de aminoácidos correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 8, 10, 12, 20, 22, 24, 30, 32, 34, 42, 50, 56, o 58; o
- 55 c) Capaz de convertir el sustrato al producto to con un exceso estereomérico porcentual de por lo menos el 99%, y que opcionalmente comprende una secuencia de aminoácidos correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 8, 10, 12, 20, 22, 24, 30, 32, o 34.
6. El polipéptido de la declaración uno, que es:
- 60 a) capaz de convertir el sustrato al producto to a una tasa que es por lo menos 15 veces mayor que el polipéptido referencial de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48, y que opcionalmente comprende una secuencia de aminoácidos correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 8, 10, 12, 20, 22, 24, 30, 32, 34, o 50; o
- 65 b) capaz de convertir el sustrato al producto a una tasa que es por lo menos 30 veces mayor que la del péptido referencial de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48, y que opcionalmente

comprende una secuencia de aminoácidos correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 6, 8, 10, 12, 20, 22, 24, 30, o 34; o

5 c) Capaz de convertir el sustrato al producto a una tasa que es por lo menos 40 veces mayor que la del polipéptido referencial de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48, y que opcionalmente comprende una secuencia de aminoácidos correspondiente a las IDENTIFICACIONES SECUENCIALES NÚMERO: 6, 8, 10, 12, y 22; o

10 d) Capaz de convertir el sustrato al producto a una tasa que es por lo menos 50 veces mayor que la del polipéptido referencial de la IDENTIFICACIÓN SECUENCIAL NÚMERO: 48, y que opcionalmente comprende una secuencia de aminoácidos correspondiente a las IDENTIFICACIONES SECUENCIALES NÚMEROS: 6, 8, 10, y 12.

15 7. Un polinucleótido que codifica a un polipéptido de acuerdo a cualquiera de las declaraciones 1 - 6, y que opcionalmente tiene una secuencia correspondiente a la IDENTIFICACIÓN SECUENCIAL NÚMERO: 5, 7, 9, 11, 19, 21, 23, 29, 31, 33, 35, 37, 39, 41, 49, 55 o 57.

20 8. Un vector de expresión que comprende el polinucleótido de la declaración 7 vinculado operacionalmente a secuencias de control adecuadas para dirigir la expresión en una célula anfitriona.

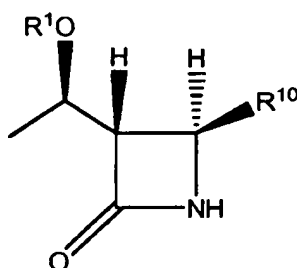
9. Una célula anfitriona que contiene el vector de expresión de la declaración 8.

25 10. Una composición que comprende una cetorreductasa de cualquiera de las declaraciones 1 - 6 y el compuesto metil-2-benzamidometil-3-oxobutirato de la fórmula (I) o el compuesto 2S,3R-metil-2-benzamidometil-3-hidroxi-butirato de la fórmula (II).

11. La composición de la declaración 10 en la cual el sustrato es el compuesto de la fórmula (I) y el compuesto de la fórmula (II).

30 12. Un método para reducir el sustrato de la fórmula (I), metil-2-benzamidometil-3-oxobutirato, al producto de la fórmula (II), 2S,3R-metil-2-benzamidometil-3-hidroxi-butirato, que comprende el contactar o incubar el sustrato con un polipéptido de cetorreductasa de cualquiera de las declaraciones 1 - 6 bajo condiciones de reacción adecuadas para reducir el sustrato al producto de la fórmula (II), y donde el producto está opcionalmente presente de con un exceso estereomérico mayor que un 99%.

35 13. Un método para la síntesis del intermediador de la fórmula (IVa).



(IVa)

50 Dónde R¹ es H o un grupo protector de hidroxilos, y R¹⁰ es un halógeno, u -Oac, donde Ac es acetato, donde un paso del método comprende contactar o reaccionar el sustrato de la fórmula (I) con las cetorreductasas de cualquiera de las declaraciones 1 - 6 bajo condiciones de reacción apropiadas para reducir o convertir el sustrato al producto de la fórmula (II).

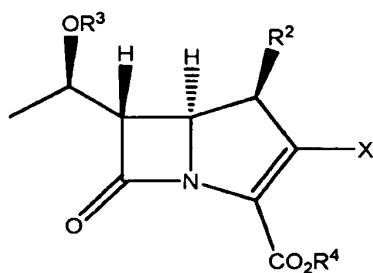
55 14. Un método para la síntesis del intermediador de la fórmula estructural (IX),

60

65

5

10



(IX)

15

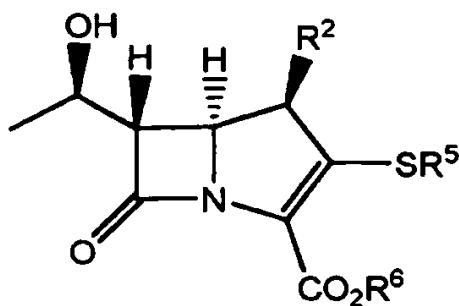
Donde R^2 es H o alquilo C1-C4 (por ejemplo, $-\text{CH}_3$); R^3 es H, o un grupo protector de hidroxilos; R^4 es H, un grupo protector de carboxis, un grupo amoniaco, un metal alcalino o un metal alcalinotérreo, Donde un paso en el método incluye contactar o reaccionar el sustrato de la fórmula (I) con las cetorreductasas de cualquiera de las declaraciones 1 - 6 bajo condiciones de reacción apropiadas para reducir o convertir el sustrato al producto de la fórmula (II).

20

15. Un método para la síntesis de un carbapenem de la fórmula estructural (V):

25

30



(V)

35

40

O sus solvatos, hidratos, sales y pro - medicamentos, donde R^2 es H o $-\text{CH}_3$; R^5 es seleccionado de un alquilo sustituido o no sustituido, un arilo sustituido o no sustituido, un heteroalquilo sustituido o no sustituido, un heterocicloalquilo sustituido o no sustituido, un heteroariloalquilo sustituido o no sustituido; y R^6 es H o un progrupo, donde un paso en el método incluye el contactar o reaccionar el sustrato de la fórmula (I) con las cetorreductasas de cualquiera de las declaraciones 1 - 6 bajo condiciones de reacción apropiadas para reducir o convertir el sustrato al producto de la fórmula (II).

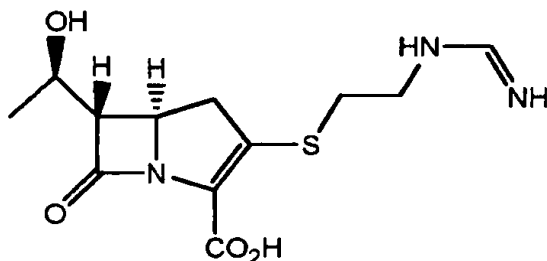
45

16. El método de la declaración 15 donde:

50

a) El carbapenem tiene la fórmula estructural (X):

55



(X); o

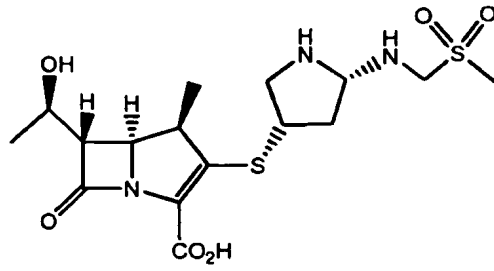
60

b) El carbapenem tiene la fórmula estructural (XI):

65

5

10



(XI); o

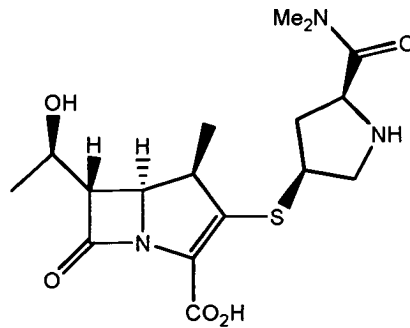
15

c) El carbapenem tiene la fórmula estructural (XII):

20

25

30



(XII); o

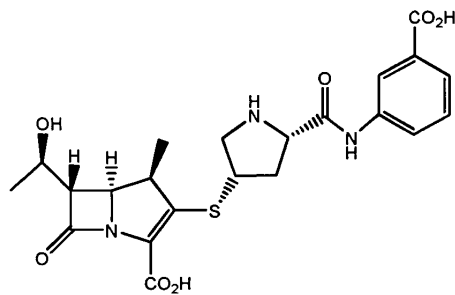
35

d) El carbapenem tiene la fórmula estructural (XIII):

40

45

50



(XIII).

55

60

65

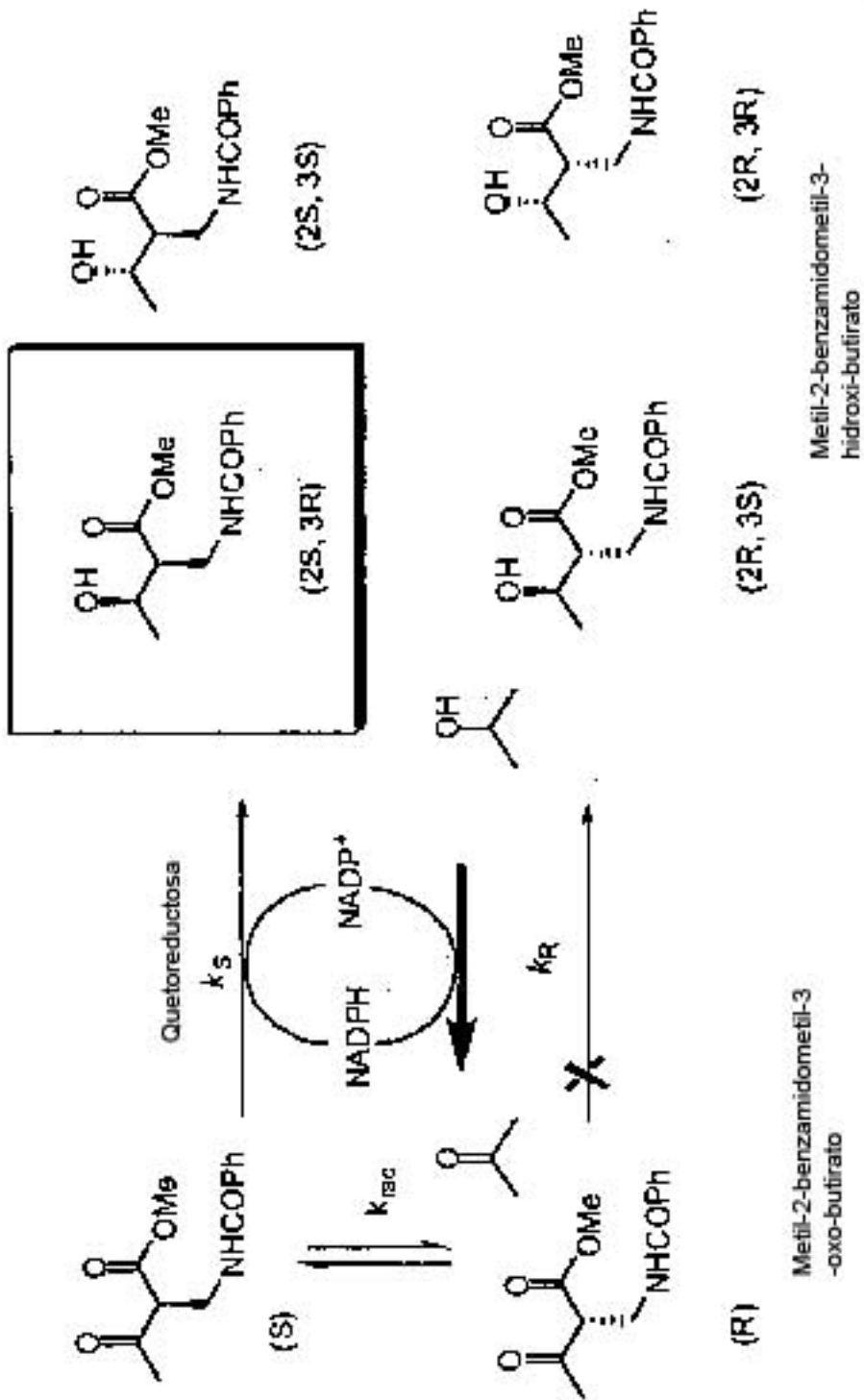


FIG. 1