(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 536 548**

(51) Int. CI.:

C12P 7/52 (2006.01)
C12P 7/42 (2006.01)
C12N 15/54 (2006.01)
C12N 9/10 (2006.01)
C12N 9/88 (2006.01)
C12N 1/21 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **14.09.2010 E 10765977 (3)**

(97) Fecha y número de publicación de la concesión europea: **01.04.2015 EP 2478109**

(54) Título: **Método para la producción enzimática de ácido 3-hidroxi-3-metilbutírico a partir de acetona y acetil-CoA**

(30) Prioridad:

15.09.2009 EP 09170312

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
26.05.2015

(73) Titular/es:

SCIENTIST OF FORTUNE S.A. (100.0%)
7a rue des Glacis
1628 Luxembourg, LU

(72) Inventor/es:

MARLIÈRE, PHILIPPE

(74) Agente/Representante:

LEHMANN NOVO, María Isabel

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Método para la producción enzimática de ácido 3-hidroxi-3-metilbutírico a partir de acetona y acetil-CoA

La presente invención se refiere a un método para la producción de ácido 3-hidroxi-3-metilbutírico (también denominado como beta-hidroxiisovalerato o HIV) a partir de acetona y acetil-coenzima A que comprende la conversión enzimática de acetona y acetil-coenzima A en ácido 3-hidroxi-3-metilbutírico. La conversión hace uso de una enzima que es capaz de catalizar la formación de un enlace covalente entre el átomo de carbono del grupo oxo (es decir, el C=O) de acetona y el grupo metilo de acetil-coenzima A. La enzima empleada en el procedimiento es una enzima con la actividad de una HMG CoA sintasa (EC 2.3.3.10). La presente invención también se refiere a organismos capaces de producir ácido 3-hidroxi-3-metilbutírico a partir de acetona y acetil-coenzima A, y al uso de la enzima y organismos mencionados anteriormente para la producción de ácido 3-hidroxi-3-metilbutírico. Finalmente, la presente invención se refiere al uso de acetona para la producción de ácido 3-hidroxi-3-metilbutírico.

El ácido 3-hidroxi-3-metilbutírico (también denominado como beta-hidroxiisovalerato o HIV; véase la Figura 1) es un metabolito del aminoácido esencial leucina, y se sintetiza en el cuerpo humano. Se puede encontrar en pequeñas cantidades en pomelo, alfalfa y bagre. También se sabe que aparece en algunos trastornos metabólicos del catabolismo de la leucina, es decir, acidemia hipovalérica. Se ha mostrado que el ácido 3-hidroxi-3-metilbutírico puede tener un efecto en el incremento del peso muscular y la resistencia (Nissen et al., J. Appl. Physiol. 81 (1996), 2095-2104). Wilson et al. (Nutrition & Metabolism 5 (2008)) propone como los mecanismos de acción lo siguiente:

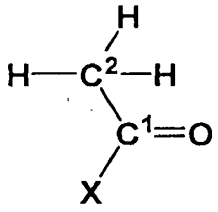
- integridad sacrolémica incrementada vía conversión mediante HMG CoA reductasa
- síntesis proteica potenciada vía la ruta de mTOR
- depresión de la degradación proteica a través de la inhibición de la ruta de ubiquitina. Se supone que el ácido 3-hidroxi-3-metilbutírico ayuda a los músculos a combatir la ruptura de las proteínas, ayuda en la reparación muscular, y respalda la resistencia incrementada. Se ha descrito que ayuda a pacientes con enfermedad pulmonar obstructiva crónica en unidades de cuidados intensivos hospitalarias, con desgaste muscular asociado con HIV y cáncer, y víctimas de traumas con lesiones graves. De este modo, es de interés comercial debido a su uso como potenciador muscular para culturismo, y como medicamento para evitar el desgaste muscular. La patente de los Estados Unidos 7026507 describe un procedimiento para preparar formulaciones sólidas de 3-hidroxi-3-metilbutirato de sodio en el que, en una primera etapa del procedimiento, se hace reaccionar 4,4-dimetiloxetan-2-ona con hidróxido sódico acuoso para formar una disolución de 3-hidroxi-3-metilbutirato de sodio, y después, si es apropiado tras la concentración, la disolución se aplica, en una etapa posterior del procedimiento, a sílice sintética, en la que, si es apropiado, se seca el producto resultante.

Sería deseable proporcionar un procedimiento para la producción de 3-hidroxi-3-metilbutirato que fuese independiente de las etapas de producción inorgánicas y que se pudiese efectuar en organismos vivos siendo de ese modo medioambientalmente bueno y barato. En este contexto, Lee et al. (Appl. Environ. Microbiol. 63 (1997), 4191-4195) describen un método para la producción de 3-hidroxi-3-metilbutirato convirtiendo ácido 3-metilbutírico en ácido 3-hidroxi-3-metilbutírico usando el microorganismo Galactomyces reessii. Además, Crouch et al. (Tetrahedron 53 (1997), 6993-7010) describen un método para la producción de 3-hidroxi-3-metilbutirato a partir de 2-oxo-4-metilpentanoato usando la enzima alfa-cetoisocaproato dioxigenasa. Sin embargo, aunque estos procedimientos permitieron la producción de 3-hidroxi-3-metilbutirato, todavía existe la necesidad de proporcionar vías eficientes alternativas y eficaces desde el punto de vista del coste para producir 3-hidroxi-3-metilbutirato, en particular mediante procesos biológicos.

La presente invención satisface esta demanda de un procedimiento alternativo para la producción de 3-hidroxi-3-metilbutirato, y proporciona un método que se basa en fuentes biológicas y permite producir 3-hidroxi-3-metilbutirato in vitro o in vivo en un microorganismo u otra especie.

En particular, la presente invención se refiere a un método para la producción de ácido 3-hidroxi-3-metilbutírico (también denominado como beta-hidroxiisovalerato o HIV) a partir de acetona y acetil-coenzima A, que comprende la conversión enzimática de acetona y acetil coenzima A en ácido 3-hidroxi-3-metilbutírico.

La acetona se representa mediante la siguiente fórmula: CH_3-(C=O)-CH_3. El compuesto que proporciona un grupo acetilo activado se caracteriza por la siguiente fórmula (I):



en la que X es S-CH2-CH2-NH-CO-CH2-CH2-NH-CO-CH(OH)-C(CH3)2-CH2-O-PO2H-O-PO2H-C10H13N5O7P (coenzima A).

La conversión hace uso de una enzima que es capaz de catalizar la formación de un enlace covalente entre el átomo de carbono del grupo oxo (es decir, el C=O) de acetona y el átomo de carbono (C^2) correspondiente al grupo metilo del compuesto que proporciona un grupo acetilo activado según la fórmula (I). Según este esquema de reacción, el grupo oxo de la acetona reacciona como un electrófilo, y el grupo metilo del compuesto que proporciona un grupo acetilo activado según la fórmula (I) reacciona como un nucleófilo. La reacción de la conversión de acetona y un compuesto que proporciona un grupo acetilo activado según la fórmula (I) se muestra en la Figura 5.

La reacción puede suceder en una etapa, es decir, 3-hidroxi-3-metilbutirato puede ser el producto directo de una reacción catalizada por la enzima descrita anteriormente. Como alternativa, la reacción puede comprender dos etapas, cuando se usa acetil CoA como el compuesto que proporciona un grupo acetilo activado, en el sentido de que primero se produce un aducto de 3-hidroxi-3-metilbutirato y el compuesto que proporciona un grupo acetilo activado, por ejemplo 3-hidroxi-3-metilbutiril-CoA, que se hidroliza subsiguientemente, por ejemplo, a 3-hidroxi-3-metilbutirato y CoA. De este modo, en la primera alternativa, la enzima cataliza la reacción completa como se muestra en la Figura 5. En la segunda alternativa, la enzima cataliza la formación de un enlace covalente entre el átomo de carbono del grupo oxo (es decir, el C=O) de acetona y el átomo de carbono (C^2) correspondiente al grupo metilo del compuesto que proporciona un grupo acetilo activado según la fórmula (I), pero X permanece en la molécula. X se elimina entonces subsiguientemente de la molécula mediante hidrólisis.

La presente invención muestra por primera vez que es posible producir 3-hidroxi-3-metilbutirato haciendo uso de una enzima que puede transferir un grupo acetilo activado a acetona. En la técnica anterior, se ha dado a conocer la producción de 3-hidroxi-3-metilbutirato a partir de ácido isovalérico mediante la bioconversión usando el hongo Galactomyces reessii. Sin embargo, considerando que el ácido isovalérico se obtiene de leucina a través de descarboxilación y que la propia leucina deriva del metabolismo a partir de la condensación global de dos moléculas de piruvato y una molécula de acetil CoA, este procedimiento de producción es energéticamente desfavorable. El procedimiento de la presente invención evita esta desventaja. El compuesto que proporciona un grupo acetilo activado es acetil CoA. Acetil CoA (también conocida como acetil Coenzima A) en estructura química es el tioéster entre la coenzima A (un tiol) y ácido acético.

La enzima empleada en el procedimiento es una enzima con la actividad de una HMG CoA sintasa (EC 2.3.3.10).

El método según la presente invención comprende la conversión enzimática de acetona y acetil CoA en 3-hidroxi-3-metilbutirato con una enzima que es capaz de catalizar la formación de un enlace covalente entre el átomo de carbono del grupo oxo (es decir, el C=O) de acetona y el átomo de carbono C^2 de acetil CoA según la fórmula (I). La enzima empleada en el procedimiento según la invención es una enzima que tiene la actividad de una HMG CoA sintasa (EC 2.3.3.10).

En particular, se ha mostrado en el contexto de la presente invención que HMG CoA sintasa puede aceptar acetona en lugar de su sustrato normal, acetoacetil-CoA, permitiendo de ese modo la conversión de acetil-CoA (o un compuesto según la fórmula (I)) y acetona en 3-hidroxi-3-metilbutirato.

En el contexto de la presente solicitud, la expresión “HMG CoA sintasa” o “una proteína/enzima que tiene la actividad de una HMG CoA sintasa” se refiere a cualquier enzima que se clasifica en el número EC EC 2.3.3.10 (antiguamente, HMG CoA sintasa se había clasificado como EC 4.1.3.5, pero se ha transferido a EC 2.3.3.10), en particular se refiere a cualquier enzima que es capaz de catalizar la reacción en la que acetil-CoA se condensa con acetoacetil-CoA para formar 3-hidroxi-3-metilglutaril-CoA (HMG-CoA) (véase la Figura 2), y la expresión también se refiere a cualquier enzima que deriva de tal HMG-CoA sintasa y que es capaz de catalizar la conversión de acetona y un compuesto que proporciona un grupo acetilo activado como se define anteriormente, preferiblemente acetil CoA, en 3-hidroxi-3-metilbutirato.

La actividad enzimática de acetil-CoA que se condensa con acetoacetil-CoA para formar 3-hidroxi-3-metilglutaril-CoA (HMG-CoA) se puede medir mediante métodos bien conocidos en la técnica. Un ensayo posible y usado preferiblemente se describe, por ejemplo, en Clinkenbeard et al. (J. Biol. Chem. 250 (1975), 3108-3116). En este ensayo, la actividad de HMG-CoA sintasa se mide monitorizando la disminución en la absorbancia a 303 nm que acompaña a la desaparición, dependiente de acetil-CoA, de la forma de enolato de acetoacetil-CoA. Preferiblemente, la actividad de HMG CoA sintasa se evalúa como se describe en el Ejemplo 3.

HMG CoA sintasa es parte de la ruta de mevalonato. Se han identificado dos rutas para la síntesis de pirofosfato de isopentenilo (IPP), es decir, la ruta de mevalonato y la ruta de gliceraldehído 3-fosfato-piruvato. HMG CoA sintasa cataliza la condensación de Claisen biológica de acetil-CoA con acetoacetil-CoA, y es un miembro de una superfamilia de enzimas acilcondensantes que incluye beta-cetotiolasas, ácido graso sintasas (proteína portadora beta-cetoacetil sintasa) y policétido sintasas.

HMG CoA sintasa se ha descrito para diversos organismos. También están disponibles las secuencias de aminoácidos y de ácidos nucleicos que codifican HMG CoA sintasas de numerosas fuentes. Generalmente, las secuencias solo comparten un bajo grado de identidad de secuencia global. Por ejemplo, las enzimas de

Staphylococcus o Streptococcus muestran solamente alrededor de 20% de identidad con aquellas de HMG CoA sintasa humana y aviar. En algunas fuentes, se da a conocer que las HMG CoA sintasas bacterianas y sus contrapartes animales muestran solamente alrededor de 10% de identidad de secuencia global (Sutherlin et al, J. Bacteriol. 184 (2002), 4065-4070). Sin embargo, los restos de aminoácidos implicados en las reacciones de acetilación y condensación están conservados entre HMG CoA sintasas bacterianas y eucariotas (Campobasso et al., J. Biol. Chem. 279 (2004), 44883-44888). Se ha determinado la estructura tridimensional de tres enzimas HMG CoA sintasas, y en principio los aminoácidos cruciales para la reacción enzimática están bien caracterizados (Campobasso et al., loc. cit.; Chun et al., J. Biol. Chem. 275 (2000), 17946-17953; Nagegowda et al., Biochem. J. 383 (2004), 517-527; Hegardt, Biochem. J. 338 (1999), 569-582). En eucariotas, existen dos formas para la HMG CoA sintasa, es decir, una forma citosólica y una forma mitocondrial. La forma citosólica desempeña un papel clave en la producción de colesterol y otros isoprenoides, y la forma mitocondrial está implicada en la producción de cuerpos cetónicos.

En principio, en el contexto de la presente invención se puede usar cualquier enzima HMG CoA sintasa, en particular procedente de organismos procariotas o eucariotas.

Las HMG CoA sintasas procariotas se describen, por ejemplo, de Staphylococcus aureus (Campobasso et al., loc. cit.; número de acceso Uniprot Q9FD87), Staphylococcus epidermidis (número de acceso Uniprot Q9FD76), Staphylococcus haemolyticus (número de acceso Uniprot Q9FD82), Enterococcus faecalis (Sutherlin et al., loc. cit.; número de acceso Uniprot Q9FD7), Enterococcus faecium (número de acceso Uniprot Q9FD66), Streptococcus pneumonia (número de acceso Uniprot Q9FD56), Streptococcus pyogenes (número de acceso Uniprot Q9FD61) y Methanobacterium thermoautotrophicum (número de acceso AE000857), Borrelia burgdorferi (número de acceso NCBI BB0683).

Además, la siguiente Tabla A enumera algunas HMG CoA sintasas conocidas de procariotas:

Tabla A

Número de acceso Swissprot/TrEmbl	Organismo
Q9YAS0	*Aeropyrum pernix*
A7Z4Y2	*Bacillus amyloliquefaciens*
P40830|2874037340	*Bacillus subtilis*
B8G795	*Chloroflexus aggregans*
A5EUV4	*Dichelobacter nodosus*
A5FM54	*Flavobacterium johnsoniae*
Q18GC4	*Haloquadratum walsbyi*
B9LS15	*Halorubrum lacusprofundi*
A9B8F0	*Herpetosiphon aurantiacus*
A2BMY8	*Hyperthermus butylicus*
Q5FLB7	*Lactobacillus acidophilus*
Q03QR0	*Lactobacillus brevis*
Q1GAH5	*Lactobacillus delbrueckii*
B2GBL1	*Lactobacillus fermentum*
B1MZ51	*Leuconostoc citreum*
Q03WZ0	*Leuconostoc mesenteroides*
A4YH99	*Metallosphaera sedula*
A5UNI8	*Methanobrevibacter smithii*
Q58941	*Methanocaldococcus jannaschii*
Q12UR3	*Methanococcoides burtonii*

Número de acceso Swissprot/TrEmbl	Organismo
A6USZ1	*Methanococcus aeolicus*
A4FWW6	*Methanococcus maripaludis*
A6UPL1	*Methanosarcina mazei*
A2STY2	*Methanocorpusculum labreanum*
Q8TVL0	*Methanopyrus_andleri*
Q8PYJ0	*Methanosarcina mazei*
Q2NHU7	*Methanosphaera stadtmanae*
Q2FPH4	*Methanospirillum hungatei*
B2HGT6	*Mycobacterium marinum*
Q3IMZ7	*Natronomonas pharaonis*
Q8EP69	*Oceanobacillus iheyensis*
Q04F95	*Oenococcus oeni*
Q03FU5	*Pediococcus pentosaceus*
Q6L233	*Picrophilds torridus*
A6G7N7	*Plesiocystis pacifica*
A4WJ12	*Pyrobaculum arsenaticum*
A7NHZ7	*Roseiflexus castenholzii*
Q8CN06	*Staphylococcus epidermidis*
Q4L958	*Staphylococcus haemolyticus*
Q4A0D6	*Staphylococcus saprophyticus*
B4U364	*Streptococcus equi*
Q8DUI5	*Streptococcus mutans*
Q4J933	*Sulfolobus acidocaldarius*
Q971K8	*Sulfolobus tokodaii*
Q9HI87	*Thermoplasma acidophilum*
Q31EW2	*Thiomicrospira crunogena*
Q51798	*Pyrococcus furiosus*
A5VJB7	*Lactobacillus reuteri*
Q7CF79	*Streptococcus pyogenes*
Q9UWU0	*Sulfolobus solfataricus*

Se describen HMG CoA sintasas eucariotas, por ejemplo, de hongos, tales como Schizosaccharomyces pombe (números de acceso U32187 y P54874), Saccharomyces cerevisiae (número de acceso P54839), plantas, tales como Arabidopsis thaliana (números de acceso X83882 y P54873), Pinus sylvestris (número de acceso X96386), y animales, tales como Caenorhabditis elegans (número de acceso P54871), Mus musculus (mitocondrial; número de acceso P54869 y Hegardt, Biochem. J. 338 (1999), 569-582), Rattus norvegicus (mitocondrial: número de acceso P22791 y Hegardt, Biochem. J. 338 (1999); citosólica: número de acceso P17425), 569-582), hámster chino

(Cricetulus griseus: número de acceso P13704), Sus scrofa (mitocondrial; número de acceso U90884 y Hegardt, Biochem. J. 338 (1999), 569-582), Homo sapiens (mitocondrial: número de acceso P54868 y Hegardt, Biochem. J. 338 (1999), 569-582; citosólica: número de acceso Q01581), Blattella germanica (forma citosólica 1; número de acceso P54961), Blattella germanica (forma citosólica 2; número de acceso P54870) y Gallus gallus (citosólica; número de acceso P23228).

En SEC ID NO: 1 a 14 se dan ejemplos de HMG CoA sintasas procedentes de diferentes organismos. SEC ID NO: 1 muestra la secuencia de la HMG CoA sintasa citoplásmica de Caenorhabditis elegans (P54871, gene bank F25B4.6); la SEC ID NO: 2 muestra la secuencia de la HMG CoA sintasa citoplásmica de Schizosaccharomyces pombe (levadura de fisión; P54874); la SEC ID NO: 3 muestra la secuencia de la HMG CoA sintasa citoplásmica de Saccharomyces cerevisiae (levadura de panadero; P54839, gene bank CAA65437.1); la SEC ID NO: 4 muestra la secuencia de la HMG CoA sintasa citoplásmica de Arabidopsis thaliana (berro de oreja de ratón; P54873); la SEC ID NO: 5 muestra la secuencia de la HMG CoA sintasa citoplásmica de Dictyostelium discoideum (moho del cieno; P54872, gene bank L2114); la SEC ID NO: 6 muestra la secuencia de la HMG CoA sintasa citoplásmica de Blattella germanica (cucaracha alemana; P54961, gene bank X73679); la SEC ID NO: 7 muestra la secuencia de la HMG CoA sintasa citoplásmica de Gallus gallus (pollo; P23228, gene bank CHKHMGCOAS); la SEC ID NO: 8 muestra la secuencia de la HMG CoA sintasa citoplásmica de Homo sapiens (ser humano; Q01581, gene bank X66435); la SEC ID NO: 9 muestra la secuencia de HMG CoA sintasa mitocondrial de Homo sapiens (ser humano; P54868, gene bank X83618); la SEC ID NO: 10 muestra la secuencia de la HMG CoA sintasa mitocondrial de Dictyostelium discoideum (moho del cieno; Q86HL5, gene bank XM_638984); la SEC ID NO: 11 muestra la secuencia de la HMG CoA sintasa de Staphylococcus epidermidis (Q9FD76); la SEC ID NO: 12 muestra la secuencia de la HMG CoA sintasa de Lactobacillus fermentum (B2GBL1); la SEC ID NO: 13 muestra la secuencia de la HMG CoA sintasa de Hyperthermus butylicus (A2BMY8); la SEC ID NO: 14 muestra la secuencia de la HMG CoA sintasa de Chloroflexus aggregans (B8G795); la SEC ID NO: 24 muestra la secuencia de la HMG CoA sintasa de Lactobacillus delbrueckii (Q1GAH5) y la SEC ID NO: 25 muestra la secuencia de la HMG CoA sintasa de Staphylococcus haemolyticus Q4L958 (198>V diferencia comparada con la proteína de tipo salvaje).

En una realización preferida de la presente invención, la HMG CoA sintasa es una enzima que comprende una secuencia de aminoácidos seleccionada del grupo que consiste en SEC ID NOs: 1 a 14, o una secuencia que es al menos n % idéntica a cualquiera de SEC ID NOs: 1 a 14 y que tiene la actividad de una HMG CoA sintasa, siendo n un número entero entre 10 y 100, preferiblemente 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98 ó 99.

Preferiblemente, el grado de identidad se determina comparando la secuencia respectiva con la secuencia de aminoácidos de una cualquiera de las SEC ID NOs mencionadas anteriormente. Cuando las secuencias que se comparan no tienen la misma longitud, el grado de identidad se refiere preferiblemente al porcentaje de restos de aminoácidos en la secuencia más corta que son idénticos a restos de aminoácidos en la secuencia más larga, o al porcentaje de restos de aminoácidos en la secuencia más larga que son idénticos a los restos de aminoácidos en la secuencia más corta. El grado de identidad de secuencia se puede determinar según métodos bien conocidos en la técnica, usando preferiblemente algoritmos de ordenador adecuados tales como CLUSTAL.

Cuando se usa el método de análisis de Clustal para determinar si una secuencia particular es, por ejemplo, 80% idéntica a una secuencia de referencia, se pueden usar ajustes por defecto, o los ajustes son preferiblemente como siguen: Matriz: blosum 30; Penalización de salto abierto: 10,0; Señalización de salto extendido: 0,05; Retraso divergente: 40; Distancia de separación de salto: 8 para comparaciones de secuencias de aminoácidos. Para comparaciones de secuencias nucleotídicas, la Penalización de salto extendido se ajusta preferiblemente a 5,0.

Preferiblemente, el grado de identidad se calcula a lo largo de toda la longitud de la secuencia.

La HMG CoA sintasa empleada en el procedimiento según la invención puede ser HMG CoA sintasa de origen natural, o puede ser HMG CoA sintasa que deriva de una HMG CoA sintasa de origen natural, por ejemplo mediante la introducción de mutaciones u otras alteraciones que, por ejemplo, alteran o mejoran la actividad enzimática, la estabilidad, etc.

La expresión “HMG CoA sintasa” o “una proteína/enzima que tiene la actividad de una HMG CoA sintasa”, en el contexto de la presente solicitud, también cubre enzimas que derivan de una HMG CoA sintasa, que son capaces de producir 3-hidroxi-3-metilbutirato mediante una conversión enzimática de acetona y acetil-coenzima A, pero que solo tienen una baja afinidad por acetoacetil-CoA como sustrato o ya no aceptan acetoacetil-CoA como sustrato. Tal modificación del sustrato preferido de una HMG CoA sintasa permite mejorar la conversión de acetona en 3-hidroxi-3-metilbutirato y reducir la producción del subproducto, por ejemplo HMG-CoA. Los métodos para modificar y/o mejorar las actividades enzimáticas deseadas de proteínas son bien conocidos por la persona experta en la técnica, e incluyen, por ejemplo, mutagénesis al azar o mutagénesis dirigida al sitio y la selección subsiguiente de enzimas que tienen las propiedades o enfoques adecuados de la denominada “evolución dirigida”.

Por ejemplo, para la manipulación mediante ingeniería genética en células procariotas, una molécula de ácido nucleico que codifica HMG CoA sintasa se puede introducir en plásmidos que permiten la mutagénesis o modificación de secuencia mediante recombinación de secuencias de ADN. Los métodos estándar (véase Sambrook

y Russell (2001), Molecular Cloning: A Laboratory Manual, CSH Press, Cold Spring Harbor, NY, USA) permiten que se lleven a cabo intercambios de bases o que se añadan secuencias naturales o sintéticas. Los fragmentos de ADN se pueden conectar entre sí aplicando adaptadores y ligadores a los fragmentos. Además, se pueden usar medidas de ingeniería que proporcionan sitios de restricción adecuados o eliminan sitios de restricción o ADN sobrantes. En esos casos, en los que son posibles las inserciones, supresiones o sustituciones, se puede usar mutagénesis *in vitro*, "reparación de cebadores", restricción o ligación. En general, como métodos de análisis se lleva a cabo un análisis de secuencia, análisis de restricción u otros métodos de bioquímica y biología molecular. Las variantes de HMG CoA sintasa resultantes se evalúan entonces para determinar su actividad enzimática, y en particular para determinar su capacidad para preferir acetona como sustrato en vez de acetoacetilCoA. Un ensayo para medir la capacidad de una HMG CoA sintasa para usar acetona como sustrato se describe en el Ejemplo 5. La formación de 3-hidroxi-3-metilbutirato se puede detectar mediante comparación con un compuesto estándar, por ejemplo tras la separación mediante cromatografía de capa fina, LC/MS y ensayo colorimétrico tras su derivatización, o mediante espectrometría de masas.

En particular, se lleva a cabo una reacción en una mezcla de reacción que contiene 40 mM de Tris-HCl pH 8, 5 a 50 mM de acetil-CoA, 100 a 500 mM de acetona, 1 $MgCl_2$ (excepto para HMG CoA sintasa mitocondrial), 0,5 mM de DTT, y enzima que varía en el intervalo de 0,2 a 8 mg/ml. Se llevan a cabo reacciones de control en ausencia de enzima y uno de los sustratos.

El progreso de la síntesis es seguido analizando alícuotas tomadas tras incrementar el período de incubación a 30 ó 37ºC. Típicamente, se retira una alícuota de 50 µl tras 48 h de incubación, se calienta durante 1 min. a 100ºC para eliminar las proteínas, se centrifuga, y el sobrenadante se transfiere a un vial limpio para la detección de HIV mediante espectrometría de masas. Una disolución de 3-hidroxi-3-metilbutirato se prepara en 40 mM de Tris-HCl pH 8, 1 mM de $MgCl_2$, 0,5 mM de DTT, se calienta como se describe anteriormente y se usa como referencia.

Las muestras se analizan en un espectrómetro de masas de cuadrupolo triple PE SCIEX API 2000 en modo de ion negativo con H_2O/acetonitrilo = 60/40 que contiene 0,1% de trietilamina como fase móvil, y el caudal fue 40 µl/min. Se mezclan 10 µl de cada sobrenadante con una cantidad igual de fase móvil y se inyectan directamente en el espectrómetro de masas. Se monitoriza la presencia de ion [3-hidroxi-3-metilbutirato-H]$^-$. La síntesis de 3-hidroxi-3-metilbutirato también se puede llevar a cabo en presencia de [2-^{14}C]acetona radiomarcada. La formación de producto se analiza tras la separación de la mezcla de reacción mediante TLC o HPLC.

En una realización preferida, la HMG CoA sintasa empleada en la presente invención es una enzima que tiene un valor K_M para acetona de 300 mM o inferior, preferiblemente de 250 mM o inferior, incluso más preferiblemente de 200 mM o inferior, y particularmente preferido de 150 mM o inferior. Se prefiere que el valor K_M se determine en las condiciones descritas en el Ejemplo 7. En otra realización preferida, la HMG CoA sintasa empleada en la presente invención tiene un valor kart para la reacción descrita de al menos 0,1 x 10^{-4} s^{-1}, preferiblemente al menos 0,2 x 10^{-4} s^{-1}, incluso más preferiblemente al menos 0,5 x 10^{-4} s^{-1}, y particularmente preferido al menos 1 x 10^{-4} s^{-1}, al menos 2 x 10^{-4} s^{-1}, al menos 3 x 10^{-4} s^{-1} o al menos 5 x 10^{-4} s^{-1}. Se prefiere que el valor k_{cat} se determina en las condiciones descritas en el Ejemplo 7.

Se sabe en la técnica que His264 de HMG CoA sintasa aviar juega un papel en la interacción de la enzima con acetoacetil-CoA, y que la variante Ala264 carece de la interacción con el oxígeno del resto tioéster de acetoacetil-CoA (Misraa et al., Biochem. 35 (1996), 9610-9616). De este modo, a fin de desarrollar variantes de HMG CoA sintasa que muestren una menor aceptación de acetoacetil-CoA como sustrato pero que acepten acetona como sustrato, es concebible mutar sistemáticamente en una HMG CoA sintasa el resto de histidina que corresponde a His264 de HMG CoA sintasa aviar descrita en Misraa et al. (lugar citado) para reducir o inhabilitar la aceptación de acetoacetil-CoA como sustrato.

Además, se pueden proporcionar variantes de HMG CoA sintasa que muestran una mayor actividad. Steussy et al. (Biochemistry 45 (2006), 14407-14414), por ejemplo, describen un mutante de la HMG CoA sintasa de Enterococcus faecalis en el que Ala110 se cambió a Gly110, y que muestra un incremento de 140 veces la velocidad de reacción global.

También se pueden llevar a cabo métodos para identificar variantes con propiedades enzimáticas mejoradas con respecto a la producción de 3-hidroxi-3-metilbutirato en presencia de un cofactor que permite una complementación estérica y/o electrónica en el sitio catalítico de la enzima/enzimas, debido al hecho de que el sustrato acetona es más corto que el sustrato natural acetoacetil-CoA de HMG CoA sintasa. Un ejemplo de tal cofactor sería la coenzima A, o una molécula estructuralmente muy relacionada tal como S-nitroso-CoA.

La versión modificada de la HMG CoA sintasa que acepta acetona como sustrato pero que tiene una baja afinidad por acetoacetil-CoA como sustrato o ya no acepta acetoacetil-CoA como sustrato se puede derivar a partir de una HMG CoA sintasa de origen natural o a partir de una HMG CoA sintasa ya modificada, optimizada o sintéticamente sintetizada.

La HMG CoA sintasa empleada en el procedimiento según la presente invención puede ser una versión natural de la proteína, o una proteína sintética, así como una proteína que se ha sintetizado químicamente o se ha producido en

un sistema biológico o mediante procedimientos recombinantes. La HMG CoA sintasa también se puede modificar químicamente, por ejemplo a fin de mejorar su/sus estabilidades, resistencias por ejemplo a temperatura, para facilitar su purificación o su inmovilización sobre un soporte. La enzima se puede usar en forma aislada, forma purificada, en forma inmovilizada, como un extracto bruto o parcialmente purificado obtenido de células que sintetizan la enzima como enzima químicamente sintetizada, como enzima recombinantemente producida, en forma de microorganismos que las producen, etc.

El procedimiento según la presente invención se puede llevar a cabo in vitro o in vivo. Se entiende que una reacción in vitro es una reacción en la que no se emplean células, es decir, una reacción acelular.

Para llevar a cabo el procedimiento in vitro, los sustratos para la reacción y la enzima se incuban en condiciones (tampón, temperatura, cofactores, etc.) que permiten que la enzima sea activa y que se produzca la conversión enzimática. La reacción se deja transcurrir durante un tiempo suficiente para producir 3-hidroxi-3-metilbutirato. La producción de 3-hidroxi-3-metilbutirato y/o 3-hidroxi-3-metilbutiril-CoA se puede detectar mediante comparación con un compuesto estándar tras la separación mediante cromatografía de capa fina, LC/MS y ensayo colorimétrico tras su derivatización.

La enzima puede estar en cualquier forma adecuada que permita que tenga lugar la reacción enzimática. Puede estar purificada o parcialmente purificada, o en forma de extractos celulares brutos o extractos parcialmente purificados. También es posible que la enzima se inmovilice sobre un portador adecuado.

Puesto que el sustrato acetona es en general más corto que el sustrato natural usado por la enzima, acetoacetil-CoA usado por HMG CoA sintasa, puede ser ventajoso añadir a la mezcla de reacción un cofactor que permita una complementación estérica y/o electrónica en el sitio catalítico de la enzima. Un ejemplo de tal cofactor sería la coenzima A o una molécula estructuralmente muy relacionada tal como S-nitroso-CoA.

Para llevar a cabo el procedimiento in vivo, se hace uso de un organismo/microorganismo (u organismos/microorganismos) adecuado o adecuados que son capaces de proporcionar los sustratos, es decir, acetona y acetil-coenzima A, y una enzima que es capaz de catalizar la formación de un enlace covalente entre el átomo de carbono del grupo oxo (es decir, el C=O) de acetona y el átomo de carbono (C^2) correspondiente al grupo metilo del compuesto que proporciona un grupo acetilo activado según la fórmula (I), que es una HMG CoA sintasa.

De este modo, en el caso de esta realización, el método según la invención se caracteriza por que la conversión de acetona y un compuesto que proporciona un grupo acetilo activado se realiza en presencia de un organismo, preferiblemente un microorganismo capaz de producir acetona y expresar una enzima que es capaz de formar un enlace covalente entre el átomo de carbono del grupo oxo (es decir, el C=O) de acetona y el átomo de carbono (C^2) correspondiente al grupo metilo de un compuesto que proporciona un grupo acetilo activado según la fórmula (I) que expresa una enzima con la actividad de una HMG CoA sintasa (EC 2.3.3.10).

La expresión “que es capaz de producir acetona”, en el contexto de la presente invención, significa que el organismo/microorganismo tiene la capacidad para producir acetona en la célula debido a la presencia de enzimas que proporcionan actividades enzimáticas que permiten la producción de acetona a partir de precursores metabólicos.

La acetona se produce mediante ciertos microorganismos, tales como Clostridium acetobutylicum, Clostridium beijerinckii, Clostridium cellulolyticum, Bacillus polymyxa y Pseudomonas putida. La síntesis de acetona está muy bien caracterizada en Clostridium acetobutylicum. Comienza con una reacción (etapa 1 de reacción) en la que dos moléculas de acetil-CoA se condensan en acetoacetil-CoA. Esta reacción es catalizada por acetil-CoA acetiltransferasa (EC 2.3.1.9). La acetoacetil-CoA se convierte entonces en acetoacetato mediante una reacción con ácido acético o ácido butírico, dando también como resultado la producción de acetil-CoA o butiril-CoA (etapa 2 de reacción). Esta reacción está catalizada por ejemplo por acetoacetilCoA transferasa (EC 2.8.3.8). La acetoacetilCoA transferasa es conocida de diversos organismos, por ejemplo de E. coli, en la que es codificada por el gen atoAD, o de Clostridium acetobutylicum, en el que es codificada por el gen ctfAB. Sin embargo, otras enzimas también pueden catalizar esta reacción, por ejemplo 3-oxoácido CoA transferasa (EC 2.8.3.5) o succinato CoA ligasa (EC 6.2.1.5).

Finalmente, el acetoacetato se convierte en acetona mediante una etapa de descarboxilación (etapa 3 de reacción) catalizada por acetoacetato descarboxilasa (EC 4.1.1.4).

Las etapas 1 y 2 de reacción descritas anteriormente y las enzimas que las catalizan no son características para la síntesis de acetona, y se pueden encontrar en diversos organismos. Por el contrario, la etapa 3 de reacción que está catalizada por acetoacetato descarboxilasa (EC 4.1.1.4) solo se encuentra en aquellos organismos que son capaces de producir acetona.

En una realización preferida, el organismo empleado en el método según la invención es un organismo, preferiblemente un microorganismo, que tiene naturalmente la capacidad para producir acetona. De este modo, preferiblemente el microorganismo pertenece al género Clostridium, Bacillus o Pseudomonas, más preferiblemente a las especies Clostridium acetobutylicum, Clostridium beijerinckii, Clostridium cellulolyticum, Bacillus polymyxa o

Pseudomonas putida.

En una realización preferida adicional, el organismo empleado en el método según la invención es un organismo, preferiblemente un microorganismo, que tiene naturalmente la capacidad para producir acetona y que es recombinante en el sentido de que se ha modificado además genéticamente para expresar una enzima como se define anteriormente. El término “recombinante”, en una realización, significa que el organismo está modificado genéticamente para contener una molécula de ácido nucleico extraña que codifica una enzima como se define anteriormente. En una realización preferida, el organismo se ha modificado genéticamente para contener una molécula de ácido nucleico extraña que codifica una enzima como se define anteriormente, es decir, una HMG CoA sintasa.

El término “extraño”, en este contexto, significa que la molécula de ácido nucleico no se produce de forma natural en dicho organismo/microorganismo. Esto significa que no aparece en la misma estructura o en la misma localización en el organismo/microorganismo. En una realización preferida, la molécula de ácido nucleico extraña es una molécula recombinante que comprende un promotor y una secuencia codificante que codifica la enzima respectiva, es decir, una HMG CoA sintasa, en la que el promotor que conduce la expresión de la secuencia codificante es heterólogo con respecto a la secuencia codificante. Heterólogo, en este contexto, significa que el promotor no es el promotor que conduce de forma natural la expresión de dicha secuencia codificante, sino que es un promotor que conduce de forma natural la expresión de una secuencia codificante diferente, es decir, deriva de otro gen, o es un promotor sintético o un promotor quimérico. Preferiblemente, el promotor es un promotor heterólogo al organismo/microorganismo, es decir, un promotor que no se produce de forma natural en el organismo/microorganismo respectivo. Incluso más preferiblemente, el promotor es un promotor inducible. Los promotores para conducir la expresión en diferentes tipos de organismos, en particular en microorganismos, son bien conocidos por la persona experta en la técnica.

En otra realización preferida, la molécula de ácido nucleico es extraña al organismo/microorganismo por cuanto la enzima codificada, es decir, la HMG CoA-sintasa no es endógena al organismo/microorganismo, es decir, no es expresada de forma natural por el organismo/microorganismo cuando no está modificado genéticamente. En otras palabras, la HMG CoA sintasa codificada es heteróloga con respecto al organismo/microorganismo.

El término “recombinante”, en otra realización, significa que el organismo está modificado genéticamente en la región reguladora que controla la expresión de la enzima como se define anteriormente que se produce de forma natural en el organismo para conducir a un incremento en la expresión de la enzima respectiva en comparación con un organismo correspondiente no modificado genéticamente. El significado de la expresión “mayor expresión” se describe posteriormente más abajo.

Tal modificación de una región reguladora se puede lograr mediante métodos conocidos por la persona experta en la técnica. Un ejemplo es intercambiar el promotor de origen natural por un promotor que permita una mayor expresión, o modificar el promotor de origen natural para mostrar una mayor expresión. De este modo, en esta realización, el organismo contiene en la región reguladora del gen que codifica una enzima como se define anteriormente una molécula de ácido nucleico extraña que no aparece de forma natural en el organismo y que conduce a una mayor expresión de la enzima en comparación con un organismo correspondiente no modificado genéticamente.

La molécula de ácido nucleico extraña puede estar presente en el organismo/microorganismo en forma extracromosómica, por ejemplo como plásmido, o puede estar integrada en el cromosoma. Se prefiere una integración estable.

En una realización adicional preferida, el organismo/microorganismo se caracteriza por que la expresión/actividad de una enzima como se define anteriormente, es decir, de una HMG CoA sintasa, es mayor en el organismo/microorganismo genéticamente modificado con la molécula de ácido nucleico extraña en comparación con el organismo/microorganismo correspondiente no modificado genéticamente. Una “mayor” expresión/actividad significa que la expresión/actividad de la enzima, es decir, de la HMG CoA sintasa, en el microorganismo genéticamente modificado es al menos 10%, preferiblemente al menos 20%, más preferiblemente al menos 30% o 50%, incluso más preferiblemente al menos 70% u 80%, y particularmente preferido al menos 90% o 100% mayor que en el organismo/microorganismo correspondiente no modificado genéticamente. En realizaciones incluso más preferidas, el incremento en la expresión/actividad puede ser al menos 150%, al menos 200% o al menos 500%. En realizaciones particularmente preferidas, la expresión es al menos 10 veces, más preferiblemente al menos 100 veces, e incluso más preferido al menos 1000 veces mayor que en el organismo/microorganismo correspondiente no modificado genéticamente.

El término “mayor” expresión/actividad también cubre la situación en la que el organismo/microorganismo correspondiente no modificado genéticamente no expresa una enzima correspondiente, es decir, una HMG CoA sintasa, de manera que la expresión/actividad correspondiente en el organismo/microorganismo no modificado genéticamente es cero.

Los métodos para medir el nivel de expresión de una proteína dada en una célula son bien conocidos por la persona experta en la técnica. En una realización, la medida del nivel de expresión se realiza midiendo la cantidad de la

proteína correspondiente. Los métodos correspondientes son bien conocidos por la persona experta en la técnica, e incluyen transferencia Western, ELISA, etc. En otra realización, la medida del nivel de expresión se realiza midiendo la cantidad del ARN correspondiente. Los métodos correspondientes son bien conocidos por la persona experta en la técnica, e incluyen, por ejemplo, transferencia Northern.

Los métodos para medir la actividad enzimática de la enzima mencionada anteriormente, HMG CoA sintasa, son conocidos en la técnica y ya se han descrito anteriormente.

En otra realización preferida, el organismo empleado en el método según la invención es un organismo genéticamente modificado, preferiblemente un microorganismo, derivado de un organismo/microorganismo que no produce de forma natural acetona pero que se ha modificado genéticamente para producir acetona, es decir, introduciendo el gen o genes necesarios para permitir la producción de acetona en el organismo/microorganismo. En principio, cualquier organismo se puede modificar genéticamente de esta manera. Las enzimas responsables de la síntesis de acetona se han descrito anteriormente. Los genes que codifican las enzimas correspondientes son conocidos en la técnica y se pueden usar para modificar genéticamente un microorganismo dado para producir acetona. Como se describe anteriormente, las etapas 1 y 2 de reacción de la síntesis de acetona se producen de forma natural en la mayoría de los organismos. Sin embargo, la etapa 3 de reacción es característica y crucial para la síntesis de acetona. De este modo, en una realización preferida, un organismo/microorganismo genéticamente modificado derivado de un organismo/microorganismo que no produce de forma natural acetona se modifica para contener una secuencia nucleotídica que codifica una enzima que cataliza la conversión de acetoacetato en acetona mediante descarboxilación, por ejemplo una acetoacetato descarboxilasa (EC 4.1.1.4). Las secuencias nucleotídicas de varios organismos que codifican esta enzima son conocidas en la técnica, por ejemplo el gen adc de Clostridium acetobutylicum (números de acceso Uniprot P23670 y P23673), Clostridium beijerinckii (Clostridium MP; Q9RPK1), Clostridium pasteurianum (número de acceso Uniprot P81336), Bradyrhizobium sp. (cepa BTAi1/ATCC BAA-1182; número de acceso Uniprot A5EBU7), Burkholderia mallei (ATCC 10399 A9LBS0), Burkholderia mallei (número de acceso Uniprot A3MAE3), Burkholderia mallei FMH A5XJB2, Burkholderia cenocepacia (número de acceso Uniprot A0B471), Burkholderia ambifaria (número de acceso Uniprot Q0b5P1), Burkholderia phytofirmans (número de acceso Uniprot B2T319), Burkholderia spec. (número de acceso Uniprot Q38ZU0), Clostridium botulinum (número de acceso Uniprot B2TLN8), Ralstonia pickettii (número de acceso Uniprot B2UIG7), Streptomyces nogalater (número de acceso Uniprot Q9EY17), Streptomyces avermitilis (número de acceso Uniprot Q82NF4), Legionella pneumophila (número de acceso Uniprot Q5ZXQ9), Lactobacillus salivarius (número de acceso Uniprot Q1WVG5), Rhodococcus spec. (número de acceso Uniprot Q0S7W4), Lactobacillus plantarum (número de acceso Uniprot Q890G0), Rhizobium leguminosarum (número de acceso Uniprot Q1M911), Lactobacillus casei (número de acceso Uniprot Q03B66), Francisella tularensis (número de acceso Uniprot Q0BLC9), Saccharopolyspora erythreae (número de acceso Uniprot A4FKR9), Korarchaeum cryptofilum (número de acceso Uniprot B1L3N6), Bacillus amyloliquefaciens (número de acceso Uniprot A7Z8K8), Cochliobolus heterostrophus (número de acceso Uniprot Q8NJ03), Sulfolobus islandicus (número de acceso Uniprot C3ML22) y Francisella tularensis subsp. holarctica (cepa OSU18).

Más preferiblemente, el organismo, preferiblemente microorganismo, se modifica genéticamente para transformarlo con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 2 de reacción mencionada anteriormente de la síntesis de acetona, es decir, la conversión de acetoacetil CoA en acetoacetato.

Incluso más preferiblemente, el organismo, preferiblemente microorganismo, se modifica genéticamente para transformarlo con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 1 de reacción mencionada anteriormente de la síntesis de acetona, es decir, la condensación de dos moléculas de acetil CoA en acetoacetil CoA.

En una realización particularmente preferida, el organismo/microorganismo se modifica genéticamente para transformarlo con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 1 de reacción mencionada anteriormente de la síntesis de acetona y con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 2 de reacción mencionada anteriormente de la síntesis de acetona, o con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 1 de reacción mencionada anteriormente de la síntesis de acetona y con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 3 de reacción mencionada anteriormente de la síntesis de acetona, o con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 2 de reacción mencionada anteriormente de la síntesis de acetona y con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 3 de reacción mencionada anteriormente de la síntesis de acetona, o con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 1 de reacción mencionada anteriormente de la síntesis de acetona y con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 2 de reacción mencionada anteriormente de la síntesis de acetona y con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 3 de reacción mencionada anteriormente de la síntesis de acetona.

Los métodos para preparar el organismo genéticamente modificado mencionado anteriormente, preferiblemente microorganismos, son bien conocidos en la técnica. De este modo, generalmente, el organismo/microorganismo se transforma con un constructo de ADN que permite la expresión de la enzima respectiva en el microorganismo. Tal constructo comprende normalmente la secuencia codificante en cuestión enlazada a secuencias reguladoras que permiten la transcripción y traducción en la célula hospedante respectiva, por ejemplo un promotor y/o potenciador

y/o terminador de la transcripción y/o sitios de unión al ribosoma, etc. La técnica anterior ya describe microorganismos que se han modificado genéticamente para que sean capaces de producir acetona. En particular, se han introducido genes procedentes de, por ejemplo, Clostridium acetobutylicum en E. coli permitiendo de ese modo la síntesis de acetona en E. coli, una bacteria que no produce de forma natural acetona (Bermejo et al., Appl. Environ. Microbiol. 64 (1998); 1079-1085; Hanai et al., Appl. Environ. Microbiol. 73 (2007), 7814-7818). En particular, Hanai et al. (lugar citado) muestran que es suficiente introducir una secuencia de ácido nucleico que codifica una acetoacetato descarboxilasa (tal como la de Clostridium acetobutylicum) a fin de lograr la producción de acetona en E. coli, indicando que las enzimas endógenas en E. coli que catalizan las etapas 1 y 2 de reacción mencionadas anteriormente (es decir, los productos de expresión de los genes atoB y atoAD de E. coli) son suficientes para proporcionar sustrato para la producción de acetona.

En una realización particularmente preferida, el organismo, preferiblemente un microorganismo, empleado en el método según la invención es un organismo/microorganismo recombinante derivado de un organismo/microorganismo que no produce de forma natural acetona pero que se ha modificado genéticamente, como se describe anteriormente, para producir acetona y que expresa una enzima que es capaz de catalizar la formación de un enlace covalente entre el átomo de carbono del grupo oxo (es decir, el C=O) de acetona y el átomo de carbono (C^2) correspondiente al grupo metilo del compuesto que proporciona un grupo acetilo activado como se define anteriormente. El término “recombinante”, en este contexto, significa preferiblemente que el organismo es recombinante en el sentido de que se ha modificado genéticamente además para expresar una enzima como se describe anteriormente. El término “recombinante”, en una realización, significa que el organismo está modificado genéticamente para contener una molécula de ácido nucleico extraña que codifica una enzima como se define anteriormente, es decir, una HMG CoA sintasa.

Con respecto a la definición de la expresión “molécula de ácido nucleico extraña”, se aplica lo mismo que ya se ha expuesto anteriormente.

El término “recombinante”, en otra realización, significa que el organismo está modificado genéticamente en la región reguladora que controla la expresión de una enzima como se define anteriormente que aparece de forma natural en el organismo para conducir a un incremento en la expresión de la enzima respectiva en comparación con un organismo correspondiente no modificado genéticamente. El significado de la expresión “mayor expresión” se describe posteriormente más abajo.

Tal modificación de una región reguladora se puede lograr mediante métodos conocidos por la persona experta en la técnica. Un ejemplo es intercambiar el promotor de origen natural por un promotor que permita una mayor expresión, o modificar el promotor de origen natural para mostrar una mayor expresión. De este modo, en esta realización, el organismo contiene en la región reguladora del gen que codifica una enzima como se define anteriormente una molécula de ácido nucleico extraña que no aparece de forma natural en el organismo y que conduce a una mayor expresión de la enzima en comparación con un organismo correspondiente no modificado genéticamente.

Preferiblemente, tal organismo/microorganismo se caracteriza por que la expresión/actividad de dicha enzima, es decir, la HMG CoA sintasa, es mayor en el organismo/microorganismo recombinante en comparación con el organismo/microorganismo correspondiente no modificado genéticamente. Una “mayor” expresión/actividad significa que la expresión/actividad de la enzima, es decir, la HMG CoA sintasa, en el organismo/microorganismo genéticamente modificado es al menos 10%, preferiblemente al menos 20%, más preferiblemente al menos 30% o 50%, incluso más preferiblemente al menos 70% u 80%, y particularmente preferido al menos 90% o 100% mayor que en el organismo/microorganismo correspondiente no modificado genéticamente. En realizaciones incluso más preferidas, el incremento en la expresión/actividad puede ser al menos 150%, al menos 200%, o al menos 500%. En realizaciones particularmente preferidas, la expresión es al menos 10 veces, más preferiblemente al menos 100 veces, e incluso más preferido al menos 1000 veces mayor que en el organismo/microorganismo correspondiente no modificado genéticamente.

El término “mayor” expresión/actividad también cubre la situación en la que el organismo/microorganismo correspondiente no modificado genéticamente no expresa dicha enzima, es decir, una HMG CoA sintasa, de forma que la expresión/actividad correspondiente en el organismo/microorganismo no modificado genéticamente es cero. Con respecto a los métodos para medir el nivel de expresión o actividad, se aplica lo mismo que ya se ha dicho anteriormente.

El término “organismo”, como se usa en el contexto de la presente invención, se refiere en general a cualquier tipo posible de organismo, en particular organismos eucariotas, organismos procariotas y arqueobacterias. El término incluye animal no humano, plantas, hongos, bacterias y arqueobacterias. El término también incluye células aisladas o agregados celulares de tales organismos, como tejido o callo.

En una realización preferida, el organismo es un microorganismo. El término “microorganismo”, en el contexto de la presente invención, se refiere a células procariotas, en particular bacterias, así como a hongos, tales como levaduras, y también a algas y arqueobacterias. En una realización preferida, el microorganismo es una bacteria. En principio, se puede usar cualquier bacteria. Las bacterias preferidas a emplear en el procedimiento según la invención son bacterias del género Bacillus, Clostridium, Pseudomonas o Escherichia. En una realización

particularmente preferida, la bacteria pertenece al género Escherichia, e incluso más preferido a la especie Escherichia coli.

En otra realización preferida, el microorganismo es un hongo, más preferiblemente un hongo del género Saccharomyces, Schizosaccharomyces, Aspergillus o Trichoderma, e incluso más preferiblemente de la especie Saccharomyces cerevisiae, Schizosaccharomyces pombe, Aspergillus niger o de la especie Trichoderma reesei.

En todavía otra realización preferida, el microorganismo es un microorganismo fotosintéticamente activo, tal como bacterias que son capaces de llevar a cabo la fotosíntesis, o microalgas.

En una realización particularmente preferida, el microorganismo es un alga, más preferiblemente un alga que pertenece a las diatomeas.

Si se usan microorganismos en el contexto del método de la presente invención, también es concebible llevar a cabo el método según la invención de una manera en la que se empleen dos tipos de microorganismos, es decir, un tipo que produce acetona y un tipo que usa la acetona producida por el primer tipo de microorganismos para convertirla con la ayuda de una enzima como se define aquí anteriormente.

Cuando el procedimiento según la invención se lleva a cabo in vivo usando microorganismos que proporcionan la actividad/actividades enzimáticas respectivas, los microorganismos se cultivan en condiciones de cultivo adecuadas que permiten la producción de la reacción o reacciones enzimáticas. Las condiciones de cultivo específicas dependen del microorganismo específico empleado, pero son bien conocidas por la persona experta en la técnica. Las condiciones de cultivo se escogen generalmente de una manera tal que permiten la expresión de los genes que codifican las enzimas para las reacciones respectivas. La persona experta en la técnica conoce diversos métodos para mejorar y ajustar de forma fina la expresión de ciertos genes en ciertas etapas del cultivo, tal como la inducción de la expresión génica por inductores químicos o mediante un desplazamiento de la temperatura.

En otra realización preferida, el organismo empleado en el método según la invención es un organismo que es capaz de fotosíntesis, tal como una planta o microalga. En principio, se puede usar cualquier planta posible, es decir, una planta monocotiledónea o una planta dicotiledónea. Es preferible usar una planta que se puede cultivar a una escala agrícolamente significativa y que permita producir grandes cantidades de biomasa. Los ejemplos son gramíneas como Lolium, cereales como centeno, cebada, avena, mijo, maíz, y otras plantas que almacenan almidón, como patata, o plantas que almacenan azúcar, como caña de azúcar o remolacha azucarera. También es concebible el uso de tabaco o de plantas de vegetales tales como tabaco, pimiento, pepino, berenjena, etc. Otra posibilidad es el uso de plantas que almacenan aceite, tales como colza, olivos, etc. También es concebible el uso de árboles, en particular árboles de crecimiento rápido tales como eucalipto, álamo o árbol del caucho (Hevea brasiliensis).

La presente invención también se refiere a un organismo, preferiblemente un microorganismo, que se caracteriza por las siguientes características:

(a) es capaz de producir acetona; y

(b) expresa una enzima que es capaz de catalizar la formación de un enlace covalente entre el átomo de carbono del grupo oxo (es decir, el C=O) de acetona y el átomo de carbono (C^2) correspondiente al grupo metilo del compuesto que proporciona un grupo acetilo activado como se define anteriormente, es decir, una enzima con la actividad de una HMG CoA sintasa (EC 2.3.3.10).

Con respecto a la fuente, naturaleza, propiedades, secuencia, etc., de la enzima, es decir, la HMG CoA sintasa, expresada en el organismo según la invención, se aplica lo mismo que lo que se ha expuesto anteriormente en relación con el método según la invención.

En una realización preferida, el organismo según la invención es un organismo, preferiblemente un microorganismo, que tiene naturalmente la capacidad para producir acetona, es decir, la característica (a) mencionada anteriormente es una característica que el organismo, preferiblemente microorganismo, muestra de forma natural. De este modo, preferiblemente el organismo es un microorganismo que pertenece al género Clostridium, Bacillus o Pseudomonas, más preferiblemente a la especie Clostridium acetobutylicum, Clostridium beijerinckii, Clostridium cellutolyticum, Bacillus polymyxa o Pseudomonas putida.

En otra realización preferida, el organismo, preferiblemente microorganismo, según la invención es un organismo/microorganismo genéticamente modificado derivado de un organismo/microorganismo que no produce de forma natural acetona pero que se ha modificado genéticamente para producir acetona, es decir, introduciendo el gen o genes necesarios para permitir la producción de acetona en el organismo/microorganismo. En principio, cualquier organismo/microorganismo se puede modificado genéticamente de esta manera. Las enzimas responsables de la síntesis de acetona se han descrito anteriormente. Los genes que codifican las enzimas correspondientes son conocidos en la técnica y se pueden usar para modificar genéticamente un organismo dado, preferiblemente microorganismo, para producir acetona.

En una realización preferida, un organismo/microorganismo modificado genéticamente derivado de un organismo/microorganismo que no produce de forma natural acetona se modifica para que contenga una secuencia nucleotídica que codifica una enzima que cataliza la conversión de acetoacetato en acetona mediante descarboxilación, por ejemplo una acetoacetato descarboxilasa (EC 4.1.1.4). Las secuencias nucleotídicas procedentes de varios organismos que codifican esta enzima son conocidas en la técnica, por ejemplo el gen adc de Clostridium acetobutylicum. Más preferiblemente, el organismo/microorganismo se modifica genéticamente para transformarlo con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 2 de reacción mencionada anteriormente de la síntesis de acetona, es decir, la conversión de acetoacetil CoA en acetoacetato.

Incluso más preferiblemente, el organismo/microorganismo se modifica genéticamente para transformarlo con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 1 de reacción mencionada anteriormente de la síntesis de acetona, es decir, la condensación de dos moléculas de acetil CoA en acetoacetatil CoA.

En una realización particularmente preferida, el organismo/microorganismo se modifica genéticamente para transformarlo con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 1 de reacción mencionada anteriormente de la síntesis de acetona y con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 2 de reacción mencionada anteriormente de la síntesis de acetona, o con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 1 de reacción mencionada anteriormente de la síntesis de acetona y con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 3 de reacción mencionada anteriormente de la síntesis de acetona, o con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 2 de reacción mencionada anteriormente de la síntesis de acetona y con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 3 de reacción mencionada anteriormente de la síntesis de acetona, o con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 1 de reacción mencionada anteriormente de la síntesis de acetona y con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 2 de reacción mencionada anteriormente de la síntesis de acetona y con una molécula de ácido nucleico que codifica una enzima capaz de catalizar la etapa 3 de reacción mencionada anteriormente de la síntesis de acetona.

Los métodos para preparar el organismo genéticamente modificado mencionado anteriormente, preferiblemente organismos/microorganismos, son bien conocidos en la técnica. De este modo, generalmente, el organismo/microorganismo se transforma con un constructo de ADN que permite la expresión de la enzima respectiva en el organismo/microorganismo. Tal constructo comprende normalmente la secuencia codificante en cuestión enlazada a secuencias reguladoras que permiten la transcripción y traducción en la célula hospedante respectiva, por ejemplo un promotor y/o potenciador y/o terminador de la transcripción y/o sitios de unión al ribosoma, etc. La técnica anterior ya describe organismos, en particular microorganismos que se han modificado genéticamente para que sean capaces de producir acetona. En particular, se han introducido genes procedentes de, por ejemplo, Clostridium acetobutylicum en E. coli permitiendo de ese modo la síntesis de acetona en E. coli, una bacteria que no produce de forma natural acetona (Bermejo et al., Appl. Environ. Microbiol. 64 (1998); 1079-1085; Hanai et al., Appl. Environ. Microbiol. 73 (2007), 7814-7818). En particular, Hanai et al. (lugar citado) muestran que es suficiente introducir una secuencia de ácido nucleico que codifica una acetoacetato descarboxilasa (tal como la de Clostridium acetobutylicum) a fin de lograr la producción de acetona en E. coli, indicando que las enzimas endógenas en E. coli que catalizan las etapas 1 y 2 de reacción mencionadas anteriormente (es decir, los productos de expresión de los genes atoB y atoAD de E. coli) son suficientes para proporcionar sustrato para la producción de acetona.

En una realización adicionalmente preferida, el organismo, preferiblemente un microorganismo, según la invención se modifica genéticamente para expresar una enzima como se define anteriormente que es capaz de catalizar la formación de un enlace covalente entre el átomo de carbono del grupo oxo (es decir, el C=O) de acetona y el átomo de carbono (C^2) correspondiente al grupo metilo del compuesto como se define anteriormente que proporciona un grupo acetilo activado. En este contexto, el término “recombinante” significa en un primer aspecto contiene una molécula de ácido nucleico extraña que codifica una enzima que es capaz de catalizar la formación de un enlace covalente entre el átomo de carbono del grupo oxo (es decir, el C=O) de acetona y el átomo de carbono (C^2) correspondiente al grupo metilo del compuesto que proporciona un grupo acetilo activado, es decir, una molécula de ácido nucleico extraña que codifica una HMG CoA sintasa. El término “extraño”, en este contexto, significa que la molécula de ácido nucleico no se produce de forma natural en dicho organismo/microorganismo. Esto significa que no aparece en la misma estructura o en la misma localización en el organismo/microorganismo. En una realización preferida, la molécula de ácido nucleico extraña es una molécula recombinante que comprende un promotor y una secuencia codificante que codifica dicha enzima, es decir, la HMG CoA sintasa, en la que el promotor que conduce la expresión de la secuencia codificante es heterólogo con respecto a la secuencia codificante. Heterólogo, en este contexto, significa que el promotor no es el promotor que conduce de forma natural la expresión de dicha secuencia codificante, sino que es un promotor que conduce de forma natural la expresión de una secuencia codificante diferente, es decir, deriva de otro gen, o es un promotor sintético o un promotor quimérico. Preferiblemente, el promotor es un promotor heterólogo al organismo/microorganismo, es decir, un promotor que no se produce de forma natural en el organismo/microorganismo respectivo. Incluso más preferiblemente, el promotor es un promotor inducible. Los promotores para conducir la expresión en diferentes tipos de organismos, en particular microorganismos, son bien conocidos por la persona experta en la técnica. En otra realización preferida, la molécula

de ácido nucleico es extraña al organismo/microorganismo por cuanto la enzima codificada, es decir, la HMG CoA sintasa, no es endógena al organismo/microorganismo, es decir, no es expresada de forma natural por el organismo/microorganismo cuando no está modificado genéticamente. En otras palabras, la enzima codificada, es decir, la HMG CoA sintasa, es heteróloga con respecto al organismo/microorganismo.

El término “recombinante”, en otro aspecto, significa que el organismo está modificado genéticamente en la región reguladora que controla la expresión de una enzima como se define anteriormente que aparece de forma natural en el organismo para conducir a un incremento en la expresión de la enzima respectiva en comparación con un organismo correspondiente no modificado genéticamente. El significado de la expresión “mayor expresión” se describe posteriormente más abajo.

Tal modificación de una región reguladora se puede lograr mediante métodos conocidos por la persona experta en la técnica. Un ejemplo es intercambiar el promotor de origen natural por un promotor que permita una mayor expresión, o modificar el promotor de origen natural para mostrar una mayor expresión. De este modo, en esta realización, el organismo contiene en la región reguladora del gen que codifica una enzima como se define anteriormente una molécula de ácido nucleico extraña que no aparece de forma natural en el organismo y que conduce a una mayor expresión de la enzima en comparación con un organismo correspondiente no modificado genéticamente.

En una realización adicional preferida, el organismo/microorganismo se caracteriza por que la expresión/actividad de dicha enzima, es decir, la HMG CoA sintasa, es mayor en el organismo/microorganismo genéticamente modificado con la molécula de ácido nucleico extraña en comparación con el organismo/microorganismo correspondiente no modificado genéticamente. Una “mayor” expresión/actividad significa que la expresión/actividad de la enzima, es decir, de la HMG CoA sintasa, en el organismo/microorganismo genéticamente modificado es al menos 10%, preferiblemente al menos 20%, más preferiblemente al menos 30% o 50%, incluso más preferiblemente al menos 70% u 80%, y particularmente preferido al menos 90% o 100% mayor que en el organismo/microorganismo correspondiente no modificado genéticamente. En realizaciones incluso más preferidas, el incremento en la expresión/actividad puede ser al menos 150%, al menos 200% o al menos 500%. En realizaciones particularmente preferidas, la expresión es al menos 10 veces, más preferiblemente al menos 100 veces, e incluso más preferido al menos 1000 veces mayor que en el organismo/microorganismo correspondiente no modificado genéticamente.

El término “mayor” expresión/actividad también cubre la situación en la que el organismo/microorganismo correspondiente no modificado genéticamente no expresa una enzima correspondiente, es decir, una HMG CoA sintasa, de manera que la expresión/actividad correspondiente en el organismo/microorganismo no modificado genéticamente es cero.

Los métodos para medir el nivel de expresión de una proteína dada en una célula son bien conocidos por la persona experta en la técnica. En una realización, la medida del nivel de expresión se realiza midiendo la cantidad de la proteína correspondiente. Los métodos correspondientes son bien conocidos por la persona experta en la técnica, e incluyen transferencia Western, ELISA, etc. En otra realización, la medida del nivel de expresión se realiza midiendo la cantidad del ARN correspondiente. Los métodos correspondientes son bien conocidos por la persona experta en la técnica, e incluyen, por ejemplo, transferencia Northern.

Los métodos para medir la actividad enzimática de la enzima mencionada anteriormente, es decir, de una HMG CoA sintasa, son conocidos en la técnica y ya se han descrito anteriormente.

El término “organismo”, como se usa en el contexto de la presente invención, se refiere en general a cualquier tipo posible de organismo, en particular organismos eucariotas, organismos procariotas y arqueobacterias. El término incluye animal no humano, plantas, hongos, bacterias y arqueobacterias. El término también incluye células aisladas o agregados celulares de tales organismos, como tejido o callo.

En una realización preferida, el organismo es un microorganismo. El término “microorganismo”, en el contexto de la presente invención, se refiere a células procariotas, en particular bacterias, así como a hongos, tales como levaduras, y también a algas y arqueobacterias. En una realización preferida, el microorganismo es una bacteria. En principio, se puede usar cualquier bacteria. Las bacterias preferidas a emplear en el procedimiento según la invención son bacterias del género Bacillus, Clostridium, Pseudomonas o Escherichia. En una realización particularmente preferida, la bacteria pertenece al género Escherichia, e incluso más preferido a la especie Escherichia coli.

En otra realización preferida, el microorganismo es un hongo, más preferiblemente un hongo del género Saccharomyces, Schizosaccharomyces, Aspergillus o Trichoderma, e incluso más preferiblemente de la especie Saccharomyces cerevisiae, Schizosaccharomyces pombe, Aspergillus niger o de la especie Trichoderma reesei.

En todavía otra realización preferida, el microorganismo es un microorganismo fotosintéticamente activo, tal como bacterias que son capaces de llevar a cabo la fotosíntesis, o microalgas.

En una realización particularmente preferida, el microorganismo es un alga, más preferiblemente un alga del género que pertenece a las diatomeas.

En otra realización preferida, el organismo según la invención es un organismo que es capaz de fotosíntesis, tal como una planta o microalga. En principio, puede ser cualquier planta posible, es decir, una planta monocotiledónea o una planta dicotiledónea. Es preferiblemente una planta que se puede cultivar a una escala agrícolamente significativa y que permita producir grandes cantidades de biomasa. Los ejemplos son gramíneas como Lolium, cereales como centeno, cebada, avena, mijo, maíz, y otras plantas que almacenan almidón, como patata, o plantas que almacenan azúcar, como caña de azúcar o remolacha azucarera. También es concebible el uso de tabaco o de plantas de vegetales tales como tabaco, pimiento, pepino, berenjena, etc. En otra realización preferida, la planta es una planta que almacena aceite, tal como colza, olivos, etc. También es concebible el uso de árboles, en particular árboles de crecimiento rápido tales como eucalipto, álamo o árbol del caucho (Hevea brasiliensis).

La presente invención también se refiere al uso de un organismo, preferiblemente un microorganismo, que se caracteriza por las siguientes características:

(a) es capaz de producir acetona; y

(b) expresa una enzima que es capaz de catalizar la formación de un enlace covalente entre el átomo de carbono del grupo oxo (es decir, el C=O) de acetona y el átomo de carbono (C^2) correspondiente al grupo metilo del compuesto que proporciona un grupo acetilo activado como se define aquí anteriormente, es decir, una enzima con la actividad de una HMG CoA sintasa (EC 2.3.3.10)

para la producción de ácido 3-hidroxi-3-metilbutírico.

Es decir, la presente invención también se refiere al uso de un organismo/microorganismo según la invención para la producción de ácido 3-hidroxi-3-metilbutírico.

La presente invención también se refiere a una composición que comprende un organismo según la presente invención.

Además, la presente invención también se refiere a una composición que comprende (i) acetona; y (ii) un compuesto que proporciona un grupo acetilo activado como se define aquí anteriormente; y (iii) una enzima que es capaz de catalizar la formación de un enlace covalente entre el átomo de carbono del grupo oxo (es decir, el C=O) de acetona y el átomo de carbono (C^2) correspondiente al grupo metilo del compuesto que proporciona un grupo acetilo activado como se define aquí anteriormente.

Además, la presente invención también se refiere al uso una enzima como se define anteriormente que es capaz de catalizar la formación de un enlace covalente entre el átomo de carbono del grupo oxo (es decir, el C=O) de acetona y el átomo de carbono (C^2) correspondiente al grupo metilo del compuesto que proporciona un grupo acetilo activado como se define aquí anteriormente, para la producción de ácido 3-hidroxi-3-metilbutírico.

Finalmente, la presente invención también se refiere al uso de acetona para la producción de ácido 3-hidroxi-3-metilbutírico, que comprende la conversión enzimática de acetona y un compuesto que proporciona un grupo acetilo activado como se define aquí anteriormente, en el que la conversión enzimática se logra mediante una enzima como se describe anteriormente en relación con el método según la invención, preferiblemente con una enzima que tiene la actividad enzimática de una HMG CoA sintasa, y más preferiblemente la conversión se logra mediante el uso de un organismo según la invención.

Figura 1: Estructura química de ácido 3-hidroxi-3-metilbutírico (también denominado como beta-hidroxiisovalerato)

Figura 2: Esquema de reacción de la reacción catalizada por HMG-CoA sintasa

Figura 3: Esquema de reacción de la reacción catalizada por HMG-CoA liasa (que cae fuera del alcance de las reivindicaciones)

Figura 4: Esquemas de reacción de las reacciones de la ruta de pksX, que incluye la reacción catalizada por la proteína PksG (que cae fuera del alcance de las reivindicaciones)

Figura 5: Esquema de reacción de la reacción de la conversión de acetona y un compuesto que contiene un grupo acetilo activado en ácido 3-hidroxi-3-metilbutírico

X representa S-CH2-CH2-NH-CO-CH2-CH2-NH-CO-CH(OH)-C(CH3)2-CH2-O-PO2H-O-PO2H-C10H13N5O7P (coenzima A).

Figura 6: Espectro de masas de 3-hidroxi-3-metilbutirato comercialmente disponible

Figura 7: Espectro de masas de la formación de 3-hidroxi-3-metilbutirato a partir de acetil-CoA y acetona en presencia de HMG-CoA sintasa de Gallus gallus (P23228).

Figura 8: Espectro de masas del ensayo de control sin enzima.

Figura 9: Gráfica de Michaelis-Menten para la reacción con la HMG CoA sintasa de S. epidermidis descrita en el Ejemplo 7.

Los siguientes ejemplos sirven para ilustrar la invención.

Ejemplo 1

Método bioinformático usado para crear bases de datos de HMG-CoA sintasas y de HMG-CoA liasas (que caen fuera del alcance de las reivindicaciones)

Se seleccionó un 12 HMG-CoA sintasas y 8 HMG-CoA liasas para crear un conjunto no redundante de proteínas con el objetivo de representar la diversidad de estas clases enzimáticas como se encuentran a lo largo de los organismos eucariotas. Estas proteínas se identificaron llevando a cabo búsquedas múltiples a base de secuencias y a base de texto en la base de datos de Universal Protein Resource Uniprot (http://www.uniprot.org/). Todas contienen características únicas tales como dominios proteicos conservados y motivos característicos para la clase enzimática de interés. A fin de cubrir de forma eficaz la diversidad de secuencias sin tener que cribar un gran conjunto de proteínas, el grupo inicial de enzimas se estrechó agrupándolas en agrupamientos de secuencias con homología mayor que 85% y seleccionando entonces una única secuencia candidato representativa de cada agrupamiento. La identidad de las secuencias proteicas osciló de 30% a 80% y de 50% a 80% entre cualesquiera dos proteínas del panel de HMG-CoA sintasas y del panel de liasas respectivamente.

Se aplicó el mismo enfoque para seleccionar las HMG-CoA sintasas y HMG-CoA liasas de organismos procariotas. El conjunto creado contenía 50 proteínas homólogas a HMG-CoA sintasas, incluyendo proteínas pksG, y 59 proteínas homólogas a HMG-CoA liasas.

Ejemplo 2

Clonación, expresión y purificación de una colección de HMG-CoA liasas y HMG-CoA sintasas (que cae fuera del alcance de las reivindicaciones)

Clonación de los genes:

Las secuencia de ácidos nucleicos que codifican HMG-CoA sintasa y liasa de organismo eucariota se optimizaron para la preferencia de codones de E. coli, y los genes se obtuvieron mediante síntesis química (GeneArt).

Los genes que codifican HMG-CoA sintasas y liasas de organismos procariotas se clonaron a partir de ADN genómico de orígenes diferentes mediante técnicas recombinantes habituales. Estos genes se insertaron entonces en los vectores pET 25b y pET 22b que contienen una etiqueta de His (Novagen), respectivamente, para organismos eucariotas y procariotas.

Sobreexpresión en E. coli:

Los plásmidos se electroporaron en bacterias de E. coli BL21 (Novagen), que entonces se extendieron sobre una cápsula de Petri con LB-agar que contiene ampicilina. Los cultivos se hicieron crecer a 30ºC en medio TAMBIÉN, que contiene 0,5 M de sorbitol, 5 mM de betaína, 100 μg/ml de ampicilina, en agitación moderada. Cuando la OD (600 nm) alcanzó 0,8, se añadió IPTG a una concentración final de 1 mM, y se llevó a cabo la expresión durante 16 horas a 20ºC con agitación moderada. Las células bacterianas se cosecharon entonces mediante centrifugación a 4ºC, 10.000 rpm, 20 minutos, y se congelaron a -80ºC.

Preparación de los extractos celulares:

Los extractos celulares se preparan resuspendieron 1,6 g de pelete celular en 5 ml de tampón de Na_2HPO_4 50 mM, que contiene 300 mM de NaCl, 5 mM de $MgCl_2$, 1 mM de DTT pH 8. Entonces se añadieron 20 μl de lysonase (Novagen) a las preparaciones, que se incubaron durante 10 min. a temperatura ambiente y 20 min. en hielo. La lisis celular se logra mediante tratamiento con ultrasonidos triple durante 5 minutos en un baño de agua ultrasónico en hielo y homogeneización del extracto entre cada pulso. Los extractos brutos se aclaran entonces mediante centrifugación a 4ºC, 10.000 rpm, 20 minutos.

Purificación de proteínas:

Los sobrenadantes transparentes se cargaron en una columna PROTINO-1000 Ni-IDA (Macherey-Nagel) que permite la inmovilización específica de proteínas que portan colas 6 histidinas. La columnas se lavaron, y las enzimas se eluyeron con 4 ml de tampón de Na_2HPO_4 50 mM, que contiene 300 mM de NaCl, 5 mM de $MgCl_2$, 1 mM de DTT, 250 mM de imidazol pH 8. Las fracciones que contienen la enzima se concentraron entonces y se desalaron en una unidad de filtración de 10 kDa Amicon Ultra-4 (Millipore) y se resuspendieron en 250 μl de Tris-HCl 40 mM pH 8, que contiene 0,5 mM de DTT. La concentración de proteína se determina mediante el método de Bradford.

La homogeneidad de las enzimas purificadas varió de 20% a 75%.

Ejemplo 3

Medida de la actividad de HMG-CoA sintasa usando los sustratos naturales acetoacetil-CoA y acetol-CoA

La actividad de la HMG-CoA sintasa se mide según Clinkenbeard et al. (J. Biol. Chem. 250 (1975), 3108-3116). La mezcla de medio de ensayo estándar para HMG-CoA sintasas contiene 40 mM de Tris-HCl pH 8, 1 mM de $MgCl_2$, 100 μM de acetoacetil-CoA, 200 μM de acetil-CoA, 0,5 mM de DTT, en un volumen total de 1 ml. Las HMG-CoA sintasas mitocondriales se evalúan en ausencia de $MgCl_2$, para evitar la inhibición observada para esta enzima (Reed et al., J. Biol. Chem. 250. (1975), 3117-3123). La reacción se inicia mediante adición de 0,02 mg/ml de enzima.

Se llevó a cabo un ensayo de control en ausencia de enzima. La actividad de HMG-CoA sintasa se mide monitorizando la disminución en la absorbancia a 303 nm que acompaña a la desaparición de la forma enolato de acetoacetil-CoA dependiente de acetil-CoA. Para explicar la desaparición no específica de acetoacetil-CoA, los resultados obtenidos en el ensayo de control que carece de enzima se restaron de los resultados obtenidos en las muestras de ensayo. El coeficiente de absorción aparente para acetoacetil-CoA en las condiciones del ensayo fue 5600 M^{-1}.

Tabla 1: Actividad fisiológica de algunas HMG-CoA sintasas purificadas o enzimas homólogas a HMG CoA sintasas

Número de acceso Uniprot	Organismo	Actividad fisiológica μmol/min · mg proteína
P54961	*Blattella germanica* (Cucaracha alemana)	0,02
P23228	*Gallus gallus* (Pollo)	0,02
Q01581	*Homo sapiens* (Humano)	0,03
P54873	*Arabidopsis thaliana*	1,19
P54871	*Caenorhabditis elegans*	0,23
P54874	*Schizosaccharomyces pombe* (levadura de fisión)	0,61
P54839	*Saccharomyces cerevisiae* (levadura de panadero)	0,28
P54872	*Dictyostelium discoideum* (moho del cieno)	0,09
Q86HL5	*Dictyostelium discoideum* (moho del cieno)	0,02
Q9M6U3	*Brassica juncea*	0,02
A5FM54	*Flavobacterium johnsoniae*	0,02
Q03WZ0	*Leuconostoc mesenteroides*	0,28
Q2NHU7	*Methanosphaera stadtmanae*	0,02
Q8CN06	*Staphylococcus epidermidis*	0,07
Q03QR0	*Lactobacillus brevis*	0,18
A6UPL1	*Methanosarcina mazei*	0,01
B2HGT6	*Mycobacterium marinum*	0,01
Q4L958	*Staphylococcus haemolyticus*	0,18
Q4A0D6	*Staphylococcus saprophyticus*	0,08
Q1GAH5	*Lactobacillus delbrueckii*	0,32

Ejemplo 4

Medida de la actividad de HMG-CoA liasa usando el sustrato natural HMG-CoA (que cae fuera del alcance de las

reivindicaciones)

La actividad de HMG-CoA liasa se mide según Mellanby J et al. (Methods of Enzymatic Analysis; Bergmeyer Ed. (1963), 454-458). El medio (1 ml) de reacción completo que contiene 40 mM de Tris-HCl pH 8, 1 mM de $MgCl_2$, 0,5 mM de DTT, 0,4 mM de HMG-CoA, 0,2 mM de NADH, 5 unidades de 3-hidroxibutirato deshidrogenasa se incubó durante 5 min. antes de añadir 0,005 mg/ml de HMG-CoA liasa, y el transcurso de la reacción se monitorizó entonces mediante la disminución en la absorbancia a 340 nm. Se llevó a cabo un ensayo de control en ausencia de enzima.

Para explicar la desaparición no específica de NADH, los resultados obtenidos en un ensayo de control que carece de enzima se restaron de los resultados obtenidos en las muestras de ensayo. Las actividades específicas se calcularon como $\Delta\mu$mol NADH/min. mg de proteína.

Tabla 2: Actividad fisiológica de algunas HMG-CoA liasas purificadas

Número de acceso Uniprot	Organismo	Actividad fisiológica μmol/min · mg proteína
A8WG57	*Danio rerio* (pez cebra) (Brachydanio rerio)	4,05
Q29448	*Bos taurus* (Bovino)	5,79
B6U7B9	*Zea mays*	13,31
A5FHS2	*Flavobacterium johnsoniae*	2,89
A1VJH1	*Polaromonas naphthalenivorans*	34,92
A9IFQ7	*Bordetella petrii*	9,84
A9IR28	*Bordetella petrii*	1,74
A1VT25	*Polaromonas naphthalenivorans*	0,39

Ejemplo 5

Producción de 3-hidroxi-3-metilbutirato

La reacción completa para la síntesis de 3-hidroxi-3-metilbutirato contenía 40 mM de Tris-HCl pH 8, 5 a 50 mM de acetil-CoA, 100 a 500 mM de acetona, 1 $MgCl_2$ (excepto para HMG-CoA sintasa mitocondrial), 0,5 mM de DTT y una enzima que varía en el intervalo de 0,2 a 8 mg/ml. Se llevaron a cabo reacciones de control en ausencia de enzima y uno de los sustratos.

El transcurso de la síntesis fue seguido analizando alícuotas tomadas tras incrementar el período de incubación a 30 ó 37ºC. Típicamente, se retiró una alícuota de 50 μl después de 48 h de incubación, se calentó durante 1 min. a 100ºC para eliminar las proteínas, se centrifugó, y el sobrenadante se transfirió a un vial limpio para la detección de HIV mediante espectrometría de masas. Se preparó una disolución de 3-hidroxi-3-metilbutirato en 40 mM de Tris-HCl pH 8, 1 mM de $MgCl_2$, 0,5 mM de DTT, se calentó como se describe previamente, y se usó como patrón.

Las muestras se analizaron en un espectrómetro de masas de cuadrupolo triple PE SCIEX API 2000 en modo de ion negativo con H_2O/acetonitrilo = 60/40 que contiene 0,1% de trietilamina como fase móvil, y el caudal fue 40 μl/min. Se mezclaron 10 μl de cada sobrenadante con una cantidad igual de fase móvil, y se inyectó directamente en el espectrómetro de masas. Se monitorizó la presencia de ion [3-hidroxi-3-metilbutirato-H]$^-$. Se observó un pico correspondiente a hidroxi-3-metilbutirato para las siguientes enzimas:

Blattella germanica (cucaracha germana) P54961 (SEC ID NO: 6)

Gallus gallus (pollo) P23228 (SEC ID NO: 7)

Homo sapiens (ser humano) Q01581 (SEC ID NO: 8)

Arabidopsis thaliana P54873 (CAA58763) (SEC ID NO: 4)

Caenorhabditis elegans P54871 (SEC ID NO: 1)

Schizosaccharomyces pombe (levadura de fisión) P54874 (SEC ID NO: 2)

Saccharomyces cerevisiae (levadura de panadero) P54839 (SEC ID NO: 3)

Dictyostelium discoideum (moho del cieno) Q86HL5 (SEC ID NO: 10)

Leuconostoc mesenteroides Q03WZ0

Staphylococcus epidermidis Q8CN06 (SEC ID NO: 11)

Lactobacillus delbrueckii Q1GAH5 (SEC ID NO: 24)

Staphylococcus haemolyticus Q4L958 (198>V diferencia comparada con la proteína de tipo salvaje) (SEC ID NO: 25)

Las Figuras 6 a 8 muestran resultados representativos para 3-hidroxi-3-metilbutirato comercialmente disponible, para la reacción que usa la CoA sintasa de Gallus gallus (P23228), y para el ensayo de control sin enzima.

Ejemplo 6

Producción de 3-hidroxi-3-metilbutiril-CoA usando liasas (que cae fuera del alcance de las reivindicaciones)

La síntesis de 3-hidroxi-3-metilbutiril-CoA se lleva a cabo en presencia de [2-^{14}C] acetona radiomarcada. La reacción completa para la síntesis de 3-hidroxi-3-metilbutiril-CoA contiene 40 mM de Tris-HCl pH 8, 5 a 50 mM de acetil-CoA, 100 a 500 mM de acetona, 1 a 10 mM de $MgCl_2$, 0,5 mM de DTT, y enzima que varía en el intervalo de 0,5 a 7 mg/ml. La formación de producto se analiza tras la separación de la mezcla de reacción mediante TLC o HPLC.

3-Hidroxi-3-metilbutiril-CoA también se analiza mediante el método de TLC (Stadtman E.R., J. Biol. Chem. 196 (1952), 535-546). Se deposita una alícuota de reacción sobre una placa de celulosa y se cromatografía en el siguiente sistema de disolvente: etanol/0,1 M de acetato de sodio pH 4,5 (1/1). Como patrones internos, se usan Co-A y acetil-CoA. El R_f dado a conocer para 3-hidroxi-3-metilbutiril-CoA es 0,88.

Ejemplo 7

Parámetros cinéticos para la reacción enzimática entre acetil-CoA y acetona en el caso de HMG sintasas

Los parámetros cinéticos se midieron usando una concentración variable de acetona y una concentración constante de acetil-CoA (10 mM), en las siguientes condiciones:

40 mM de Tris-HCl pH 8

2 mM de $MgCl_2$

0-1 M de acetona

El pH final se ajustó a 8.

La reacción se inició mediante adición de 3 mg de enzima purificada a 1 ml de la mezcla de reacción. La mezcla se incubó entonces sin agitación a 37ºC durante 40 h.

Análisis de la producción de 3-hidroxi-3-metilbutirato

Se aplicaron condiciones termoquímicas que conducen a la descomposición de 3-hidroxi-3-metilbutirato en isobuteno (Pressman et al., JACS, 1940, 2069-2080): el pH de las mezclas de reacción se ajustó en primer lugar a pH 4 usando HCl 6N, y las muestras se transfirieron entonces a viales de cromatografía de gases (Interchim). Los viales se cerraron herméticamente y se incubaron a 110ºC durante 4 horas, conduciendo así a la descomposición de 3-hidroxi-3-metilbutirato en isobuteno.

La curva de calibración se preparó en las mismas condiciones usando 3-hidroxi-3-metilbutirato comercial.

Se recogió un mililitro de gas del espacio de cabeza y se inyectó en un cromatógrafo de gases HP5890 (HP) equipado con un detector FID y una columna CP SilicaPlot (Varian). Como patrón, se usó isobuteno comercial. Para la señal de isobuteno, se calculó la cantidad de 3-hidroxi-3-metilbutirato presente inicialmente en la muestra.

Los parámetros cinéticos para varias de las HMG-CoA sintasas estudiadas se presentan en la siguiente Tabla.

Organismo	K_M para acetona, mM	k_{cat} x 10^{-4}, s^{-1}	k_{cat}/K_M x 10^{-6}, mM^{-1} x s^{-1}
Gallus gallus	250	5	2
Staphylococcus epidermidis	200	0,6	0,3
Schizosaccharomyces pombe	200	0,2	0,1

Para la enzima procedente de S. epidermidis, la Figura 9 muestra una gráfica de Michaelis-Menten correspondiente.

LISTADO DE SECUENCIAS

<110> MARLIÈRE, Philippe

<120> Método para la producción de ácido 3-hidroxi-3-metilbutírico a partir de acetona y un compuesto acetílico activado

<130> R1889 PCT S3

<150> EP 09 17 0312.4

<151> 15-09-2009

<160> 25

<170> PatentIn version 3.5

<210> 1

<211> 462

<212> PRT

<213> *Caenorhabditis elegans*

<400> 1

```
Met Ser Leu Gly Gln Leu Ser Tyr Thr Pro Val Thr Asp Val Gly Ile
1               5                  10                  15

Gly Ala Ile Glu Leu Tyr Phe Pro Gln Asn Phe Val Asp Gln Asn Asp
            20                  25                  30

Leu Glu Lys Phe Asn Asn Val Ser Ser Gly Lys Tyr Thr Ile Gly Leu
        35                  40                  45

Gly Gln Gln Gln Met Gly Phe Cys Ser Asp Asn Glu Asp Ile Val Ser
    50                  55                  60

Ile Ser Leu Thr Val Thr Arg Lys Leu Ile Glu Thr Tyr Lys Ile Ser
65                  70                  75                  80

Thr Asp Ser Ile Gly Cys Leu Val Val Gly Thr Glu Thr Met Ile Asp
                85                  90                  95

Lys Ser Lys Ser Val Lys Thr Ala Leu Met Asp Leu Phe Pro Gly Asn
            100                 105                 110

Ser Asp Ile Glu Gly Val Asp Ile Lys Asn Ala Cys Phe Gly Gly Ala
        115                 120                 125

Gln Ala Leu Leu His Ala Ile Asp Trp Val Thr Val Asn His Pro Leu
    130                 135                 140

Asp Lys Lys Asn Ala Ile Val Val Val Ala Asp Ile Ala Ile Tyr Glu
145                 150                 155                 160

Glu Gly Pro Ala Arg Cys Thr Gly Gly Ala Gly Ala Ile Ala Phe Leu
                165                 170                 175
```

```
Ile Cys Pro Asp Ala Ser Ile Pro Ile Asp Arg Gln Phe Ser Ala Cys
            180                 185                 190

His Met Lys Asn Thr Trp Asp Phe Phe Lys Pro Ile Thr Pro Ile Pro
        195                 200                 205

Ser Glu Tyr Pro Val Val Asp Gly Ser Leu Ser Leu Ser Ser Tyr Leu
    210                 215                 220

Glu Ala Val Arg Met Thr Tyr Thr Tyr Phe Ile Ser Lys Val Asn Arg
225                 230                 235                 240

His Thr Thr Gly Ile Asp Gly Leu Asn Ser Phe Asp Gly Val Phe Leu
                245                 250                 255

His Ser Pro Phe Thr Lys Met Val Gln Lys Gly Leu Ala Val Met Asn
            260                 265                 270

Tyr Thr Asp Ser Gln Leu Arg His Lys Gln Leu Asn Gly Asn Gly Val
        275                 280                 285

Asp His Lys Leu Asp Glu Asn Asp Arg Ala Gly Leu Ala Lys Met Ile
    290                 295                 300

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

Phe Asn Arg Arg Ile Gly Asn Met Tyr Thr Pro Ser Leu Phe Ala Gln
                325                 330                 335

Leu Leu Ala Tyr Leu Ala Ala Asp Asp Cys Val Thr Gly Glu Lys Ser
            340                 345                 350

Ile Leu Phe Phe Ala Tyr Gly Ser Gly Leu Ala Ser Ala Ile Phe Pro
        355                 360                 365

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

Ile Arg Ala Ile Lys Arg Leu Asp Asp Arg Ile Gln Phe Thr Pro Glu
385                 390                 395                 400

Glu Phe Thr Glu Thr Leu Gln Lys Arg Glu Val Phe Leu Arg Ser Lys
                405                 410                 415

Glu Ile Pro Lys Ser Pro Ser Glu Thr Ser Leu Phe Pro Asn Thr Tyr
            420                 425                 430

Phe Leu Asp Asn Met Asp Lys Leu Tyr Arg Arg Ser Tyr Thr Leu His
        435                 440                 445

Glu Glu Pro Asn Gly Val Gln Asn Gly Asn Gly Ile His His
    450                 455                 460
```

<210> 2

<211> 447

<212> PRT

<213> *Schizosaccharomyces pombe*

<400> 2

```
Met Ser Phe Asp Arg Lys Asp Ile Gly Ile Lys Gly Leu Val Leu Tyr
1               5                  10                  15

Thr Pro Asn Gln Tyr Val Glu Gln Ala Ala Leu Glu Ala His Asp Gly
            20                  25                  30

Val Ser Thr Gly Lys Tyr Thr Ile Gly Leu Gly Leu Thr Lys Met Ala
        35                  40                  45

Phe Val Asp Asp Arg Glu Asp Ile Tyr Ser Phe Gly Leu Thr Ala Leu
    50                  55                  60

Ser Gln Leu Ile Lys Arg Tyr Gln Ile Asp Ile Ser Lys Ile Gly Arg
65                  70                  75                  80

Leu Glu Val Gly Thr Glu Thr Ile Ile Asp Lys Ser Lys Ser Val Lys
                85                  90                  95

Ser Val Leu Met Gln Leu Phe Gly Asp Asn His Asn Val Glu Gly Ile
            100                 105                 110

Asp Cys Val Asn Ala Cys Tyr Gly Gly Val Asn Ala Leu Phe Asn Thr
        115                 120                 125

Ile Asp Trp Ile Glu Ser Ser Ala Trp Asp Gly Arg Asp Gly Ile Val
    130                 135                 140

Val Ala Gly Asp Ile Ala Leu Tyr Ala Lys Gly Asn Ala Arg Pro Thr
145                 150                 155                 160

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

Val Phe Glu Pro Gly Leu Arg Gly Thr Tyr Met Gln His Ala Tyr Asp
            180                 185                 190

Phe Tyr Lys Pro Asp Leu Thr Ser Glu Tyr Pro Tyr Val Asp Gly His
        195                 200                 205

Phe Ser Leu Glu Cys Tyr Val Lys Ala Leu Asp Gly Ala Tyr Ala Asn
    210                 215                 220

Tyr Asn Val Arg Asp Val Ala Lys Asn Gly Lys Ser Gln Gly Leu Gly
```

```
225                 230                 235                 240

Leu Asp Arg Phe Asp Tyr Cys Ile Phe His Ala Pro Thr Cys Lys Gln
                245                 250                 255

Val Gln Lys Ala Tyr Ala Arg Leu Leu Tyr Thr Asp Ser Ala Ala Glu
            260                 265                 270

Pro Ser Asn Pro Glu Leu Glu Gly Val Arg Glu Leu Leu Ser Thr Leu
        275                 280                 285

Asp Ala Lys Lys Ser Leu Thr Asp Lys Ala Leu Glu Lys Gly Leu Met
    290                 295                 300

Ala Ile Thr Lys Glu Arg Phe Asn Lys Arg Val Ser Pro Ser Val Tyr
305                 310                 315                 320

Ala Pro Thr Asn Cys Gly Asn Met Tyr Thr Ala Ser Ile Phe Ser Cys
                325                 330                 335

Leu Thr Ala Leu Leu Ser Arg Val Pro Ala Asp Glu Leu Lys Gly Lys
            340                 345                 350

Arg Val Gly Ala Tyr Ser Tyr Gly Ser Gly Leu Ala Ala Ser Phe Phe
        355                 360                 365

Ser Phe Val Val Lys Gly Asp Val Ser Glu Ile Ala Lys Lys Thr Asn
    370                 375                 380

Leu Val Asn Asp Leu Asp Asn Arg His Cys Leu Thr Pro Thr Gln Tyr
385                 390                 395                 400

Glu Glu Ala Ile Glu Leu Arg His Gln Ala His Leu Lys Lys Asn Phe
                405                 410                 415

Thr Pro Lys Gly Ser Ile Glu Arg Leu Arg Ser Gly Thr Tyr Tyr Leu
            420                 425                 430

Thr Gly Ile Asp Asp Met Phe Arg Arg Ser Tyr Ser Val Lys Pro
        435                 440                 445
```

<210> 3

<211> 491

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 3

```
Met Lys Leu Ser Thr Lys Leu Cys Trp Cys Gly Ile Lys Gly Arg Leu
1               5                   10                  15

Arg Pro Gln Lys Gln Gln Gln Leu His Asn Thr Asn Leu Gln Met Thr
            20                  25                  30
```

```
Glu Leu Lys Lys Gln Lys Thr Ala Glu Gln Lys Thr Arg Pro Gln Asn
        35                  40                  45

Val Gly Ile Lys Gly Ile Gln Ile Tyr Ile Pro Thr Gln Cys Val Asn
    50                  55                  60

Gln Ser Glu Leu Glu Lys Phe Asp Gly Val Ser Gln Gly Lys Tyr Thr
65                  70                  75                  80

Ile Gly Leu Gly Gln Thr Asn Met Ser Phe Val Asn Asp Arg Glu Asp
                85                  90                  95

Ile Tyr Ser Met Ser Leu Thr Val Leu Ser Lys Leu Ile Lys Ser Tyr
            100                 105                 110

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

Leu Ile Asp Lys Ser Lys Ser Val Lys Ser Val Leu Met Gln Leu Phe
    130                 135                 140

Gly Glu Asn Thr Asp Val Glu Gly Ile Asp Thr Leu Asn Ala Cys Tyr
145                 150                 155                 160

Gly Gly Thr Asn Ala Leu Phe Asn Ser Leu Asn Trp Ile Glu Ser Asn
                165                 170                 175

Ala Trp Asp Gly Arg Asp Ala Ile Val Val Cys Gly Asp Ile Ala Ile
            180                 185                 190

Tyr Asp Lys Gly Ala Ala Arg Pro Thr Gly Gly Ala Gly Thr Val Ala
        195                 200                 205

Met Trp Ile Gly Pro Asp Ala Pro Ile Val Phe Asp Ser Val Arg Ala
    210                 215                 220

Ser Tyr Met Glu His Ala Tyr Asp Phe Tyr Lys Pro Asp Phe Thr Ser
225                 230                 235                 240

Glu Tyr Pro Tyr Val Asp Gly His Phe Ser Leu Thr Cys Tyr Val Lys
                245                 250                 255

Ala Leu Asp Gln Val Tyr Lys Ser Tyr Ser Lys Lys Ala Ile Ser Lys
            260                 265                 270

Gly Leu Val Ser Asp Pro Ala Gly Ser Asp Ala Leu Asn Val Leu Lys
        275                 280                 285

Tyr Phe Asp Tyr Asn Val Phe His Val Pro Thr Cys Lys Leu Val Thr
    290                 295                 300
```

Lys Ser Tyr Gly Arg Leu Leu Tyr Asn Asp Phe Arg Ala Asn Pro Gln
305 310 315 320

Leu Phe Pro Glu Val Asp Ala Glu Leu Ala Thr Arg Asp Tyr Asp Glu
325 330 335

Ser Leu Thr Asp Lys Asn Ile Glu Lys Thr Phe Val Asn Val Ala Lys
340 345 350

Pro Phe His Lys Glu Arg Val Ala Gln Ser Leu Ile Val Pro Thr Asn
355 360 365

Thr Gly Asn Met Tyr Thr Ala Ser Val Tyr Ala Ala Phe Ala Ser Leu
370 375 380

Leu Asn Tyr Val Gly Ser Asp Asp Leu Gln Gly Lys Arg Val Gly Leu
385 390 395 400

Phe Ser Tyr Gly Ser Gly Leu Ala Ala Ser Leu Tyr Ser Cys Lys Ile
405 410 415

Val Gly Asp Val Gln His Ile Ile Lys Glu Leu Asp Ile Thr Asn Lys
420 425 430

Leu Ala Lys Arg Ile Thr Glu Thr Pro Lys Asp Tyr Glu Ala Ala Ile
435 440 445

Glu Leu Arg Glu Asn Ala His Leu Lys Lys Asn Phe Lys Pro Gln Gly
450 455 460

Ser Ile Glu His Leu Gln Ser Gly Val Tyr Tyr Leu Thr Asn Ile Asp
465 470 475 480

Asp Lys Phe Arg Arg Ser Tyr Asp Val Lys Lys
485 490

<210> 4

<211> 461

<212> PRT

<213> *Arabidopsis thaliana*

<400> 4

Met Ala Lys Asn Val Gly Ile Leu Ala Met Asp Ile Tyr Phe Pro Pro
1 5 10 15

Thr Cys Val Gln Gln Glu Ala Leu Glu Ala His Asp Gly Ala Ser Lys
20 25 30

Gly Lys Tyr Thr Ile Gly Leu Gly Gln Asp Cys Leu Ala Phe Cys Thr
35 40 45

```
Glu Leu Glu Asp Val Ile Ser Met Ser Phe Asn Ala Val Thr Ser Leu
    50                  55                  60

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

Gly Ser Glu Thr Val Ile Asp Lys Ser Lys Ser Ile Lys Thr Phe Leu
                85                  90                  95

Met Gln Leu Phe Glu Lys Cys Gly Asn Thr Asp Val Glu Gly Val Asp
            100                 105                 110

Ser Thr Asn Ala Cys Tyr Gly Gly Thr Ala Ala Leu Leu Asn Cys Val
        115                 120                 125

Asn Trp Val Glu Ser Asn Ser Trp Asp Gly Arg Tyr Gly Leu Val Ile
    130                 135                 140

Cys Thr Asp Ser Ala Val Tyr Ala Glu Gly Pro Ala Arg Pro Thr Gly
145                 150                 155                 160

Gly Ala Ala Ala Ile Ala Met Leu Ile Gly Pro Asp Ala Pro Ile Val
                165                 170                 175

Phe Glu Ser Lys Leu Arg Ala Ser His Met Ala His Val Tyr Asp Phe
            180                 185                 190

Tyr Lys Pro Asn Leu Ala Ser Glu Tyr Pro Val Val Asp Gly Lys Leu
        195                 200                 205

Ser Gln Thr Cys Tyr Leu Met Ala Leu Asp Ser Cys Tyr Lys His Leu
    210                 215                 220

Cys Asn Lys Phe Glu Lys Ile Glu Gly Lys Glu Phe Ser Ile Asn Asp
225                 230                 235                 240

Ala Asp Tyr Ile Val Phe His Ser Pro Tyr Asn Lys Leu Val Gln Lys
                245                 250                 255

Ser Phe Ala Arg Leu Leu Tyr Asn Asp Phe Leu Arg Asn Ala Ser Ser
            260                 265                 270

Ile Asp Glu Ala Ala Lys Glu Lys Phe Thr Pro Tyr Ser Ser Leu Thr
        275                 280                 285

Leu Asp Glu Ser Tyr Gln Ser Arg Asp Leu Glu Lys Val Ser Gln Gln
    290                 295                 300

Ile Ser Lys Pro Phe Tyr Asp Ala Lys Val Gln Pro Thr Thr Leu Ile
305                 310                 315                 320
```

Pro Lys Glu Val Gly Asn Met Tyr Thr Ala Ser Leu Tyr Ala Ala Phe
325 330 335

Ala Ser Leu Ile His Asn Lys His Asn Asp Leu Ala Gly Lys Arg Val
340 345 350

Val Met Phe Ser Tyr Gly Ser Gly Ser Thr Ala Thr Met Phe Ser Leu
355 360 365

Arg Leu Asn Asp Asn Lys Pro Pro Phe Ser Ile Ser Asn Ile Ala Ser
370 375 380

Val Met Asp Val Gly Gly Lys Leu Lys Ala Arg His Glu Tyr Ala Pro
385 390 395 400

Glu Lys Phe Val Glu Thr Met Lys Leu Met Glu His Arg Tyr Gly Ala
405 410 415

Lys Asp Phe Val Thr Thr Lys Glu Gly Ile Ile Asp Leu Leu Ala Pro
420 425 430

Gly Thr Tyr Tyr Leu Lys Glu Val Asp Ser Leu Tyr Arg Arg Phe Tyr
435 440 445

Gly Lys Lys Gly Glu Asp Gly Ser Val Ala Asn Gly His
450 455 460

<210> 5

<211> 482

<212> PRT

<213> *Dictyostelium discoideum*

<400> 5

Met Thr Lys Pro Glu Asn Ile Gly Ile His Gly Ile Glu Val Tyr Phe
1 5 10 15

Pro Ser Thr Tyr Val Ala Gln Glu Asp Leu Glu Lys Phe Asp Gly Val
20 25 30

Ser Gln Gly Lys Tyr Thr Leu Gly Leu Gly Gln Thr Asn Met Ala Phe
35 40 45

Cys Gly Asp Arg Glu Asp Ile Tyr Ser Leu Ser Leu Asn Ala Val Asn
50 55 60

Asn Leu Met Asp Lys Phe Asn Val Asp Pro Asn Ser Ile Gly Arg Leu
65 70 75 80

Glu Val Gly Thr Glu Thr Val Ile Asp Lys Ser Lys Ser Val Lys Thr
85 90 95

```
Val Leu Met Asp Leu Phe Ala Lys His Gly Asn Thr Ser Ile Asp Gly
            100                 105                 110

Ile Asp Thr Ile Asn Ala Cys Tyr Gly Gly Thr Ser Ala Leu His Asn
        115                 120                 125

Ala Leu Gln Trp Met Glu Ser Ser Tyr Trp Asp Gly Arg Asn Ala Ile
    130                 135                 140

Val Val Ala Gly Asp Ile Ala Val Tyr Glu Lys Gly Pro Ala Arg Pro
145                 150                 155                 160

Thr Gly Gly Ala Gly Val Val Ala Met Leu Ile Gly Pro Asn Ala Pro
                165                 170                 175

Ile Thr Phe Glu Ser Gly Leu Arg Gly Val His Met Glu Asn Val Tyr
            180                 185                 190

Asp Phe Tyr Lys Pro Asp Met Asp Ser Glu Tyr Pro Arg Val Asp Gly
        195                 200                 205

Lys Leu Ser Ile Ser Cys Tyr Phe Arg Ala Ile Asp Asn Cys Tyr Asn
    210                 215                 220

Arg Tyr Ala Lys Ala Phe Glu Lys Lys Tyr Gly Lys Ser Phe Ser Leu
225                 230                 235                 240

Asp Gln Val Asp Phe Ala Leu Phe His Ser Pro Tyr Asn Lys Leu Val
                245                 250                 255

Gln Lys Ser Phe Gly Arg Met Leu Tyr Asn Asp Phe Leu Asn Asn Pro
            260                 265                 270

Asn Asp Ser Arg Tyr Ala Ser Leu Glu Ala Tyr Lys Asn Val Lys Pro
        275                 280                 285

Glu Asp Thr Tyr Phe Asp Ser Val Leu Glu Lys Ala Leu Ser Ala Ile
    290                 295                 300

Thr Lys Asn Asp Tyr Ala Thr Lys Val Ala Pro Thr Thr Leu Leu Ala
305                 310                 315                 320

Lys Gln Leu Gly Asn Thr Tyr Cys Gly Ser Thr Tyr Ser Gly Leu Leu
                325                 330                 335

Ser Leu Leu Asp Glu Lys Ser Asn Asp Leu Val Gly Lys Arg Val Leu
            340                 345                 350

Thr Phe Ser Tyr Gly Ser Gly Leu Ala Ala Ser Ala Phe Ser Phe Lys
        355                 360                 365
```

Val Glu Lys Pro Ile Asn His Ile Val Glu Lys Val Asp Leu Lys Asn
370 375 380

Arg Leu Ala Lys Arg Val Arg Val Glu Pro Glu Ile Phe Thr Glu Lys
385 390 395 400

Leu Ser Leu Arg Glu Thr Arg His Asn Leu Lys Asn Tyr Val Pro Ser
405 410 415

Asp Glu Thr Thr Asn Met Phe Pro Gly Ser Phe Tyr Leu Ser Ser Val
420 425 430

Asp Asn Ala Gly Ile Arg Lys Tyr Asp Arg Thr Tyr Ser Thr Ser Ala
435 440 445

Val Leu Gly Ala Phe Gln Arg Arg Gln Gln Ile Ser Gln Ser Thr Ile
450 455 460

Lys Ser Leu Asn Leu Phe Arg Ala Thr Lys Ser Val Leu Ser Ile Leu
465 470 475 480

Lys Lys

<210> 6

<211> 453

<212> PRT

<213> *Blattella germanica*

<400> 6

Met Trp Pro Ser Asp Val Gly Ile Val Ala Leu Glu Leu Ile Phe Pro
1 5 10 15

Ser Gln Tyr Val Asp Gln Val Asp Leu Glu Val Tyr Asp Asn Val Ser
20 25 30

Ala Gly Lys Tyr Thr Val Gly Leu Gly Gln Ala Arg Met Gly Phe Cys
35 40 45

Thr Asp Arg Glu Asp Ile Asn Ser Leu Cys Leu Thr Val Val Ser Arg
50 55 60

Leu Met Glu Arg Trp Ser Ile Pro Tyr Ser Gln Ile Gly Arg Leu Glu
65 70 75 80

Val Gly Thr Glu Thr Leu Leu Asp Lys Ser Lys Ser Val Lys Thr Val
85 90 95

Leu Met Gln Leu Phe Lys Asp Asn Thr Asp Ile Glu Gly Val Asp Thr
100 105 110

Val Asn Ala Cys Tyr Gly Gly Thr Ser Ala Leu Phe Asn Ala Ile Ser

```
        115                 120                 125

Trp Val Glu Ser Ser Ser Trp Asp Gly Arg Tyr Ala Leu Val Val Ala
    130                 135                 140

Gly Asp Ile Ala Val Tyr Ala Lys Gly Ser Ala Arg Pro Thr Gly Gly
145                 150                 155                 160

Ala Gly Ala Val Ala Met Leu Val Gly Ala Asn Ala Pro Leu Val Phe
                165                 170                 175

Asp Arg Gly Val Arg Ser Ser His Met Gln His Ala Tyr Asp Phe Tyr
            180                 185                 190

Lys Pro Asp Leu Ser Ser Leu Tyr Pro Thr Val Asp Gly Lys Leu Ser
        195                 200                 205

Ile Gln Cys Tyr Leu Ser Ala Leu Asp His Cys Tyr Gln Leu Tyr Cys
    210                 215                 220

Ser Lys Ile Gln Lys Gln Leu Gly Glu Lys Phe Asp Ile Glu Arg Leu
225                 230                 235                 240

Asp Ala Val Leu Phe His Ala Pro Tyr Cys Lys Leu Val Gln Lys Ser
                245                 250                 255

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

Arg Thr Thr Lys Tyr Ser Ser Leu Glu Ala Leu Lys Gly Val Lys Leu
        275                 280                 285

Glu Asp Thr Tyr Phe Asp Arg Glu Val Glu Lys Ala Val Met Thr Tyr
    290                 295                 300

Ser Lys Asn Met Phe Glu Glu Lys Thr Lys Pro Ser Leu Leu Leu Ala
305                 310                 315                 320

Asn Gln Val Gly Asn Met Tyr Thr Pro Ser Leu Tyr Gly Gly Leu Val
                325                 330                 335

Ser Leu Leu Val Ser Lys Ser Ala Gln Glu Leu Ala Gly Lys Arg Val
            340                 345                 350

Ala Leu Phe Ser Tyr Gly Ser Gly Leu Ala Ser Ser Met Phe Ser Leu
        355                 360                 365

Arg Ile Ser Ser Asp Ala Ser Ala Lys Ser Ser Leu Gln Arg Leu Val
    370                 375                 380

Ser Asn Leu Ser His Ile Lys Pro Gln Leu Asp Leu Arg His Lys Val
```

385 390 395 400

Ser Pro Glu Glu Phe Ala Gln Thr Met Glu Thr Arg Glu His Asn His
405 410 415

His Lys Ala Pro Tyr Thr Pro Glu Gly Ser Ile Asp Val Leu Phe Pro
420 425 430

Gly Thr Trp Tyr Leu Glu Ser Val Asp Ser Leu Tyr Arg Arg Ser Tyr
435 440 445

Lys Gln Val Pro Gly
450

<210> 7

<211> 522

<212> PRT

<213> *Gallus gallus*

<400> 7

Met Pro Gly Ser Leu Pro Val Asn Thr Glu Ser Cys Trp Pro Lys Asp
1 5 10 15

Val Gly Ile Val Ala Leu Glu Ile Tyr Phe Pro Ser Gln Tyr Val Asp
20 25 30

Gln Thr Glu Leu Glu Lys Tyr Asp Gly Val Asp Ala Gly Lys Tyr Thr
35 40 45

Ile Gly Leu Gly Gln Ser Lys Met Gly Phe Cys Ser Asp Arg Glu Asp
50 55 60

Ile Asn Ser Leu Cys Leu Thr Val Val Gln Lys Leu Met Glu Arg Asn
65 70 75 80

Ser Leu Ser Tyr Asp Cys Ile Gly Arg Leu Glu Val Gly Thr Glu Thr
85 90 95

Ile Ile Asp Lys Ser Lys Ser Val Lys Thr Val Leu Met Gln Leu Phe
100 105 110

Glu Glu Ser Gly Asn Thr Asp Val Glu Gly Ile Asp Thr Thr Asn Ala
115 120 125

Cys Tyr Gly Gly Thr Ala Ala Leu Phe Asn Ala Ile Asn Trp Ile Glu
130 135 140

Ser Ser Ser Trp Asp Gly Arg Tyr Ala Leu Val Val Ala Gly Asp Ile
145 150 155 160

Ala Val Tyr Ala Thr Gly Asn Ala Arg Pro Thr Gly Gly Ala Gly Ala
165 170 175

Val Ala Met Leu Val Gly Ser Asn Ala Pro Leu Ile Phe Glu Arg Gly
180 185 190

Leu Arg Gly Thr His Met Gln His Ala Tyr Asp Phe Tyr Lys Pro Asp
195 200 205

Met Val Ser Glu Tyr Pro Val Val Asp Gly Lys Leu Ser Ile Gln Cys
210 215 220

Tyr Leu Ser Ala Leu Asp Arg Cys Tyr Ser Val Tyr Arg Asn Lys Ile
225 230 235 240

His Ala Gln Trp Gln Lys Glu Gly Thr Asp Arg Gly Phe Thr Leu Asn
245 250 255

Asp Phe Gly Phe Met Ile Phe His Ser Pro Tyr Cys Lys Leu Val Gln
260 265 270

Lys Ser Val Ala Arg Leu Leu Leu Asn Asp Phe Leu Ser Asp Gln Asn
275 280 285

Ala Glu Thr Ala Asn Gly Val Phe Ser Gly Leu Glu Ala Phe Arg Asp
290 295 300

Val Lys Leu Glu Asp Thr Tyr Phe Asp Arg Asp Val Glu Lys Ala Phe
305 310 315 320

Met Lys Ala Ser Ala Glu Leu Phe Asn Gln Lys Thr Lys Ala Ser Leu
325 330 335

Leu Val Ser Asn Gln Asn Gly Asn Met Tyr Thr Pro Ser Val Tyr Gly
340 345 350

Cys Leu Ala Ser Leu Leu Ala Gln Tyr Ser Pro Glu His Leu Ala Gly
355 360 365

Gln Arg Ile Ser Glu Phe Ser Tyr Gly Ser Gly Phe Ala Ala Thr Leu
370 375 380

Tyr Ser Ile Arg Val Thr Gln Asp Ala Thr Pro Gly Ser Ala Leu Asp
385 390 395 400

Lys Ile Thr Ala Ser Leu Ser Asp Leu Lys Ala Arg Leu Asp Ser Arg
405 410 415

Lys Cys Ile Ala Pro Asp Val Phe Ala Glu Asn Met Lys Ile Arg Gln
420 425 430

Glu Thr His His Leu Ala Asn Tyr Ile Pro Gln Cys Ser Val Glu Asp
435 440 445

Leu Phe Glu Gly Thr Trp Tyr Leu Val Arg Val Asp Glu Lys His Arg
450 455 460

Arg Thr Tyr Ala Arg Arg Pro Val Met Gly Asp Gly Pro Leu Glu Ala
465 470 475 480

Gly Val Glu Val Val His Pro Gly Ile Val His Glu His Ile Pro Ser
485 490 495

Pro Ala Lys Lys Val Pro Arg Ile Pro Ala Thr Thr Glu Ser Glu Gly
500 505 510

Val Thr Val Ala Ile Ser Asn Gly Val His
515 520

<210> 8

<211> 520

<212> PRT

<213> *Homo sapiens*

<400> 8

Met Pro Gly Ser Leu Pro Leu Asn Ala Glu Ala Cys Trp Pro Lys Asp
1 5 10 15

Val Gly Ile Val Ala Leu Glu Ile Tyr Phe Pro Ser Gln Tyr Val Asp
20 25 30

Gln Ala Glu Leu Glu Lys Tyr Asp Gly Val Asp Ala Gly Lys Tyr Thr
35 40 45

Ile Gly Leu Gly Gln Ala Lys Met Gly Phe Cys Thr Asp Arg Glu Asp
50 55 60

Ile Asn Ser Leu Cys Met Thr Val Val Gln Asn Leu Met Glu Arg Asn
65 70 75 80

Asn Leu Ser Tyr Asp Cys Ile Gly Arg Leu Glu Val Gly Thr Glu Thr
85 90 95

Ile Ile Asp Lys Ser Lys Ser Val Lys Thr Asn Leu Met Gln Leu Phe
100 105 110

Glu Glu Ser Gly Asn Thr Asp Ile Glu Gly Ile Asp Thr Thr Asn Ala
115 120 125

Cys Tyr Gly Gly Thr Ala Ala Val Phe Asn Ala Val Asn Trp Ile Glu
130 135 140

Ser Ser Ser Trp Asp Gly Arg Tyr Ala Leu Val Val Ala Gly Asp Ile
145 150 155 160

```
Ala Val Tyr Ala Thr Gly Asn Ala Arg Pro Thr Gly Gly Val Gly Ala
                165                 170                 175

Val Ala Leu Leu Ile Gly Pro Asn Ala Pro Leu Ile Phe Glu Arg Gly
            180                 185                 190

Leu Arg Gly Thr His Met Gln His Ala Tyr Asp Phe Tyr Lys Pro Asp
        195                 200                 205

Met Leu Ser Glu Tyr Pro Ile Val Asp Gly Lys Leu Ser Ile Gln Cys
    210                 215                 220

Tyr Leu Ser Ala Leu Asp Arg Cys Tyr Ser Val Tyr Cys Lys Lys Ile
225                 230                 235                 240

His Ala Gln Trp Gln Lys Glu Gly Asn Asp Lys Asp Phe Thr Leu Asn
                245                 250                 255

Asp Phe Gly Phe Met Ile Phe His Ser Pro Tyr Cys Lys Leu Val Gln
            260                 265                 270

Lys Ser Leu Ala Arg Met Leu Leu Asn Asp Phe Leu Asn Asp Gln Asn
        275                 280                 285

Arg Asp Lys Asn Ser Ile Tyr Ser Gly Leu Glu Ala Phe Gly Asp Val
    290                 295                 300

Lys Leu Glu Asp Thr Tyr Phe Asp Arg Asp Val Glu Lys Ala Phe Met
305                 310                 315                 320

Lys Ala Ser Ser Glu Leu Phe Ser Gln Lys Thr Lys Ala Ser Leu Leu
                325                 330                 335

Val Ser Asn Gln Asn Gly Asn Met Tyr Thr Ser Ser Val Tyr Gly Ser
            340                 345                 350

Leu Ala Ser Val Leu Ala Gln Tyr Ser Pro Gln Gln Leu Ala Gly Lys
        355                 360                 365

Arg Ile Gly Val Phe Ser Tyr Gly Ser Gly Leu Ala Ala Thr Leu Tyr
    370                 375                 380

Ser Leu Lys Val Thr Gln Asp Ala Thr Pro Gly Ser Ala Leu Asp Lys
385                 390                 395                 400

Ile Thr Ala Ser Leu Cys Asp Leu Lys Ser Arg Leu Asp Ser Arg Thr
                405                 410                 415

Gly Val Ala Pro Asp Val Phe Ala Glu Asn Met Lys Leu Arg Glu Asp
            420                 425                 430
```

Thr His His Leu Val Asn Tyr Ile Pro Gln Gly Ser Ile Asp Ser Leu
435 440 445

Phe Glu Gly Thr Trp Tyr Leu Val Arg Val Asp Glu Lys His Arg Arg
450 455 460

Thr Tyr Ala Arg Arg Pro Thr Pro Asn Asp Asp Thr Leu Asp Glu Gly
465 470 475 480

Val Gly Leu Val His Ser Asn Ile Ala Thr Glu His Ile Pro Ser Pro
485 490 495

Ala Lys Lys Val Pro Arg Leu Pro Ala Thr Ala Ala Glu Pro Glu Ala
500 505 510

Ala Val Ile Ser Asn Gly Val Trp
515 520

<210> 9

<211> 508

<212> PRT

<213> *Homo sapiens*

<400> 9

Met Gln Arg Leu Leu Thr Pro Val Lys Arg Ile Leu Gln Leu Thr Arg
1 5 10 15

Ala Val Gln Glu Thr Ser Leu Thr Pro Ala Arg Leu Leu Pro Val Ala
20 25 30

His Gln Arg Phe Ser Thr Ala Ser Ala Val Pro Leu Ala Lys Thr Asp
35 40 45

Thr Trp Pro Lys Asp Val Gly Ile Leu Ala Leu Glu Val Tyr Phe Pro
50 55 60

Ala Gln Tyr Val Asp Gln Thr Asp Leu Glu Lys Tyr Asn Asn Val Glu
65 70 75 80

Ala Gly Lys Tyr Thr Val Gly Leu Gly Gln Thr Arg Met Gly Phe Cys
85 90 95

Ser Val Gln Glu Asp Ile Asn Ser Leu Cys Leu Thr Val Val Gln Arg
100 105 110

Leu Met Glu Arg Ile Gln Leu Pro Trp Asp Ser Val Gly Arg Leu Glu
115 120 125

Val Gly Thr Glu Thr Ile Ile Asp Lys Ser Lys Ala Val Lys Thr Val
130 135 140

```
Leu Met Glu Leu Phe Gln Asp Ser Gly Asn Thr Asp Ile Glu Gly Ile
145                 150                 155                 160

Asp Thr Thr Asn Ala Cys Tyr Gly Gly Thr Ala Ser Leu Phe Asn Ala
                165                 170                 175

Ala Asn Trp Met Glu Ser Ser Ser Trp Asp Gly Arg Tyr Ala Met Val
            180                 185                 190

Val Cys Gly Asp Ile Ala Val Tyr Pro Ser Gly Asn Ala Arg Pro Thr
        195                 200                 205

Gly Gly Ala Gly Ala Val Ala Met Leu Ile Gly Pro Lys Ala Pro Leu
    210                 215                 220

Ala Leu Glu Arg Gly Leu Arg Gly Thr His Met Glu Asn Val Tyr Asp
225                 230                 235                 240

Phe Tyr Lys Pro Asn Leu Ala Ser Glu Tyr Pro Ile Val Asp Gly Lys
                245                 250                 255

Leu Ser Ile Gln Cys Tyr Leu Arg Ala Leu Asp Arg Cys Tyr Thr Ser
            260                 265                 270

Tyr Arg Lys Lys Ile Gln Asn Gln Trp Lys Gln Ala Gly Ser Asp Arg
        275                 280                 285

Pro Phe Thr Leu Asp Asp Leu Gln Tyr Met Ile Phe His Thr Pro Phe
    290                 295                 300

Cys Lys Met Val Gln Lys Ser Leu Ala Arg Leu Met Phe Asn Asp Phe
305                 310                 315                 320

Leu Ser Ala Ser Ser Asp Thr Gln Thr Ser Leu Tyr Lys Gly Leu Glu
                325                 330                 335

Ala Phe Gly Gly Leu Lys Leu Glu Asp Thr Tyr Thr Asn Lys Asp Leu
            340                 345                 350

Asp Lys Ala Leu Leu Lys Ala Ser Gln Asp Met Phe Asp Lys Lys Thr
        355                 360                 365

Lys Ala Ser Leu Tyr Leu Ser Thr His Asn Gly Asn Met Tyr Thr Ser
    370                 375                 380

Ser Leu Tyr Gly Cys Leu Ala Ser Leu Leu Ser His His Ser Ala Gln
385                 390                 395                 400

Glu Leu Ala Gly Ser Arg Ile Gly Ala Phe Ser Tyr Gly Ser Gly Leu
                405                 410                 415
```

Ala Ala Ser Phe Phe Ser Phe Arg Val Ser Gln Asp Ala Ala Pro Gly
420 425 430

Ser Pro Leu Asp Lys Leu Val Ser Ser Thr Ser Asp Leu Pro Lys Arg
435 440 445

Leu Ala Ser Arg Lys Cys Val Ser Pro Glu Glu Phe Thr Glu Ile Met
450 455 460

Asn Gln Arg Glu Gln Phe Tyr His Lys Val Asn Phe Ser Pro Pro Gly
465 470 475 480

Asp Thr Asn Ser Leu Phe Pro Gly Thr Trp Tyr Leu Glu Arg Val Asp
485 490 495

Glu Gln His Arg Arg Lys Tyr Ala Arg Arg Pro Val
500 505

<210> 10

<211> 468

<212> PRT

<213> *Dictyostelium discoideum*

<400> 10

Met Lys Lys Thr Lys Asp Ile Gly Ile Cys Ala Ile Asp Ile Tyr Phe
1 5 10 15

Pro Gln Thr Tyr Val Asn Gln Ser Glu Leu Lys Lys Tyr Asp Lys Val
20 25 30

Ser Asn Gly Lys Tyr Thr Ile Gly Leu Gly Gln Thr Asn Met Ser Phe
35 40 45

Val Gly Asp Arg Glu Asp Ile Val Ser Met Ala Met Thr Ser Val Lys
50 55 60

Met Met Met Ser Lys Tyr Ser Ile Asp Tyr Gln Ser Ile Gly Arg Leu
65 70 75 80

Glu Val Gly Thr Glu Thr Ile Ile Asp Lys Ser Lys Ser Val Lys Ser
85 90 95

Ser Ile Met Ser Leu Phe Gln Glu Tyr Gly Asn Thr Ser Leu Glu Gly
100 105 110

Val Asp Thr Leu Asn Ala Cys Tyr Gly Gly Thr Asn Ala Leu Phe Asn
115 120 125

Ser Leu Gln Trp Ile Glu Ser Ser Tyr Trp Asp Gly Arg Tyr Ala Leu
130 135 140

Val Val Thr Gly Asp Ile Ala Val Tyr Ser Lys Gly Ala Ala Arg Pro

145 150 155 160

Thr Gly Gly Ala Gly Val Val Thr Met Leu Ile Gly Pro Asn Ala Thr
165 170 175

Leu Ile Phe Asp Gln Ser Leu Arg Gly Thr His Met Glu Asn Val Asn
180 185 190

Asp Phe Tyr Lys Pro Asp Leu Ser Ser Glu Tyr Pro Tyr Val Asp Gly
195 200 205

Lys Leu Ser Ile Glu Cys Tyr Leu Arg Ala Leu Asp Lys Cys Tyr Leu
210 215 220

Glu Tyr Lys Lys Lys Phe Glu Ser Ile Asn Asp Asp Asn Lys Phe Ser
225 230 235 240

Met Asp Ser Phe Asp Tyr Val Cys Phe His Ser Pro Tyr Asn Arg Leu
245 250 255

Val Gln Lys Ser Tyr Ala Arg Leu Ile Tyr Asn Asp Phe Leu Gln Asn
260 265 270

Pro Asn Asn Pro Lys Tyr Gln Asp Leu Leu Pro Phe Lys Asp Leu Ser
275 280 285

Thr Gly Lys Asp Ser Tyr Ile Asn Ser Lys Leu Asp Gln Ile Thr Leu
290 295 300

Lys Leu Ser Leu Asp Asp Phe Lys Thr Lys Val Asn Pro Ser Thr Leu
305 310 315 320

Leu Ser Lys Glu Cys Gly Asn Ser Tyr Cys Gly Ser Val Tyr Ser Gly
325 330 335

Ile Leu Ser Leu Leu Ser Asn Val Asn Asp Leu Asn Asn Lys Lys Val
340 345 350

Leu Val Phe Ser Tyr Gly Ser Gly Leu Ala Ala Ser Leu Phe Ser Phe
355 360 365

Arg Ile Asn Asn Asn Lys Asn Arg Asn Asn Asn Asn Asn Asn Asn Asn
370 375 380

Cys Phe Phe Lys Thr Thr Asn Asp Ile Gly Lys Ile Ser Asn Ile Lys
385 390 395 400

Glu Arg Leu Ser Asn Arg Val Lys Val Ser Pro Glu Glu Phe Thr Arg
405 410 415

Ile Leu Asp Ile Arg Glu Lys Ser His Gln Met Val Gly Ala Arg Thr

420 425 430

Pro Ile Asp Thr Leu Asp Tyr Ile Ser Ala Gly Thr Phe Tyr Leu Glu
435 440 445

Lys Ile Asp Glu Lys Leu Ile Arg His Tyr Lys Ser Lys Pro Ile Ile
450 455 460

Ser Ser Lys Leu
465

<210> 11

<211> 388

<212> PRT

<213> *Staphylococcus epidermidis*

<400> 11

Met Asn Ile Gly Ile Asp Lys Ile Ser Phe Tyr Val Pro Lys Tyr Tyr
1 5 10 15

Val Asp Met Ala Lys Leu Ala Glu Ala Arg Gln Val Asp Pro Asn Lys
20 25 30

Phe Leu Ile Gly Ile Gly Gln Thr Glu Met Thr Val Ser Pro Val Asn
35 40 45

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

Glu Glu Asp Lys Lys Asn Ile Gly Met Val Ile Val Ala Thr Glu Ser
65 70 75 80

Ala Ile Asp Asn Ala Lys Ala Ala Ala Val Gln Ile His His Leu Leu
85 90 95

Gly Ile Gln Pro Phe Ala Arg Cys Phe Glu Met Lys Glu Ala Cys Tyr
100 105 110

Ala Ala Thr Pro Ala Ile Gln Leu Ala Lys Asp Tyr Leu Ala Gln Arg
115 120 125

Pro Asn Glu Lys Val Leu Val Ile Ala Ser Asp Thr Ala Arg Tyr Gly
130 135 140

Ile His Ser Gly Gly Glu Pro Thr Gln Gly Ala Gly Ala Val Ala Met
145 150 155 160

Met Ile Ser His Asp Pro Ser Ile Leu Lys Leu Asn Asp Asp Ala Val
165 170 175

Ala Tyr Thr Glu Asp Val Tyr Asp Phe Trp Arg Pro Thr Gly His Gln
180 185 190

```
Tyr Pro Leu Val Ala Gly Ala Leu Ser Lys Asp Ala Tyr Ile Lys Ser
        195                 200                 205

Phe Gln Glu Ser Trp Asn Glu Tyr Ala Arg Arg His Asn Lys Thr Leu
    210                 215                 220

Ala Asp Phe Ala Ser Leu Cys Phe His Val Pro Phe Thr Lys Met Gly
225                 230                 235                 240

Gln Lys Ala Leu Asp Ser Ile Ile Asn His Ala Asp Glu Thr Thr Gln
                245                 250                 255

Asp Arg Leu Asn Ser Ser Tyr Gln Asp Ala Val Asp Tyr Asn Arg Tyr
            260                 265                 270

Val Gly Asn Ile Tyr Thr Gly Ser Leu Tyr Leu Ser Leu Ile Ser Leu
        275                 280                 285

Leu Glu Thr Arg Asp Leu Lys Gly Gly Gln Thr Ile Gly Leu Phe Ser
    290                 295                 300

Tyr Gly Ser Gly Ser Val Gly Glu Phe Phe Ser Gly Thr Leu Val Asp
305                 310                 315                 320

Gly Phe Lys Glu Gln Leu Asp Val Glu Arg His Lys Ser Leu Leu Asn
                325                 330                 335

Asn Arg Ile Glu Val Ser Val Asp Glu Tyr Glu His Phe Phe Lys Arg
            340                 345                 350

Phe Asp Gln Leu Glu Leu Asn His Glu Leu Glu Lys Ser Asn Ala Asp
        355                 360                 365

Arg Asp Ile Phe Tyr Leu Lys Ser Ile Asp Asn Asn Ile Arg Glu Tyr
    370                 375                 380

His Ile Ala Glu
385
```

<210> 12

<211> 389

<212> PRT

<213> *Lactobacillus fermentum*

<400> 12

```
Met Lys Ile Gly Ile Asp Lys Leu Ala Phe Ala Thr Thr Pro Tyr Tyr
1               5                   10                  15

Leu Ala Met Glu Asp Leu Ala Gln Gly Arg Asn Val Asp Pro Asn Lys
            20                  25                  30
```

```
Tyr Leu Ile Gly Ile Gly Gln Ser Lys Gln Ala Val Val Pro Pro Thr
        35                  40                  45

Gln Asp Val Val Thr Leu Ala Ala Ala Ala Ala Asp Lys Leu Leu Asp
    50                  55                  60

Pro Val Glu Arg Asp Gln Val Ser Thr Val Ile Val Ala Thr Glu Ser
65                  70                  75                  80

Gly Ile Asp Asn Ser Lys Ala Ala Ala Val Tyr Val Lys His Leu Leu
                85                  90                  95

Lys Leu Ser Asp Phe Thr Arg Ala Val Glu Val Lys Glu Ala Cys Tyr
            100                 105                 110

Ser Ala Thr Ala Ala Leu Gln Phe Ala Arg Gly Leu Val Ala Leu Asn
        115                 120                 125

Pro Gln Glu Lys Ile Leu Val Ile Ala Ser Asp Ile Ala Arg Tyr Gly
    130                 135                 140

Leu Glu Thr Gly Gly Glu Val Thr Gln Gly Ala Gly Ala Val Ala Met
145                 150                 155                 160

Leu Ile Thr Ala Asn Pro Arg Val Leu Ala Ile Glu Pro Thr Ser Val
                165                 170                 175

Ala Tyr Thr Lys Asp Val Met Asp Phe Trp Arg Pro Leu Tyr Ala Glu
            180                 185                 190

Glu Ala Leu Val Asn Gly Lys Tyr Ser Thr Asn Val Tyr Ile Asp Phe
        195                 200                 205

Phe Lys Gln Cys Trp Thr Arg Tyr Gln Gln Leu Ala Gly Tyr Gly Leu
    210                 215                 220

Glu Asp Phe Ala Ala Leu Ala Phe His Leu Pro Phe Thr Lys Met Gly
225                 230                 235                 240

Lys Lys Ala Leu Glu Ala Glu Leu Gly Asp Arg Asp Asp Gln Val Ala
                245                 250                 255

Thr Arg Leu Arg Ala Asn Leu Thr Ala Gly Gln Glu Ala Cys Arg Gln
            260                 265                 270

Val Gly Asn Leu Tyr Thr Gly Ser Leu Tyr Leu Gly Leu Met Ser Leu
        275                 280                 285

Leu Thr Glu Gly Asp Val Lys Pro Gly Glu Arg Ile Gly Leu Phe Ser
    290                 295                 300
```

Tyr Gly Ser Gly Ala Glu Gly Glu Phe Phe Ala Gly Ile Leu Gln Pro
305 310 315 320

Gly Tyr Gln Glu Gly Leu Gly Asp Leu Asn Glu Gln Leu Ala Ala Arg
325 330 335

Thr Gln Val Ser Leu Ala Glu Tyr Glu Asp Leu Phe Asn Gln Gln Leu
340 345 350

Gly Leu Lys Glu Glu Asp Val Thr Phe Lys Thr Pro Ala Ala Gly Gln
355 360 365

Arg Phe Val Leu Val Gly Gln Lys Asp His Gln Arg Gln Tyr Arg Asp
370 375 380

Leu Ala Glu Arg Asp
385

<210> 13

<211> 351

<212> PRT

<213> *Hyperthermus butylicus*

<400> 13

Met Pro Arg Gly Ser Gly Ile Val Gly Trp Gly Gly Tyr Val Pro Arg
1 5 10 15

Tyr Arg Ile Lys Ala Ala Glu Ile Val Arg Val Trp Gly Trp Glu Pro
20 25 30

Ser Val Pro Ala Gly Leu Gly Val Lys Glu Lys Ala Val Glu Asn Val
35 40 45

Asp Glu Asp Ser Val Thr Met Gly Tyr Glu Ala Ala Arg Asn Ala Ile
50 55 60

Ala Arg Ala Asn Val Asp Pro Arg Glu Ile Lys Ala Val Phe Phe Gly
65 70 75 80

Thr Glu Ser Lys Pro Tyr Ala Val Lys Pro Ser Ala Thr Ile Ile Ala
85 90 95

Glu Ala Leu Gly Ile Thr Pro Glu Thr Met Ala Ser Asp Leu Glu Phe
100 105 110

Ala Cys Arg Ala Ala Ser Glu Gly Leu Arg Ala Ser Leu Ala Leu Val
115 120 125

Glu Ala Gly Tyr Met Lys Tyr Ala Leu Val Val Ala Ser Asp Thr Ala
130 135 140

```
Gln Ala Asn Pro Gly Asp Val Leu Glu Phe Thr Ala Ala Ser Gly Ala
145                 150                 155                 160

Ala Ala Phe Val Val Gly Pro Ala Ser Glu Ser Val Ala Val Leu Glu
                165                 170                 175

Gly Val Tyr Thr Tyr Val Thr Asp Thr Pro Asp Phe Trp Arg Gly Gln
            180                 185                 190

His Ser Arg Tyr Pro Met His Gly Glu Ala Phe Thr Gly Glu Pro Ala
        195                 200                 205

Tyr Phe His His Ile Glu Ser Ala Val Lys Gly Leu Met Glu Lys Leu
    210                 215                 220

Gly Leu Lys Pro Glu Asp Phe Asp Tyr Ala Val Phe His Gln Pro Asn
225                 230                 235                 240

Gly Lys Phe Pro Leu Arg Val Gly Ala Arg Leu Gly Phe Pro Lys Glu
                245                 250                 255

Lys Ile Leu Pro Gly Leu Leu Thr Pro Ile Ile Gly Asn Thr Tyr Asn
            260                 265                 270

Ala Ser Ala Leu Leu Gly Phe Ala Arg Ile Leu Asp Gln Ala Lys Pro
        275                 280                 285

Gly Gln Arg Ile Leu Val Ala Pro Phe Gly Ser Gly Ala Gly Ser Asp
    290                 295                 300

Ala Tyr Ser Phe Ile Val Thr Asp Arg Ile Glu Glu Ala Arg Asn Arg
305                 310                 315                 320

Ala Pro Lys Val Asp Asp Tyr Val Asn Trp Lys Arg Tyr Ile Asp Tyr
                325                 330                 335

Ala Met His Ala Arg Met Arg Lys Leu Tyr Asp Arg Arg Pro Val
            340                 345                 350
```

<210> 14

<211> 349

<212> PRT

<213> *Chloroflexus aggregans*

<400> 14

```
Met Met Lys Pro Asn Gln Pro Val Gly Ile Ile Gly Tyr Gly Val Tyr
1               5                   10                  15

Ile Pro Arg Tyr Arg Ile Ala Ala Arg Glu Ile Ala Arg Ile Trp Thr
            20                  25                  30

Asp Gly Gln Asn Gly Val Pro Val Glu Ala Lys Ser Val Pro Gly Pro
```

```
        35                  40                  45

Asp Glu Asp Thr Ile Thr Met Ala Ile Glu Ala Ala Arg Asn Ala Leu
    50                  55                  60

Val Arg Ala Asp Ile Pro Ala Ser Ala Leu Gly Ala Val Trp Ile Gly
65                  70                  75                  80

Ser Glu Ser His Pro Tyr Ser Val Lys Pro Ser Gly Thr Val Val Ala
                85                  90                  95

Asp Ala Leu Gly Ala Gly Pro Trp Val Ser Ala Ala Asp Trp Glu Phe
            100                 105                 110

Ala Cys Lys Ala Gly Ser Glu Ala Leu Thr Ala Ala Met Ala Leu Val
        115                 120                 125

Gly Ser Gly Met Gln Arg Tyr Ala Leu Ala Ile Gly Ala Asp Thr Ala
    130                 135                 140

Gln Gly Arg Pro Gly Asp Ala Leu Glu Tyr Thr Ala Ser Ala Gly Ala
145                 150                 155                 160

Ala Ala Leu Ile Val Gly Pro Ala Thr Glu Ala Leu Ala Thr Ile Asp
                165                 170                 175

Ala Thr Val Ser Tyr Val Thr Asp Thr Pro Asp Phe Tyr Arg Arg Ala
            180                 185                 190

Asp Arg Pro Tyr Pro Val His Gly Asn Arg Phe Thr Gly Glu Pro Ala
        195                 200                 205

Tyr Phe His Gln Ile Gln Ser Ala Ala Ser Glu Leu Leu Arg Gln Leu
    210                 215                 220

Asn Arg Thr Ala Ala Asp Phe Thr Tyr Ala Val Phe His Gln Pro Asn
225                 230                 235                 240

Ala Lys Phe Pro Gln Thr Val Ala Lys Arg Leu Gly Phe Thr Asp Ala
                245                 250                 255

Gln Ile Ala Pro Gly Leu Leu Ser Pro Gln Ile Gly Asn Thr Tyr Ser
            260                 265                 270

Gly Ala Ala Leu Leu Gly Leu Cys Ala Ile Leu Asp Val Ala Lys Pro
        275                 280                 285

Gly Asp Thr Ile Phe Val Thr Ser Tyr Gly Ser Gly Ala Gly Ser Asp
    290                 295                 300

Ala Tyr Ala Leu Thr Val Thr Glu Ala Ile Val Glu Arg Arg Glu Arg
305                 310                 315                 320

Ala Pro Leu Thr Ala Ala Tyr Leu Ala Arg Lys Val Met Ile Asp Tyr
                325                 330                 335

Ala Met Tyr Ala Lys Trp Arg Gly Lys Leu Val Met Gly
            340                 345
```

<210> 15

<211> 420

<212> PRT

<213> *Bacillus subtilis*

<400> 15

```
Met Val Ser Ala Gly Ile Glu Ala Met Asn Val Phe Gly Gly Thr Ala
1               5                  10                  15

Tyr Leu Asp Val Met Glu Leu Ala Lys Tyr Arg His Leu Asp Thr Ala
            20                  25                  30

Arg Phe Glu Asn Leu Leu Met Lys Glu Lys Ala Val Ala Leu Pro Tyr
        35                  40                  45

Glu Asp Pro Val Thr Phe Gly Val Asn Ala Ala Lys Pro Ile Ile Asp
    50                  55                  60

Ala Leu Ser Glu Ala Glu Lys Asp Arg Ile Glu Leu Leu Ile Thr Cys
65                  70                  75                  80

Ser Glu Ser Gly Ile Asp Phe Gly Lys Ser Leu Ser Thr Tyr Ile His
                85                  90                  95

Glu Tyr Leu Gly Leu Asn Arg Asn Cys Arg Leu Phe Glu Val Lys Gln
            100                 105                 110

Ala Cys Tyr Ser Gly Thr Ala Gly Phe Gln Met Ala Val Asn Phe Ile
        115                 120                 125

Leu Ser Gln Thr Ser Pro Gly Ala Lys Ala Leu Val Ile Ala Ser Asp
    130                 135                 140

Ile Ser Arg Phe Leu Ile Ala Glu Gly Gly Asp Ala Leu Ser Glu Asp
145                 150                 155                 160

Trp Ser Tyr Ala Glu Pro Ser Ala Gly Ala Gly Ala Val Ala Val Leu
                165                 170                 175

Val Gly Glu Asn Pro Glu Val Phe Gln Ile Asp Pro Gly Ala Asn Gly
            180                 185                 190

Tyr Tyr Gly Tyr Glu Val Met Asp Thr Cys Arg Pro Ile Pro Asp Ser
        195                 200                 205
```

```
Glu Ala Gly Asp Ser Asp Leu Ser Leu Met Ser Tyr Leu Asp Cys Cys
    210                 215                 220

Glu Gln Thr Phe Leu Glu Tyr Gln Lys Arg Val Pro Gly Ala Asn Tyr
225                 230                 235                 240

Gln Asp Thr Phe Gln Tyr Leu Ala Tyr His Thr Pro Phe Gly Gly Met
                245                 250                 255

Val Lys Gly Ala His Arg Thr Met Met Arg Lys Val Ala Lys Val Lys
            260                 265                 270

Thr Ser Gly Ile Glu Thr Asp Phe Leu Thr Arg Val Lys Pro Gly Leu
        275                 280                 285

Asn Tyr Cys Gln Arg Val Gly Asn Ile Met Gly Ala Ala Leu Phe Leu
    290                 295                 300

Ala Leu Ala Ser Thr Ile Asp Gln Gly Arg Phe Asp Thr Pro Lys Arg
305                 310                 315                 320

Ile Gly Cys Phe Ser Tyr Gly Ser Gly Cys Cys Ser Glu Phe Tyr Ser
                325                 330                 335

Gly Ile Thr Thr Pro Gln Gly Gln Glu Arg Gln Arg Thr Phe Gly Ile
            340                 345                 350

Glu Lys His Leu Asp Arg Arg Tyr Gln Leu Ser Met Glu Glu Tyr Glu
        355                 360                 365

Leu Leu Phe Lys Gly Ser Gly Met Val Arg Phe Gly Thr Arg Asn Val
    370                 375                 380

Lys Leu Asp Phe Glu Met Ile Pro Gly Ile Met Gln Ser Thr Gln Glu
385                 390                 395                 400

Lys Pro Arg Leu Phe Leu Glu Glu Ile Ser Glu Phe His Arg Lys Tyr
                405                 410                 415

Arg Trp Ile Ser
            420
```

<210> 16

<211> 388

<212> PRT

<213> *Mycobacterium marinum*

<400> 16

```
Met Val Ser Ile Gly Ile His Asp Leu Ser Ile Ala Thr Ala His Tyr
1               5                   10                  15
```

```
Val Leu Asp His Ala Thr Leu Ala Glu His His Gly Val Asp Val Asn
            20                  25                  30

Lys Tyr Leu Ile Gly Leu Gly Gln Gln Gln Met Ser Ile Val Ala Pro
        35                  40                  45

Asp Glu Asp Ile Val Thr Leu Ala Ala Ala Ala Ala Asp Pro Ile Ile
    50                  55                  60

Lys Arg His Gly Ser Gln Lys Ile Arg Thr Ile Val Ile Gly Thr Glu
65                  70                  75                  80

Thr Gly Val Asp Gln Ser Lys Ser Ala Gly Ile Trp Val Ser Ser Leu
                85                  90                  95

Leu Gly Leu Pro Ser Ser Ala Arg Val Leu Glu Val Lys Gln Ala Cys
            100                 105                 110

Tyr Gly Ala Thr Gly Ala Leu Gln Leu Ala Leu Ala Leu Val His Arg
        115                 120                 125

Asp Pro Thr Gln Gln Val Leu Val Ile Ala Ala Asp Val Ala Arg Tyr
    130                 135                 140

Asp Leu Asp Ser Pro Gly Glu Pro Thr Gln Gly Ala Ala Ala Ala Ala
145                 150                 155                 160

Met Leu Val Ser Ala Asp Pro Ala Leu Leu Arg Leu Glu Glu Pro Thr
                165                 170                 175

Gly Ile Tyr Thr Ala Asp Ile Met Asp Phe Trp Arg Pro Asn Tyr Arg
            180                 185                 190

Ser Thr Ala Leu Val Asp Gly Lys Ala Ser Val Thr Ala Tyr Met Glu
        195                 200                 205

Ala Ala Ser Gly Ala Trp Lys Asp Tyr Thr Glu Arg Gly Gly Arg Ala
    210                 215                 220

Phe Gly Glu Phe Ala Ala Phe Cys Tyr His Gln Pro Phe Thr Lys Met
225                 230                 235                 240

Ala Tyr Lys Ala His Lys Gln Leu Ala Ala Glu Ala Gly Glu Asp Ala
                245                 250                 255

Ser Gly Ala Ala Val Gln Ala Ala Val Gly Asn Thr Val Glu Tyr Asn
            260                 265                 270

Arg Arg Ile Gly Asn Ser Tyr Thr Ala Ser Leu Tyr Leu Ala Leu Ala
        275                 280                 285
```

Ala Leu Leu Asp Gln Ala Asp Asp Leu Ser Asp Gln Pro Ile Ala Met
290 295 300

Leu Ser Tyr Gly Ser Gly Cys Val Ala Glu Leu Phe Ala Gly Thr Val
305 310 315 320

Thr Pro Gly Tyr Gln Gln His Leu Arg Thr Asp Gln His Arg Ala Ala
325 330 335

Leu Glu Thr Arg Ile Pro Leu Ser Tyr Glu His Tyr Arg Arg Leu His
340 345 350

Asn Leu Thr Leu Pro Thr Asn Gly Asn His His Ser Leu Pro Val Glu
355 360 365

Thr Ser Arg Pro Phe Arg Leu Thr Ala Ile Ser Glu His Lys Arg Met
370 375 380

Tyr Gly Ala Val
385

<210> 17

<211> 435

<212> PRT

<213> *Zea mays*

<400> 17

Met Leu Ala Ala Ser Thr Lys Val Gly Ser Arg Leu Ala Ser Pro His
1 5 10 15

Ala Ser Leu Ser Ala Gly Ala Ala Ala Ala Ala Leu Ala Ser Ser Pro
20 25 30

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

Asn His His Ala Ala Asp Ala Pro Pro Pro Leu Pro Cys Ser Ser Ser
50 55 60

Gly Asp Ser Arg Glu Tyr Tyr Gln Trp Lys Arg Leu Val Asn Gln Arg
65 70 75 80

Gln Ser Thr Leu His Val Gly Glu Val Pro Ala Ala Leu Gly His His
85 90 95

Val Phe Gly Ala Gly Cys Ser Ser Arg Lys Gln His Ile Tyr Arg Tyr
100 105 110

Phe Ser Ser Ser Ser His Gln Gly Ser Ile Trp Ala Arg Ser Lys Ile
115 120 125

```
Leu His Asp Leu Pro Gly Tyr Val Lys Ile Val Glu Val Gly Pro Arg
    130                 135                 140

Asp Gly Leu Gln Asn Glu Lys Asp Ile Val Pro Thr Pro Val Lys Val
145                 150                 155                 160

Glu Leu Ile Arg Arg Leu Ala Thr Ser Gly Leu Pro Val Val Glu Ala
                165                 170                 175

Thr Ser Phe Val Ser Pro Lys Trp Val Pro Gln Leu Ala Asp Ala Lys
            180                 185                 190

Asp Val Met Glu Ala Val Arg Thr Ile Gly Gly Val Arg Phe Pro Val
        195                 200                 205

Leu Thr Pro Asn Leu Lys Gly Phe Glu Ala Ala Ile Ala Ala Gly Ala
    210                 215                 220

Lys Glu Ile Ala Ile Phe Ala Ser Ala Ser Glu Gly Phe Ser Lys Ser
225                 230                 235                 240

Asn Ile Asn Cys Thr Ile Lys Glu Ser Ile Ala Arg Tyr Asn Asp Val
                245                 250                 255

Ala Leu Ala Ala Lys Glu Lys Glu Ile Pro Val Arg Gly Tyr Val Ser
            260                 265                 270

Cys Val Val Gly Cys Pro Val Asp Gly Pro Val Pro Pro Ser Asn Val
        275                 280                 285

Ala Tyr Val Ala Lys Glu Leu Tyr Asp Met Gly Cys Tyr Glu Val Ser
    290                 295                 300

Leu Gly Asp Thr Ile Gly Val Gly Thr Pro Gly Thr Val Val Pro Met
305                 310                 315                 320

Leu Glu Ala Ala Ile Ser Val Val Pro Val Glu Lys Leu Ala Val His
                325                 330                 335

Phe His Asp Thr Tyr Gly Gln Ser Leu Ser Asn Ile Leu Ile Ser Leu
            340                 345                 350

Gln Met Gly Val Ser Val Val Asp Ser Ser Val Ala Gly Leu Gly Gly
        355                 360                 365

Cys Pro Tyr Ala Lys Gly Ala Ser Gly Asn Val Ala Thr Glu Asp Val
    370                 375                 380

Val Tyr Met Leu Asn Gly Leu Gly Val Lys Thr Gly Val Asp Leu Gly
385                 390                 395                 400

Lys Val Met Ala Ala Gly Glu Phe Ile Cys Arg His Leu Gly Arg Gln
                405                 410                 415

Ser Gly Ser Lys Ala Ala Thr Ala Leu Ser Lys Val Thr Ala Asn Ala
            420                 425                 430

Ser Lys Leu
        435
```

<210> 18

<211> 335

<212> PRT

<213> *Danio rerio* (Brachydanio rerio)

<400> 18

```
Met Gly Asn Val Ser Ser Ala Val Lys His Cys Leu Ser Tyr Glu Thr
1               5                  10                  15

Phe Leu Arg Asp Tyr Pro Trp Leu Pro Arg Leu Leu Trp Glu Glu Lys
            20                  25                  30

Cys Ser Glu Leu Pro Lys Leu Pro Val Tyr Val Lys Ile Val Glu Val
        35                  40                  45

Gly Pro Arg Asp Gly Leu Gln Asn Glu Lys Glu Ile Val Pro Thr Glu
    50                  55                  60

Val Lys Ile Gln Leu Ile Asp Leu Leu Ser Gln Thr Gly Leu Pro Val
65                  70                  75                  80

Ile Glu Ala Thr Ser Phe Val Ser Ser Lys Trp Val Ala Gln Met Ala
                85                  90                  95

Asp His Thr Ala Val Leu Lys Gly Ile Lys Arg Ser Pro Asp Val Arg
            100                 105                 110

Tyr Pro Val Leu Thr Pro Asn Ile Gln Gly Phe Gln Ala Ala Val Ala
        115                 120                 125

Ala Gly Ala Asn Glu Val Ala Val Phe Gly Ser Ala Ser Glu Thr Phe
    130                 135                 140

Ser Arg Lys Asn Ile Asn Cys Ser Ile Glu Glu Ser Leu Gln Arg Phe
145                 150                 155                 160

Glu Gln Val Val Ser Ala Ala Lys Gln Glu Gly Ile Pro Val Arg Gly
                165                 170                 175

Tyr Val Ser Cys Ala Leu Gly Cys Pro Tyr Glu Gly Gln Val Lys Pro
            180                 185                 190

Ser Gln Val Thr Lys Val Ala Lys Arg Leu Phe Glu Leu Gly Cys Tyr
```

195 200 205

Glu Val Ser Leu Gly Asp Thr Ile Gly Val Gly Thr Ala Gly Ser Met
210 215 220

Ala Glu Met Leu Ser Asp Val Leu Thr Glu Val Pro Ala Gly Ala Leu
225 230 235 240

Ala Val His Cys His Asp Thr Tyr Gly Gln Ala Leu Pro Asn Ile Leu
245 250 255

Ile Ala Leu Gln Met Gly Val Ser Val Val Asp Ala Ser Val Ala Gly
260 265 270

Leu Gly Gly Cys Pro Phe Ala Lys Gly Ala Ser Gly Asn Val Ser Thr
275 280 285

Glu Asp Leu Leu Tyr Met Leu His Gly Leu Gly Ile Glu Thr Gly Val
290 295 300

Asp Leu Leu Lys Val Met Glu Ala Gly Asp Phe Ile Cys Lys Ala Leu
305 310 315 320

Asn Arg Lys Thr Asn Ser Lys Val Ser Gln Ala Thr Arg Asn Asn
325 330 335

<210> 19

<211> 325

<212> PRT

<213> *Bos taurus*

<400> 19

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

Thr Leu Arg Ala Val Ser Thr Ser Ser Val Gly Thr Phe Pro Lys Gln
20 25 30

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

Asn Ile Val Pro Thr Pro Val Lys Ile Lys Leu Ile Asp Met Leu Ser
50 55 60

Glu Ala Gly Leu Pro Val Val Glu Ala Thr Ser Phe Val Ser Pro Lys
65 70 75 80

Trp Val Pro Gln Met Ala Asp His Ala Glu Val Leu Lys Gly Ile Gln
85 90 95

Lys Phe Pro Gly Val Asn Tyr Pro Val Leu Thr Pro Asn Phe Lys Gly
100 105 110

```
Phe Gln Ala Ala Val Ala Ala Gly Ala Lys Glu Val Ala Ile Phe Gly
        115                 120                 125

Ala Ala Ser Glu Leu Phe Thr Lys Lys Asn Ile Asn Cys Ser Ile Asp
    130                 135                 140

Glu Ser Leu Gln Arg Phe Asp Glu Ile Leu Lys Ala Ala Arg Ala Ala
145                 150                 155                 160

Gly Ile Ser Val Arg Gly Tyr Val Ser Cys Val Leu Gly Cys Pro Tyr
                165                 170                 175

Glu Gly Lys Ile Ser Pro Ala Lys Val Ala Glu Val Thr Lys Lys Leu
            180                 185                 190

Tyr Ser Met Gly Cys Tyr Glu Ile Ser Leu Gly Asp Thr Ile Gly Val
        195                 200                 205

Gly Thr Pro Gly Ala Met Lys Asp Met Leu Ser Ala Val Leu Gln Glu
    210                 215                 220

Val Pro Val Thr Ala Leu Ala Val His Cys His Asp Thr Tyr Gly Gln
225                 230                 235                 240

Ala Leu Ala Asn Thr Leu Thr Ala Leu Gln Met Gly Val Ser Val Met
                245                 250                 255

Asp Ser Ser Val Ala Gly Leu Gly Gly Cys Pro Tyr Ala Gln Gly Ala
            260                 265                 270

Ser Gly Asn Leu Ala Thr Glu Asp Leu Val Tyr Met Leu Ala Gly Leu
        275                 280                 285

Gly Ile His Thr Gly Val Asn Leu Gln Lys Leu Leu Glu Ala Gly Ala
    290                 295                 300

Phe Ile Cys Gln Ala Leu Asn Arg Arg Thr Asn Ser Lys Val Ala Gln
305                 310                 315                 320

Ala Thr Cys Lys Leu
                325
```

<210> 20

<211> 325

<212> PRT

<213> *Homo sapiens*

<400> 20

```
Met Ala Ala Met Arg Lys Ala Leu Pro Arg Arg Leu Val Gly Leu Ala
1               5                   10                  15
```

```
Ser Leu Arg Ala Val Ser Thr Ser Ser Met Gly Thr Leu Pro Lys Arg
            20                  25                  30

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

Asn Ile Val Ser Thr Pro Val Lys Ile Lys Leu Ile Asp Met Leu Ser
    50                  55                  60

Glu Ala Gly Leu Ser Val Ile Glu Thr Thr Ser Phe Val Ser Pro Lys
65                  70                  75                  80

Trp Val Pro Gln Met Gly Asp His Thr Glu Val Leu Lys Gly Ile Gln
                85                  90                  95

Lys Phe Pro Gly Ile Asn Tyr Pro Val Leu Thr Pro Asn Leu Lys Gly
            100                 105                 110

Phe Glu Ala Ala Val Ala Ala Gly Ala Lys Glu Val Val Ile Phe Gly
        115                 120                 125

Ala Ala Ser Glu Leu Phe Thr Lys Lys Asn Ile Asn Cys Ser Ile Glu
    130                 135                 140

Glu Ser Phe Gln Arg Phe Asp Ala Ile Leu Lys Ala Ala Gln Ser Ala
145                 150                 155                 160

Asn Ile Ser Val Arg Gly Tyr Val Ser Cys Ala Leu Gly Cys Pro Tyr
                165                 170                 175

Glu Gly Lys Ile Ser Pro Ala Lys Val Ala Glu Val Thr Lys Lys Phe
            180                 185                 190

Tyr Ser Met Gly Cys Tyr Glu Ile Ser Leu Gly Asp Thr Ile Gly Val
        195                 200                 205

Gly Thr Pro Gly Ile Met Lys Asp Met Leu Ser Ala Val Met Gln Glu
    210                 215                 220

Val Pro Leu Ala Ala Leu Ala Val His Cys His Asp Thr Tyr Gly Gln
225                 230                 235                 240

Ala Leu Thr Asn Thr Leu Met Ala Leu Gln Met Gly Val Ser Val Val
                245                 250                 255

Asp Ser Ser Val Ala Gly Leu Gly Gly Cys Pro Tyr Ala Gln Gly Ala
            260                 265                 270

Ser Gly Asn Leu Ala Thr Glu Asp Leu Val Tyr Met Leu Glu Gly Leu
        275                 280                 285

Gly Ile His Thr Gly Val Asn Leu Gln Lys Leu Leu Glu Ala Gly Asn
    290                 295                 300

Phe Ile Cys Gln Ala Leu Asn Arg Lys Thr Ser Ser Lys Val Ala Gln
305                 310                 315                 320

Ala Thr Cys Lys Leu
                325
```

<210> 21

<211> 299

<212> PRT

<213> *Pseudomonas putida* Q88H25

<400> 21

```
Met Ser Leu Pro Lys His Val Arg Leu Val Glu Val Gly Pro Arg Asp
1               5                  10                  15

Gly Leu Gln Asn Glu Ala Gln Pro Ile Ser Val Ala Asp Lys Val Arg
            20                  25                  30

Leu Val Asn Asp Leu Thr Glu Ala Gly Leu Ala Tyr Ile Glu Val Gly
        35                  40                  45

Ser Phe Val Ser Pro Lys Trp Val Pro Gln Met Ala Gly Ser Ala Glu
    50                  55                  60

Val Phe Ala Gly Ile Gln Gln Arg Pro Gly Val Thr Tyr Ala Ala Leu
65                  70                  75                  80

Ala Pro Asn Leu Arg Gly Phe Glu Asp Ala Leu Ala Ala Gly Val Lys
                85                  90                  95

Glu Val Ala Val Phe Ala Ala Ala Ser Glu Ala Phe Ser Gln Arg Asn
            100                 105                 110

Ile Asn Cys Ser Ile Ser Glu Ser Leu Lys Arg Phe Glu Pro Ile Met
        115                 120                 125

Asp Ala Ala Arg Ser His Gly Met Arg Val Arg Gly Tyr Val Ser Cys
    130                 135                 140

Val Leu Gly Cys Pro Tyr Glu Gly Lys Val Ser Ala Glu Gln Val Ala
145                 150                 155                 160

Pro Val Ala Arg Ala Leu His Asp Met Gly Cys Tyr Glu Val Ser Leu
                165                 170                 175

Gly Asp Thr Ile Gly Thr Gly Thr Ala Gly Asp Thr Arg Arg Leu Phe
            180                 185                 190
```

Glu Val Val Ser Ala Gln Val Pro Arg Glu Gln Leu Ala Gly His Phe
195 200 205

His Asp Thr Tyr Gly Gln Ala Leu Ala Asn Val Tyr Ala Ser Leu Leu
210 215 220

Glu Gly Ile Ser Val Phe Asp Ser Ser Val Ala Gly Leu Gly Gly Cys
225 230 235 240

Pro Tyr Ala Lys Gly Ala Thr Gly Asn Ile Ala Ser Glu Asp Val Val
245 250 255

Tyr Leu Leu Gln Gly Leu Gly Ile Glu Thr Gly Ile Asp Leu Gly Leu
260 265 270

Leu Ile Ala Ala Gly Gln Arg Ile Ser Gly Val Leu Gly Arg Asp Asn
275 280 285

Gly Ser Arg Val Ala Arg Ala Cys Ser Ala Gln
290 295

<210> 22

<211> 312

<212> PRT

<213> *Acinetobacter baumannii* B7H4C6

<400> 22

Met Thr Ala Phe Ser Asp Leu Leu Val Val Gln Glu Val Ser Pro Arg
1 5 10 15

Asp Gly Leu Gln Ile Glu Pro Thr Trp Val Pro Thr Asp Lys Lys Ile
20 25 30

Asp Leu Ile Asn Gln Leu Ser Thr Met Gly Phe Ser Arg Ile Glu Ala
35 40 45

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

Glu Val Phe Thr Gly Ile Thr Arg His Lys Asp Ile Ile Tyr Val Gly
65 70 75 80

Leu Ile Pro Asn Leu Lys Gly Ala Leu Arg Ala Val Glu Ala Asn Ala
85 90 95

Asn Glu Leu Asn Leu Val Leu Ser Ala Ser Gln Thr His Asn Leu Ala
100 105 110

Asn Met Arg Met Thr Lys Ala Gln Ser Phe Ala Gly Phe Thr Glu Ile
115 120 125

Val Glu Gln Leu Gln Gly Lys Thr Gln Phe Asn Gly Thr Val Ala Thr

130 135 140

Thr Phe Gly Cys Pro Phe Glu Gly Lys Ile Ser Glu Arg Glu Val Phe
145 150 155 160

Ser Leu Val Glu His Tyr Leu Lys Leu Gly Ile His Asn Ile Thr Leu
165 170 175

Ala Asp Thr Thr Gly Met Ala Asn Pro Val Gln Val Lys Arg Ile Val
180 185 190

Ser His Val Leu Ser Leu Ile Ser Pro Glu Gln Leu Thr Leu His Phe
195 200 205

His Asn Thr Arg Gly Leu Gly Leu Thr Asn Val Leu Ala Ala Tyr Glu
210 215 220

Val Gly Ala Arg Arg Phe Asp Ala Ala Leu Gly Gly Leu Gly Gly Cys
225 230 235 240

Pro Phe Ala Pro Gly Ala Ser Gly Asn Ile Cys Thr Glu Asp Leu Val
245 250 255

Asn Met Cys Glu Glu Ile Gly Ile Pro Thr Thr Ile Asp Leu Asp Ala
260 265 270

Leu Ile Gln Leu Ser Arg Thr Leu Pro Ala Leu Leu Gly His Asp Thr
275 280 285

Pro Ser Gln Leu Ala Lys Ala Gly Arg Asn Thr Asp Leu His Pro Ile
290 295 300

Pro Asp Tyr Ile Lys Ser Leu Asn
305 310

<210> 23

<211> 286

<212> PRT

<213> *Thermus thermophilus* Q72IH0

<400> 23

Met Lys Ala Ser Val Arg Trp Val Glu Cys Pro Arg Asp Ala Trp Gln
1 5 10 15

Gly Phe Ser Arg Phe Ile Pro Thr Glu Glu Lys Val Ala Phe Leu Asn
20 25 30

Glu Leu Leu Glu Ala Gly Phe Ala His Leu Asp Leu Thr Ser Phe Val
35 40 45

Ser Pro Lys Trp Val Pro Gln Met Gln Asp Ala Glu Glu Val Leu Lys
50 55 60

```
Ala Leu Pro Pro Pro Asn Gly Arg Thr Tyr Leu Ala Ile Val Ala Asn
65                  70                  75                  80

Glu Lys Gly Leu Glu Arg Ala Leu Ala Ala Pro Asn Leu Thr His Val
                85                  90                  95

Gly Tyr Pro Phe Ser Leu Ser Glu Thr Phe Gln Gln Arg Asn Thr Asn
            100                 105                 110

Arg Ser Ile Glu Ala Ser Trp Pro Leu Val Gly Ala Met Val Glu Arg
        115                 120                 125

Thr Glu Gly Arg Leu Gly Leu Val Val Tyr Leu Ser Met Ala Phe Gly
    130                 135                 140

Asn Pro Tyr Gly Asp Pro Trp Ser Val Glu Ala Val Leu Glu Ala Leu
145                 150                 155                 160

Ala Arg Leu Lys Glu Met Gly Val Arg Glu Ile Ala Leu Ala Asp Thr
                165                 170                 175

Tyr Gly Val Ala Glu Pro Glu Arg Ile His Glu Val Leu Lys Ala Ala
            180                 185                 190

Val Ala Arg Phe Gly Pro Glu Gly Leu Gly Ala His Leu His Ala Arg
        195                 200                 205

Pro Glu Gly Ala Leu Ala Lys Val Glu Ala Val Leu Ala Ala Gly Val
    210                 215                 220

Thr Trp Leu Glu Gly Ala Leu Ala Gly Val Gly Gly Cys Pro Phe Ala
225                 230                 235                 240

Gly Asp Glu Leu Val Gly Asn Leu Pro Thr Glu Val Val Leu Pro His
                245                 250                 255

Leu Glu Lys Arg Gly Leu Ala Thr Gly Val Asp Leu Ser Arg Leu Pro
            260                 265                 270

Leu Leu Ala Glu Glu Ala Ala Arg Leu Lys Ala Leu Tyr Ala
        275                 280                 285
```

<210> 24

<211> 387

<212> PRT

<213> *Lactobacillus delbrueckii*

<400> 24

```
Met Asp Ile Gly Ile Asp Gln Ile Gly Phe Tyr Thr Pro Asn Lys Phe
1               5                   10                  15
```

```
Val Asp Met Val Asp Leu Ala Asn Ala Arg Asn Gln Asp Pro Asn Lys
          20                  25                  30

Phe Leu Ile Gly Ile Gly Gln Asp Arg Met Ala Val Ala Asp Lys Thr
       35                  40                  45

Gln Asp Ala Val Ser Met Gly Ile Asn Ala Thr Ala Glu Tyr Leu Asp
    50                  55                  60

Gln Val Asp Leu Glu Gln Leu Gly Leu Leu Ile Phe Ala Thr Glu Ser
65                  70                  75                  80

Gly Ile Asp Gln Ser Lys Ser Ala Ser Leu Phe Val Lys Glu Ala Leu
                85                  90                  95

Asn Leu Pro Ala Arg Ile Arg Thr Phe Glu Ile Lys Glu Ala Cys Phe
           100                 105                 110

Ala Leu Thr Ala Ser Leu Gln Val Ala Arg Asp Tyr Val Arg Ala His
       115                 120                 125

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

Leu Ala Thr Ala Gly Glu Val Thr Gln Gly Ala Gly Ala Ile Ser Met
145                 150                 155                 160

Leu Ile Lys Glu Asn Pro Ala Ile Ile Ala Leu Glu Asp Gly His Thr
                165                 170                 175

Ser His Ser Glu Asn Ile Asn Asp Phe Trp Arg Pro Asn Asn Leu Ala
           180                 185                 190

Thr Ala Val Val Asp Gly His Tyr Ser Arg Asp Val Tyr Leu Asp Phe
       195                 200                 205

Phe Lys Ser Thr Phe Lys Pro Phe Leu Ala Glu Lys Gln Leu Gln Val
    210                 215                 220

Ser Asp Phe Ala Gly Ile Cys Tyr His Leu Pro Tyr Thr Lys Met Gly
225                 230                 235                 240

Tyr Lys Ala His Lys Ile Ala Ile Glu Gly Gln Asp Asp Glu Thr Val
                245                 250                 255

Lys Arg Leu Ser Asp Asn Phe Gln Leu Ser Ala Lys Tyr Ser Arg Gln
           260                 265                 270

Val Gly Asn Ile Tyr Thr Ala Ser Leu Tyr Met Ser Val Leu Ser Leu
       275                 280                 285
```

```
Leu Glu Asn Gly Asp Leu Glu Ala Gly Asp Arg Ile Gly Phe Phe Ser
    290                 295                 300

Tyr Gly Ser Gly Ala Met Ala Glu Phe Phe Ser Gly Lys Val Val Ala
305                 310                 315                 320

Gly Tyr Gln Lys Arg Leu Arg Pro Ala Leu His Ala Arg Met Leu Lys
                325                 330                 335

Glu Arg Ile Arg Leu Gly Val Gly Gln Tyr Glu Asp Ile Phe Thr Glu
            340                 345                 350

Gly Leu Glu Ala Leu Pro Glu Asn Val Glu Phe Thr Ser Asp Ala Asn
        355                 360                 365

His Gly Thr Trp Tyr Leu Ala Gly Gln Glu Gly Tyr Val Arg Gln Tyr
    370                 375                 380

Lys Gln Lys
385
```

<210> 25

<211> 388

<212> PRT

<213> *Staphylococcus haemolyticus*

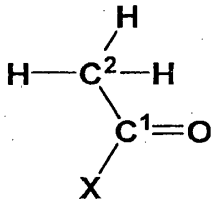
<400> 25

```
Met Ser Ile Gly Ile Asp Lys Ile Asn Phe Tyr Val Pro Lys Tyr Tyr
1               5                   10                  15

Val Asp Met Ala Lys Leu Ala Glu Ala Arg Gln Val Asp Pro Asn Lys
            20                  25                  30

Phe Leu Ile Gly Ile Gly Gln Thr Gln Met Ala Val Ser Pro Val Ser
        35                  40                  45

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

Asp Asp Asp Lys Lys His Ile Gly Met Val Ile Val Ala Thr Glu Ser
65                  70                  75                  80

Ala Ile Asp Asn Ala Lys Ala Ala Ala Val Gln Ile His Asn Leu Leu
                85                  90                  95

Gly Val Gln Pro Phe Ala Arg Cys Phe Glu Met Lys Glu Ala Cys Tyr
            100                 105                 110

Ala Ala Thr Pro Ala Ile Gln Leu Ala Lys Asp Tyr Ile Glu Lys Arg
        115                 120                 125
```

```
Pro Asn Glu Lys Val Leu Val Ile Ala Ser Asp Thr Ala Arg Tyr Gly
    130                 135                 140

Ile Gln Ser Gly Gly Glu Pro Thr Gln Gly Ala Gly Ala Val Ala Met
145                 150                 155                 160

Leu Ile Ser Asn Asn Pro Ser Ile Leu Glu Leu Asn Asp Asp Ala Val
                165                 170                 175

Ala Tyr Thr Glu Asp Val Tyr Asp Phe Trp Arg Pro Thr Gly His Lys
            180                 185                 190

Tyr Pro Leu Val Ala Gly Ala Leu Ser Lys Asp Ala Tyr Ile Lys Ser
        195                 200                 205

Phe Gln Glu Ser Trp Asn Glu Tyr Ala Arg Arg Glu Asp Lys Thr Leu
    210                 215                 220

Ser Asp Phe Glu Ser Leu Cys Phe His Val Pro Phe Thr Lys Met Gly
225                 230                 235                 240

Lys Lys Ala Leu Asp Ser Ile Ile Asn Asp Ala Asp Glu Thr Thr Gln
                245                 250                 255

Glu Arg Leu Thr Ser Gly Tyr Glu Asp Ala Val Tyr Tyr Asn Arg Tyr
            260                 265                 270

Val Gly Asn Ile Tyr Thr Gly Ser Leu Tyr Leu Ser Leu Ile Ser Leu
        275                 280                 285

Leu Glu Asn Arg Ser Leu Lys Gly Gly Gln Thr Ile Gly Leu Phe Ser
    290                 295                 300

Tyr Gly Ser Gly Ser Val Gly Glu Phe Phe Ser Ala Thr Leu Val Glu
305                 310                 315                 320

Gly Tyr Glu Lys Gln Leu Asp Ile Glu Gly His Lys Ala Leu Leu Asn
                325                 330                 335

Glu Arg Gln Glu Val Ser Val Glu Asp Tyr Glu Ser Phe Phe Lys Arg
            340                 345                 350

Phe Asp Asp Leu Glu Phe Asp His Ala Thr Glu Gln Thr Asp Asp Asp
        355                 360                 365

Lys Ser Ile Tyr Tyr Leu Glu Asn Ile Gln Asp Asp Ile Arg Gln Tyr
    370                 375                 380

His Ile Pro Lys
385
```

REIVINDICACIONES

1. Un método para la producción de ácido 3-hidroxi-3-metilbutírico, que comprende la conversión enzimática de acetona y un compuesto que proporciona un grupo acetilo activado caracterizado por la siguiente fórmula (I):



en la que X es S-CH2-CH2-NH-CO-CH2-CH2-NH-CO-CH(OH)-C(CH3)2-CH2-O-PO2H-O-PO2H-C10H13N5O7P (coenzima A),

usando una enzima que es capaz de catalizar la formación de un enlace covalente entre el átomo de carbono del grupo oxo (es decir, el C=O) de acetona y el átomo de carbono (C^2) correspondiente al grupo metilo de dicho compuesto que proporciona un grupo acetilo activado según la fórmula (I), y en el que dicha enzima es una enzima que tiene la actividad de una HMG CoA sintasa (EC 2.3.3.10).

2. El método de la reivindicación 1, que se caracteriza por que la conversión se realiza en presencia de un organismo capaz de producir acetona y que expresa la enzima como se define en la reivindicación 1.

3. Un organismo recombinante que se caracteriza por las siguientes características:

(a) es capaz de producir acetona; y

(b) expresa una enzima que es capaz de catalizar la formación de un enlace covalente entre el átomo de carbono del grupo oxo (es decir, el C=O) de acetona y el átomo de carbono (C^2) correspondiente al grupo metilo del compuesto que proporciona un grupo acetilo activado según la fórmula (I), en la que X es coenzima A, y en el que dicha enzima es una enzima que tiene la actividad de una HMG CoA sintasa (EC 2.3.3.10).

4. El organismo de la reivindicación 3, en el que el organismo es un organismo que tiene naturalmente la capacidad para producir acetona.

5. El organismo de la reivindicación 4, en el que el organismo es un microorganismo que pertenece al género Clostridium, Bacillus o Pseudomonas.

6. El microorganismo de la reivindicación 5, en el que el microorganismo es un microorganismo que pertenece a las especies Clostridium acetobutylicum, Clostridium beijerinckii, Clostridium cellulolyticum, Bacillus polymyxa o Pseudomonas putida.

7. El organismo de la reivindicación 3, en el que el organismo deriva de un organismo que no produce naturalmente acetona y está modificado genéticamente para que sea capaz de producir acetona mediante la introducción de un gen o genes necesarios para permitir la producción de acetona en el organismo.

8. El organismo de la reivindicación 7, en el que el organismo es un organismo capaz de fotosíntesis.

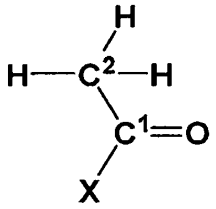
9. Uso de un organismo como se define en una cualquiera de las reivindicaciones 3 a 8, para la producción de ácido 3-hidroxi-3-metilbutírico.

10. Una composición que comprende un organismo de una cualquiera de las reivindicaciones 3 a 8.

11. Una composición que comprende

(i) acetona; y

(ii) un compuesto que proporciona un grupo acetilo activado caracterizado por la siguiente fórmula (I):

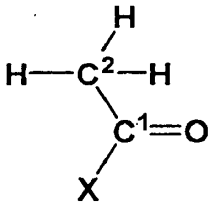


en la que X es S-CH2-CH2-NH-CO-CH2-CH2-NH-CO-CH(OH)-C(CH3)2-CH2-O-PO2H-O-PO2H-C10H13N5O7P (coenzima A); y

(iii) una enzima que es capaz de catalizar la formación de un enlace covalente entre el átomo de carbono del grupo oxo (es decir, el C=O) de acetona y el átomo de carbono (C^2) correspondiente al grupo metilo del compuesto definido en (ii), en el que dicha enzima es una enzima que tiene la actividad enzimática de una HMG CoA sintasa (EC 2.3.3.10)

12. Uso de una enzima como se define en la reivindicación 1, para la producción de ácido 3-hidroxi-3-metilbutírico.

13. Uso de acetona para la producción de ácido 3-hidroxi-3-metilbutírico, que comprende la conversión enzimática de acetona y un compuesto que proporciona un grupo acetilo activado caracterizado por la siguiente fórmula (I):



en la que X es S-CH2-CH2-NH-CO-CH2-CH2-NH-CO-CH(OH)-C(CH3)2-CH2-O-PO2H-O-PO2H-C10H13N5O7P (coenzima A), y en el que la conversión enzimática se efectúa mediante una enzima que tiene la actividad enzimática de una HMG CoA sintasa (EC 2.3.3.10).

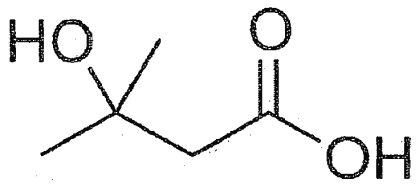


Figura 1

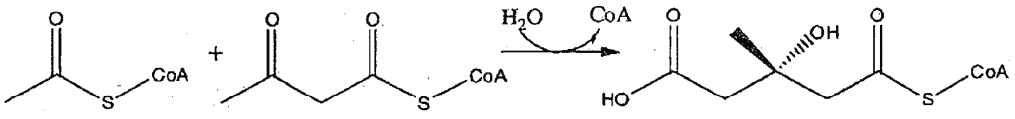


Figura 2

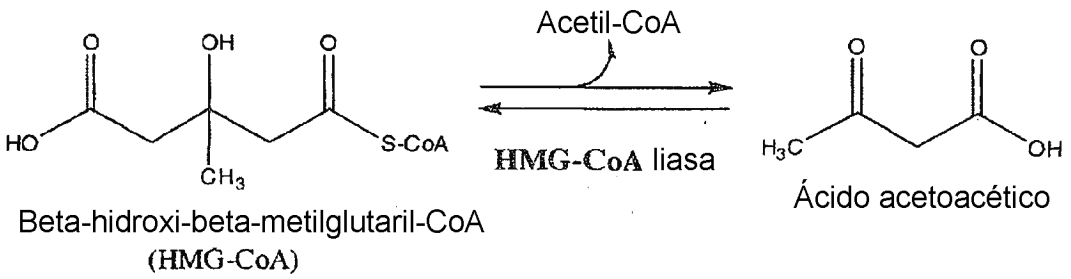


Figura 3

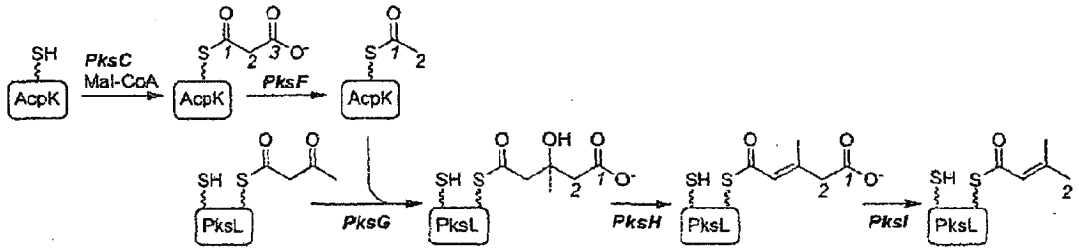


Figura 4

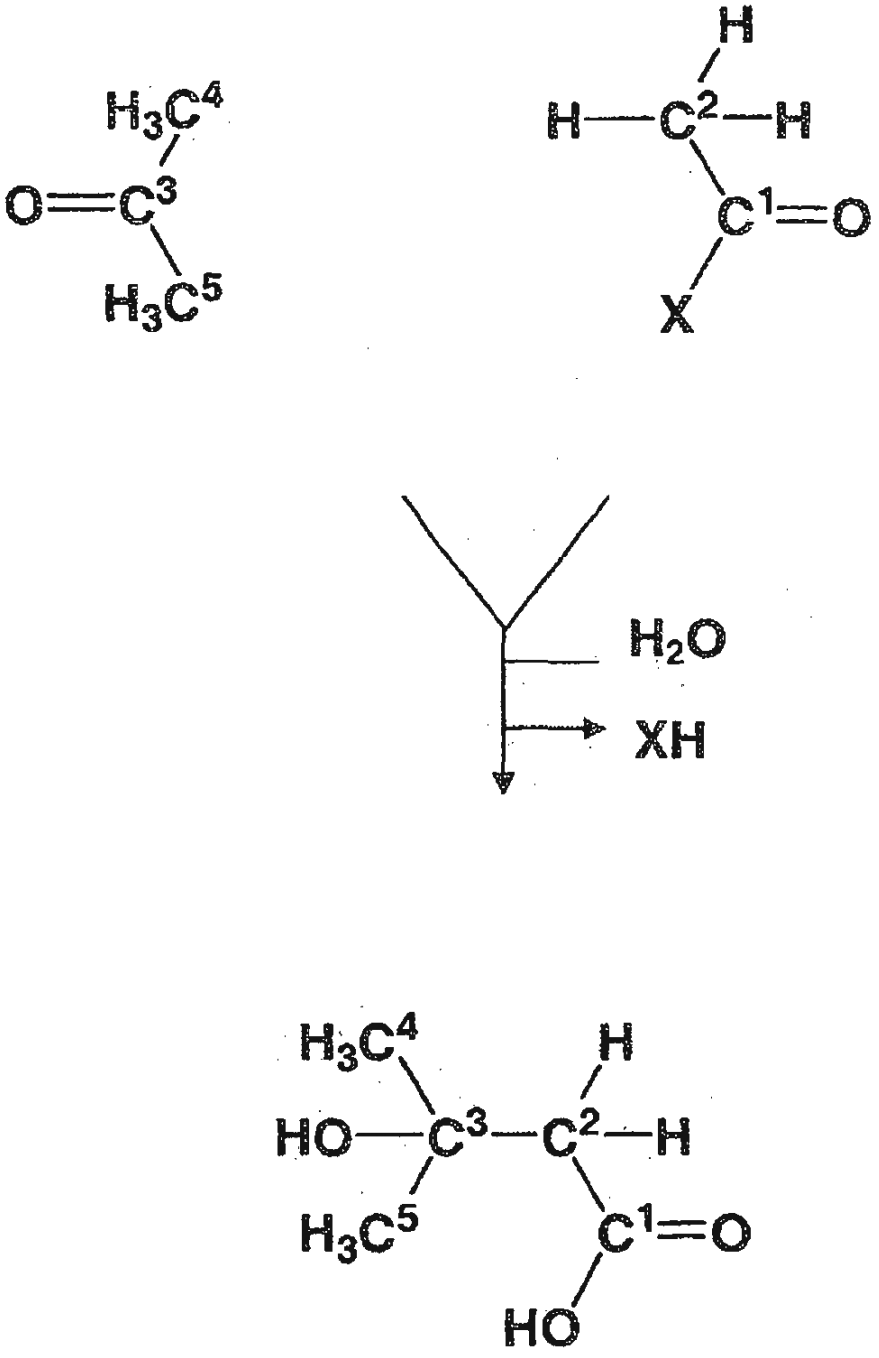


Figura 5

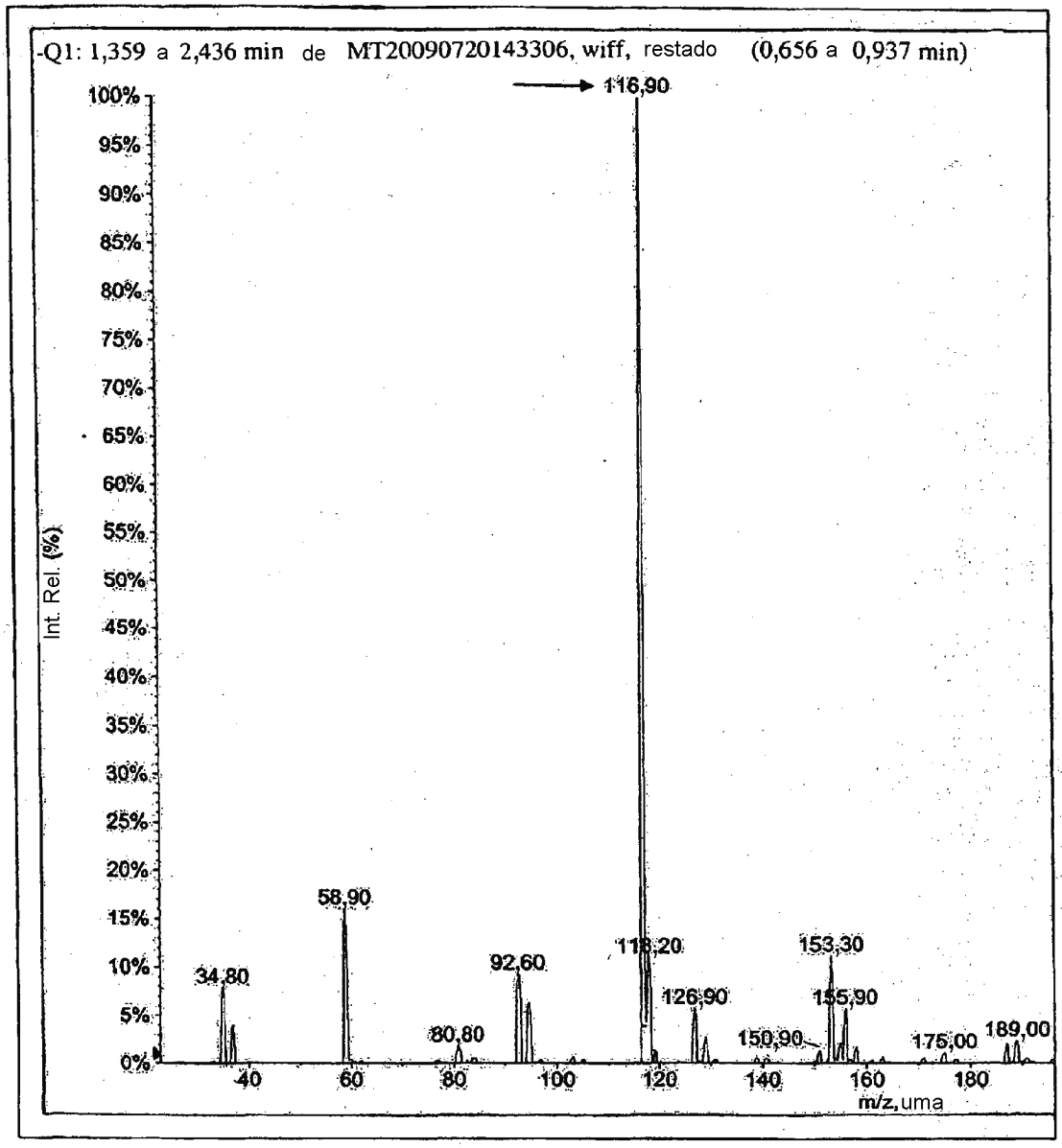

-Q1: 1,359 a 2,436 min de MT20090720143306, wiff, restado (0,656 a 0,937 min)
Int. Rel. (%)
m/z, uma
116,90
58,90
34,80
92,60
80,80
118,20
126,90
153,30
155,90
150,90
175,00
189,00

Figura 6

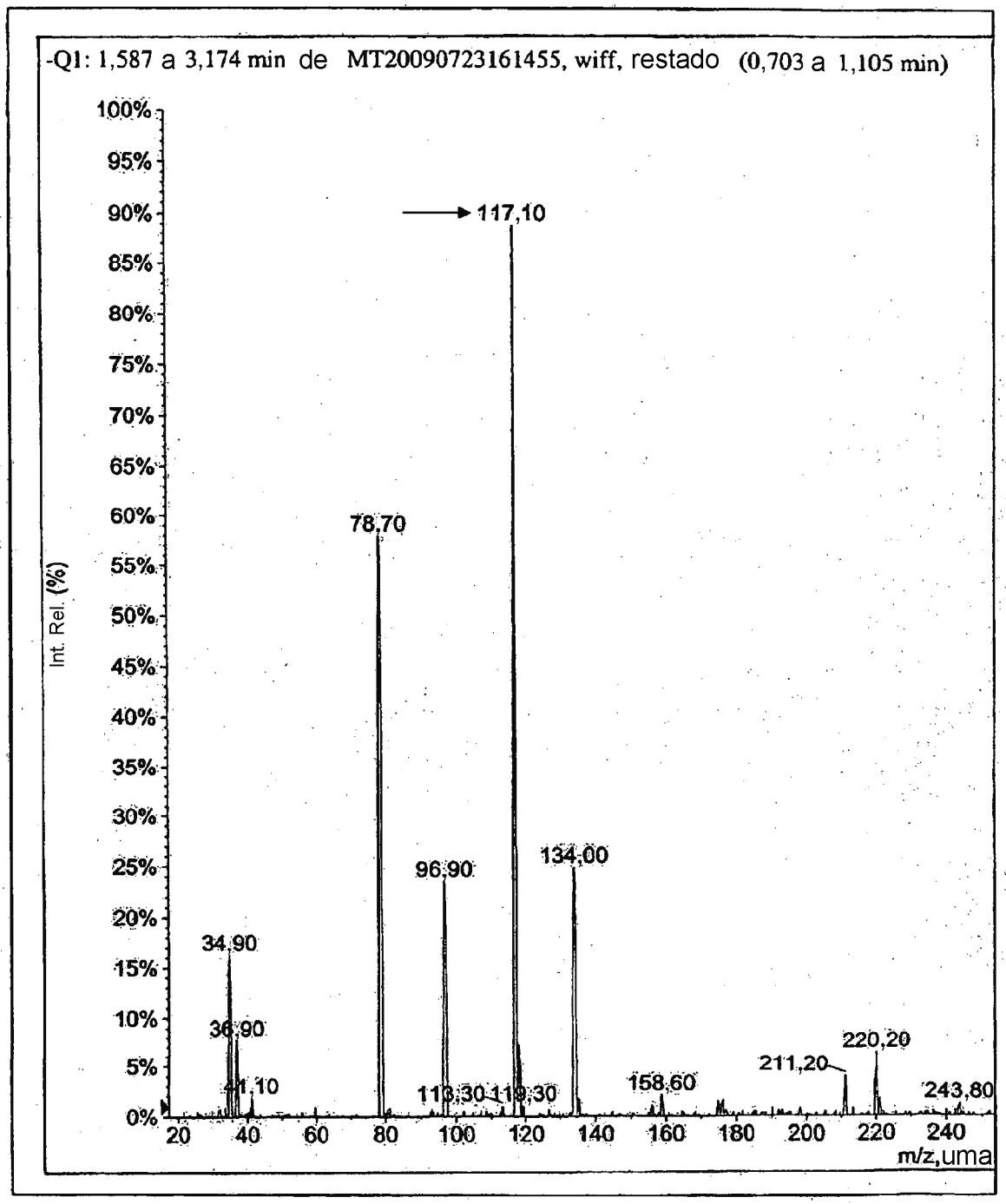

-Q1: 1,587 a 3,174 min de MT20090723161455, wiff, restado (0,703 a 1,105 min)
Int. Rel. (%)
100%
95%
90%
85%
80%
75%
70%
65%
60%
55%
50%
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
117,10
78,70
96,90
134,00
34,90
36,90
41,10
113,30
119,30
158,60
211,20
220,20
243,80
20
40
60
80
100
120
140
160
180
200
220
240
m/z,uma

Figura 7

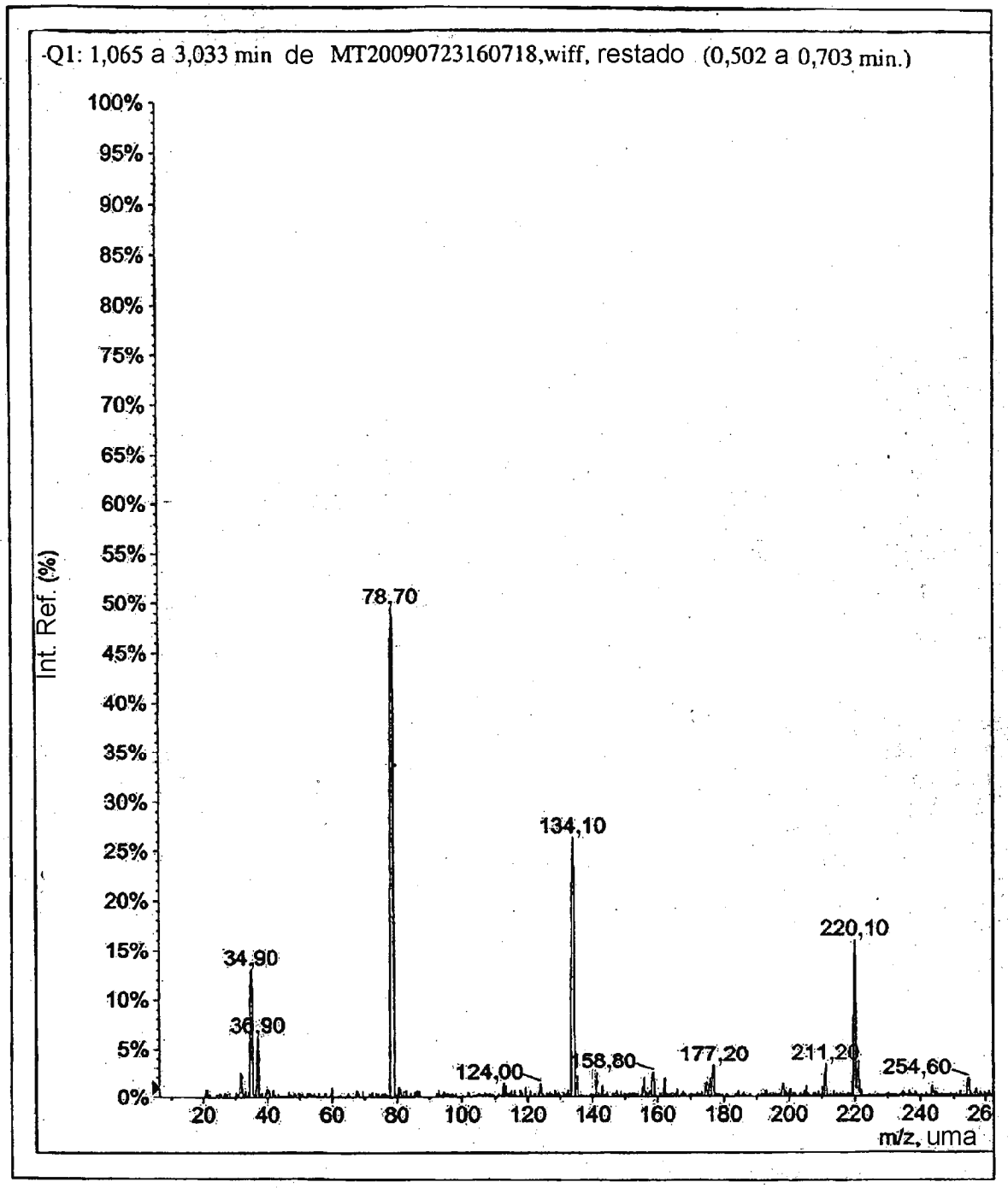

Q1: 1,065 a 3,033 min de MT20090723160718,wiff, restado (0,502 a 0,703 min.)
Int. Ref. (%)
100%
95%
90%
85%
80%
75%
70%
65%
60%
55%
50%
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
78,70
134,10
220,10
34,90
36,90
124,00
158,80
177,20
211,20
254,60
20
40
60
80
100
120
140
160
180
200
220
240
m/z, uma

Figura 8

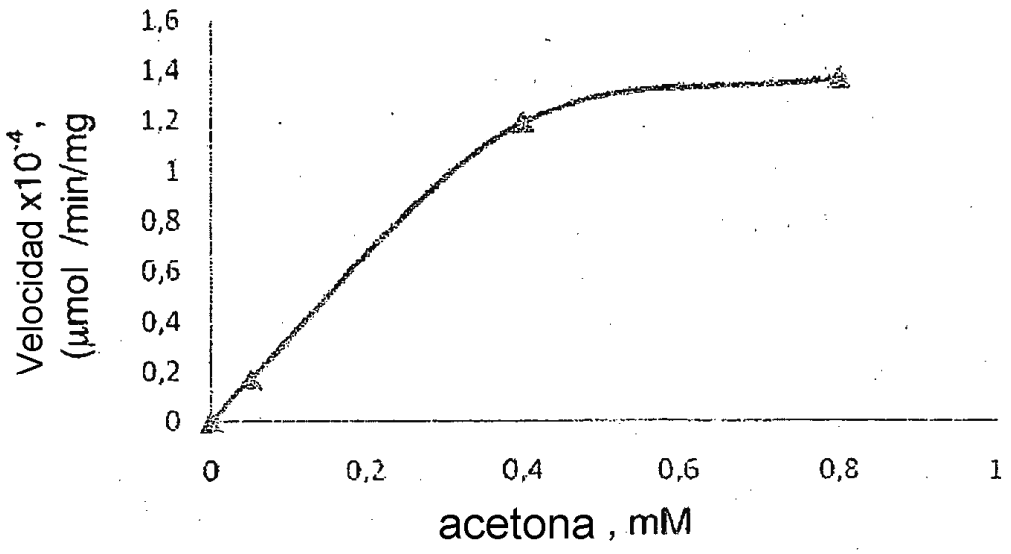

Velocidad x10-4 , (µmol /min/mg
1,6
1,4
1,2
1
0,8
0,6
0,4
0,2
0
0
0,2
0,4
0,6
0,8
1
acetona , mM

Figura 9