



①9



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

①1 Número de publicación: **2 356 132**

⑤1 Int. Cl.:

C12P 17/10 (2006.01)
C07D 209/18 (2006.01)
C12N 1/00 (2006.01)
C12N 9/10 (2006.01)
C12N 9/88 (2006.01)
C12N 9/02 (2006.01)
C12N 15/52 (2006.01)
C12N 15/63 (2006.01)

①2

TRADUCCIÓN DE PATENTE EUROPEA

T3

⑨6 Número de solicitud europea: **03747304 .8**

⑨6 Fecha de presentación : **23.04.2003**

⑨7 Número de publicación de la solicitud: **1503985**

⑨7 Fecha de publicación de la solicitud: **09.02.2005**

⑤4 Título: **Polipéptidos y rutas biosintéticas.**

③0 Prioridad: **23.04.2002 US 374831 P**

④5 Fecha de publicación de la mención BOPI:
05.04.2011

④5 Fecha de la publicación del folleto de la patente:
05.04.2011

⑦3 Titular/es: **CARGILL, INCORPORATED**
P.O. Box 5624
Minneapolis, Minnesota 55440-5624, US

⑦2 Inventor/es: **Abraham, Timothy, W.;**
Cameron, Douglas, C.;
Dalluge, Joseph;
Hicks, Paula, M.;
Hobson, Russell, J.;
McFarlan, Sara, C.;
Millis, Jim;
Rosazza, John;
Weiner, David P. y
Zhao, Lishan

⑦4 Agente: **Curell Aguilá, Marcelino**

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

CAMPO DE LA INVENCIÓN

La presente invención proporciona polipéptidos y rutas biosintéticas que son útiles en la producción de indol-3-piruvato, 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutárico (PM) y/o monatina.

ANTECEDENTES*Indol piruvato*

El indol-3-piruvato es un potente antioxidante que se cree que contrarresta el estrés oxidativo en tejidos con altas concentraciones de oxígeno (Politi *et al.*, "Recent advances in Tryptophan Research", editado por G. A. Filippini *et al.*, Plenum Press, Nueva York, 1996, págs. 291-8). El indol piruvato además es un compuesto intermedio en una serie de reacciones para producir ácido indol-acético (IAA), la hormona auxina del crecimiento vegetal primario (factor activador de crecimiento difusible). El IAA es activo en cantidades inferiores al microgramo en una gama de procesos fisiológicos que incluyen la dominancia apical, tropismos, alargamiento del brote, inducción de la división de las células del cambium e iniciación de la raíz. Las auxinas sintéticas se utilizan en horticultura para provocar enraizamiento y para favorecer el cuajado y el desarrollo del fruto. A altas concentraciones las auxinas sintéticas son herbicidas eficaces contra las plantas de hoja ancha. Las auxinas naturales producidas por fermentación pueden considerarse más adecuadas desde un punto de vista ambiental que los herbicidas producidos químicamente. Los reguladores del crecimiento se vendieron en todo el mundo en 1999 por valor de 0,4 miles de millones de libras (1,4 miles de millones de dólares US).

Algunos ejemplos de patentes sobre el ácido indolacético y de derivados del mismo incluyen: la patente US nº 5.843.782 Micropropagation of rose plants, auxina utilizada en el medio de cultivo y la patente US nº 5.952.231 Micropropagation of rose plants.

Además de las utilidades relacionadas con las plantas, el ácido indolacético es útil en aplicaciones farmacéuticas, por ejemplo, la patente US nº 5.173.497 "Method of preparing alpha-oxopyrrolo[2,3-B]indole acetic acids and derivatives" propone la utilización de estos compuestos en el tratamiento del deterioro de la memoria tal como el asociado a la enfermedad de Alzheimer y a la demencia senil. El mecanismo propuesto en la patente US nº 5.173.497 es que estos compuestos inhiben la polipéptido acetilcolinesterasa y aumentan las concentraciones de acetilcolina en el cerebro.

El indol-3-carbinol se produce a partir del indol-3-acético por oxidación catalizada por peroxidasa, y puede convertirse fácilmente en diindolilmetano. Se informa de que ambos compuestos eliminan toxinas y favorecen la producción de hormonas beneficiosas para la salud de las mujeres.

Derivados de triptófano

El D-triptófano clorado se ha identificado como un edulcorante no nutritivo, y existe un interés creciente en dedicarse también a otros derivados. La monatina es un edulcorante natural que es similar en composición al aminoácido triptófano. Puede extraerse de la corteza de las raíces del arbusto sudafricano, *Sclerochiton ilicifolius*, y es prometedor en la industria alimenticia y de bebidas como edulcorante de gran intensidad. Algunos ejemplos de patentes sobre monatina incluyen: la patente US nº 5.994.559 Synthesis of monatin-A high intensity natural sweetener, la patente US nº 5.975.298 3-(1-amino-1,3-dicarboxy-3-hydroxy-but-4-yl)-indole compounds, la patente US nº 5.128.164 Composition for human consumption containing 3-(1-amino-1,3-dicarboxy-3-hydroxy-but-4-yl)-indole compounds; y la patente US nº 5.128.482 Process for the production of 3-1(1-amino-1,3-dicarboxy-3-hydroxy-but-4-yl) indole.

Algunos de los precursores de monatina descritos en la presente memoria pueden ser también útiles como edulcorantes sintéticos o como productos intermedios en la síntesis de derivados de monatina.

Anteriormente, fueron descritos procedimientos para producir monatina (Nakamura *et al.*, *Organic Letters*, 2, 19: 2967-2970, 2000). Además, una cepa de *E. coli* que sobreexpresa un gen para aldolasa (Patil *et al.*, *J. Bacteriol.*, 171, 1: 1402-107, 1992), se han descrito células de *E. coli* que expresan aminoácido deshidrogenasas o aminoácido transferasas heterólogas (Galkin *et al.*, *Applied and Environmental Microbiology*, 63, 12:4651-4656, 1997) y las células de levadura que expresan glutamato deshidrogenasas (Deluna *et al.*, *J. Biol. Chem.*, 274, 47:43775-43783, 2001). Además, la patente EP 1 445 323 describe la producción biosintética de monatina a partir de triptófano con indol-3-piruvato y 2-hidroxiácido 2(indol-3-il-metil)-4-cetoglutárico como productos intermedios.

SUMARIO

La presente invención se refiere a las formas de realización caracterizadas en las reivindicaciones. Por lo tanto se refiere a un procedimiento de producción de monatina, que comprende convertir el triptófano y/o el ácido

indol-3-láctico en monatina, en el que el procedimiento comprende:

- (a) poner en contacto el triptófano o el ácido indol-3-láctico con un polipéptido, en el que, cuando el triptófano está en contacto con el primer polipéptido, el primer polipéptido se selecciona de entre triptófano aminotransferasa (CE 2.6.1.27), triptófano deshidrogenasa (CE 1.4.1.19), tirosina aminotransferasa (aromática) (CE 2.6.1.5), triptofan-fenilpiruvato transaminasa (CE 2.6.1.28), sustrato múltiple aminotransferasa (CE 2.6.1.-), aspartato aminotransferasa (CE 2.6.1.1), triptófano oxidasa, L-aminoácido oxidasa (CE 1.4.3.2), D-amino-ácido deshidrogenasa (CE 1.4.99.1), D-aminoácido oxidasa (CE 1.4.3.3), D-triptófano aminotransferasa, D-alanina aminotransferasa (CE 2.6.1.21), glutamato deshidrogenasa (CE 1.4.1.2-4), fenilalanina deshidrogenasa (CE 1.4.1.20), o una combinación de las mismas, en la que, cuando el ácido indol-3-láctico se pone en contacto con el primer polipéptido, el primer polipéptido se selecciona de entre indol-lactato deshidrogenasa (CE 1.1.1.110), R-4 hidroxifenil-lactato deshidrogenasa (CE 1.1.1.222), 3-(4)-hidroxifenilpiruvato reductasa (CE 1.1.1.237), lactato deshidrogenasa (CE 1.1.1.27, 1.1.1.28, 1.1.2.3), (3-imidazol-5-il) lactato deshidrogenasa (CE 1.1.1.111), y lactato oxidasa (CE 1.1.3.-), o una combinación de las mismas, en la que el primer polipéptido convierte el triptófano o el ácido indol-3-láctico en indol-3-piruvato;
- (b) poner en contacto el indol-3-piruvato con un segundo polipéptido y una fuente de carbono C3, en el que el segundo polipéptido es una enzima CE 4.1.2 o CE 4.1.3., en la que el segundo polipéptido convierte el indol-3-piruvato y la fuente de carbono C3 en 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico; y
- (c) poner en contacto el 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico con un tercer polipéptido, en el que el tercer polipéptido se selecciona de entre tirosina aminotransferasa (aromático) (CE 2.6.1.5), triptófano deshidrogenasa (CE 1.4.1.19), triptófano-fenilpiruvato transaminasa (CE 2.6.1.28), triptófano aminotransferasa (CE 2.6.1.27), aspartato aminotransferasa (CE 2.6.1.1), D-alanina aminotransferasa (CE 2.6.1.21), D-triptófano aminotransferasa, glutamato deshidrogenasa (CE 1.4.1.2-4), fenilalanina deshidrogenasa (CE 1.4.1.20), sustrato múltiple aminotransferasa (CE 2.6.1.-), D-aminoácido deshidrogenasa (CE 1.4.99.1), una lisina épsilon aminotransferasa, o una combinación de las mismas, en la que el tercer polipéptido convierte el 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico en monatina.

La presente invención se refiere además, al procedimiento de producción de monatina, en el que, cuando el primer polipéptido que convierte triptófano en indol-3-piruvato es una aminoácido oxidasa (CE 1.4.3.-), se utiliza además una catalasa (CE 1.11.1.6).

La presente invención se refiere además, al procedimiento de producción de monatina, en el que, cuando el primer polipéptido que convierte triptófano en indol-3-piruvato es una aminoácido oxidasa (CE 1.4.3.-), se utiliza además una catalasa (CE 1.11.1.6) en el que la aminoácido oxidasa es una triptófano oxidasa..

La presente invención se refiere además, al procedimiento de producción de monatina, en el que, la enzima CE 4.1.3 es la 4-hidroxi-2-oxoglutarato glioxilato-liasa (CE 4.1.3.16), la 4-hidroxi-4-metil-2-oxoglutarato piruvato-liasa (CE 4.1.3.17) o una combinación de las mismas.

La presente invención se refiere además, al procedimiento de producción de monatina, en el que el segundo polipéptido es una secuencia de aminoácidos seleccionada de entre el grupo constituido por:

- (a) SEC. ID. nº 66, nº de registro en Genbank: CAC46344, nº de registro en Genbank: CAB 14127.1, nº de registro en Genbank: AAC74920.1 o nº de registro en Genbank: CAC47463.1;
- (b) una secuencia de aminoácidos que comprende por lo menos el 90% de identidad de la secuencia con la SEC. ID. nº 66, nº de registro en Genbank: CAC46344, nº de registro en Genbank: CAB 14127.1, nº de registro en Genbank: AAC74920.1 o nº de registro en Genbank: CAC47463.1; y
- (c) secuencias de aminoácidos que se diferencian en la SEC. ID. nº 66, nº de registro en Genbank: CAC46344, nº de registro en Genbank: CAB 14127.1, nº de registro en Genbank: AAC74920.1 o nº de registro en Genbank: CAC47463.1 en menos de 50 sustituciones de aminoácidos conservadoras, en el que la secuencia de aminoácidos convierte el indol-3-piruvato y el piruvato en 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico.

La presente invención se refiere además, al procedimiento de producción de monatina, en el que, cuando el primer polipéptido es una aspartato aminotransferasa, el 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico se pone en contacto con el aspartato para producir oxaloacetato y monatina.

La presente invención se refiere además, al procedimiento de producción de monatina, en el que, cuando el tercer polipéptido es una aspartato aminotransferasa, el 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico se pone en contacto con el aspartato para producir oxaloacetato y monatina y que comprende además la puesta en contacto del oxaloacetato con una descarboxilasa.

La presente invención se refiere además, al procedimiento de producción de monatina, en el que, cuando el tercer polipéptido es una lisina epsilon aminotransferasa el 2-hidroxiácido 2-(indol-3-ilmetil)-cetoglutárico se pone en contacto además con lisina.

La presente invención se refiere además, a un procedimiento de producción de monatina, en el que, el primer y/o tercer polipéptidos se seleccionan de entre el grupo constituido por:

- (a) una secuencia de aminoácidos mostrada en las SEC. ID. nº: 2, 4, 6, 10, 12, 14, 32, nº de registro en Genbank: NP_388848.1, nº de registro en Genbank: ZP00005082.1 o nº de registro en Genbank: ACC74014.1;
- (b) una secuencia de aminoácidos que comprende por lo menos 90% de identidad de secuencias con la secuencia mostrada en las SEC. ID. nº 2, 4, 6, 8, 10, 12, 14, 32, nº de registro en Genbank: NP_388848.1, nº de registro en Genbank: ZP00005082.1 o nº de registro en Genbank: ACC74014.1; y
- (c) secuencias de aminoácidos que se diferencian de las SEC. ID. nº 2, 4, 6, 8, 10, 12, 14, 32, nº de registro en Genbank: NP_388848.1, nº de registro en Genbank: ZP00005082.1 o nº de registro en Genbank: ACC74014.1 en menos de 50 sustituciones conservadoras de aminoácidos, en las que la secuencia de aminoácidos convierte el triptófano en indol-3-piruvato.

La presente invención se refiere además, al procedimiento de producción de monatina, en el que, el triptófano se produce a partir del indol utilizando un polipéptido procedente de CE 4.1.99.1.

La presente invención se refiere además, al procedimiento de producción de monatina, en el que, la fuente de carbono C3 se selecciona de entre piruvato, fosfoenolpiruvato, alanina, serina, cisteína o una combinación de los mismos.

La presente invención se refiere al procedimiento de producción de monatina, en el que, la fuente de carbono C3 se selecciona de entre piruvato, fosfoenolpiruvato, alanina, serina, cisteína o una combinación de los mismos, en el que el piruvato es producido por:

- (a) alanina y un polipéptido que puede realizar la transaminación de alanina;
- (b) serina y un polipéptido que puede realizar la beta-eliminación de serina;
- (c) cisteína y un polipéptido que puede realizar la beta-eliminación de cisteína;
- (d) aspartato y un polipéptido que puede realizar las reacciones de beta-liasa con un aminoácido dicarboxílico;
- (e) lactato y lactato-oxidasa (CE 1.1.3.-) o una lactato-deshidrogenasa (CE 1.1.1.27, 1.1.1.28, 1.1.2.3); o
- (f) una combinación de los mismos.

La presente invención se refiere al procedimiento de producción de monatina, en el que, el procedimiento comprende además reducir una cantidad de peróxido de hidrógeno producido cuando el triptófano se convierte en indol-3-piruvato poniendo en contacto el peróxido de hidrógeno con una catalasa.

La presente invención se refiere al procedimiento de producción de monatina, en el que se producen por lo menos 10 g/l de monatina.

La presente invención se refiere además a un procedimiento de producción de monatina, en el que, el procedimiento comprende además la detección del 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutárico por espectrometría de masas.

La presente invención se refiere además a un procedimiento de producción de monatina, que comprende convertir el indol-3-piruvato en monatina en presencia de piruvato y un primer y un segundo polipéptidos, en el que el primer polipéptido es una aldolasa, y en el que el segundo polipéptido se selecciona de entre tirosina aminotransferasa (aromática) (CE 2.6.1.5), triptófano deshidrogenasa (CE 1.4.1.19), triptófano-fenilpiruvato transaminasa (CE 2.6.1.28), triptófano aminotransferasa (CE 2.6.1.27), aspartato aminotransferasa (CE 2.6.1.1), D-alanina aminotransferasa (CE 2.6.1.21), D-triptófano aminotransferasa, glutamato deshidrogenasa (CE 1.4.1.2-4), fenilalanina deshidrogenasa (CE 1.4.1.20), sustrato múltiple aminotransferasa (CE 2.6.1.-), D-aminoácido deshidrogenasa (CE 1.4.99.1) o una combinación de las mismas.

La invención proporciona varias rutas biosintéticas para la preparación de monatina a partir de glucosa, triptófano, ácido indol-3-láctico y/o mediante precursores de monatina tales como indol-3-piruvato y 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutárico. Se dan a conocer las secuencias de polipéptidos y de ácidos nucleicos que pueden utilizarse para preparar monatina, indol-3-piruvato y 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutárico.

La monatina puede producirse mediante indol-3-piruvato, 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico (precursor de monatina, PM, la forma alfa-ceto de monatina), ácido indol-3-láctico, triptófano y/o glucosa (figura 1). Se dan a conocer procedimientos de producción o preparación de monatina o sus intermedios mostrados en las figuras 1 a 3 y 11 a 13 que implican convertir un sustrato en un primer producto, y a continuación convertir el primer producto en un segundo producto y así sucesivamente hasta que se crea el producto final deseado.

Las figuras 1 a 3 y 11 a 13 presentan productos intermedios potenciales y productos finales en casillas. Por ejemplo, una conversión de un producto en otro, tal como glucosa en triptófano, triptófano en indol-3-piruvato, indol-3-piruvato en PM, PM en monatina, o ácido indol-3-láctico (indol-lactato) en indol-3-piruvato, puede realizarse utilizando estos procedimientos. Estas conversiones pueden facilitarse químicamente o biológicamente. El término "convertir" se refiere a la utilización ya sea de medios químicos o de polipéptidos en una reacción que cambia un primer intermedio en un segundo intermedio. La expresión "conversión química" se refiere a las reacciones que los polipéptidos no facilitan activamente. La expresión "conversión biológica" se refiere a las reacciones que los polipéptidos facilitan activamente. Las conversiones pueden tener lugar *in vivo* o *in vitro*. Cuando se utilizan conversiones biológicas los polipéptidos y/o las células pueden inmovilizarse sobre soportes tal como por enlace químico sobre soportes de polímero. La conversión puede realizarse utilizando cualquier reactor conocido por cualquier experto en la materia, por ejemplo en un reactor discontinuo o continuo.

Se proporcionan además procedimientos que incluyen poner en contacto un primer polipéptido con un sustrato y elaborar un primer producto, y a continuación poner en contacto el primer producto creado con un segundo polipéptido y crear un segundo producto, y a continuación poner en contacto el segundo producto creado con un tercer polipéptido y crear un tercer producto, por ejemplo monatina. Los polipéptidos utilizados en los productos producidos se presentan en las figuras 1 a 3 y 11 a 13.

Se dan a conocer los polipéptidos, y sus secuencias de codificación, que pueden utilizarse y realizar las conversiones representadas en las figuras 1 a 3 y 11 a 13. En algunos ejemplos, los polipéptidos que tienen una o más mutaciones puntuales que permiten que se modifique la especificidad de sustrato y/o la actividad del polipéptido, se utilizan para preparar monatina.

Se dan a conocer células aisladas y recombinantes que producen monatina. Estas células pueden ser cualquier célula, tal como vegetal, animal, bacteriana, de levadura, de alga, de arquea o fúngica.

En un ejemplo específico, las células dadas a conocer incluyen una o más de las actividades siguientes, por ejemplo dos o más o tres o más de las actividades siguientes: triptófano aminotransferasa (CE 2.6.1.27), tirosina aminotransferasa (aromática) (CE 2.6.1.5), sustrato múltiple aminotransferasa (CE 2.6.1.-), aspartato aminotransferasa (CE 2.6.1.1), triptófano deshidrogenasa (CE 1.4.1.19), triptófano-fenilpiruvato transaminasa (CE 2.6.1.28), L-aminoácido oxidasa (CE 1.4.3.2), triptófano oxidasa (sin número CE, Hadar *et al.*, *J. Bacteriol.* 125:1096-1104, 1976 y Furuya *et al.*, *Biosci. Biotechnol. Biochem.* 64: 1486-93, 2000), D-aminoácido deshidrogenasa (CE 1.4.99.1), D-aminoácido oxidasa (CE 1.4.3.3), D-alanina aminotransferasa (CE 2.6.1.21), sintasa/liasa (4.1.3.-), tal como 4-hidroxi-4-metil-2-oxoglutarato aldolasa (CE 4.1.3.17) o 4-hidroxi-2-oxoglutarato aldolasa (CE 4.1.3.16), sintasa/liasa (4.1.2.-), D-triptófano aminotransferasa (Kohiba y Mito, Proceedings of the 8th International Symposium on Vitamin B₆ and Carbonyl Catalysis, Osaka, Japón, 1990), fenilalanina deshidrogenasa (CE 1.4.1.20) y/o glutamato deshidrogenasa (CE 1.4.1.2, 1.4.1.3, 1.4.1.4).

En otro ejemplo, las células incluyen una o más, por ejemplo dos o más, o tres o más, de las actividades siguientes: indol-lactato deshidrogenasa (CE 1.1.1.110), R-4-hidroxifenil-lactato deshidrogenasa (CE1.1.1.222), 3-(4)-hidroxifenilpiruvato reductasa (CE1.1.1.237), lactato deshidrogenasa (CE1.1.1.27, 1.1.1.28, 1.1.2.3), (3-imidazol-5-il) lactato deshidrogenasa (CE 1.1.1.111), lactato oxidasa (CE 1.1.3.-), sintasa/liasa (CE 4.1.3.-), tal como 4-hidroxi-4-metil-2-oxoglutarato aldolasa (CE 4.1.3.17) o 4-hidroxi-2-oxoglutarato aldolasa (CE 4.1.3.16), sintasa/liasa (4.1.2.-), triptófano deshidrogenasa (CE 1.4.1.19), triptófano-fenilpiruvato transaminasa (CE 2.6.1.28), triptófano amino-transferasa (CE 2.6.1.27), tirosina aminotransferasa (aromática) (CE 2.6.1.5), sustrato múltiple aminotransferasa (CE 2.6.1.-), aspartato aminotransferasa (CE 2.6.1.1), fenilalanina deshidrogenasa (CE 1.4.1.20), glutamato deshidrogenasa (CE 1.4.1.2, 1.4.1.3, 1.4.1.4), D-aminoácido deshidrogenasa (CE 1.4.99.1), D-triptófano aminotransferasa y/o D-alanina aminotransferasa (CE 2.6.1.21).

Además, las células dadas a conocer incluyen una o más de las siguientes actividades, por ejemplo dos o más o tres o más de las siguientes actividades: triptófano aminotransferasa (CE 2.6.1.27), tirosina aminotransferasa (aromática) (CE 2.6.1.5), sustrato múltiple aminotransferasa (CE 2.6.1.-), aspartato aminotransferasa (CE 2.6.1.1), triptófano deshidrogenasa (CE 1.4.1.19), triptófano-fenilpiruvato transaminasa (CE 2.6.1.28), L-aminoácido oxidasa (CE 1.4.3.2), triptófano oxidasa (sin número CE), D-aminoácido deshidrogenasa (CE 1.4.99.1), D-aminoácido oxidasa (CE 1.4.3.3), D-alanina aminotransferasa (CE 2.6.1.21), indol-lactato deshidrogenasa (CE 1.1.1.110), R-4-hidroxifenil-lactato deshidrogenasa (CE1.1.1.222), 3-(4)-hidroxifenilpiruvato reductasa (CE1.1.1.237), lactato deshidrogenasa (CE1.1.1.27, 1.1.1.28, 1.1.2.3), (3-imidazol-5-il) lactato deshidrogenasa (CE 1.1.1.111), lactato oxidasa (CE 1.1.3.-), sintasa/liasa (CE 4.1.3.-), tal como 4-hidroxi-4-metil-2-oxoglutarato aldolasa (CE 4.1.3.17) o 4-hidroxi-2-oxoglutarato aldolasa (CE 4.1.3.16), sintasa/liasa (4.1.2.-), glutamato deshidrogenasa (CE 1.4.1.2, 1.4.1.3, 1.4.1.4), fenilalanina deshidrogenasa (CE 1.4.1.20) y/o D-triptófano aminotransferasa.

La monatina puede producirse por un procedimiento que incluye poner en contacto el triptófano y/o el ácido indol-3-láctico, con un primer polipéptido, en el que el primer polipéptido convierte el triptófano y/o el ácido indol-3-láctico en indol-3-piruvato (bien la forma D o la L del triptófano o el ácido indol-3-láctico puede utilizarse como sustrato que se convierte en indol-3-piruvato; un experto en la materia apreciará que los polipéptidos seleccionados para esta etapa presentan teóricamente la especificidad apropiada), poner en contacto el indol-3-piruvato resultante con un segundo polipéptido, en el que el segundo polipéptido convierte el indol-3-piruvato en 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico (PM) y poner en contacto el PM con un tercer polipéptido, en el que el tercer polipéptido convierte el PM en monatina. Los polipéptidos ejemplificativos que pueden utilizarse para esta conversión se muestran en las figuras 2 y 3.

Otro aspecto de la invención proporciona composiciones tales como el PM, células que contienen por lo menos dos polipéptidos, o algunas veces por lo menos tres o por lo menos cuatro polipéptidos, que son codificados en por lo menos una secuencia de ácido nucleico exógeno.

Estos y otros aspectos de la invención resultan evidentes a partir de la descripción detallada y de los ejemplos ilustrativos siguientes.

BREVE DESCRIPCIÓN DE LAS FIGURAS

La figura 1 presenta las rutas biosintéticas utilizadas para producir monatina y/o indol-3-piruvato. Una ruta produce indol-3-piruvato por medio de triptófano, mientras que otra produce indol-3-piruvato por medio de ácido indol-3-láctico. La monatina se produce posteriormente por medio de un 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico (PM) intermedio.

Los compuestos mostrados en casillas son sustratos y productos producidos en las rutas biosintéticas.

Las composiciones junto a las flechas son cofactores, o reactantes que pueden utilizarse durante la conversión de un sustrato en un producto. El cofactor o reactante utilizado dependerá del polipéptido utilizado para la etapa específica de la ruta biosintética. El cofactor PLP (piridoxal 5'-fosfato) puede catalizar reacciones independientes de un polipéptido, y por consiguiente, proporcionando únicamente PLP puede permitir la evolución del sustrato a producto.

La figura 2 es un diagrama con mayor detalle de la ruta biosintética que utiliza el PM intermedio. Los sustratos de cada etapa en la ruta se muestran en casillas. Los polipéptidos que permiten la conversión entre los sustratos se listan junto a las flechas entre los sustratos. Cada polipéptido está descrito por su denominación corriente y un número de clase enzimática (CE).

La figura 3 presenta un diagrama con mayor detalle de la ruta biosintética de la conversión del ácido indol-3-láctico en indol-3-piruvato. Los sustratos se muestran en casillas y los polipéptidos que permiten la conversión entre los sustratos se listan junto a la flecha entre los sustratos. Cada polipéptido está descrito por su denominación corriente y un número de clase enzimática (CE).

La figura 4 presenta una posible reacción para preparar el PM por medios químicos.

Las figuras 5A y 5B son cromatogramas que presentan la identificación por LC/MS de monatina producida mediante enzimas.

La figura 6 es un espectro de masas con electroatomización de monatina sintetizada mediante enzimas.

Las figuras 7A y 7B representan cromatogramas de los análisis LC/MS/MS del ión derivado de monatina producida en una mezcla enzimática.

La figura 8 es un cromatograma que muestra la cuantificación por espectroscopia de masas de alta resolución de monatina producida mediante enzimas.

Las figuras 9A a 9C son cromatogramas que presentan la separación quiral del (A) R-triptófano, (B) S-triptófano y (C) monatina producida mediante enzimas.

La figura 10 es un gráfico de barras que muestra la cantidad relativa de monatina producida en células bacterianas después de la inducción con IPTG. El (-) indica una falta de adición del sustrato (no se añadió triptófano ni piruvato).

Las figuras 11 y 12 son diagramas esquemáticos que muestran las rutas utilizadas para aumentar el rendimiento de monatina producida a partir de triptófano o de indol-3-piruvato.

La figura 13 es un diagrama esquemático que presenta una ruta que puede utilizarse para aumentar el

rendimiento de monatina producida a partir de triptófano o de indol-3-piruvato.

LISTADO DE SECUENCIAS

- 5 Las secuencias de ácidos nucleicos y de aminoácidos en el listado de secuencias adjunto se muestran utilizando abreviaturas en letras normales para las bases nucleotídicas y código de tres letras para los aminoácidos. Solamente se muestra una cadena de cada secuencia de ácido nucleico, pero la cadena complementaria se sobreentiende que está incluida por cualquier referencia a la cadena presentada.
- 10 Las SEC. ID. nº 1 y nº 2 representan las secuencias de ácido nucleico y de aminoácidos de una aminotransferasa procedente de *Sinorhizobium meliloti*, respectivamente (gen *tatA*, denominado una tirosina aminotransferasa o aromática en la bibliografía).
- 15 Las SEC. ID. nº 3 y nº 4 representan las secuencias de ácidos nucleicos y de aminoácidos de una tirosina aminotransferasa procedente de *Rhodobacter sphaeroides* (2.4.1), respectivamente (por homología con la *tatA* (SEC. ID. nº 1 y nº 2) se predijo que era una "aspartato aminotransferasa" por programa informático genómico).
- 20 Las SEC. ID. nº 5 y nº 6 representan las secuencias de ácido nucleico y de aminoácidos de una aminotransferasa de *Rhodobacter sphaeroides* (35053), respectivamente (nuevas, clonadas basadas en la secuencia de 2.4.1 de las SEC. ID. nº 3 y nº 4).
- 25 Las SEC. ID. nº 7 y nº 8 representan las secuencias de ácido nucleico y de aminoácidos de una sustrato amplio aminotransferasa (*bsat*) de *Leishmania major*, respectivamente.
- 30 Las SEC. ID. nº 9 y nº 10 representan las secuencias de ácido nucleico y de aminoácidos de una aminotransferasa aromática (*araT*) de *Bacillus subtilis*, respectivamente.
- 35 Las SEC. ID. nº 11 y nº 12 representan nuevas secuencias de ácidos nucleicos y de aminoácidos de una aminotransferasa aromática (*araT*) de *Lactobacillus amylovorus*, respectivamente (identificada por homología como una aminotransferasa aromática).
- 40 Las SEC. ID. nº 13 y nº 14 representan las secuencias de ácidos nucleicos y de aminoácidos de una sustrato múltiple aminotransferasa (*msa*) de *R. sphaeroides* (35053), respectivamente (identificada como una sustrato múltiple aminotransferasa por homología con el nº de registro AAAE01000093.1, 14743 a 16155 pb y nº de registro ZP00005082.1).
- 45 Las SEC. ID. nº 15 y nº 16 presentan los cebadores utilizados para clonar la secuencia de D-alanina aminotransferasa (*dat*) de *B. subtilis*.
- 50 Las SEC. ID. nº 17 y nº 18 representan los cebadores utilizados para clonar la secuencia *tatA* de *S. meliloti*.
- 55 Las SEC. ID. nº 19 y nº 20 representan los cebadores utilizados para clonar la secuencia *araT* de aminotransferasa de *B. subtilis*.
- 60 Las SEC. ID. nº 21 y nº 22 representan los cebadores utilizados para clonar las secuencias de sustrato múltiple aminotransferasa de *Rhodobacter sphaeroides* (2.4.1 y 35053).
- 65 Las SEC. ID. nº 23 y nº 24 representan los cebadores utilizados para clonar la secuencia *bsat* de *Leishmania major*.
- Las SEC. ID. nº 25 y nº 26 representan los cebadores utilizados para clonar la secuencia *araT* de *Lactobacillus amylovorus*.
- Las SEC. ID. nº 27 y nº 28 representan los cebadores utilizados para clonar las secuencias *tatA* de *R. sphaeroides* (tanto la 2.4.1 como la 35053).
- Las SEC. ID. nº 29 y nº 30 representan los cebadores utilizados para clonar la secuencia *aspC* de *E. coli* (secuencia génica nº de registro en Genbank: AE000195.1, secuencia de proteínas nº de registro en Genbank: AAC74014.1).
- Las SEC. ID. nº 31 y nº 32 representan las secuencias de ácidos nucleicos y de aminoácidos de la aminotransferasa aromática (*tyrB*) de *E. coli*, respectivamente.
- Las SEC. ID. nº 33 y nº 34 representan los cebadores utilizados para clonar la secuencia *tyrB* de *E. coli*.

Las SEC. ID. nº 35 a 40 representan los cebadores utilizados para clonar polipéptidos con actividad de 4-hidroxi-2-oxoglutarato aldolasa (KHG) (CE 4.1.3.16).

Las SEC. ID. nº 41 y nº 42 representan las secuencias de ácidos nucleicos de triptofanasa (*trnA*) de *E. coli* y tirosina fenol-liasa (*tpf*) de *Citrobacter freundii*, que codifica las proteínas P00913 (GI: 401195) y P31013 (GI: 401201), respectivamente.

Las SEC. ID. nº 43 a 46 representan los cebadores utilizados para clonar polipéptidos de triptofanasa y de polipéptidos de β -tirosinasa (tirosina fenol-liasa).

Las SEC. ID. nº 47 a 54 representan los cebadores utilizados para mutar los polipéptidos de triptofanasa y polipéptidos de β -tirosinasa.

Las SEC. ID. nº 55 a 64 representan los cebadores utilizados para clonar los polipéptidos con actividad de 4-hidroxi-4-metil-2-oxoglutarato aldolasa (CE 4.1.3.17).

Las SEC. ID. nº 65 y nº 66 representan las secuencias de ácidos nucleicos y de aminoácidos de 4-hidroxi-4-metil-2-oxoglutarato aldolasa (*proA*) de *C. testosteroni*, respectivamente.

Las SEC. ID. nº 67 a 68 presentan los cebadores utilizados para clonar 4-hidroxi-4-metil-2-oxoglutarato aldolasa (*proA*) de *C. testosteroni* en un operón con *aspC* de *E. coli* en el pET30 Xa/LIC.

Las SEC. ID. nº 69 a 72 representan los cebadores utilizados para clonar *aspC* de *E. coli* y *proA* de *C. testosteroni* en pESC-his.

Las SEC. ID. nº 73 y nº 74 representan las secuencias añadidas al extremo 5' de los cebadores utilizados para clonar los genes dados a conocer en la presente memoria.

DESCRIPCIÓN DETALLADA DE VARIAS FORMAS DE REALIZACIÓN

Abreviaturas y términos

Para describir mejor la presente invención y para guiar a los expertos en la materia en la práctica de la presente invención se proporcionan las explicaciones siguientes de los términos y procedimientos. Tal como se utiliza en la presente memoria, "que comprende" significa que incluye y las formas singulares "un", "una" o "el" o "la" incluyen las referencias del plural a menos que el contexto lo indique claramente de otra manera. Por ejemplo, la referencia a "que comprende una proteína" incluye una o varias de dichas proteínas, y la referencia a "que comprende las células" incluye la referencia a una o más células y equivalentes de las mismas conocidas por los expertos en la materia y así sucesivamente.

A menos que se explique de otra manera, todos los términos técnicos y científicos utilizados en la presente memoria tienen el mismo significado normalmente entendido por cualquier experto en la materia a la que pertenece esta invención. Aunque los procedimientos o materiales similares o equivalentes a los descritos en la presente memoria pueden utilizarse en la práctica o experimentación de la presente invención, los procedimientos y materiales adecuados se describen a continuación. Los materiales, procedimientos y ejemplos se proporcionan únicamente a título de ejemplo limitativo. Otras propiedades y ventajas de la invención resultan evidentes en la siguiente descripción detallada y en las reivindicaciones.

ADNc (ADN complementario): Pieza de ADN que carece de segmentos internos, no codificadores (intrones) y secuencias reguladoras que determinan la transcripción. El ADNc puede sintetizarse en el laboratorio por transcripción inversa a partir del ARN mensajero extraído de las células.

Sustitución conservadora: Una o más sustituciones de aminoácido (por ejemplo 2, 5 ó 10 restos) para los restos de aminoácidos que tienen similares propiedades bioquímicas. Por lo general, las sustituciones conservadoras tienen poco a ningún impacto sobre la actividad del polipéptido resultante. Por ejemplo, teóricamente, un polipéptido de triptófano aminotransferasa que incluye una o más sustituciones conservadoras conserva actividad de triptófano aminotransferasa. Puede producirse un polipéptido que contenga una o más sustituciones conservadoras manipulando la secuencia nucleotídica que codifica este polipéptido utilizando, por ejemplo, procedimientos convencionales tales como la mutagénesis dirigida o la RCP.

Variantes de sustitución son aquellas en las que por lo menos un resto en la secuencia de aminoácidos se ha eliminado y un resto diferente se ha insertado en su lugar. Ejemplos de aminoácidos que pueden ser sustituidos por un aminoácido original en una proteína y que se consideran como sustituciones conservadoras incluyen: Ala sustituido con ser o thr; arg sustituido por gln, his o lys; asn sustituido por glu, gln, lys, his, asp; asp sustituido por asn, glu o gln; cys sustituido por ser o ala; gln sustituido por asn, glu, lys, his, asp o arg; glu sustituido con asn, gln, lys, o asp; gly sustituido por pro; his sustituido por asn, lys, gln, arg, tyr; ile sustituido por leu, met, val, phe; leu

sustituida por ile, met, val, phe; lys sustituida por asn, glu, gln, his, arg; met sustituida por ile, leu, val, phe; fe sustituida por trp, tyr, met, ile o leu; ser sustituida por thr, ala; thr sustituida por ser o ala; trp sustituida por fe, tyr; tyr sustituida por his, fe o trp; y val sustituida por met, ile y leu.

Más información acerca de las sustituciones conservadoras puede encontrarse, entre otros sitios en Ben-Bassat *et al.*, (*J. Bacteriol.*, 169:751-7, 1987), O'Reagan *et al.*, (*Gene* 77:237-51, 1989), Sahin-Toth *et al.*, (*Protein Sci.*, 3:240-7, 1994), Hochuli *et al.*, (*Bio/Technology*, 6:1321-5, 1988), documento WO 00/67796 (Curd *et al.*) y en los libros de textos normales de genética y de biología molecular.

Exógeno: El término "exógeno" tal como se lo utiliza en la presente memoria haciendo referencia al ácido nucleico y a una célula concreta se refiere a cualquier ácido nucleico que no se origine de esta célula específica como se encuentra en la naturaleza. De esta manera, el ácido nucleico artificial se considera que es exógeno a una célula una vez introducido en la célula. El ácido nucleico que es natural también puede ser exógeno para una célula específica. Por ejemplo, un cromosoma completo aislado de una célula de la persona X es un ácido nucleico exógeno con respecto a una célula de la persona Y una vez que el cromosoma se ha introducido en la célula de Y.

Funcionalmente equivalente: Que tiene una función equivalente. En el contexto de una enzima, las moléculas funcionalmente equivalentes incluyen diferentes moléculas que conservan la función de la enzima. Por ejemplo, pueden proporcionarse equivalentes funcionales mediante alteraciones de la secuencia en una secuencia enzimática, en la que el péptido con una o más alteraciones de la secuencia conserva una función del péptido inalterado, de modo que conserva su actividad enzimática. En un ejemplo específico, un equivalente funcional de triptófano aminotransferasa conserva la capacidad para convertir el triptófano en indol-3-piruvato.

Los ejemplos de alteraciones de la secuencia comprenden de manera no limitativa, sustituciones conservadoras, sustracciones, mutaciones, desplazamientos del marco e inserciones. En un ejemplo, un polipéptido dado se une a un anticuerpo, y un equivalente funcional es un polipéptido que se une al mismo anticuerpo. De este modo un equivalente funcional incluye péptidos que tienen la misma especificidad de enlace que un polipéptido, y que puede utilizarse como reactivo en lugar del polipéptido. En un ejemplo un equivalente funcional incluye un polipéptido en el que la secuencia de enlace es discontinua, en el que el anticuerpo se une a un epítipo lineal. Por lo tanto, si la secuencia del péptido es MPELANDLGL (aminoácidos 1 a 10 de la SEC. ID. nº 12) un equivalente funcional incluye epítopos discontinuos, que pueden aparecer de la manera siguiente (** = cualquier número de aminoácidos intervinientes): NH₂-**M**P**E**L**A**N**D**L**G**L-COOH. En este ejemplo, el polipéptido es funcionalmente equivalente a los aminoácidos 1 a 10 de la SEC. ID. nº 12 si la estructura tridimensional del polipéptido es tal que puede unirse a un anticuerpo monoclonal que se une a aminoácidos 1 a 10 de la SEC. ID. nº 12.

Hibridación: El término "hibridación" tal como se utiliza en la presente memoria, se refiere a un procedimiento de experimentación para complementariedad en la secuencia de nucleótidos de dos moléculas de ácido nucleico, basado en la capacidad del ADN monocatenario complementario y/o del ARN para formar una molécula de doble cadena. Pueden utilizarse técnicas de hibridación del ácido nucleico para obtener un ácido nucleico aislado dentro del alcance de la invención. En resumen, cualquier ácido nucleico que tenga alguna homología con una secuencia indicada en SEC. ID. nº 11 puede utilizarse como sonda para identificar un ácido nucleico similar por hibridación en condiciones de moderada a alta severidad. Una vez identificado, el ácido nucleico luego puede purificarse, secuenciarse y analizarse para determinar si está comprendida dentro del alcance de la presente invención.

La hibridación puede realizarse por análisis Southern o Northern para identificar una secuencia de ADN o de ARN, respectivamente, que se hibrida a una sonda. La sonda puede marcarse con biotina, digoxigenina, un polipéptido o un radioisótopo tal como ³²P. El ADN o ARN que va a analizarse puede separarse por electroforesis en un gel de agarosa o de poliacrilamida, transferirse a nitrocelulosa, nilón u otra membrana adecuada e hibridarse con la sonda utilizando técnicas convencionales bien conocidas en la materia tales como las descritas en los apartados 7.39 a 7.52 de Sambrook *et al.*, (1989) Molecular Cloning, 2ª edición, Cold Spring Harbor Laboratory, Plainview, NY. Por lo general, una sonda tiene por lo menos aproximadamente 20 nucleótidos de longitud. Por ejemplo una sonda que corresponde a una secuencia de 20 nucleótidos contiguos indicada en SEC. ID. nº 11 puede utilizarse para identificar un ácido nucleico idéntico o similar. Además, pueden utilizarse sondas más largas o más cortas de 20 nucleótidos.

La invención proporciona además secuencias aisladas de ácido nucleico que son por lo menos de 12 bases de longitud (por ejemplo, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50, 60, 100, 250, 500, 750, 1000, 1500, 2000, 3000, 4000 ó 5000 bases de longitud) y se hibridan, en condiciones de hibridación, a la cadena transcrita o complementaria de un ácido nucleico que tiene la secuencia indicada en la SEC. ID. nº 11. Las condiciones de hibridación pueden ser condiciones de hibridación moderadamente o muy severas.

En aras de la presente invención, condiciones de hibridación moderadamente severas significa que la hibridación se realiza a aproximadamente 42°C en una solución de hibridación que contiene KPO₄ 25 mM (pH 7,4), 5 x SSC, 5 x solución de Denhart, 50 µg/ml de ADN de esperma de salmón desnaturalizado y sometido a ultrasonidos, formamida al 50%, sulfato de dextrano al 10% y 1 a 15 ng/ml de sonda (aproximadamente 5 x

10^7 cpm/ μ g), mientras que los lavados se llevan a cabo aproximadamente a 50°C con una solución de lavado que contiene 2 x SSC y dodecilsulfato sódico al 0,1%.

Condiciones de hibridación muy severas significa que la hibridación se realiza a aproximadamente 42°C en una solución de hibridación que contiene KPO_4 25 mM (pH 7,4), 5 x SSC, 5 x solución de Denhart, 50 μ g/ml de ADN de esperma de salmón desnaturalizado y sometido a ultrasonidos, formamida al 50%, sulfato de dextrano al 10% y 1 a 15 ng/ml de sonda (aproximadamente 5×10^7 cpm/ μ g), mientras que los lavados se llevan a cabo aproximadamente a 65°C con una solución de lavado que contiene 0,2 x SSC y dodecilsulfato sódico al 0,1%.

Aislado: El término "aislado" tal como se utiliza en la presente memoria, haciendo referencia a un ácido nucleico se refiere a un ácido nucleico natural que no está inmediatamente contiguo a ambas secuencias con las que está inmediatamente contiguo (una en el extremo 5' y la otra en el extremo 3') en el genoma natural del organismo del que procede. Por ejemplo, un ácido nucleico puede ser, sin limitación, una molécula de ADN recombinante de cualquier longitud, con tal que una de las secuencias de ácido nucleico que se encuentra normalmente flanqueando inmediatamente esta molécula de ADN recombinante en un genoma natural se elimine o falte. Por lo tanto, un ácido nucleico aislado comprende de manera no limitativa, un ADN recombinante que existe como molécula independiente (por ejemplo, un ADNc o un fragmento de ADN genómico producido por RCP o un tratamiento de endonucleasa de restricción) independiente de otras secuencias así como un ADN recombinante que esté incorporado en un vector, un plásmido que se replica de forma autónoma, un virus (por ejemplo, un retrovirus, adenovirus o virus herpético) o en el ADN genómico de un procarionte o eucarionte. Además, un ácido nucleico aislado puede incluir una molécula de ADN recombinante que forma parte de un híbrido o de una secuencia de ácido nucleico de fusión.

El término "aislado" tal como se utiliza en la presente memoria con relación al ácido nucleico incluye también cualquier ácido nucleico artificial ya que las secuencias de ácido nucleico artificiales no se encuentran en la naturaleza y no tienen secuencias inmediatamente contiguas en un genoma natural. Por ejemplo, el ácido nucleico artificial tal como un ácido nucleico modificado genéticamente se considera que es un ácido nucleico aislado. El ácido nucleico modificado genéticamente puede utilizar técnicas de clonación molecular corriente o de síntesis química de ácido nucleico. El ácido nucleico artificial aislado puede existir independiente de otras secuencias, o incorporado en un vector, un plásmido que se replica de manera autónoma, un virus (por ejemplo, un retrovirus, adenovirus o virus herpético) o el ADN genómico de un procarionte o eucarionte. Además, el ácido nucleico artificial puede incluir una molécula de ácido nucleico que forma parte de una secuencia de ácido nucleico híbrido o de fusión.

Para los expertos en la materia resultará evidente que un ácido nucleico que existe entre centenares a millones de otras moléculas de ácido nucleico en, por ejemplo, bancos de ADNc o genómicos, o en rodajas de gel que contienen una enzima de restricción de ADN genómico no se considera un ácido nucleico aislado.

Ácido nucleico: La expresión "ácido nucleico" tal como se utiliza en la presente memoria, comprende tanto el ARN como el ADN comprendiendo de manera no limitativa, ADNc, ADN genómico y ADN sintético (por ejemplo, sintetizado químicamente). El ácido nucleico puede ser bicatenario o monocatenario. Cuando es monocatenario, el ácido nucleico puede ser la cadena transcrita o la cadena complementaria. Además, el ácido nucleico puede ser circular o lineal.

Operativamente unida: Una primera secuencia de ácido nucleico está "operativamente unida" con una segunda secuencia de ácido nucleico siempre que la primera secuencia de ácido nucleico esté colocada en una relación funcional con la segunda secuencia de ácido nucleico. Por ejemplo, un activador está operativamente unido a una secuencia de codificación si el activador afecta la transcripción de la secuencia de codificación. Generalmente, las secuencias de ADN operativamente unidas son contiguas y, cuando es necesario unir dos zonas de codificación del polipéptido, en el mismo marco de lectura.

Modificaciones de péptidos: La presente invención incluye enzimas, así como formas de realización sintéticas de las mismas. Además, los análogos (moléculas orgánicas no peptídicas), derivados (moléculas peptídicas químicamente funcionalizadas que comienzan con las secuencias peptídicas dadas a conocer) y las variantes (homólogos) con la actividad enzimática deseada pueden utilizarse en los procedimientos descritos en la presente memoria. Los péptidos dados a conocer en la presente memoria incluyen una secuencia de aminoácidos, que pueden ser L- y/o D- aminoácidos, naturales y distintos.

Los péptidos pueden modificarse mediante varias técnicas químicas para producir derivados que tienen esencialmente la misma actividad que los péptidos inalterados, y opcionalmente tienen otras propiedades deseables, por ejemplo, grupos de ácido carboxílico de la proteína, si el terminal carboxilo o la cadena lateral, pueden suministrarse en forma de sal de un catión farmacéuticamente aceptable o esterificarse para formar un éster C1-C16, o convertirse en una amida de fórmula NR_1R_2 , en la que R1 y R2 son cada una independientemente H o alquilo C1-C16 o combinarse para formar un anillo heterocíclico, tal como un anillo de 5 ó 6 elementos. Los grupos amino del péptido, ya sean de cadena con terminal amino o lateral, pueden estar en forma de una sal de adición de ácido farmacéuticamente aceptable, tal como el HCl, HBr, acético, benzoico, toluensulfónico, maleico,

tartático y otras sales orgánicas, o puede modificarse a alquilo C1-C16 o dialquilamino o transformarse más hasta una amida.

Los grupos hidroxilo de las cadenas peptídicas laterales pueden convertirse en alcohoxi C1-C16 o en éster C1-C16 utilizando técnicas muy reconocidas. Los anillos de fenilo y fenólico de las cadenas laterales del péptido pueden estar sustituidas con uno o más átomos de halógeno, tales como F, Cl, Br o I, o con alquilo C1-C16, alcohoxi C1-C16, ácidos carboxílicos y ésteres de los mismos o amidas de dichos ácidos carboxílicos. Los grupos metileno de las cadenas laterales del péptido pueden prolongarse a alquilenos C2-C4 homólogos. Los tioles pueden protegerse con cualquiera de los numerosos grupos protectores bien reconocidos tales como los grupos acetamida. Los expertos en la materia reconocerán también los métodos para introducir estructuras cíclicas en los péptidos de la presente invención para seleccionar y proporcionar limitaciones de configuración a la estructura que dan como resultado un aumento de estabilidad. Por ejemplo, una cisteína con terminal C o N puede añadirse al péptido, de modo que cuando se oxide el péptido contenga un enlace disulfuro, generando un péptido cíclico. Otros procedimientos de ciclación de péptidos incluyen la formación de tioéteres y de amidas y ésteres con terminal carboxilo y amino.

Las formas de realización peptidomiméticas y organomiméticas están asimismo comprendidas dentro del alcance de la presente invención, por lo que la disposición tridimensional de los constituyentes químicos de dicho péptido- y organomiméticos simula la disposición tridimensional del eje central del péptido y de las cadenas laterales del aminoácido componente, dando como resultado dichos péptido- y organomiméticos de las proteínas de la presente invención con actividad enzimática detectable. Para aplicaciones de modelado por ordenador, un farmacóforo es una definición tridimensional idealizada de los requisitos estructurales para la actividad biológica. Pueden diseñarse péptido- y organomiméticos para ajustar cada farmacóforo con el programa informático de modelado por ordenador actual (utilizando el diseño de fármacos asistidos por ordenador o CADD). Véase Walters, "Computer-Assisted Modeling of Drugs", en Klegerman & Groves (ediciones), Pharmaceutical Biotechnology, 1993, Interpharm Press: Buffalo Grove, IL, págs. 165-174 y Cap. 102 en Munson (edición), Principles of Pharmacology, 1995, Chapman & Hall, para descripciones y técnicas utilizadas en CADD. Están incluidos también dentro del alcance de la invención los miméticos preparados utilizando dichas técnicas. En un ejemplo, un mimético simula la actividad enzimática generada por una enzima o una variante, fragmento o fusión de la misma.

Sondas y cebadores: Las sondas y cebadores de ácido nucleico pueden prepararse fácilmente basándose en las secuencias de aminoácidos y en las secuencias de ácidos nucleicos proporcionadas en la presente memoria. Una "sonda" incluye un ácido nucleico aislado que contiene un marcador detectable o una molécula indicadora. Los marcadores ejemplificativos comprenden de manera no limitativa, isótopos radioactivos, ligandos, agentes quimioluminiscentes y polipéptidos. Los procedimientos de marcado y de dirección en la selección de marcadores apropiados para varios propósitos se exponen, por ejemplo, en Sambrook *et al.*, (ed.), Molecular Cloning: A Laboratory Manual, 2ª ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY., 1989 y Ausubel *et al.*, (ed.) Current Protocols in Molecular Biology, Greene Publishing y Wiley-Interscience, Nueva York (con actualizaciones periódicas), 1987.

Los "cebadores" son por lo general moléculas de ácido nucleico que tienen diez o más nucleótidos (por ejemplo, moléculas de ácido nucleico que tienen entre aproximadamente 10 nucleótidos y aproximadamente 100 nucleótidos). Un cebador puede hibridarse con una cadena de ácido nucleico diana complementario por hibridación del ácido nucleico para formar un híbrido entre el cebador y la cadena de ácido nucleico diana, y a continuación prolongarse a lo largo de la cadena de ácido nucleico diana mediante, por ejemplo, un polipéptido de ADN polimerasa. Los pares de cebadores pueden utilizarse para la ampliación de una secuencia de ácido nucleico, por ejemplo, mediante la reacción en cadena de la polimerasa (RCP) u otros procedimientos de ampliación de ácidos nucleicos conocidos en la técnica.

Los procedimientos de preparación y utilización de sondas y cebadores se describen, por ejemplo, en referencias tales como Sambrook *et al.*, (ed.), Molecular Cloning: A Laboratory Manual, 2ª ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY., 1989 y Ausubel *et al.*, (ed.) Current Protocols in Molecular Biology, Greene Publishing y Wiley-Interscience, Nueva York (con actualizaciones periódicas), 1987; e Innis *et al.*, (ediciones), PCR Protocols: A Guide to Methods and Applications, Academic Press: San Diego, 1990. Los pares de cebador de RCP pueden proceder de una secuencia conocida, por ejemplo, utilizando programas informáticos pensados con esta finalidad tales como Primer (Versión 0,5, © 1991, Whitehead Institute for Biomedical Research, Cambridge, Mass.). Un experto en la materia valorará que la especificidad de una sonda o cebador concreto aumente con la longitud, pero que una sonda o cebador pueda variar de tamaño desde una secuencia completa a secuencias tan cortas como de cinco nucleótidos consecutivos. Así, por ejemplo, un cebador de 20 nucleótidos consecutivos puede hibridarse a una diana con una especificidad mayor que un cebador correspondiente de sólo 15 nucleótidos. Así, con objeto de obtener mayor especificidad, pueden seleccionarse sondas y cebadores que comprenden, por ejemplo, 10, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750, 1800, 1850, 1900, 2000, 2050, 2150, 2200, 2250, 2300, 2350, 2400, 2450, 2500, 2550, 2600, 2650, 2700, 2750, 2800, 2850, 2900, 3000, 3050, 3100, 3150, 3200, 3250, 3300, 3350, 3400, 3450, 3500, 3550, 3600, 3650, 3700, 3750, 3800, 3850, 3900, 4000, 4050, 4100, 4150, 4200, 4250, 4300, 4350, 4400, 4450, 4500, 4550, 4600, 4650, 4700, 4750, 4800, 4850, 4900, 5000, 5050, 5100, 5150, 5200, 5250, 5300, 5350,

5400, 5450, o más nucleótidos consecutivos.

Activador: Matriz de secuencias de control de ácido nucleico que dirige la transcripción de un ácido nucleico. Un activador incluye las secuencias necesarias de ácido nucleico próximas al punto de partida de la transcripción, tales como, en el caso de un activador tipo II de polimerasa, un elemento TATA. Un activador puede incluir elementos distales de potenciador o represor que pueden estar situados hasta varios pares de miles de bases desde el punto de partida de la transcripción.

Purificado: El término "purificado" tal como se utiliza en la presente invención, no requiere pureza absoluta; mejor dicho, pretende ser un término relativo. Así, por ejemplo, una preparación purificada de polipéptido o ácido nucleico puede ser aquella en la que el polipéptido o ácido nucleico en cuestión, respectivamente, está a una concentración mayor que la que estaría en su medio natural el polipéptido o ácido nucleico en un organismo. Por ejemplo, una preparación de polipéptido puede considerarse purificada si el contenido en polipéptido en la preparación representa por lo menos el 50%, 60%, 70%, 80%, 85%, 90%, 92%, 95%, 98% o 99% del contenido en proteína total de la preparación.

Recombinante: Un ácido nucleico "recombinante" es el que tiene (1) una secuencia que no es natural en el organismo en el que se expresa o (2) una secuencia hecha por una combinación artificial de dos secuencias diferentes separadas, más cortas. Esta combinación artificial se realiza con frecuencia por síntesis química o, más frecuentemente, por manipulación artificial de los segmentos aislados de ácidos nucleicos, por ejemplo, por técnicas de ingeniería genética. "Recombinante" se utiliza también para describir moléculas de ácidos nucleicos que han sido manipuladas artificialmente, pero contienen las mismas secuencias reguladoras y las zonas de codificación que se encuentran en el organismo en el que se aisló el ácido nucleico.

Identidad de secuencia: La similitud entre las secuencias de aminoácidos se expresa en términos de similitud entre las secuencias, denominada de otra manera identidad de secuencia. La identidad de secuencia se mide frecuentemente en términos de porcentaje de identidad (o similitud u homología); cuanto mayor es el porcentaje, más similares son las dos secuencias. Los homólogos o variantes de un péptido, tales como la SEC. ID. nº 12, posee un grado relativamente alto de identidad de secuencia cuando se alinea utilizando métodos normalizados.

Los métodos de alineación de las secuencias para comparación son bien conocidos en la técnica. Varios programas y algoritmos de alineación se describen en: Smith y Waterman, *Adv. Appl. Math.*, 2:482, 1981; Needleman y Wunsch, *J. Mol. Biol.*, 48:443-53, 1970; Pearson y Lipman, *Proc. Natl. Acad. Sci. USA.*, 85:2444-8, 1988; Higgins y Sharp, *Gene*, 73:237-44, 1988; Higgins y Sharp, *CABIOS*, 5:151-3, 1989; Corpet *et al.*, *Nucleic Acids Research*, 16:10881-90; y Altschul *et al.*, *Nature Genet.*, 6:119-29, 1994.

La herramienta básica local de búsqueda de alineación local del NCBI (Basic Local Alignment Search Tool) (BLAST™) (Altschul *et al.*, *J. Mol. Biol.*, 215:403-10, 1990) está disponible en varios proveedores, incluyendo el National Center for Biotechnology Information (NCBI, Bethesda, MD) y en Internet para utilización en relación con los programas de análisis de secuencias blastp, blastn, blastx, tblastn y tblastx.

Las variantes de un péptido están caracterizadas por lo general por la posesión de al menos el 50% de identidad de secuencia contada sobre la alineación completa con la secuencia de aminoácidos que utiliza el Blast 2.0 del NCBI, blastp incompleto ajustado para parámetros por defecto. Para las comparaciones de secuencias de aminoácidos mayores de aproximadamente 30 aminoácidos, se emplea la función de secuencias Blast 2 utilizando la matriz BLOSUM62 por defecto ajustada para parámetros por defecto (coste por existencia de hueco de 11 y un coste de 1 por hueco residual). Cuando se alinean péptidos cortos (menores de aproximadamente 30 aminoácidos), la alineación se realiza utilizando la función de las secuencias Blast 2, utilizando la matriz PAM30 ajustada para parámetros por defecto (penalizaciones por hueco abierto 9, por ampliación de hueco 1). Las proteínas con aún mayor similitud a las secuencias de referencia presentarán un aumento del porcentaje de identidades cuando se evalúan por este procedimiento, tal como por lo menos del 80%, por lo menos del 90%, por lo menos del 95%, por lo menos del 98% o incluso por lo menos del 99% de identidad de secuencias. Cuando se está comparando la identidad de secuencias en menos de la secuencia completa, los homólogos y variantes por lo general poseerán por lo menos 80% de identidad de secuencias en intervalos cortos de 10 a 20 aminoácidos, y pueden poseer identidades de secuencias de por lo menos 85%, por lo menos del 90%, por lo menos del 95% o 98% dependiendo de su similitud con la secuencia de referencia. Los métodos para determinar la identidad de secuencias en dichos intervalos cortos se describen en la página web que está mantenida por el National Center for Biotechnology Information in Bethesda, Maryland. Un experto en la materia apreciará que estos intervalos de identidad de secuencias se proporcionan únicamente como orientación; es completamente posible que pudieran obtenerse homólogos muy significativos que no estén comprendidos en los intervalos proporcionados.

Pueden utilizarse procedimientos similares para determinar el porcentaje de identidad de secuencia de una secuencia de ácido nucleico. En un ejemplo específico, una secuencia homóloga se alinea a una secuencia natural, y el número de emparejamientos se determina haciendo el recuento del número de posiciones en las que un nucleótido idéntico o un resto de aminoácido está presente en ambas secuencias. El porcentaje de identidad de

secuencia se determina dividiendo el número de emparejamientos ya sea por la longitud de la secuencia indicada en la secuencia identificada (por ejemplo, SEC. ID. nº 11) o por una longitud articulada (por ejemplo, 100 nucleótidos consecutivos o restos de aminoácidos procedentes de una secuencia indicada en una secuencia identificada), seguido de multiplicar el valor resultante por 100. Por ejemplo, una secuencia de ácido nucleico que tiene 1166 emparejamientos cuando se alinea con la secuencia indicada en la SEC. ID. nº 11, es el 75,0 por ciento idéntica a la secuencia indicada en la SEC. ID. nº 11 (es decir, $1166 \div 1554 \times 100 = 75,0$). Se observa que el valor de la identidad de la secuencia por ciento está redondeado a la décima más próxima. Por ejemplo, 75,11, 75,12, 75,13 y 75,14 se redondean a la baja a 75,1 en tanto que 75,15, 75,16, 75,17, 75,18 y 75,19 se redondean a la alta a 75,2. Se indica también que el valor de la longitud será siempre un número entero. En otro ejemplo, una secuencia diana que contiene una zona de 20 nucleótidos que se alinea con 20 nucleótidos consecutivos de una secuencia idéntica como la siguiente contiene una zona que comparte el 75 por ciento de identidad de secuencia con esta secuencia identificada (es decir, $15 \div 20 \times 100 = 75$).

	1	20
Secuencia diana:	AGGTCGTGTACTGTCAGTCA	
Secuencia identificada:	ACGTGGTGAAGTCCAGTGA	

Agente aglutinante específico: Agente que puede unirse específicamente a alguno de los polipéptidos descritos en la presente memoria. Los ejemplos comprenden de manera no limitativa anticuerpos policlonales, anticuerpos monoclonales (incluyendo los anticuerpos monoclonales humanizados) y fragmentos de anticuerpos monoclonales tales como los fragmentos Fab, F(ab')₂ y Fv también como cualquier otro agente capaz de unirse específicamente a un epítipo de dichos polipéptidos.

Los anticuerpos contra los polipéptidos suministrados en la presente memoria (o fragmentos, variantes o fusiones de los mismos) pueden utilizarse para purificar o identificar dichos polipéptidos. Las secuencias de aminoácidos y de ácidos nucleicos suministrados en la presente memoria permiten la producción de agentes aglutinantes a base de anticuerpos específicos que reconocen los polipéptidos descritos en la presente memoria.

Pueden producirse anticuerpos monoclonales o policlonales contra los polipéptidos, fragmentos de los polipéptidos o variantes de los mismos. De manera óptima, los anticuerpos producidos contra uno o más epítopos en un antígeno de polipéptido detectarán específicamente este polipéptido, es decir, los anticuerpos producidos contra un polipéptido específico reconocerán y se unirán a este polipéptido específico, y no reconocerán sustancialmente ni se unirán a otros polipéptidos. La determinación de que un anticuerpo se une específicamente a un polipéptido específico se hace mediante alguno de los numerosos procedimientos de inmunoanálisis habituales; por ejemplo inmunotransferencia Western (véase, por ejemplo, Sambrook *et al.*, (ed.), Molecular Cloning: A Laboratory Manual, 2ª ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY., 1989).

Para determinar que una preparación de anticuerpos dada (tal como una preparación producida en un ratón contra un polipéptido que tiene la secuencia de aminoácidos indicada en la SEC. ID. nº 12) detecta específicamente el polipéptido apropiado (por ejemplo, un polipéptido que tiene la secuencia de aminoácidos indicada en la SEC. ID. nº 12) por inmunotransferencia Western, puede extraerse de las células toda la proteína celular y separarse por SDS-electroforesis en gel de poliacrilamida.

Toda la proteína celular separada puede transferirse a continuación a una membrana (por ejemplo, de nitrocelulosa) e incubarse la preparación de anticuerpo con la membrana. Después de lavar la membrana para eliminar los anticuerpos unidos de manera no específica, puede detectarse la presencia de anticuerpos unidos específicamente utilizando un anticuerpo secundario apropiado (por ejemplo, un anticuerpo antirratón) conjugado con un polipéptido tal como fosfatasa alcalina ya que la aplicación de 5-bromo-4-cloro-3-indolil fosfato/nitroazul de tetrazolio da como resultado la producción de un compuesto coloreado densamente de azul por la fosfatasa alcalina inmunolocalizada.

A partir de células transfectadas, células transformadas o células naturales pueden obtenerse polipéptidos sustancialmente puros, adecuados para su utilización como inmunógenos. Las concentraciones de polipéptidos en la preparación final pueden ajustarse, por ejemplo, por concentración en un dispositivo con filtro Amicon, al nivel de pocos microgramos por mililitro. Además, pueden utilizarse como inmunógenos los polipéptidos de tamaño comprendido entre polipéptidos completos y polipéptidos que presentan solamente nueve restos de aminoácidos. Dichos polipéptidos pueden producirse en cultivo celular, pueden sintetizarse químicamente utilizando procedimientos convencionales o pueden obtenerse cortando grandes polipéptidos en polipéptidos más pequeños que pueden purificarse. Los polipéptidos que presentan solamente nueve restos de aminoácidos de longitud pueden ser inmunógenos cuando se presentan a un sistema inmunitario en el contexto de una molécula del complejo de histocompatibilidad mayor (CHM) tal como una molécula del CHM de clase I o CHM de clase II. Por consiguiente, los polipéptidos que tienen por lo menos 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 900, 1000, 1050, 1100, 1150,

1200, 1250, 1300, 1350 o más restos de aminoácidos consecutivos de cualquier secuencia de aminoácido dada a conocer en la presente memoria pueden utilizarse como inmunógenos para producir anticuerpos.

Pueden prepararse anticuerpos monoclonales contra algunos de los polipéptidos dados a conocer en la presente memoria a partir de hibridomas murinos según el procedimiento clásico de Kohler y Milstein (*Nature* 256:495-7, 1975) o un procedimiento modificado del mismo.

El antisuero policlonal que contiene anticuerpos contra los epítomos heterogéneos de cualquier polipéptido dado a conocer en la presente memoria puede prepararse inmunizando animales adecuados con el polipéptido (o un fragmento del mismo), que puede estar inalterado o modificado para aumentar la inmunogenicidad. Un protocolo de inmunización eficaz para conejos puede encontrarse en Vaitukaitis *et al.* (*J. Clin. Endocrinol. Metab.*, 33:988-91, 1971).

Pueden utilizarse fragmentos de anticuerpo en lugar de anticuerpos completos y pueden expresarse fácilmente en células hospedadoras procarióticas. Los procedimientos para preparar y utilizar fracciones inmunológicamente eficaces de anticuerpos monoclonales, denominados también "fragmentos de anticuerpos", son muy conocidos e incluyen los descritos en Better & Horowitz (*Methods Enzymol.*, 178:476-96, 1989), Glockshuber *et al.*, (*Biochemistry*, 29:1362-7, 1990), patente US nº 5.648.237 ("Expression of Functional Antibody Fragments"), patente US nº 4.946.778 ("Single Polypeptide Chain Binding Molecules"), patente US nº 5.455.030 ("Immunotherapy Using Single Chain Polypeptide Binding Molecules") y las referencias citadas en éstas.

Transformada: Una célula "transformada" es una célula en la que se introducido una molécula de ácido nucleico, por ejemplo, por técnicas de biología molecular. La transformación comprende todas las técnicas mediante las cuales puede introducirse una molécula de ácido nucleico en dicha célula incluyendo, sin limitación, transfección con un vector vírico, conjugación, transformación con un vector plásmido e introducción del ADN desnudo por electroporación, lipofección y aceleración con cañón de partículas.

Variantes, fragmentos o proteínas de fusión: Las proteínas dadas a conocer, incluyen variantes, fragmentos y fusiones de las mismas. Las secuencias de ADN (por ejemplo la SEC. ID nº 11) que codifica para una proteína (por ejemplo la SEC. ID. nº 12), la proteína de fusión, o un fragmento o variante de una proteína, puede modificarse genéticamente para permitir que la proteína se exprese en células eucarióticas, bacterias, insectos y/o plantas. Para obtener la expresión, la secuencia de ADN puede alterarse y unirse operativamente a otras secuencias reguladoras. El producto final, que contiene las secuencias reguladoras y la proteína, se denomina vector. Este vector puede introducirse en células eucarióticas, bacterias, de insecto y/o vegetales. Una vez dentro de la célula el vector permite producir la proteína.

Una proteína de fusión que incluye una proteína, tal como una triptófano aminotransferasa (o variante, poliformismo, mutante o fragmento de la misma), por ejemplo la SEC. ID. nº 12, unida a otras secuencias de aminoácidos que no inhiben la actividad deseada de la proteína, por ejemplo la capacidad para convertir el triptófano en indol-3-piruvato. En un ejemplo, las demás secuencias de aminoácidos no son más de aproximadamente 10, 12, 15, 20, 25, 30 ó 50 aminoácidos de longitud.

El experto en la materia valorará que una secuencia de ADN pueda alterarse de numerosas maneras sin afectar la actividad biológica de la proteína codificada. Por ejemplo, puede utilizarse la RCP para producir variaciones en la secuencia del ADN que codifica una proteína. Dichas variantes pueden ser variantes optimizadas por preferencia de codón en una célula hospedadora utilizada para expresar la proteína, u otros cambios de frecuencia que facilitan la expresión.

Vector: Molécula de ácido nucleico introducida en una célula, que produce de este modo una célula transformada. Un vector puede incluir secuencias de ácido nucleico que le permiten replicarse en la célula, tal como un origen de replicación. Un vector puede incluir uno o más genes marcadores seleccionables y otros elementos genéticos conocidos en la técnica.

Resumen de las rutas biosintéticas

Como se muestra en las figuras 1 a 3 y 11 a 13, pueden utilizarse muchas rutas biosintéticas para producir monatina o sus productos intermedios tales como indol-3-piruvato o PM. Para la conversión de cada sustrato (glucosa, triptófano, ácido indol-3-láctico, indol-3-piruvato y PM) a cada producto (triptófano, indol-3-piruvato, PM y monatina) pueden utilizarse varios polipéptidos diferentes. Además, estas reacciones pueden realizarse *in vivo*, *in vitro* o mediante de una combinación de reacciones *in vivo* y reacciones *in vitro* tales como las reacciones *in vitro* que incluyen reacciones químicas no enzimáticas. Por consiguiente, las figuras 1 a 3 y 11 a 13 se proporcionan a título de ejemplo y muestran múltiples rutas diferentes que pueden utilizarse para obtener los productos deseados.

Glucosa en triptófano

Muchos organismos pueden sintetizar triptófano a partir de glucosa. El(los) montaje(s) que contiene(n)

el(los) gen(es) necesario(s) para producir monatina, PM y/o indol-3-piruvato a partir de glucosa y/o triptófano pueden clonarse en dichos organismos. Se muestra en la presente que el triptófano puede convertirse en monatina.

En otros ejemplos, se modifica genéticamente un organismo utilizando conocidos polipéptidos para producir triptófano o sobreproducir triptófano. Por ejemplo, la patente US nº 4.371.614 describe una cepa de *E. coli* transformada con un plásmido que contiene un operón de triptófano natural.

Los valores máximos de triptófano dados a conocer en la patente US nº 4.371.614 son de aproximadamente 230 ppm. Asimismo, el documento WO 8701130 describe una cepa de *E. coli* que ha sido modificada genéticamente para producir triptófano y expone el aumento de producción fermentativa de L-triptófano. Los expertos en la materia reconocerán que los organismos capaces de producir triptófano a partir de glucosa son también capaces de utilizar otras fuentes de carbono y energía que pueden convertirse en glucosa o fructosa-6-fosfato, con similares resultados. Las fuentes de carbono y energía ejemplificativas comprenden de manera no limitativa, sacarosa, fructosa, almidón, celulosa o glicerol.

Triptófano en indol-3-piruvato

Pueden utilizarse varios polipéptidos para convertir triptófano en indol-3-piruvato. Los péptidos ejemplificativos incluyen elementos de las clases enzimáticas (CE) 2.6.1.27, 1.4.1.19, 1.4.99.1, 2.6.1.28, 1.4.3.2, 1.4.3.3, 2.6.1.5, 2.6.1.-, 2.6.1.1, y 2.6.1.21. Estas clases incluyen polipéptidos denominados triptófano aminotransferasa (denominados también L-fenilalanina-2-oxoglutarato aminotransferasa, triptófano transaminasa, 5-hidroxitriptófano-cetaglutarico transaminasa, hidroxitriptófano aminotransferasa, L-triptófano aminotransferasa, L-triptófano transaminasa y L-triptófano-2-oxoglutarato aminotransferasa) que convierten el L-triptófano y 2-oxoglutarato en indol-3-piruvato y L-glutamato; D-triptófano aminotransferasa que convierte D-triptófano y un 2-oxoácido en indol-3-piruvato y un aminoácido; triptófano deshidrogenasa (denominado también NAD(P)-L-triptófano deshidrogenasa, L-triptófano deshidrogenasa, L-Trp-deshidrogenasa, TDH y L-triptófano:NAD(P) óxidorreductasa (desaminante)) que convierte L-triptófano y NAD(P) en indol-3-piruvato, NH₃ y NAD(P)H; D-aminoácido deshidrogenasa, que convierte D-aminoácidos y FAD en indol-3-piruvato, NH₃ y FADH₂; triptófano-fenilpiruvato transaminasa (denominada también L-triptófano- α -cetoisocaproato aminotransferasa y L-triptófano:fenilpiruvato aminotransferasa) que convierte L-triptófano y fenilpiruvato en indol-3-piruvato y L-fenilalanina; L-aminoácido oxidasa (denominada también ofio-aminoácido oxidasa y L-aminoácido:oxígeno oxidorreductasa (desaminante)) que convierte un L-aminoácido y H₂O y O₂ en un 2-oxoácido, NH₃ y H₂O₂; D-aminoácido oxidasa (denominada también ofio-aminoácido oxidasa y D-aminoácido:oxígeno oxidorreductasa desaminante)) que convierte una D-aminoácido, H₂O y O₂ en un 2-oxoácido, NH₃ y H₂O₂; y triptófano oxidasa que convierte L-triptófano y H₂O y O₂ en indol-3-piruvato y NH₃ y H₂O₂. Estas clases también contienen tirosina (aromática) aminotransferasa, aspartato aminotransferasa, D-aminoácido (o D-alanina) aminotransferasa y aminotransferasa amplia (sustrato múltiple) que tienen múltiples actividades de aminotransferasa, algunas de las cuales pueden convertir el triptófano y un 2-oxoácido en indol-3-piruvato y un aminoácido.

En el Ejemplo 1 se describen a continuación once miembros de la clase aminotransferasa que presentan dicha actividad, incluyendo una nueva aminotransferasa mostrada en las SEC. ID. nº 11 y nº 12. Por consiguiente, la presente invención proporciona secuencias de ácido nucleico y de proteínas aisladas que tienen por lo menos 80%, por lo menos 85%, por lo menos 90%, por lo menos 95% y por lo menos 98% o incluso por lo menos 99% de identidad de secuencias con las SEC. ID. nº 11 y nº 12. También están comprendidos en la presente invención los fragmentos y las fusiones de las SEC. ID. nº 11 y nº 12 que conservan actividad de aminotransferasa o codifican una proteína que tiene actividad de aminotransferasa. A título de ejemplo los fragmentos comprenden de manera no limitativa por lo menos 10, 12, 15, 20, 25, 50, 100, 200, 500 ó 1000 nucleótidos contiguos de la SEC. ID. nº 11 o por lo menos 6, 10, 15, 20, 25, 50, 75, 100, 200, 300 ó 350 aminoácidos contiguos de la SEC. ID. nº 12. Las secuencias dadas a conocer (y las variantes, fragmentos y fusiones de las mismas) pueden formar parte de un vector. El vector puede utilizarse para transformar células hospedadoras, produciendo de este modo células recombinantes que pueden producir indol-3-piruvato a partir del triptófano, y en algunos ejemplos pueden producir además PM y/o monatina.

Se conocen las L-aminoácido oxidasas (1.4.3.2), y las secuencias pueden aislarse de varias fuentes diferentes, tales como *Vipera lebetine* (sp P81375), *Ophiophagus hannah* (sp P81383), *Agkistrodon rhodostoma* (sp P81382), *Crotalus atrox* (sp P56742), *Burkholderia cepacia*, *Arabidopsis thaliana*, *Caulobacter crescentus*, *Chlamydomonas reinhardtii*, *Mus musculus*, *Pseudomonas syringae* y *Rhodococcus* str. Además, se describen las triptófano oxidasas en la bibliografía y pueden aislarse, por ejemplo, de *Coprinus* sp. SF-1, col china con hernia de la col, *Arabidopsis thaliana* y de hígado de mamífero. En el Ejemplo 3 se expone a continuación un miembro de la clase L-aminoácido de la clase oxidasa que puede convertir el triptófano en indol-3-piruvato, así como fuentes alternativas para la clonación molecular. Muchos genes de D-aminoácido oxidasa están disponibles en las bases de datos para la clonación molecular.

Las triptófano deshidrogenasas son conocidas, y pueden aislarse, por ejemplo, de la espinaca, *Pisum sativum*, *Prosopis juliflora*, guisante, mezquite, trigo, maíz, tomate, tabaco, *Chromobacterium violaceum* y *Lactobacilli*. Se conocen muchas secuencias génicas de D-aminoácido deshidrogenasa.

Como se muestra en las figuras 11 a 13, si se utiliza una aminoácido oxidasa, tal como triptófano oxidasa, para convertir el triptófano en indol-3-piruvato, puede añadirse catalasa para reducir e incluso eliminar la presencia de peróxido de hidrógeno.

Indol-3-lactato a indol-3-piruvato

La reacción para convertir el indol-3-lactato en indol-3-piruvato puede ser catalizada por varios polipéptidos, tales como los miembros de las clases 1.1.1.110, 1.1.1.27, 1.1.1.28, 1.1.2.3, 1.1.1.222, 1.1.1.237, 1.1.3.- ó 1.1.1.111 de polipéptidos. La clase 1.1.1.110 de polipéptidos incluye las indol-lactato deshidrogenasas (denominadas también ácido indol-láctico:NAD⁺ oxidorreductasa). Las clases 1.1.1.27, 1.1.1.28 y 1.1.2.3 incluyen lactato-deshidrogenasas (denominadas también ácido láctico deshidrogenasas, lactato:NAD⁺ oxidorreductasa). La clase 1.1.1.222 contiene (R)-4-hidroxifenil-lactato deshidrogenasa (denominada también D-lactato aromático deshidrogenasa, R-lactato aromático deshidrogenasa, y R-3-(4-hidroxifenil)lactato:NAD(P)⁺2-oxidorreductasa) y la clase 1.1.1.237 contiene 3-(4-hidroxifenilpiruvato) reductasa (denominada también hidroxifenilpiruvato reductasa y 4-hidroxifenil-lactato:NAD⁺ oxidorreductasa). La clase 1.1.3.- contiene lactato oxidasas y la clase 1.1.1.111 contiene (3-imidazol-5-il)lactato deshidrogenasas (denominadas también (S)-3-(imidazol-5-il)lactato:NAD(P)⁺ oxidorreductasa). Es probable que varios de los polipéptidos en estas clases permitan la producción de indol-3-piruvato a partir de ácido indol-3-láctico. Los ejemplos de esta conversión se proporcionan en el Ejemplo 2.

Las reacciones químicas pueden utilizarse también para convertir el ácido indol-3-láctico en indol-3-piruvato. Dichas reacciones químicas incluyen una etapa de oxidación que puede realizarse utilizando varios procedimientos, por ejemplo: oxidación con aire utilizando un catalizador B2 (China Chemical Reporter, v 13, n 28, p 18 (1), 2002), permanganato y perclorato diluidos o peróxido de hidrógeno en presencia de catalizadores metálicos.

Indol-3-piruvato en 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico (PM)

Pueden utilizarse varios polipéptidos conocidos para convertir el indol-3-piruvato en PM. Las clases de polipéptidos ejemplificativas incluyen 4.1.3.-, 4.1.3.16, 4.1.3.17 y 4.1.2.-. Estas clases incluyen las carbono-carbonosintas/lisas, tales como las aldolasas que catalizan la condensación de dos sustratos de ácido carboxílico. La clase CE 4.1.3.- de péptido son las sintasas/lisas que forman enlaces carbono-carbono utilizando sustratos de oxo-ácido (tal como indol-3-piruvato) como electrófilo, aunque CE 4.1.2.- son sintasas-lisas que forman enlaces carbono-carbono utilizando sustratos de aldehído (tal como benzaldehído) como electrófilo.

Por ejemplo, el polipéptido descrito en la patente EP 1 045 029 (CE 4.1.3.16, 4-hidroxi-2-oxoglutarato glioxilato-liasa denominada también 4-hidroxi-2-oxoglutarato aldolasa, 2-oxo-4-hidroxi-2-oxoglutarato aldolasa o KHG aldolasa) convierte el ácido glioxílico y el piruvato en ácido 4-hidroxi-2-cetoglutarico y el polipéptido 4-hidroxi-4-metil-2-oxoglutarato aldolasa (CE 4.1.3.17, denominado también 4-hidroxi-4-metil-2-oxoglutarato piruvato-liasa o ProA aldolasa), condensa dos cetoácidos tales como dos piruvatos a 4-hidroxi-4-metil-2-oxoglutarato. Las reacciones que utilizan estas lisas se describen en la presente memoria.

Las figuras 1 a 2 y 11 a 13 representan diagramas esquemáticos de estas reacciones en los que una molécula de 3 carbonos (C3) está combinada con indol-3-piruvato. Muchos miembros de CE 4.1.2.- y 4.1.3.-, particularmente los polipéptidos que utilizan PLP, pueden utilizar moléculas de C3 que son aminoácidos tales como serina, cisteína y alanina o derivados de los mismos. Las condensaciones aldólicas catalizadas por representantes de CE 4.1.2.- y 4.1.3.- requieren que la molécula de tres carbonos de esta ruta sea piruvato o un derivado de piruvato. Sin embargo, otros compuestos pueden actuar como fuente de carbono de C3 y convertirse en piruvato. La alanina puede ser transaminada por muchas transaminasas que utilizan PLP, incluyendo muchas de las mencionadas anteriormente, para dar piruvato. El piruvato y el amoníaco pueden obtenerse por reacciones de beta-eliminación (tal como las catalizadas por triptófano o β -tirosinasa) de L-serina, L-cisteína y derivados de serina y cisteína con suficientes grupos salientes, tales como O-metil-L-serina, O-bencil-L-serina, S-metilcisteína, S-bencilcisteína, S-alkil-L-cisteína, O-acil-L-serina y 3-cloro-L-alanina. El aspartato puede servir como fuente de piruvato en las reacciones de beta-liasa mediadas por PLP tales como las catalizadas por triptofanasa (CE 4.1.99.1) y/o β -tirosinasa (CE 4.1.99.2, denominada también tirosina-fenol liasa). La velocidad de las reacciones con beta-liasa puede aumentarse realizando la mutagénesis dirigida en los polipéptidos (4.1.99.1-2) descritos por Mouratou *et al.* (J. Biol. Chem., 274:1320-5, 1999) y en el Ejemplo 8. Estas modificaciones permiten a los polipéptidos aceptar sustratos de aminoácido dicarboxílico. El lactato puede actuar también como fuente de piruvato, y se oxida a piruvato mediante la adición de lactato deshidrogenasa y un cofactor oxidado o lactato oxidado y oxígeno. Los ejemplos de estas reacciones se describen a continuación. Por ejemplo, como se muestra en la figura 2 y en las figuras 11 a 13, las ProA aldolasa puede ponerse en contacto con indol-3-piruvato cuando se utiliza piruvato como molécula de C3.

El PM puede generarse también utilizando reacciones químicas, tales como las condensaciones aldólicas proporcionadas en el Ejemplo 5.

PM en monatina

La conversión de PM en monatina puede ser catalizada por uno o más de las siguientes: triptófano

aminotransferasa (2.6.1.27), triptófano deshidrogenasas (1.4.1.19), D-aminoácido deshidrogenasas (1.4.99.1), glutamato deshidrogenasas (1.4.1.2-4), fenilalanina deshidrogenasas (CE 1.4.1.20), triptófano-fenilpiruvato transaminasas (2.6.1.28) o más en general miembros de la familia de la aminotransferasa (2.6.1.-) tal como aspartato aminotransferasa (CE 2.6.1.1), tirosina aminotransferasa (aromática) (2.6.1.5.), D-triptófano aminotransferasa o D-alanina (2.6.1.21) aminotransferasa (figura 2). Se describen a continuación (Ejemplo 1) once miembros de la clase aminotransferasa, incluyendo un nuevo miembro de la clase mostrada en las SEC. ID. nº 11 y nº 12 y las reacciones que demuestran la actividad de las enzimas aminotransferasa y deshidrogenasa se proporcionan en el Ejemplo 7.

Esta reacción puede realizarse además utilizando reacciones químicas. La aminación del cetoácido (PM) se realiza por aminación reductora utilizando amoníaco y cianoborohidruro sódico.

Las figura 11 a 13 muestran polipéptidos adicionales que pueden utilizarse para convertir PM en monatina, así como para proporcionar rendimientos aumentados de monatina a partir de indol-3-piruvato o triptófano. Por ejemplo, si se utiliza aspartato como donante de amino, puede utilizarse aspartato aminotransferasa para convertir el aspartato en oxaloacetato (figura 11). El oxaloacetato se convierte en piruvato y óxido de carbono mediante una descarboxilasa, tal como oxaloacetato descarboxilasa (figura 11). Además, si se utiliza lisina como donante de amino, puede utilizarse lisina épsilon aminotransferasa para convertir la lisina en alisina (figura 12). La alisina se convierte espontáneamente en 1-piperidina 6-carboxilato (figura 12). Si un polipéptido capaz de catalizar reacciones de aminación reductoras (por ejemplo, como glutamato deshidrogenasa) se utiliza para convertir PM en monatina, un polipéptido que puede reciclar NAD(P)H y/o producir un producto volátil (figura 13) puede utilizarse, tal como formiato deshidrogenasa.

Consideraciones adicionales en el diseño de las rutas biosintéticas

Dependiendo de qué polipéptidos se utilicen para generar indol-3-piruvato, PM y/o monatina, cofactores, sustratos y/o polipéptidos adicionales pueden proporcionarse para la producción de células que aumenten la formación del producto.

Eliminación de peróxido de hidrógeno

El peróxido de hidrógeno (H_2O_2) es un producto que, si se genera puede ser tóxico para la producción de células y puede dañar los polipéptidos o compuestos intermedios producidos. La L-aminoácido oxidasa descrita anteriormente genera H_2O_2 como producto, por consiguiente, si se utiliza la L-aminoácido oxidasa, puede eliminarse el H_2O_2 resultante o sus niveles disminuidos para disminuir la lesión potencial a la célula o producto.

Pueden utilizarse catalasas para reducir la concentración de H_2O_2 en la célula (figura 11 a 13). La célula de producción puede expresar un gen o una secuencia de ADNc que codifica una catalasa (CE 1.11.1.6), que cataliza la descomposición del peróxido de hidrógeno en agua y gas oxígeno. Por ejemplo, una catalasa puede expresarse en un vector transfectado en la célula de la producción. Ejemplos de catalasas que pueden utilizarse comprenden de manera no limitativa: trIQ9EV50 (*Staphylococcus xylosus*), trIQ9KBE8 (*Bacillus halodurans*), trIQ9URJ7 (*Candida albicans*), trIP77948 (*Streptomyces coelicolor*), trIQ9RBJ5 (*Xanthomonas campestris*) (números de registro en SwissProt). Los reactores biocatalíticos que utilizan L-aminoácido oxidasa, D-aminoácido oxidasa o triptófano oxidasa pueden contener también un polipéptido de catalasa.

Modulación de la disponibilidad de PLP (piridoxal-5'-fosfato)

Como se muestra en la figura 1, el PLP puede utilizarse en una o más de las etapas de biosíntesis descritas en la presente memoria. La concentración de PLP puede complementarse de modo que PLP no llegue a ser una limitación en la eficacia general de la reacción.

La ruta biosintética para la vitamina B₆ (precursor de PLP) ha sido estudiada a fondo en *E. coli* y algunas de las proteínas se han cristalizado (Laber *et al.*, *FEBS Letters*, 449:45-8, 1999). Dos de los genes (*epd* o *gapB* y *serC*) se requieren en otras rutas metabólicas, mientras tres genes (*pdxA*, *pdxB* y *pdxJ*) son únicos para la biosíntesis de piridoxal fosfato. Uno de los materiales de partida en la ruta de *E. coli* es el 1-desoxi-D-xilulosa-5-fosfato (DXP). La síntesis de este precursor de metabolitos centrales de los carbonos 2 y 3 comunes está catalizada por el polipéptido 1-desoxi-D-xilulosa-5-fosfato sintasa (DSX). El otro precursor es un derivado de treonina formado a partir del azúcar de 4 carbonos, D-eritrosa 4-fosfato. Los genes requeridos para la conversión en fosfo-4-hidroxil-L-treonina (HTP) son *epd*, *pdxB* y *serC*. La última reacción para la formación de PLP es una condensación intramolecular compleja y la reacción de cierre del anillo entre DXP y HTP, catalizada por los productos génicos de *pdxA* y *pdxJ*.

Si el PLP se convierte en un nutriente restringido durante la fermentación para producir monatina, el aumento de expresión de uno más de los genes de la serie de reacciones en una célula hospedadora para producción puede utilizarse para aumentar el rendimiento de monatina. Un organismo hospedador puede contener muchas copias de sus genes de la ruta natural o copias de genes de la ruta artificial pueden incorporarse en el

genoma del organismo. Además, muchas copias de los genes naturales de la serie de reacciones pueden clonarse en el organismo hospedador.

Una serie de reacciones natural que se conserva en todos los organismos recicla varios derivados de la vitamina B₆ a la forma activa PLP. Los polipéptidos implicados en esta serie de reacciones son la *pdxK* cinasa, *pdxH* oxidasa y *pdxY* cinasa. La sobreexpresión de uno o más de estos genes puede aumentar la disponibilidad de PLP.

Las concentraciones de vitamina B₆ pueden elevarse por eliminación o represión de la regulación metabólica de los genes naturales de la serie de reacciones biosintéticas en el organismo hospedador. El PLP reprime los polipéptidos implicados en la biosíntesis del derivado del precursor treonina en la cepa 238-7 de la bacteria *Flavobacterium* sp. Esta cepa bacteriana, liberada de control metabólico sobreproduce derivados de piridoxal y puede excretar hasta 20 mg/l de PLP. La manipulación genética del organismo hospedador que produce monatina en un modo similar producirá el aumento de producción de PLP sin sobreexpresión de los genes de las series de reacciones biosintéticas.

Utilización de amonio

Las reacciones de triptofanasa pueden ser dirigidas en la dirección de la síntesis (producción de triptófano a partir del indol) haciendo que el amoniaco esté más disponible o por eliminación del agua. Las reacciones reductoras de aminación, tales como las catalizadas por glutamato deshidrogenasa, pueden también ser dirigidas hacia adelante por un exceso de amonio.

El amoniaco puede estar disponible en forma de carbonato amónico o de fosfato amónico en un sistema tamponado de carbonato o fosfato. El amoniaco puede suministrarse también en forma de piruvato amónico o formiato amónico. Alternativamente, el amoniaco puede suministrarse si la reacción está acoplada a una reacción que genere amoniaco, tal como glutamato deshidrogenasa o triptófano deshidrogenasa. Puede generarse amoniaco por adición de los sustratos naturales de CE 4.1.99.-(tirosina o triptófano) que hidrolizará al fenol o indol, piruvato y NH₃. Esto permite también un aumento de rendimiento del producto de síntesis sobre la cantidad en equilibrio normal permitiendo a la enzima hidrolizar su sustrato preferido.

Eliminación de productos y subproductos

La conversión de triptófano en indol-3-piruvato mediante una triptófano aminotransferasa puede afectar desfavorablemente la velocidad de producción de indol-3-piruvato porque la reacción produce glutamato y requiere el cosustrato 2-oxoglutarato (α -cetoglutarato). El glutamato puede originar inhibición de la aminotransferasa, y la reacción consumirá grandes cantidades de cosustrato. Por otra parte, grandes concentraciones de glutamato son perjudiciales para los procesos de separación corriente abajo.

El polipéptido glutamato deshidrogenasa (GLDH) convierte el glutamato en 2-oxoglutarato, reciclando de este modo el cosustrato en la reacción catalizada por triptófano aminotransferasa. GLDH genera además equivalentes reductores (NADH o NADPH) que pueden utilizarse para generar energía para la célula (ATP) en condiciones aerobias. La utilización de glutamato por GLDH reduce además la formación de subproductos. Además, la reacción genera amoniaco, que puede servir como fuente de nitrógeno para la célula o como un sustrato en una aminación reductora para la etapa final mostrada en la figura 1. Por consiguiente, puede utilizarse una célula de producción que sobreexpresa un polipéptido GLDH para aumentar el rendimiento y reducir el coste del medio y/o de los procesos de separación.

En la ruta de triptófano a monatina, el donante amino de la etapa tres (por ejemplo, glutamato o aspartato), puede volver a convertirse en el receptor amino requerido para la etapa 1 (por ejemplo, 2-oxo-glutarato u oxalacetato), si se utiliza una aminotransferasa de la clases enzimáticas apropiadas. La utilización de dos transaminasas independientes para esta serie de reacciones, en la que el sustrato de una transaminasa no inhibe completamente la actividad de la otra transaminasa, puede aumentar la eficacia de esta serie de reacciones.

Muchas de las reacciones en las series de reacciones descritas son reversibles y, por consiguiente alcanzarán un equilibrio entre sustratos y productos. El rendimiento de la series de reacciones puede aumentarse por eliminación continua de los productos procedentes de los polipéptidos. Por ejemplo, la secreción de monatina en el caldo de fermentación utilizando una permeasa u otra proteína de transporte, o la cristalización selectiva de monatina en una corriente del reactor biocatalítico con reciclo simultáneo de sustratos aumentará el rendimiento de la reacción.

La eliminación de subproductos por reacciones enzimáticas adicionales o por sustitución de grupos del donante amino es otro modo de aumentar el rendimiento de la reacción. Varios ejemplos se exponen en el Ejemplo 13 y se muestran en las figuras 11 a 13.

Idealmente se produce un subproducto que no está disponible para reaccionar en la dirección inversa, ya sea por cambio de fase (evaporación) o por conversión espontánea en un producto final no reactivo, tal como el

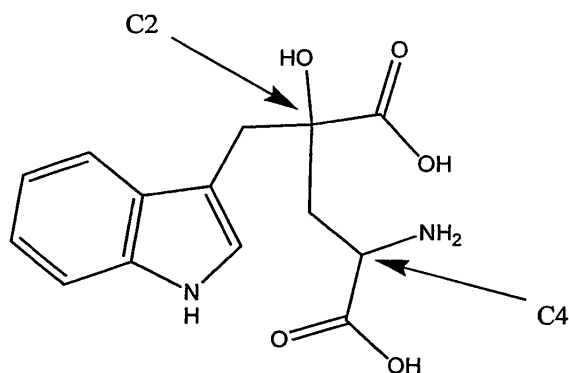
La mezcla de indol puede modularse aumentando la producción de precursores de triptófano y/o alterando las rutas catabólicas que implican indol-3-piruvato y/o triptófano. Por ejemplo, la producción de ácido indol-3-acético a partir de indol-3-piruvato puede reducirse o eliminarse, eliminando funcionalmente el gen que codifica CE 4.1.1.74 en la célula hospedadora. La producción de indol a partir de triptófano puede reducirse o eliminarse suprimiendo funcionalmente el gen que codifica CE 4.1.99.1 en la célula hospedadora. Alternativamente, se utilizó un exceso de indol como sustrato en un proceso *in vitro* o *in vivo* en combinación con cantidades aumentadas del gen que codifica CE 4.1.99.1 (Kawasaki *et al.*, *J. Ferm. and Bioeng.*, 82:604-6, 1996). Pueden hacerse modificaciones genéticas para aumentar el nivel de productos intermedios tales como D-eritrosa-4-fosfato y corismato.

La producción de triptófano está regulada en la mayoría de los organismos. Un mecanismo es por inhibición por retroalimentación de determinadas enzimas en la serie de reacciones; a medida que aumentan las concentraciones de triptófano, la velocidad de producción del triptófano disminuye. Así, cuando se utiliza una célula hospedadora modificada genéticamente para producir monatina mediante un intermedio de triptófano, puede utilizarse un organismo que no sea sensible a las concentraciones de triptófano. Por ejemplo, una cepa de *Catharanthus roseus* que es resistente a la inhibición del crecimiento por varios análogos de triptófano se seleccionó por exposición repetida a altas concentraciones de 5-metilriptófano (Schallenberg y Berlin, Z. Naturforsch., 34:541-5, 1979). La actividad de la cepa resultante de triptófano sintasa se efectuó menos por inhibición de la producción, debido probablemente a mutaciones en el gen. Asimismo, puede optimizarse una célula hospedadora para la producción de monatina.

La producción de triptófano puede optimizarse mediante la utilización de la evolución dirigida para producir polipéptidos que son menos sensibles a la inhibición del producto. Por ejemplo, puede realizarse el cribado en placas que no contienen triptófano en el medio pero con altas concentraciones de análogos de triptófano no metabolizables. Las patentes US nº 5.756.345; nº 4.742.007 y nº 4.371.614 describen procedimientos utilizados para aumentar la producción de triptófano en un organismo de fermentación. La última etapa de la biosíntesis del triptófano es la adición de serina a indol, con lo cual la disponibilidad de la serina podría aumentarse para aumentar la producción de triptófano.

La cantidad de monatina producida por un organismo de fermentación puede aumentarse al aumentar la cantidad de piruvato producido por el organismo hospedador. Determinadas levaduras, tales como *Trichosporon cutaneum* (Wang et al., Lett. Appl. Microbiol., 35:338-42, 2002) y *Torulopsis glabrata* (Li et al., Appl. Microbiol. Biotechnol., 57:451-9, 2001) sobreproducen piruvato y pueden utilizarse para practicar los procedimientos dados a conocer en la presente memoria. Además, pueden realizarse modificaciones genéticas en organismos para activar la producción de ácido pirúvico, tal como en la cepa W1485*lip2* de *E. coli* (Kawasaki et al., J. Ferm. and Bioeng., 82:604-6, 1996).

Las características de sabor de monatina pueden alterarse controlando la estereoquímica (quiralidad) de la molécula. Por ejemplo, pueden desearse diferentes isómeros de monatina en diferentes mezclas de concentraciones para diferentes sistemas de alimentación. La quiralidad puede controlarse mediante una combinación de pH y polipéptidos.



19

polipéptidos dependerán de la vía biosintética utilizada para la producción de monatina.

Cuando se forma monatina utilizando la serie de reacciones representada en la figura 2, puede considerarse lo siguiente. En una reacción biocatalítica la quiralidad del carbono 2 es determinada por la enzima que convierte indol-3-piruvato en PM. Muchas enzimas (por ejemplo, desde CE 4.1.2.-, 4.1.3.-) puede convertir el indol-3-piruvato en PM, de este modo, se puede seleccionar la enzima que forma el isómero deseado. Alternativamente, la enantioespecificidad de la enzima que convierte indol-3-piruvato en PM puede modificarse mediante la utilización de la evolución dirigida o de anticuerpos catalíticos que pueden modificarse genéticamente para catalizar la reacción deseada. Una vez se ha producido PM (por condensación enzimática o química) el grupo amino puede añadirse de manera estereoespecífica utilizando una transaminasa, tal como las descritas en la presente memoria.

La configuración R o S del carbono 4 puede generarse dependiendo de si se utiliza una D- o L-aminoácido aromático transferasa. La mayoría de las aminotransferasas son específicas para el isómero L, sin embargo existen D-triptófano aminotransferasas en determinadas plantas (Kohiba y Mito, Proceedings of the 8th International Symposium on Vitamin B₆ and Carbonyl Catalysis, Osaka, Japón, 1990). Además, se han identificado las D-alanina aminotransferasas (2.6.1.21), D-metionina-piruvato aminotransferasas (2.6.1.41) y tanto (R)-3-amino-2-metilpropanoato aminotransferasa (2.6.1.61) como (S)-3-amino-2-metilpropanoato aminotransferasa (2.6.1.22). Determinadas aminotransferasas pueden aceptar solamente el sustrato para esta reacción con una configuración específica en el carbono 2. Por consiguiente, incluso si la conversión de PM no es estereoespecífica, la estereoquímica del producto final puede controlarse mediante la selección apropiada de una transaminasa. Como las reacciones son reversibles, el PM sin reaccionar (isómero indeseado) pueden volver a reciclarse en sus constituyentes y una mezcla racémica de PM puede reformarse.

Activación de sustratos

En las reacciones dadas a conocer en la presente memoria pueden utilizarse sustratos fosforilados, tales como fosfoenolpiruvato (PEP). Los sustratos fosforilados pueden ser más favorables desde el punto de vista energético y, por consiguiente, pueden utilizarse para aumentar las velocidades y/o los rendimientos de la reacción. En las condensaciones aldólicas, la adición de un grupo fosfato estabiliza el tautómero enólico del sustrato nucleofílico, haciéndole más reactivo. En otras reacciones, un sustrato fosforilado proporciona con frecuencia un grupo saliente mejor. Asimismo, pueden activarse sustratos por conversión a derivados de CoA o derivados de pirofosfato.

EJEMPLO 1

Clonación y expresión de triptófano aminotransferasas

Este ejemplo describe procedimientos que se utilizaron para clonar las triptófano aminotransferasas, que pueden utilizarse para convertir el triptófano en indol-3-piruvato.

Generalidades experimentales

Se clonaron en *E. coli* once genes que codifican aminotransferasas. Estos genes eran de D-alanina aminotransferasa de *Bacillus subtilis* (*dat*, nº de registro de Genbank Y14082.1 pb 28622-29470 y nº de registro NP_388848.1 en Genbank, secuencia de ácido nucleico y secuencia de aminoácidos, respectivamente), tirosina aminotransferasa (*tatA*/SEC. ID. nº 1 y nº 2, secuencia de ácido nucleico y de secuencia de aminoácidos, respectivamente) de *Sinorhizobium meliloti* (denominado también *Rhizobium meliloti*), tirosina aminotransferasa (*tatA* confirmada por homología, SEC. ID. nº 3 y nº 4, secuencia de ácido nucleico y secuencia de aminoácidos, respectivamente) de la cepa 2.4.1 de *Rhodobacter sphaeroides*, tirosina aminotransferasa de 35053 de *R. sphaeroides* (confirmada por homología, SEC. ID. nº 5 y nº 6, secuencia de ácido nucleico y secuencia de aminoácidos, respectivamente), sustrato amplio aminotransferasa de *Leishmania major* (*bsat*, confirmado por homología con los fragmentos peptídicos de *L. mexicana*, SEC. ID. nº 7 y nº 8, secuencia de ácido nucleico y secuencia de aminoácidos, respectivamente), aminotransferasa aromática de *Bacillus subtilis* (*araT*, confirmado por homología, SEC. ID. nº 9 y nº 10, secuencia de ácido nucleico y secuencia de aminoácidos, respectivamente), aminotransferasa aromática de *Lactobacillus amylovorus* (*araT*, confirmado por homología, SEC. ID. nº 11 y nº 12, secuencia de ácido nucleico y secuencia de aminoácidos, respectivamente), sustrato múltiple aminotransferasa 35053 de *R. sphaeroides* (confirmada por homología, SEC. ID. nº 13 y nº 14, secuencia de ácido nucleico y secuencia de aminoácidos, respectivamente), sustrato múltiple aminotransferasa de la cepa 2.4.1 de *Rhodobacter sphaeroides* (*msa*, confirmada por homología, nº de registro AAEE01000093.1, en Genbank de 14743-16155 pb y nº de registro ZP00005082.1 en Genbank, secuencia de ácido nucleico y secuencia de aminoácidos, respectivamente), aspartato aminotransferasa de *Escherichia coli* (*aspC*, nº de registro AE000195.1 en Genbank de 2755-1565 pb y nº de registro AAC74014.1 en Genbank, secuencia de ácido nucleico y secuencia de aminoácidos, respectivamente) y tirosina aminotransferasa de *E. coli* (*tyrB*, SEC. ID. nº 31 y nº 32, secuencia de ácido nucleico y secuencia de aminoácidos, respectivamente).

Los genes se clonaron, expresaron y se determinó su actividad en la conversión de triptófano a indol-3-piruvato, junto con las enzimas disponibles en el mercado. Los once clones tenían actividad.

Identificación de cepas bacterianas que pueden contener polipéptidos con la actividad deseada

Ninguno de los genes en la base de datos del NCBI (National Center for Biotechnology Information) se denominaban triptófano aminotransferasas. Sin embargo, se han identificado organismos que tienen esta actividad enzimática. Se ha medido la actividad de L-triptófano aminotransferasa (TAT) en extractos celulares o en proteínas purificadas de las siguientes procedencias: cepa Rhizobacteriana de *Festuca octoflora*, mitocondrias y citosol de guisante, células cresta de gallo del girasol, biovariedad trifoli de *Rhizobium leguminosarum*, *Erwinia herbicola* pv gypsophilae, *Pseudomonas syringae* pv. savastanoi, *Agrobacterium tumefaciens*, *Azospirillum lipoferum* y *brasilense*, *Enterobacter cloacae*, *Enterobacter agglomerans*, *Bradyrhizobium elkanii*, *Candida maltosa*, *Azotobacter vinelandii*, cerebro de rata, hígado de rata, *Sinorhizobium meliloti*, *Pseudomonas Fluorescens* CHA0, *Lactococcus lactis*, *Lactobacillus casei*, *Lactobacillus helveticus*, plantones de trigo, cebada, *Phaseolus aureus* (frijol mungo), *Saccharomyces uvarum* (calsbergensis), *Leishmania* sp., maíz, tallos de tomate, plantas de guisante, tabaco, cerdo, *Clostridium sporogenes* y *Streptomyces griseus*.

Aislamiento del ADN genómico para clonación

Se cultivó *S. meliloti* (nº 9930 del ATCC) en medio TY a 25°C, pH 7,2. Se cultivaron las células a una densidad óptica a 600 nm (D.O.₆₀₀) de 1,85 y se utilizó un inóculo al 2% para las preparaciones del ADN genómico. El kit tip 20/G genómico de Qiagen (Valencia, CA) se utilizó para el aislamiento del ADN genómico.

Se cultivó *Bacillus subtilis* 6051 (ATCC) a 30°C en caldo de cultivo Nutrient de Bereto (Difco; Detroit, MI). Se utilizó el protocolo tip 20/G genómico de Qiagen para aislar el ADN genómico con los siguientes cambios: se duplicaron las concentraciones de proteasa K y lisozima y los tiempos de incubación se aumentaron 2 a 3 veces.

El ADN genómico ATCC 50122 de *Leishmania major* fue suministrado por IDI, Inc. (Quebec, Canadá) en tampón TE, pH 8,0, 17 ng/μl.

Se preparó 2.4.1 de *Rhodobacter sphaeroides* (suministrado por el profesor Sam Kaplan, University of Texas, Houston), *R. sphaeroides* 35053 (nº de ATCC) y se preparó ADN de *L. amylovorus* por extracción convencional con fenol. Se recolectaron las células en fase logarítmica tardía, se volvieron a poner en suspensión en tampón TEN (Tris-HCl 10 mM, pH 7,5, EDTA 1 mM, NaCl 100 mM) y se lisaron mediante adición de 0,024 ml de laurilsarcosina sódica por ml de suspensión celular. Después de extraer al menos tres veces con un volumen igual de fenol saturado con tampón TE (Tris-HCl 10 mM, pH 7,5, EDTA 1 mM), la solución de ADN se extrajo una vez con cloroformo:octanol 9:1 y tres veces con cloroformo. Se precipitó el ADN por adición de 0,1 volúmenes de acetato sódico 3 M, pH 6,8 y 2 volúmenes de etanol. Se recolectó el precipitado por centrifugación y se lavó una vez con etanol al 70%. Por último se disolvió el ADN en 0,10 ml de agua destilada.

Se aisló ADN genómico de *Escherichia coli* de la cepa DH10B (Invitrogen) y se preparó utilizando el kit Genomic-tip™ (500/G) de Qiagen. En 30 ml de este cultivo de la cepa en LB a una D.O.₆₅₀ de 1,87, se obtuvieron 0,3 mg de ADN purificado. El ADN purificado se disolvió en tampón de elución (EB) de Qiagen a una concentración de 0,37 μg/μl.

Protocolo de reacción en cadena de la polimerasa

Se diseñaron cebadores con prominencias compatibles para el vector pET30 Xa/LIC (Novagen, Madison, WI). El vector pET tiene una prominencia monocatenaria de 12 bases en lado 5' de la zona Xa/LIC y una prominencia monocatenaria de 15 bases en el lado 3' de la zona Xa/LIC. Se diseñó el plásmido para la clonación independiente de ligadura, con las etiquetas His con terminal N y S y una etiqueta His en terminal C opcional. La secuencia de reconocimiento de la Xa proteasa (IEGR) se fija directamente enfrente del codón de iniciación del gen de interés, de modo que pueden eliminarse las etiquetas de la proteína de fusión.

Se añadieron las siguientes secuencias a los extremos 5' de las secuencias específicas para el organismo al diseñar los cebadores: cebador directo, 5' GGTATTGAGGGTTCGC (SEC ID nº: 73); cebador inverso: 5' AGAGGAGAGTTAGAGCC (SEC ID nº: 74).

Bacillus subtilis *dat* cebadores: terminal N: 5'-GGTATTGAGGGTTCGCATGAAGGTTTTAGTCAATGG-3' y terminal C: 5'-AGAGGAGAGTTAGAGCCTTATGAAATGCTAGCAGCCT-3' (SEC ID nº: 15 y 16).

Sinorhizobium meliloti *tatA* cebadores: terminal N: 5'-GGTATTGAGGGTTCGCATGTTTCGACGCCCTCGCCCG y terminal C: 5'-AGAGGAGAGTTAGAGCCTCAGAGACTGGTGAACCTGC (SEC ID nº: 17 y 18).

Bacillus subtilis *araT* cebadores: terminal N: 5'-GGTATTGAGGGTTCGCATGGAACATTGCTGAATCC y

terminal C: 5'-AGAGGAGAGTTAGAGCCTTAAACGCCGTTGTTTATCG (SEC ID nº: 19 y 20).

Rhodobacter sphaeroides msa (both 2.4.1. y 35053): terminal N:

5'-GGTATTGAGGGTCGCATGCGCGAGCCTCTTGCCCT y terminal C:

5'-AGAGGAGAGTTAGAGCCTCAGCCGGGAAGCTCCGGG (SEC ID nº: 21 y 22).

Leishmania major bsat: terminal N: 5'-GGTATTGAGGGTCGCATGTCCACGCAGGCGGCCAT y terminal C:

5'-AGAGGAGAGTTAGAGCCTCACTCACGATTCACATTGC (SEC ID nº: 23 y 24).

Lactobacillus amylovorus araT: terminal N: 5'-GGTATTGAGGGTCGCATGCCAGAATTAGCTAATGA y

terminal C: 5'-AGAGGAGAGTTAGAGCCTTATTCGTCCTCTTGTAATAA (SEC ID nº: 25 y 26).

Rhodobacter sphaeroides tatA (both 2.4.1 y 35053 strains): terminal N:

5'-GGTATTGAGGGTCGCATGCGCTCTACGACGGCTCC y terminal C:

5'-AGAGGAGAGTTAGAGCCTCAGCCGCGCAGCACCTTGG (SEC ID nº: 27 y 28).

Escherichia coli aspC: terminal N: 5'-GGTATTGAGGGTCGCATGTTTGAGAACATTACCGC-3' y terminal C:

5'-AGAGGAGAGTTAGAGCCTTACAGCACTGCCACAATCG-3' (SEC ID nº: 29 y 30).

Escherichia coli tyrB: terminal N: 5'-GGTATTGAGGGTCGCGTGTTCAAAAAGTTGACGC y terminal C:

5'-AGAGGAGAGTTAGAGCCTTACATCACC GCAGCAAACG-3' (SEC ID nº: 33 y 34).

Se amplió el gen procedente de *S. meliloti* (*tatA*) utilizando el siguiente protocolo de la RCP. En una reacción con 50 µl se utilizaron 0,1 a 0,5 µg de plantilla, 1,5 µM de cada cebador, 0,4 mM de cada dNTP, 3,5 U de Polimerasa High Fidelity Expand (Roche, Indianapolis, IN) y 1 x tampón Expand™ con Mg. El programa del termociclador utilizado incluía un comienzo en caliente a 96°C durante 5 minutos, seguido de 29 repeticiones de las etapas siguientes: 94°C durante 30 segundos, 55°C durante 2 minutos y 72°C durante 2,5 minutos. Después de 29 repeticiones se mantuvo la muestra a 72°C durante 10 minutos y a continuación se almacenó a 4°C. Este protocolo de la RCP produjo un producto de 1199 pb.

Las secuencias de los genes procedentes de *R. sphaeroides* (*msa* y *tatA*), *araT* de *L. amylovorus*, y *araT* de *Bacillus*, se ampliaron utilizando el siguiente protocolo de RCP. En una reacción con 50 µl se utilizaron 0,1 a 0,5 µg de plantilla, 1,5 µM de cada cebador, 0,4 mM de cada dNTP, 3,5 U de Polimerasa High Fidelity Expand™ y 1 x tampón Expand™ con Mg. El programa del termociclador utilizado incluía un comienzo en caliente a 96°C durante 5 minutos, seguido de 29 repeticiones de las etapas siguientes: 94°C durante 30 segundos, 40-60°C durante 1 minuto, 45 segundos (gradiente del termociclador) y 72°C durante 2 minutos 15 segundos. Después de 29 repeticiones se mantuvo la muestra a 72°C durante 10 minutos y a continuación se almacenó a 4°C.

Para cada gen *msa* de *R. sphaeroides* las temperaturas de hibridación de 42°C y 48°C produjeron múltiples productos, salvo una banda distinta a aproximadamente 1464 pb. Para *araT* de *L. amylovorus*, las temperaturas de hibridación de 42°C, 48°C y 56°C proporcionaron productos individuales con bandas intensas a 1173 pb. Para *araT* de *B. subtilis*, las temperaturas de hibridación de 40°C, 45°C, 50°C y 55°C generaron productos individuales intensos (1173 pb), procedentes tanto de ADN genómico como de colonias. Para *bsat* de *L. major*, la temperatura de hibridación de 55°C dio el producto más limpio (1239 pb). Para los genes *tatA* de *Rhodobacter*, las temperaturas de hibridación entre 50 y 55°C dieron productos limpios en el tamaño correcto (1260 pb). Tanto para los genes de *E. coli* como para el gen *dat* de *B. subtilis*, se utilizó una temperatura de hibridación entre 55 y 60°C, y el tiempo de hibridación se acortó hasta 45 segundos. Se obtuvieron productos limpios de tamaños correctos (aproximadamente 1,3 kb para los genes de *E. coli*, 850 pb para el gen *dat*).

Clonación

Los productos de la RCP se purificaron en geles de 0,8 o 1% TAE-agarosa utilizando el kit de extracción en gel de Qiagen (Valencia, CA). Los productos de la RCP se cuantificaron por comparación con patrones en un gel de agarosa, y a continuación se trataron con T4 ADN polimerasa siguiendo los protocolos recomendados por el fabricante para la Clonación Independiente de la Ligadura (Novagen, Madison, WI).

En resumen, se trataron 0,2 pmoles de producto purificado por RCP con 1 U de T4 ADN polimerasa en presencia de dGTP durante 30 minutos a 22°C. La polimerasa elimina las bases sucesivas de los extremos 3' del producto de la RCP. Cuando la polimerasa encuentra un resto de guanina, la actividad en 5' a 3' de polimerasa de la enzima contrarresta la actividad de exonucleasa para impedir eficazmente más escisiones. Esto crea prominencias monocatenarias que son compatibles con el vector pET30 Xa/LIC. La polimerasa se inactiva por incubación a 75°C durante 20 minutos.

El vector y la inserción tratada se hibridaron tal como recomienda Novagen. Aproximadamente 0,02 pmoles de la inserción tratada y 0,01 pmoles del vector se incubaron durante 5 minutos a 22°C, se añadió EDTA 6,25 mM (concentración final) y se repitió la incubación a 22°C. La reacción de hibridación (1 µl) se añadió a células competentes aisladas NovaBlue™ (Novagen, Madison, WI) y se incubaron en hielo durante 5 minutos. Después del mezclado, las células se transformaron por choque térmico durante 30 segundos a 42°C. Se colocaron las células

en hielo durante 2 minutos y se dejaron recuperar en 250 µl de SOC a temperatura ambiente durante 30 minutos a 37°C con agitación a 225 rpm. Se colocaron las células en placas con LB que contenía kanamicina (25-50 µg/ml).

Se purificó el ADN plásmido utilizando el kit spin miniprep de Qiagen y se analizaron las inserciones correctas con las enzimas de restricción *Xho*I y *Xba*I. Las secuencias de los plásmidos que parecían tener la inserción correcta se verificaron mediante secuenciado del ADN de terminación de la cadena didesoxi.

Las SEC. ID. nº 1 a nº 14 y nº 31 y nº 32 presentan secuencias nucleótidos y secuencias de aminoácidos correspondientes de las aminotransferasas recombinantes, algunos cambios de las secuencias del Genbank eran sustituciones conservadoras imperceptibles o generadas en la secuencia proteica. Las SEC. ID. nº 11 y nº 12 son nuevas secuencias.

Expresión y análisis génicos

El ADN plásmido, verificado por análisis de secuencias, se subclonó en hospedadores de expresión de *E. coli* BLR(DE3) o BL21(DE3) (Novagen, Madison, WI). Los cultivos se desarrollaron y los plásmidos se aislaron utilizando el kit miniprep de Qiagen y se analizó mediante enzimas de restricción para confirmar la identidad.

Inicialmente la inducción se realizó con *araT* de *L. amylovorus*, *araT* de *B. subtilis* y *tatA* de *S. meliloti* tanto en células BLR(DE3) como BL21(DE3). Se realizó un estudio temporal con cultivos desarrollados en kanamicina que contenía LB (30 mg/l) a una D.O.₆₀₀ entre 0,5 y 0,8 y se indujo con IPTG 1 mM (isopropil tiogalactósido) y se tomaron muestras a las 0, 1, 2 y 4 horas tras la inducción. Las células procedentes de 2,0 ml se volvieron a poner en suspensión en 0,10 ml de Tris-HCl 120 mM, pH 6,8, que contienen dodecilsulfato sódico al 10%, 2-mercaptoetanol al 10% y glicerol al 20%, calentado a 95°C durante 10 min., se enfriaron y se diluyeron con 0,10 ml de H₂O. Se analizaron por SDS-PAGE alícuotas de estas muestras de la proteína celular completa utilizando un gel en gradiente 4 al 15%. No hubo diferencias significativas en la cantidad de proteína expresada entre las 2 y 4 horas de inducción, ni entre las células BLR(DE3) y BL21(DE3).

Se prepararon también extractos celulares procedentes de las muestras de 4 horas poniendo en suspensión los sedimentos celulares procedentes de 2 ml de cultivo en 0,25 ml de reactivo BugBuster™ de Novagen que contenía 0,25 µl de benzonasa nucleasa, incubando a temperatura ambiente durante 20 minutos con agitación suave y centrifugando a 16.000 x g para eliminar el residuo celular. Se cargaron los sobrenadantes (extractos celulares) en geles con gradiente del 4 al 15% para el análisis de las proteínas celulares solubles.

Los tres clones (*araT* de *L. amylovorus* (SEC. ID. nº 11 y nº 12), *araT* de *B. subtilis* (SEC. ID. nº 9 y nº 10) y *tatA* de *S. meliloti* (SEC. ID. nº 1 y nº 2) presentaban la proteína soluble que correspondía al tamaño correcto (aproximadamente 45 kDa). El producto génico *araT* de *B. subtilis* se sobreexpresó al nivel más alto y/o fue más soluble que los otros dos productos génicos.

En los procedimientos de expresión posteriores, el ADN plásmido procedente de los clones positivos se subclonó en BL21(DE3) debido a las mejores características de crecimiento de este hospedador. Se repitió la inducción utilizando IPTG 1 mM con los cultivos desarrollados en LB que contienen kanamicina a razón de 50 mg/l que se induce cuando la D.O.₆₀₀ alcanzó aproximadamente 0,8. Se recolectaron las células después de 4 horas de cultivo a 37°C, se centrifugaron a 3.000 rpm durante 10 minutos (4°C), se lavaron con tampón TEGGP (Tris-HCl 50 mM (pH 7,0), EDTA 0,5 mM, 100 mg/l de glutatión, 5% de glicerol con una mezcla de inhibidor de proteasa completo de Roche) y se congeló instantáneamente en etanol a -80°C.

Las muestras se volvieron a poner en suspensión en 5 ml/g de peso de células húmedas de reactivo BugBuster™ (Novagen) que contenía 5 µl/ml de el conjunto nº 3 de la mezcla de inhibidor de proteasa (Calbiochem-Novabiochem Corp., San Diego, CA) y 1 µl/ml de benzonasa nucleasa. Se incubaron las muestras a temperatura ambiente durante 20 minutos en un agitador orbital. Se eliminó el residuo celular insoluble por centrifugación a 16.000 x g durante 20 minutos a 4°C.

Los extractos celulares se analizaron por SDS-PAGE, y se ensayó la actividad de triptófano aminotransferasa siguiendo la producción del ácido indol-pirúvico que utiliza el siguiente protocolo. Se realizaron reacciones con un ml en tetraborato sódico 50 mM (pH 8,5), EDTA 0,5 mM, arseniato sódico 0,5 mM, piridoxal fosfato 50 µM, α-cetoglutarato 5 mM y L-triptófano 5 mM. Las reacciones se iniciaron con la adición de extractos exentos de células o enzima purificada y se incubaron 30 minutos a 30°C. Se añadió TCA al 20% (200 µl) para interrumpir la reacción, y la proteína precipitada se separó por centrifugación. Se midió la absorbancia a 327 nm y se comparó con una curva patrón de indol-3-piruvato recién preparada en el tampón de ensayo. Se realizaron también reacciones de control sin el sustrato triptófano o utilizando extractos exentos de células procedentes de clones transformados con pET30a solo.

Debido al historial de las aminotransferasas de *E. coli* natural en extractos celulares, se purificaron proteínas recombinantes utilizando cartuchos His-Bind siguiendo los protocolos del fabricante (Novagen, Madison, WI). Las fracciones eluyentes se desalaron en columnas PD-10 (Amersham Biosciences, Piscataway, NJ) y se eluyeron en Tris 50 mM, pH 7,0. Se analizaron las proteínas purificadas por SDS-PAGE y se determinó su actividad de aminotransferasa.

Los resultados de la inducción a 37°C con IPTG 1 mM (4 horas) demuestran que *bsat* de *L. major*, *tatA* de *S. meliloti*, *aspC* de *E. coli* y ambos clones *tatA* de *R. sphaeroides* tienen niveles significativos de actividad de triptófano aminotransferasa. La proteína *araT* de *B. subtilis* se sobreexpresó y era soluble, pero presentó poca actividad enzimática. El producto génico *araT* de *L. amylovorus* parecía ser soluble en el extracto celular, pero la purificación que utiliza un cartucho His-Bind produjo solo pequeñas cantidades de proteína con peso molecular correcto. Los productos génicos de *msa* eran insolubles y además los experimentos de expresión se realizaron a 24°C para minimizar la formación de inclusión corporal. Varias concentraciones de IPTG entre 10 µM y 1 mM se utilizaron para maximizar la cantidad de proteína soluble.

La Tabla 1 lista las actividades específicas medidas en microgramos de indol-3-piruvato (I3P) formado por miligramo de proteína por minuto. En algunos casos, cantidades muy pequeñas de proteína recombinante presentaban grandes concentraciones de actividad por encima del intervalo lineal eficaz del ensayo. En estos casos un signo ">" precede al número de actividad específico.

Tabla 1: Actividades específicas de los clones en extractos celulares (EC) y enzimas purificadas (P) y comerciales

Enzima	Actividad específica (µg I3P/mg de proteína/min)	Notas
<i>L. major bsat</i> CE	>49,3	
<i>L. major bsat</i> P	>4280	
<i>S. meliloti tat</i> ACE	>28,6	
<i>S. meliloti tat</i> AP	>931	
2.4.1 <i>tat</i> ACE	>41,2	
2.4.1 <i>tat</i> AP	1086	
35053 <i>tat</i> ACE	>62,3	
35053 <i>tat</i> AP	>486	
<i>L. amylovorus ara</i> TCE	1,26	
<i>L. amylovorus ara</i> TP	0	proteína pequeña después del cartucho His-Bind
<i>B. subtilis ara</i> TCE	0	indetectable
<i>B. subtilis ara</i> TP	1,5-4,5	
2.4.1 <i>msa</i> CE	2,05	muy poca proteína soluble
2.4.1 <i>msa</i> P	0	nada de proteína después del cartucho His-Bind
35053 <i>msa</i> CE	3,97	Muy poca proteína soluble
35053 <i>msa</i> P	0	Sin proteína después del cartucho His-Bind
<i>E. coli asp</i> C(P)	800	

Tabla 1 (Continuación)

Enzima	Actividad específica (µg I3P/mg de proteína/min)	Notas
<i>E. coli tyr</i> B(P)	1	no muy soluble
<i>B. subtilis</i> D-aminotransf.(P)	2,7	utilizando D-triptófano como sustrato
Transaminasa de amplio espectro	22	nº T7684 en cat. Sigma
Porcina de tipo II-A	1,5	Sigma G7005
Porcina de tipo I	1	Sigma G2751

Una alineación que compara todas las proteínas recombinantes clonadas ilustra que no existen muchas áreas muy conservadas entre las secuencias *araT*, *tatA*, *bsat* y *msa*. Una alineación de proteínas recombinantes con máxima actividad: homólogos del producto génico *tatA* de *Rhodobacter*, sustrato amplio aminotransferasa de *L. major* y la tirosina aminotransferasa de *Sinorhizobium meliloti* presentó varias regiones conservadas, sin embargo son solo aproximadamente 30 a 43% idénticas en concentración de proteína. La disponibilidad de la (D-alanina) aminotransferasa, D-específica (D-alanina) de amplio espectro puede utilizarse en la producción de otros estereoisómeros de monatina.

EJEMPLO 2**Conversión del indol-3-lactato en indol-3-piruvato**

Como se muestra en las figuras 1 y 3, el ácido indol-3-láctico puede utilizarse para producir indol-3-piruvato. La conversión entre el ácido láctico y el piruvato es una reacción reversible, como es la conversión entre el indol-3-piruvato y el indol-3-lactato. La oxidación del indol-3-lactato fue seguida típicamente debido a la alta cantidad de fondo a 340 nm del indol-3-piruvato.

La mezcla analítica patrón contenía fosfato potásico 100 mM, pH 8,0, NAD^+ 0,3 mM, 7 unidades de lactato deshidrogenasa (LDH) (Sigma-L2395, St. Louis, MO) y sustrato 2 mM en 0,1 ml. El ensayo se realizó por duplicado en una placa de microvalorador transparente a UV, utilizando un lector de placas SpectraMax Plus de Molecular Devices. Se mezclaron el polipéptido y el tampón y se pipetearon en los pocillos que contenían el ácido indol-3-láctico y NAD^+ y se leyó la absorbancia a 340 nm de cada pocillo a intervalos de 9 segundos después de un breve mezclado. La reacción se mantuvo a 25°C durante 5 minutos. El aumento de absorbancia a 340 nm siguió a la producción de NADH a partir de NAD^+ . Se realizaron controles negativos por separado sin NAD^+ y sin sustrato. D-LDH procedente de *Leuconostoc mesenteroides* (nº L2395 en catálogo de Sigma) parecía presentar más actividad con los sustratos derivados de indol que L-LDH de *Bacillus stearothermophilus* (nº L5275 en catálogo de Sigma).

Se utilizaron procedimientos similares con el ácido D-láctico y NAD^+ o NADH y piruvato, sustratos naturales de los polipéptidos D-LDH. La $V_{\max}$ para la reducción del piruvato fue de 100 a 1.000 veces superior que la $V_{\max}$ para la oxidación del lactato. La $V_{\max}$ para la reacción de oxidación del indol-3-láctico con 3-LDH fue de aproximadamente una quinta parte de la del ácido láctico. La presencia de indol-3-piruvato se determinó también siguiendo el cambio de absorbancia a 327 (derivado de enol-borato) utilizando tampón de borato sódico 50 mM que contenía EDTA 0,5 mM y arseniato sódico 0,5 mM. Se observaron pequeños pero repetibles cambios de absorbancia en comparación con las referencias negativas tanto para los polipéptidos L como D-LDH.

Además, las lactato deshidrogenasas de amplia especificidad (enzimas con actividad asociada a las CE 1.1.1.27, CE 1.1.1.28 y/o CE 1.1.2.3) pueden clonarse y utilizarse para preparar indol-3-piruvato a partir de ácido indol-3-láctico. Las fuentes de deshidrogenasas de amplia especificidad incluyen *E. coli*, *Neisseria gonorrhoeae* y *Lactobacillus plantarum*.

Alternativamente, puede producirse indol-3-piruvato poniendo en contacto indol-3-lactato con extractos celulares procedentes de *Clostridium sporogenes* que contienen una indol-lactato deshidrogenasa (CE 1.1.1.110); o extractos celulares de *Trypanosoma cruzi epimastigotes* que contienen *p*-hidroxifenil-lactato deshidrogenasa (CE 1.1.1.222) conocida por tener actividad sobre indol-3-piruvato; o *Pseudomonas acidovorans* o extractos celulares de *E. coli* que contienen una imidazol-5-il lactato deshidrogenasa (CE 1.1.1.111); o *Coleus blumei*, que contiene una hidroxifenilpiruvato reductasa (CE 1.1.1.237); o *Candida maltosa* que contiene una D-lactato aromático deshidrogenasa (CE 1.1.1.222). Las referencias que describen dichas actividades incluyen, Nowicki *et al.* (*FEMS Microbiol. Lett.*, 71:119-24, 1992) Jean y DeMoss (*Canadian J. Microbiol.* 14, 1968), Coote y Hassall (*Biochem. J.*, 111:237-9, 1969), Cortese *et al.*, (*C. R. Seances Soc. Biol. Fil.*, 162 390-5, 1968), Petersen y Alfermann (*Z. Naturforsch. C: Biosci.*, 43 501-4, 1988) y Bhatnagar *et al.*, (*J. Gen. Microbiol.*, 135:353-60, 1989). Además, una lactato oxidasa tal como la procedente de *Pseudomonas sp.* (Gu *et al.*, *J. Mol. Catalysis B: Enzymatic.*, 18:299-305, 2002), puede utilizarse para la oxidación de indol-3-láctico a indol-3-piruvato.

EJEMPLO 3**Conversión de L-triptófano en indol-3-piruvato utilizando L-aminoácido oxidasa**

Este ejemplo describe los procedimientos utilizados para convertir el triptófano en indol-3-piruvato mediante una oxidasa (CE 1.4.3.2), como alternativa a la utilización de una triptófano aminotransferasa tal como se describe en el Ejemplo 1. La L-aminoácido oxidasa se purificó a partir de *Crotalus durissus* (Sigma, St. Louis, MO, nº de catálogo A2805). Los números de registro de las L-aminoácido oxidasas para la clonación molecular incluyen: CAD21325.1, AAL14831, NP_490275, BAB78253, A38314, CAB71136, JE0266, T08202, S48644, CAC00499, P56742, P81383, 093364, P81382, P81375, S62692, P23623, AAD45200, AAC32267, CAA88452, AP003600, y Z48565.

Se llevaron a cabo reacciones en tubos de microcentrifugadora en un volumen total de 1 ml, se incubaron durante 10 minutos en agitación a 37°C. La mezcla de reacción contenía L-triptófano 5 mM, tampón fosfato sódico 100 mM, pH 6,6, arseniato sódico 0,5 mM, EDTA 0,5 mM, tetraborato sódico 25 mM, 0,016 mg de catalasa (83 U, Sigma C-3515), 0,008 mg de FAD (Sigma) y 0,005 a 0,125 unidades de L-aminoácido oxidasa. Las referencias negativas contenían todos los componentes excepto triptófano, y los blancos contenían todos los componentes excepto la oxidasa. Se utilizó catalasa para eliminar el peróxido de hidrógeno formado durante la desanimación oxidativa. Se utilizaron tetraborato y arseniato sódico para estabilizar la forma enol-borato del indol-3-piruvato, que presenta una absorbancia máxima a 327 nm. Se prepararon patrones de indol-3-piruvato a concentraciones entre 0,1 y 1 mM en la mezcla de reacción.

La L-aminoácido oxidasa adquirida tenía una actividad específica de 540 µg de indol-3-piruvato formado por minuto por mg de proteína. Esta es del mismo orden de magnitud que la actividad específica de las enzimas triptófano aminotransferasa.

5

EJEMPLO 4

Conversión de indol-3-piruvato en 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico con una aldolasa

10

Este ejemplo describe los procedimientos que pueden utilizarse para convertir el indol-3-piruvato en el precursor de monatina (PM) 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico utilizando una aldolasa (liasa) (figura 2). Las condensaciones aldólicas son reacciones que forman enlaces Carbono-carbono entre el carbono β de un aldehído o cetona y el carbono carbonilo de otro aldehído o cetona. Se forma un carbanión en el carbono adyacente al grupo carbonilo de un sustrato, y sirve como nucleófilo que ataca al carbono carbonilo del segundo sustrato (carbono electrófilo). Más frecuentemente, el sustrato electrófilo es un aldehído, de manera que la mayoría de las aldolasas están comprendidas en la categoría CE 4.1.2. Muy a menudo, el sustrato nucleófilo es el piruvato. Es menos frecuente que las aldolasas catalicen la condensación entre dos cetoácidos o dos aldehídos.

15

20

Sin embargo, se han identificado aldolasas que catalizan la condensación de dos ácidos carboxílicos. Por ejemplo, el documento EP 1 045 029 describe la producción del ácido L-4-hidroxi-2-cetoglutarico a partir del ácido glioxílico y piruvato utilizando el cultivo de *Pseudomonas* (CE 4.1.3.16). Además, la 4-hidroxi-4-metil-2-oxoglutarato aldolasa (4-hidroxi-4-metil-2-oxoglutarato piruvato liasa, CE 4.1.3.17) puede catalizar la condensación de dos cetoácidos. Por lo tanto, se utilizaron polipéptidos de aldolasa similares para catalizar la condensación de indol-3-piruvato con piruvato.

25

Clonación

30

Las 4-hidroxi-4-metil-2-oxoglutarato piruvato liasas (ProA aldolasa, CE 4.1.3.17) y 4-hidroxi-2-oxoglutarato glioxilato liasa (KHG aldolasa, CE 4.1.3.16) catalizan reacciones muy similares a la reacción con aldolasa de la figura 2. Se diseñaron cebadores con prominencias compatibles para el vector pET30 Xa/LIC (Novagen, Madison, WI). El diseño de estos cebadores se describió anteriormente en el Ejemplo 1.

Se diseñaron los siguientes cebadores para la clonación con pET30 Xa/LIC:

35

1. Gen *proA* para *Pseudomonas straminea* (nº de registro en GenBank: 12964663 Version: 12964663) y gen *proA* para *Comamonas testosteroni* (SEC. ID. nº 65 y nº 66, secuencia de ácidos nucleicos y secuencia de aminoácidos, respectivamente) directo 5'-GGTATTGAGGGTCGCATGTACGAAGTGGGAGTTGT-3' e inverso 5'-AGAGGAGAGTTAGAGCCTTAGTCAATATATTTTCAGGC-3' (SEC ID nº: 55 y 56).

40

2. Gen 1021 SMc00502 para *Sinorhizobium meliloti* (homólogo a *proA*, nº. de registro en GenBank: 15074579 y nº CAC46344, secuencia de ácidos nucleicos y secuencia de aminoácidos, respectivamente) directo 5'-GGTATTGAGGGTCGCATGAGCGTGGTTCACCGGAA-3' e inverso 5'-AGAGGAGAGTTAGAGCCTCAATCGATATATTTTCAGTC-3' (SEC ID nº: 61 y 62).

45

3. Gen LB126 fldZ para *Sphingomonas sp.* (nº 7573247 de registro en Genbank versión: 7543247, códigos para una supuesta aciltransferasa) directo 5'-GGTATTGAGGGTCGCATGTCCGGCATCGTTGTCCA-3' e inverso 5'-AGAGGAGAGTTAGAGCCTCAGACATATTTTCAGTCCCA-3' (SEC ID nº: 57 y 58).

50

4. Gen *pcmE* para *Arthrobacter keyseri* (nº AF331043 de registro en Genbank versión: AF331043.1, códigos para una oxalocitramalato aldolasa) directo 5'-GGTATTGAGGGTCGCATGCGACTGAACAACCTCGG-3' e inverso 5'-AGAGGAGAGTTAGAGCCTCAGTTCTCCACGTATTCCA-3' (SEC ID nº: 59 y 60).

55

5. Gen YPO0082 para la cepa CO92 de *Yersinia pestis* (nº 15978115 de registro en Genbank versión: 15978115, códigos para una posible transferasa) directo 5'-GGTATTGAGGGTCGCATGAGCCTGGTTAATATGAA-3' e inverso 5'-AGAGGAGAGTTAGAGCCTTATGACTTTAACGCGTTGA-3' (SEC ID nº: 63 y 64).

60

6. Gen *khg* para *Bacillus subtilis* (nº Z99115.1, nº GI:2634478, nº 126711 a 127301 y nº CAB14127.1 de registro en Genbank, secuencia de ácidos nucleicos y secuencia de aminoácidos, respectivamente) directo 5'-GGTATTGAGGGTCGCATGGAGTCCAAAGTCGTTGA-3' e inverso 5'-AGAGGAGAGTTAGAGCCTTACACTTGGAAAACAGCCT-3' (SEC ID nº: 35 y 36).

65

7. Gen *khg* para *E. coli* (nº AE000279.1 1331-1972 y nº AAC74920.1 de registro en Genbank, secuencia de ácidos nucleicos y secuencia de aminoácidos, respectivamente) directo 5'-GGTATTGAGGGTCGCATGAAAACTGGAACAAG-3' e inverso 5'-AGAGGAGAGTTAGAGCCTTACAGCTTAGCGCCTTCTA-3' (SEC ID nº: 37 y 38).

8. Gen *khg* para *S. meliloti* (nº AL591792.1, nº GI:15075850, nº 65353 a nº 64673 y nº CAC47463.1 de registro en Genbank, secuencia de ácidos nucleicos y secuencia de aminoácidos, respectivamente) directo 5'-GGTATTGAGGGTCGCATGCGAGGGGCATTATTCAA-3' e inverso 5'-AGAGGAGAGTTAGAGCCTCAGCCCTTGAGCGCGAAG-3' (SEC ID nº: 39 y 40).

El ADN genómico procedente de los organismos descritos en los apartados 1 y 2 y 6 a 8, anteriormente, se purificó utilizando el protocolo genómico-tip de Qiagen. Utilizando técnicas similares puede purificarse el ADN genómico procedente de organismos descritos en los apartados 3 a 5.

Se cultivó *Pseudomonas straminea* (ATCC 33636) a 30°C en caldo de cultivo nutritivo y medio de hidroxibenzoato. Se cultivó *Comamonas testosteroni* (ATCC 49249) a 26°C en caldo de cultivo nutritivo y medio de hidroxibenzoato. *Sphingomonas* sp. LB126 (Flemish Institute of Technological Research, VITO, B-2400 Mol, Belgica) se cultiva según el procedimiento descrito por Wattiau *et al.*, (*Research in Microbiol.*, 152:861-72, 2001). *Arthrobacter keyseri* (Gulf Ecology Division, National Health and Environmental Effects Research Laboratory, U.S. Environmental Protection Agency, Gulf Breeze, FL 32561, US.) se cultiva según el protocolo descrito por Eaton (*J. Bacteriol.*, 183:3869-3703, 2001). *Sinorhizobium meliloti* 1021 (ATCC 51124) se cultivó a 26°C en medio ATCC TY y medio de hidroxibenzoato. La cepa CO92 (ATCC) de *Yersinia pestis* se cultiva a 26°C en medio 739 de agar-agar con sangre de caballo del ATCC. *Bacillus subtilis* 6051 (ATCC) se cultivó a 30°C en caldo de cultivo Nutrient de Bereto (Difco; Detroit, MI). Se aisló ADN genómico de *E. coli* procedente de la cepa DH10B (Invitrogen) tal como se describe en el Ejemplo 1.

Se utilizaron los protocolos de la RCP, de clonación y de identificación descritos en el Ejemplo 1 para clonar las secuencias *proA* de *C. testosteroni* y de *S. meliloti*, así como las secuencias *khg* de *E. coli*, *B. subtilis* y *S. meliloti*. Pueden utilizarse los mismos procedimientos para clonar las demás secuencias descritas anteriormente.

Se secuenciaron clones positivos utilizando el secuenciado de terminación de la cadena didesoxi (Seqwright, Houston, TX) con la etiqueta S y los cebadores del termiador T7 (Novagen) y cebadores internos de Integrated DNA Technologies, Inc. (Coralville, IA).

Ensayos de expresión y actividad

El ADN plásmido (verificado por análisis de secuencias) se subclonó en el hospedador de expresión BL21(DE3) (Novagen). Los cultivos se desarrollaron en medio LB con 50 mg/l de kanamicina, los plásmidos se aislaron utilizando un kit spin plasmid miniprep de Qiagen y se analizaron posteriormente con enzimas de restricción para confirmar la identidad. Se hicieron experimentos de inducción con los montajes BL21(DE3) cultivados en medio LB que contenía 50 mg/l de kanamicina a 37°C. Se indujo a la expresión de la proteína utilizando IPTG 0,1 mM después que la D.O.₆₀₀ alcanzó aproximadamente 0,6. Se cultivaron las células durante 4 horas a 30°C y se recolectaron por centrifugación. Las células se lisaron a continuación utilizando reactivo Bugbuster™ (Novagen) y las proteínas recombinantes con etiqueta His se purificaron utilizando los cartuchos His-Bind tal como se describió anteriormente (Ejemplo 1). Las proteínas purificadas se desalaron en columnas PD-10 disponibles y se eluyeron en tampón Tris-HCl 50 mM, pH 7,3 con MgCl₂ 2 mM.

Se analizaron las proteínas por SDS-PAGE en geles con gradiente 4 al 15% para detectar las concentraciones de proteína soluble al P.M. predicho de la proteína de fusión recombinante.

Se analizó la actividad de las proteínas utilizando indol-3-piruvato y piruvato sódico como sustratos. La mezcla analítica contenía Tris-HCl 100 mM (pH 7- pH 8,9), MgCl₂ 0-8 mM, fosfato potásico 3 mM (pH 8) y 6 mM de cada sustrato en 1 ml. Se comenzó la reacción añadiendo cantidades variables de polipéptido (por ejemplo de 10 a 100 µg) y se incubó entre 25°C y 37°C durante 30 minutos, se filtró y se congeló a continuación a -80°C.

Resultados de actividad con los productos del gen *proA*

Tanto los montajes del gen *proA* para *C. testosteroni* como de *SMc00502* para *S. meliloti* presentaban altos niveles de expresión cuando se inducían con IPTG. Las proteínas recombinantes eran muy solubles, determinadas por análisis SDS-PAGE de proteína total y muestras de extracto celular. El producto génico para *C. testosteroni* se purificó hasta > 95% de pureza. Debido a que el rendimiento del producto génico para *S. meliloti* fue muy bajo después de la purificación por afinidad utilizando un cartucho His-Bind, se utilizó extracto celular para el análisis enzimático.

Ambas aldolasas recombinantes catalizaron la formación de PM a partir de indol-3-piruvato y piruvato. La presencia tanto de fosfato de magnesio bivalente como de potasio se requirieron para la actividad enzimática. Ningún producto se manifestó cuando estaba ausente indol-3-piruvato, piruvato o fosfato potásico. Una pequeña cantidad del producto se formó también en ausencia de enzima (por lo general de un orden de magnitud menos que cuando la enzima estaba presente).

El pico del producto eluyó de la columna C18 de fase inversa poco después que el patrón indol-3-piruvato, el espectro de masas de este pico presentaba un ión original ($[M+H]^+$) provocado por colisión de 292,1, el ión original esperado para el producto PM. Los fragmentos derivados mayores presentes en el espectro de masas incluían aquellos con $m/z = 158$ (ión carbonio de 1*H*-indol-3-carbaldehído), 168 (ión carbonio de 3-buta-1,3-dienil-1*H*-indol), 274 (292 - H_2O), 256 (292 - 2 H_2O), 238 (292 - 3 H_2O), 228 (292 - CH_4O_3) y 204 (pérdida de piruvato). El producto presentaba además un espectro UV característico de otros compuestos que contienen indol, tal como triptófano, con la λ_{max} de 279 a 280 y un pequeño lomo a aproximadamente 290 nm.

La cantidad de PM producida por la aldolasa de *C. testosteroni* aumentó con un incremento en la temperatura de reacción desde la temperatura ambiente hasta 37°C, cantidad de sustrato y cantidad de magnesio. La actividad sintética de la enzima disminuyó al aumentar el pH, el producto máximo observado fue a pH 7. Referida a patrones de triptófano, la cantidad de PM producida en un ensayo patrón utilizando 20 µg de proteína purificada fue de aproximadamente 10 a 40 µg por ml de reacción.

Debido al alto grado de homología de las secuencias de codificación de aldolasa de *proA* de *S. meliloti* y de *C. testosteroni* con otros genes descritos anteriormente, cabe esperar que todos los productos génicos recombinantes puedan catalizar esta reacción. Por otra parte, cabe esperar que las aldolasas que tienen treonina (T) en las posiciones 59 y 87, arginina (R) en 119, aspartato (D) en 120 e histidina (H) en 31 y 71 (referidas al sistema de numeración de *C. testosteroni*) tengan similar actividad.

Resultados de actividad con los productos del gen *khg*

Tanto los montajes génicos *khg* de *B. subtilis* como de *E. coli* tenían altos niveles de expresión de la proteína cuando se inducían con IPTG, mientras que el *khg* de *S. meliloti* tenía un nivel de expresión mas bajo. Las proteínas recombinantes eran muy solubles, como se interpreta mediante los análisis SDS-PAGE proteínas totales y de extractos celulares. Los productos génicos *khg* de *B. subtilis* y de *E. coli* se purificaron hasta > 95% de pureza; el rendimiento del producto génico de *S. meliloti* no fue tan elevado después de la purificación por afinidad utilizando un cartucho His-Bind.

No existen pruebas de que el magnesio y el fosfato se necesiten para la actividad de esta enzima. Sin embargo, la bibliografía da cuenta de la realización de ensayos en tampón de fosfato sódico, y la enzima supuestamente es bifuncional y tiene actividad sobre sustratos fosforilados tales como 2-ceto-3-desoxi-6-fosfogluconato (KDPG). Se realizaron análisis enzimáticos como se describió anteriormente, y en algunos casos se omitió el fosfato. Los resultados indican que las aldolasas de KHG recombinantes produjeron PM, pero no eran tan activas como las *proA* aldolasas. En algunos casos el nivel de PM producido por KHG fue casi idéntico a la cantidad producida por magnesio y fosfato solos. El fosfato no parecía aumentar las actividades de KHG. La enzima de *Bacillus* tenía la mayor actividad, aproximadamente 20 al 25% de actividad mayor que el magnesio y fosfato solos, determinada por RSM (véase el Ejemplo 10). La enzima *Sinorhizobium* tenía la mínima cantidad de actividad, lo que puede asociarse a problemas de plegamiento y solubilidad indicados en la expresión. Las tres enzimas tienen el glutamato de la zona activa (posición 43 en el sistema de numeración de *B. subtilis*) así como la lisina requerida para la formación de la base de Schiff con piruvato (posición 130); sin embargo, la enzima de *B. subtilis* contiene una treonina en la posición 47, un resto de la zona activa, en lugar de arginina. La KHG de *B. subtilis* es más pequeña y parece estar en un grupo distinto de las enzimas de *S. meliloti* y de *E. coli*, con otras enzimas que presentan la treonina de la zona activa. Las diferencias en la zona activa pueden ser la razón para el aumento de actividad de la enzima de *B. subtilis*.

Mejora de la actividad de aldolasa

Los anticuerpos catalíticos pueden ser tan eficaces como las aldolasas naturales, aceptan un amplia gama de sustratos y pueden utilizarse para catalizar la reacción mostrada en la figura 2.

Las aldolasas pueden mejorarse también por evolución dirigida, por ejemplo como se describió anteriormente para la aldolasa de KDPG (muy homóloga a KHG descrita anteriormente) producida por reordenación del ADN y RCP propensa a error para eliminar el requisito del fosfato e invertir la enantioselectividad. Los péptidos de KDPG aldolasa son útiles en las reacciones bioquímicas, ya que son muy específicos para el sustrato del donante (en la presente memoria, piruvato) pero son relativamente flexibles con respecto al sustrato del receptor (es decir indol-3-piruvato) (Koeller & Wong, *Nature*, 409:232-9, 2001). La KHG aldolasa tiene actividad para la condensación del piruvato con numerosos ácidos carboxílicos. Las versiones para mamíferos de la KGH aldolasa se cree que tienen especificidad más amplia que las versiones bacterianas, incluyendo mayor actividad sobre 4-hidroxi-4-metil-2-oxoglutarato y aceptación de ambos estereoisómeros de 4-hidroxi-2-cetoglutarato. Las fuentes bacterianas parecen tener una preferencia de 10 veces para el isómero R. Existen casi 100 homólogos de KHG disponibles en las bases de datos genómicas y se ha demostrado actividad en *Pseudomonas*, *Paracoccus*, *Providencia*, *Sinorhizobium*, *Morganella*, *E. coli* y tejidos de mamíferos. Estas enzimas pueden utilizarse como punto de partida para ajustar la enantioespecificidad que se desea para la producción de monatina.

Las aldolasas que utilizan piruvato y otro sustrato que es un cetoácido y/o tiene un grupo hidrófobo

voluminoso como indol pueden "producirse" para ajustar la especificidad, velocidad y selectividad del polipéptido. Además de las KHG y proA aldolasas demostradas en la presente memoria, los ejemplos de estas enzimas incluyen, pero no están limitados a: KDPG aldolasa y los polipéptidos afines (KDPH); la transcarboxibenzalpiruvato hidratasa-aldolasa de *Nocardioidea* st., 4-(2-carboxifenil)-2-oxobut-3-enoato aldolasa (2'-carboxibenzalpiruvato aldolasa) que condensa piruvato y 2-carboxibenzaldehído (sustrato que contiene anillo aromático); trans-O-hidroxibencilidenpiruvato hidratasa-aldolasa de *Pseudomonas putida* y *Sphingomonas aromatocivorans*, que también utiliza piruvato y un aldehído que contiene aromáticos como sustratos; 3-hidroxiaspartato aldolasa (eritro-3-hidroxi-L-aspartato glioxilato liasa), que utiliza 2-oxoácidos como sustratos y se cree que está en el organismo de *Micrococcus denitrificans*; benzoin aldolasa (benzalaldehído liasa) que utiliza sustratos que contienen grupos bencilo; dihidroneopterina aldolasa; L-treo-3-fenilserina benzaldehído-liasa (fenilserina aldolasa) que condensa glicina con benzaldehído; 4-hidroxi-2-oxovalerato aldolasa; 1,2-dihidroxibencilpiruvato aldolasa y 2-hidroxibenzalpiruvato aldolasa.

Un polipéptido con la actividad deseada puede seleccionarse cribando clones de interés que utilizan los procedimientos siguientes. Se transforman auxótrofos de triptófano con vectores que llevan los clones de interés en una casete de expresión y se cultivan en un medio que contiene pequeñas cantidades de monatina o PM. Ya que las reacciones de aminotransferasas y aldolasas son reversibles, las células pueden producir triptófano a partir de una mezcla racémica de monatina. Asimismo, los organismos (tanto recombinantes como naturales) pueden cribarse por su capacidad para utilizar PM o monatina como fuente de carbono y energía. Una fuente de aldolasas dianas son los bancos de expresión de varias cepas de *Pseudomonas* y rhizobacterianas. Las *Pseudomonas* tienen muchas rutas catabólicas poco frecuentes para la degradación de moléculas aromáticas y además contienen muchas aldolasas; mientras que las rhizobacterias contienen aldolasas, son conocidas por desarrollarse en la rizosfera vegetal y tienen muchos de los genes descritos para la construcción de una ruta biosintética para la monatina.

EJEMPLO 5

Síntesis química del precursor de monatina

El Ejemplo 4 describió un procedimiento de utilización de una aldolasa para convertir el indol-3-piruvato en el 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico, precursor de monatina (PM). Este ejemplo describe un procedimiento alternativo para sintetizar químicamente el PM.

El PM se forma utilizando una condensación aldólica típica (figura 4). En resumen, una reacción aldólica típica conlleva la generación de un carbanión del éster piruvato utilizando una base fuerte, tal como LDA (diisopropilamida de litio), hexametildisilazano de litio o butil-litio. El carbanión que se genera reacciona con el indol-piruvato para formar el producto acoplado.

Los grupos protectores que pueden utilizarse para proteger el nitrógeno del indol comprenden de manera no limitativa: t-butiloxicarbonilo (Boc) y benciloxicarbonilo (Cbz). Los grupos bloqueadores para los ácidos carboxílicos comprenden de manera no limitativa, ésteres de alquilo (por ejemplo ésteres de metilo, etilo o bencilo). Cuando se utilizan dichos grupos protectores, no es posible controlar la estereoquímica del producto que se forma. Sin embargo, si R2 y/o R3 son grupos protectores quirales (figura 4) tal como (S)-2-butanol, mentol o una amina quiral, esto puede favorecer la formación de un enantiómero de PM sobre el otro.

EJEMPLO 6

Conversión de triptófano o indol-3.piruvato en monatina

Un procedimiento *in vitro* que utiliza dos enzimas, una aminotransferasa y una aldolasa, produjeron monatina a partir de triptófano y piruvato. En la primera etapa, el alfa-cetoglutarato fue el receptor del grupo amino del triptófano en una reacción de transaminación que genera indol-3-piruvato y glutamato. Una aldolasa catalizó la segunda reacción en la que el piruvato se hizo reaccionar con indol-3-piruvato, en presencia de Mg^{2+} y fosfato, generando el derivado alfa-ceto de monatina (PM), 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico. La transferencia del grupo amino desde el glutamato formado en la primera reacción produjo el producto deseado, monatina. La purificación y caracterización del producto demostró que el isómero formado era S,S-monatina. Se describen sustratos, enzimas y condiciones alternativas así como las mejoras que se hicieron en este procedimiento.

Enzimas

La aldolasa, 4-hidroxi-4-metil-2-oxoglutarato piruvato liasa (ProA aldolasa, gen *proA*) (CE 4.1.3.17) de *Comamonas testosteroni* se clonó, expresó y purificó tal como se describe en el Ejemplo 4. Las 4-hidroxi-2-oxoglutarato glioxilato liasas (KHG aldolasas) (CE 4.1.3.16) de *B. subtilis*, *E. coli* y *S. meliloti* se clonaron, expresaron y purificaron tal como se describe en el Ejemplo 4.

Las aminotransferasas utilizadas junto con las aldolasas para producir monatina eran la L-aspartato

aminotransferasa codificada por el gen *aspC* de *E. coli*, la tirosina aminotransferasa codificada por el gen *tyrB* de *E. coli*, la enzima *tatA* de *S. melliloti*, la aminotransferasa de amplio sustrato codificada por el gen *bsat* de *L. major* o la glutámico-oxaloacética transaminasa del corazón de cerdo (tipo IIa). La clonación, expresión y purificación de las proteínas que no son de mamífero se describen en el Ejemplo 1. La glutámico-oxaloacética transaminasa del corazón del cerdo (tipo IIa) se adquirió en Sigma (nº G7005).

Procedimiento que utiliza *ProA* aldolasa y *L*-aspartato aminotransferasa

La mezcla de reacción contenía acetato amónico 50 mM, pH 8,0, $MgCl_2$ 4 mM, fosfato potásico 3 mM, piridoxal fosfato 0,05 mM, piruvato amónico 100 mM, triptófano 50 mM, alfa-cetoglutarato 10 mM, 160 mg de ProA aldolasa recombinante de *C. testosteroni* (extracto celular impurificado, aldolasa al ~30%), 233 mg de L-aspartato aminotransferasa de *E. coli* recombinante (extracto celular impurificado, ~40% de aminotransferasa) en un litro. Todos los componentes excepto las enzimas se mezclaron y se incubaron a 30°C hasta que se disolvió el triptófano. Las enzimas se añadieron a continuación y la solución de reacción se incubó a 30°C con agitación suave (100 rpm) durante 3,5 horas. 0,5 y 1 hora después de la adición de las alícuotas en enzimas de triptófano sólido (50 mmoles cada una) se añadieron a la reacción. Todo el triptófano añadido no se disolvió, pero la concentración se mantuvo en 50 mM o superior. Después de 3,5 horas se filtró el triptófano sólido. El análisis de la mezcla de reacción por LC/MS utilizando una cantidad definida de triptófano como patrón demostró que la concentración de triptófano en la solución era de 60,5 mM y la concentración de monatina era de 5,81 mM (1,05 g).

Se utilizaron los procedimientos siguientes para purificar el producto final. El 90% de la solución transparente se aplicó a una columna de resina AG50W-X8 de BioRad (225 ml; capacidad de enlace de 1,7 meq/ml). La columna se lavó con agua, recolectando fracciones de 300 ml hasta que la absorbancia a 280 nm fue < 5% de la primera fracción a través del flujo. La columna se eluyó a continuación con acetato amónico 1 M, pH 8,4, recolectando 4 fracciones de 300 ml. Las 4 fracciones contenían monatina y se evaporaron hasta 105 ml utilizando un roto-evaporador con un baño de agua tibia. Se formó un precipitado a medida que se reducía el volumen y se filtró a lo largo del proceso de evaporación.

Por análisis de las fracciones de la columna por LC/MS se demostró que el 99% del triptófano y de la monatina se unió a la columna. El precipitado que se formó durante el proceso de evaporación contenía > 97% de triptófano y < 2% de monatina. La relación del triptófano a producto en el sobrenadante fue aproximadamente de 2:1.

Se aplicó el sobrenadante (7 ml) a una columna de 100 ml de DEAE Fast Flow Sepharose (Amersham Biosciences) previamente convertida a la forma acetato lavando con 0,5 l de NaOH 1 M, 0,2 l de agua, 1,0 l de acetato amónico 1,0 M, pH 8,4 y 0,5 l de agua. El sobrenadante se cargó a razón de < 2 ml/min y la columna se lavó con agua a razón de 3 a 4 ml/min hasta que la absorbancia a 280 nm fue ~0. Se eluyó la monatina con acetato amónico 100 mM, pH 8,4, recogiendo 4 fracciones de 100 ml.

El análisis de las fracciones demostró que la relación de triptófano a monatina en las fracciones flujo a través era 81:15 y la relación en las fracciones del eluyente era 7:93. Suponiendo que el coeficiente de extinción a 280 nm de monatina es el mismo que el del triptófano, las fracciones eluyentes contenían 0,146 mmoles de producto. La extrapolación a 1 l de reacción total produciría ~2,4 mmoles (~710 mg) de monatina, para una recuperación del 68%.

Las fracciones eluyentes de la columna DEAE Sepharose se evaporaron hasta < 20 ml. Una alícuota del producto se purificó más durante la aplicación a una columna C_8 de preparación en fase inversa utilizando las mismas condiciones cromatográficas que las descritas en el Ejemplo 10 para la caracterización de la monatina a escala analítica. Se empleó el programa informático Fractionlynx™ de Waters para activar la recolección automática de fracciones de monatina basado en la detección de la $m/z = 293$ ion. Se recolectó la fracción procedente de la columna C_8 con el ión molecular correspondiente protonado para monatina, se evaporó a sequedad y a continuación se disolvió en un pequeño volumen de agua. Esta fracción se utilizó para la caracterización del producto.

El producto resultante se caracterizó utilizando los procedimientos siguientes.

Espectroscopía UV/visible. Las determinaciones espectroscópicas por UV/visible de monatina producidas mediante enzimas se llevaron a cabo utilizando un espectrofotómetro UV/visible 100 Bio de Cary. El producto purificado, disuelto en agua, presentó un máximo de absorción de 280 nm con un codo a 288 nm, características típicas de los compuestos que contienen indol.

Análisis LC/MS. Los análisis de las mezclas para derivados de monatina de las reacciones bioquímicas *in vitro* se realizaron tal como se describe en el Ejemplo 10. Un análisis LC/MS típico de monatina en una mezcla sintética enzimática *in vitro* se ilustra en la figura 5. El panel inferior de figura 5 ilustra un cromatograma iónico seleccionado para el ión molecular protonado de monatina a $m/z = 293$. Esta identificación de monatina en la mezcla fue corroborada por el espectro de masas ilustrados en la figura 6. El análisis del producto purificado por

LC/MS presentó un solo pico con un ión molecular de 293 y absorbancia a 280 nm. El espectro de masas fue idéntico al mostrado en la figura 6.

Análisis por MS/MS. Los experimentos con ión derivado por LC/MS/MS, tal como se describe en el Ejemplo 10, se realizaron también en monatina. Un espectro de masas de ión derivado de monatina se ilustra en la figura 7. Se hicieron asignaciones estructurales por tanteo de todos los fragmentos de iones marcados en la figura 7. Éstas incluyen los fragmentos de iones de $m/z = 275$ ($293 - H_2O$), 257 ($293 - (2 \times H_2O)$), 230 ($275 - COOH$), 212 ($257 - COOH$), 168 (ión carbonio de 3-buta-1,3-dienil-1*H*-indol), 158 (ión carbonio de 1*H*-indol-3-carbaldehído), 144 (ión carbonio de 3-etil-1*H*-indol), 130 (ión carbonio de 3-metileno-1*H*-indol) y 118 (ión indol carbonio). Muchos de éstos son los mismos que los obtenidos para PM (Ejemplo 4), como era de esperar si procedían de la fracción indol de la molécula. Algunos son 1 unidad de masa mayores que los observados para PM debido a la presencia de un grupo amino en lugar de una cetona.

Análisis MS de alta resolución. La figura 8 ilustra el espectro de masas obtenido para la monatina purificada empleando un espectrómetro de masas de cuadrupolo híbrido/tiempo de vuelo Q-Star de Applied Biosystems-Perkin Elmer. La masa medida para la monatina protonada utilizando triptófano como patrón de calibración de masa interno fue de 293,1144. La masa calculada de la monatina protonada basada en la composición elemental $C_{14}H_{17}N_2O_5$ es de 293,1137. Este es un error de medición de la masa inferior a 2 partes por millón (ppm), que proporciona pruebas concluyentes de la composición elemental de la monatina producida por vía enzimática.

Espectroscopía de RMN. Se realizaron experimentos de RMN en un instrumento de 500 MHz Inova de Varian. La muestra de monatina (~3 mg) se disolvió en 0,5 ml de D_2O . Inicialmente, el disolvente (D_2O) se utilizó como referencia interna a 4,78 ppm. Como el pico para el agua era grande, la 1H -RMN se desarrolló con supresión del pico para el agua. Posteriormente, debido a la amplitud del pico de agua, se utilizó el protón en el C-2 de monatina como pico de referencia, y se fijó en el valor publicado de 7,192 ppm.

Para la ^{13}C -RMN, una serie inicial de varios cientos de exploraciones indicó que la muestra estaba demasiado diluida para obtener un espectro de ^{13}C en el tiempo asignado. Por consiguiente, se realizó un experimento de coherencia cuántica múltiple heteronuclear (HMQC), que permitió la correlación de los hidrógenos y los carbonos a los que estaban unidos, y proporcionando además información de los desplazamientos químicos de los carbonos.

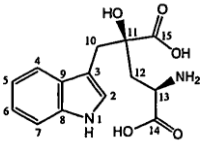
Un resumen de los datos de 1H y de HMQC se muestra en las Tablas 2 y 3. Por comparación con los valores publicados, los datos de la RMN indicaban que la monatina producida por vía enzimática era bien (S,S), (R,R) o una mezcla de ambos.

Análisis quiral por LC/MS. Para demostrar que el producto monatina *in vitro* era un isómero, y no una mezcla de enantiómeros (R,R) y (S,S) se realizaron análisis quirales LC/MS utilizando los instrumentos descritos en el Ejemplo 10.

Se realizaron separaciones quirales por LC utilizando una columna cromatográfica quiral Chirobiotic T (Advances Separations Technology) a temperatura ambiente. La separación y detección, basadas en los protocolos publicados del vendedor, se optimizaron para los isómeros R- (D) y S- (L) de triptófano. La fase LC constaba de A) agua que contenía ácido trifluoroacético al 0,05% (v/v); B) metanol que contenía ácido trifluoroacético al 0,05% (v/v). La elución fue isocrática en 70% de A y 30% de B. El caudal fue de 1,0 ml/min y la absorbancia de PDA se controló desde 200 nm hasta 400 nm. Los parámetros instrumentales utilizados para el análisis LC/MS quiral del triptófano y monatina son idénticos a los descritos en el Ejemplo 10 para el análisis LC/MS. Se utilizó la recolección de los espectros de masas para la zona m/z 150 a 400. Los cromatogramas iónicos seleccionados para los iones moleculares protonados ($[M+H]^+ = 205$ tanto para el R-triptófano como para el S-triptófano y $[M+H]^+ = 293$ para monatina) mostraban identificación directa de estos analitos en las mezclas.

Los cromatogramas de R-triptófano y S-triptófano y monatina, separados por cromatografía quiral y controlados por MS, se muestran en la figura 9. El único pico en el cromatograma de monatina indica que el compuesto es un isómero con un tiempo de retención casi idéntico al de S-triptófano.

Tabla 2: Datos de ^1H RMN

						
	Cargill		Vleggaar <i>et al.</i> ¹		Takeshi <i>et al.</i> ²	
Atom	δH	J(HH) Hz	δH	J(HH) Hz	δH	J(HH) Hz
2	7,192 (1H, s)		7,192 (s)		7,18 (s)	
4	7,671 (d)	7,99	7,686 (d)	7,9	7,67 (d)	8,0
5	7,104 (dd)	7,99	7,102 (dd)	8,0,8,0	7,11 (dd)	7,5, 7,5
6	7,178 (dd)	*	7,176 (dd)	8,0,8,0	7,17 (dd)	7,5, 7,5
7	7,439 (d)	7,99	7,439 (d)	8,1	7,43 (d)	8,0
10a	3,242(d)	14,5	3,243 (d)	14,3	3,24 (d)	14,5
10b	3,033 (d)	14,5	3,051 (d)	14,3	3,05 (d)	14,5
12	2,626 (dd)	15,5, 1,5	2,651 (dd)	15,3, 1,7	2,62 (dd)	15,5, 1,8
	2,015 (dd)	15,0, 12,0	2,006 (dd)	15,3, 11,7	2,01 (dd)	15,5, 12,0
13	3,571 (dd)	10,75*, 1,5	3,168 (dd)	11,6, 1,8	3,57 (dd)	12,0, 1,8

¹ Vleggaar *et al.* (J.C.S. Perkin Trans. 1:3095-8, 1992).² Takeshi and Shusuke (JP2002060382, 2002-02-26).Tabla 3: Datos de ^{13}C RMN (procedente del espectro de HMQC)

	Cargill	Vleggaar <i>et al.</i> ¹
Atom	δC	δC
2	126,1	126,03
3	*	110,31
4	120,4	120,46
5	120,2	120,25
6	122,8	122,74
7	112,8	112,79
8	*	137,06
9	*	129,23
10a	36,4	36,53
12	39,5	39,31
13	54,9	54,89
14	*	175,30
15	*	181,18

¹ Vleggaar *et al.* (J.C S. Perkin Trans. 1:3095-8, 1992).

Polarimetría. La rotación óptica se midió en un polarímetro Autopol III de Rudolph. La monatina se preparó como una solución de 14,6 mg/ml en agua. La rotación específica esperada ($[\alpha]_D^{20}$) para S,S-monatina (forma salina) es -49,6 para una solución de 1 g/ml en agua (Vleggaar *et al.*). La $[\alpha]_D^{20}$ observada fue -28,1 para la monatina purificada, producida por vía enzimática lo que indica que era el isómero S,S.

Mejoras

Se optimizaron las condiciones de reacción, incluyendo las concentraciones en reactivo y enzima, y se produjeron rendimientos de 10 mg/ml utilizando la mezcla reactiva siguiente:

acetato amónico 50 mM, pH 8,3, MgCl_2 2 mM, piruvato (sal de sodio o amonio) 200 mM, alfa-cetoglutarato (sal sódica) 5 mM, piridoxal fosfato 0,05 mM,

agua desaireada hasta conseguir un volumen final de 1 ml después de la adición de las enzimas,

fosfato potásico 3 mM, 50 $\mu\text{g/ml}$ de ProA aldolasa recombinante (extracto celular, concentración total de proteínas de 167 $\mu\text{g/ml}$), 1000 $\mu\text{g/ml}$ de L-aspartato aminotransferasa codificada por el gen *aspC* de *E. coli* (extracto celular, concentración total de proteínas de 2500 $\mu\text{g/ml}$) y triptófano sólido para proporcionar una concentración > 60 mM (saturado; algo sin disolver en toda la reacción). La mezcla se incubó a 30°C durante 4 horas con agitación o mezclado suave.

Sustituciones

La concentración de alfa-cetoglutarato puede reducirse a 1 mM y enriquecerse con aspartato 9 mM con un rendimiento de monatina equivalente. Pueden utilizarse receptores de aminoácidos alternativos en la primera etapa, tal como oxaloacetato.

Cuando se utiliza aminotransferasa de amplio sustrato de *L. major* recombinante en lugar de la L-aspartato aminotransferasa de *E. coli*, se consiguieron rendimientos similares de monatina. Sin embargo, se detectó también por análisis LC/MS un segundo producto no identificado (3 al 10% del producto principal) con una masa molecular de 292. Se produjeron concentraciones de monatina entre 0,1 y 0,5 mg/ml cuando se añadía la enzima codificada por *tyrB* de *E. coli*, la enzima codificada por *tatA* de *S. meliloti* o la glutámico-oxaloacético transaminasa del corazón de cerdo (tipo IIa) como aminotransferasa. Al comenzar la reacción a partir de indol-3-piruvato puede efectuarse una aminación reductora durante la última etapa con glutamato deshidrogenasa y NADH (como en el Ejemplo 7).

Las KHG aldolasas de *B. subtilis*, *E. coli* y *S. meliloti* se utilizaron también con la L-aspartato aminotransferasa de *E. coli* para producir monatina por vía enzimática. Se utilizaron las siguientes condiciones de reacción: $\text{NH}_4\text{-OAc}$ 50 mM, pH 8,3, MgCl_2 2 mM, piruvato 200 mM, glutamato 5 mM, piridoxal fosfato 0,05 mM, agua desaireada hasta conseguir un volumen final de 0,5 ml después de la adición de las enzimas, fosfato potásico 3 mM, 20 $\mu\text{g/ml}$ de KHG aldolasa recombinante de *B. subtilis* (purificada), aprox. 400 $\mu\text{g/ml}$ de L-aspartato aminotransferasa (AspC) de *E. coli* sin purificar del extracto celular e indol-3-piruvato 12 mM. Las reacciones se incubaron a 30°C durante 30 minutos con agitación. La cantidad de monatina producida utilizando la enzima de *B. subtilis* fue de 80 mg/ml, y aumentó al aumentar las cantidades de aldolasa. Si el indol-3-piruvato y el glutamato se sustituían saturando cantidades de triptófano y alfa-cetoglutarato 5 mM, la producción de monatina aumentaba hasta 360 ng/ml. Se repitieron las reacciones con 30 $\mu\text{g/ml}$ de cada una de las tres enzimas KHG en Tris 50 mM, pH 8,3, saturando cantidades de triptófano, y se dejaron procesar durante una hora para aumentar la detección. La enzima de *Bacillus* tenía actividad máxima como en el Ejemplo 4, produciendo aproximadamente 4.000 ng/ml de monatina. La KHG de *E. coli* produjo 3.000 ng/ml de monatina y la enzima de *S. meliloti* produjo 2.300 ng/ml.

EJEMPLO 7

Interconversión entre PM y monatina

La aminación de PM para formar monatina puede ser catalizada por aminotransferasas tales como las identificadas en los Ejemplos 1 y 6, o por deshidrogenasas que requieren un cofactor reductor tal como NADH o NADPH. Estas reacciones son reversibles y pueden medirse en ambas direcciones. La direccionalidad, al utilizar una enzima deshidrogenasa, puede controlarse en gran medida mediante la concentración de sales amónicas.

Actividad de deshidrogenasa. La desaminación oxidativa de la monatina se controló siguiendo el aumento de absorbancia a 340 nm ya que NAD(P)^+ se convirtió en más NAD(P)H cromofórico. La monatina se produjo por vía enzimática y se purificó como se describe en el Ejemplo 6.

Una mezcla analítica típica contenía Tris-HCl 50 mM, pH 8,0 a 8,9, NAD^+ o NADP^+ 0,33 mM, 2 a 22 unidades de glutamato deshidrogenasa (Sigma) y sustrato 10 a 15 mM en 0,2 ml. El análisis se realizó por duplicado en una placa de microvaloración transparente a UV, en un lector de placas SpectraMax Plus de Molecular Devices. Una mezcla de la enzima, tampón y una mezcla de NAD(P)^+ se pipetearon en los pocillos que contenían el sustrato y se controló el aumento de absorbancia a 340 nm a intervalos de 10 segundos después de un breve mezclado. Se incubó la reacción a 25°C durante 10 minutos. Se realizaron controles negativos sin adición de sustrato y se utilizó el glutamato como referencia positiva. La glutamato deshidrogenasa tipo III del hígado bovino (Sigma nº G-7882) catalizó la conversión de la monatina en el precursor de monatina a una velocidad de conversión de aproximadamente una centésima de la velocidad de conversión del glutamato en alfa-cetoglutarato.

Actividad de transaminación. Se realizaron ensayos de monatina aminotransferasa con la aspartato aminotransferasa (AspC) de *E. coli*, la tirosina aminotransferasa (TyrB) de *E. coli*, la aminotransferasa de amplio sustrato (BSAT) de *L. major* y las dos glutamato-oxaloacetato aminotransferasas porcinas disponibles en el mercado descritas en el Ejemplo 1. Tanto el oxaloacetato como el alfa-cetoglutarato se probaron como receptores amino. La mezcla de ensayo contenía (en 0,5 ml) Tris-HCl 50 mM, pH 8,0, PLP 0,05 mM, receptor amino 5 mM, monatina 5 mM y 25 μg de aminotransferasa. Los ensayos se incubaron a 30°C durante 30 minutos y las reacciones se interrumpieron mediante la adición de 0,5 ml de alcohol isopropílico. La pérdida de monatina fue controlada por LC/MS (Ejemplo 10). La mayor cantidad de actividad se apreció con BSAT de *L. major* con oxaloacetato como receptor de amino, seguido de la misma enzima con alfa-cetoglutarato como receptor amino. La actividad relativa con oxaloacetato fue: BSAT > AspC > porcino tipo IIa > porcino tipo I = TyrB. La actividad relativa con alfa-cetoglutarato fue: BSAT > AspC > porcino tipo I > porcino tipo IIa > TyrB.

EJEMPLO 8**Producción de monatina a partir de triptófano y otras fuentes de C3 aparte de piruvato**

Tal como se describió anteriormente en el Ejemplo 6, el indol-3-piruvato o el triptófano pueden convertirse en monatina utilizando piruvato como molécula de C3. Sin embargo, en algunas circunstancias el piruvato puede no ser una materia prima deseable. Por ejemplo, el piruvato puede ser más costoso que otras fuentes de carbono de C3 o puede tener efectos perjudiciales en las fermentaciones si se añade al medio. La alanina puede ser transaminada por muchas enzimas con PLP para producir piruvato.

Las enzimas de tipo triptofanasa realizan reacciones de beta-eliminación a velocidades más rápidas que otras enzimas con PLP tales como las aminotransferasas. Las enzimas de esta clase (4.1.99.-) pueden producir amoniaco y piruvato procedente de aminoácidos tales como L-serina, L-cisteína y derivados de serina y cisteína con grupos prominentes adecuados tales como O-metil-L-serina, O-bencil-L-serina, S-metil-cisteína, S-bencil-cisteína, S-alkil-L-cisteína, O-acil-L-serina o 3-cloro-L-alanina.

Los procedimientos para producir monatina que utilizan CE 4.1.99.-polipéptidos pueden ser mejorados mutando la β -tirosinasa (TPL) o triptofanasa según el procedimiento de Mouratou *et al.* (*J. Biol. Chem.*, 274:1320-5, 1999). Mouratou *et al.* describen la capacidad para convertir la β -tirosinasa en una aminoácido dicarboxílico β -liasa, que no se ha descrito por ser natural. El cambio de especificidad fue acompañado por la conversión de valina (V) 283 en arginina (R) y arginina (R) 100 en treonina (T). Estos cambios de aminoácidos permiten que la liasa acepte un aminoácido dicarboxílico para la reacción de desaminación hidrolítica (tal como aspartato). Por consiguiente, puede utilizarse también aspartato como fuente de piruvato para reacciones de condensación aldólica posteriores.

Además, las células o reactores enzimáticos pueden suministrarse con lactato y una enzima que convierte lactato en piruvato. Los ejemplos de enzimas que pueden catalizar esta reacción incluyen lactato deshidrogenasa y lactato oxidasa.

Aislamiento de ADN genómico

Los polipéptidos de la triptofanasa se han descrito anteriormente, por ejemplo en Mouratou *et al.* (*J.B.C.*, 274:1320-5, 1999). Para aislar los genes que codifican los polipéptidos de la triptofanasa se utilizó como plantilla ADN genómico procedente de DH10B de *E. coli* para la RCP tal como se describe en el Ejemplo 1.

El gen para la tirosina-fenol liasa se aisló de *C. freundii* (nº 8090 del catálogo ATCC, denominación ATCC 13316; NCTC 9750) y se cultivó en agar nutritivo (Difco 0001) y caldo de cultivo nutritivo (Difco 0003) a 37°C hasta una D.O. de 2,0. El ADN genómico se purificó utilizando el kit Genomic-tip™ 100/G de Qiagen.

Amplificación por RCP de secuencias de codificación

Se diseñaron cebadores con prominencias compatibles para el vector pET 30 Xa/LIC (Novagen, Madison, WI) tal como se describió anteriormente en el Ejemplo 1.

tna de *E. coli* (SEC. ID. nº 41). Cebador con terminal N para la clonación de pET30 Xa/LIC: 5'-GGT ATT GAG GGT CGC ATG GAA AAC TTT AAA CAT CT-3' (SEC. ID. nº 43). Cebador con terminal C para la clonación de pET30 Xa/LIC: 5'-AGA GGA GAG TTA GAG CCT TAA ACT TCT TTA AGT TTT G-3' (SEC. ID. nº 44).

tpl de *C. freundii* (SEC. ID. nº 42). Cebador con terminal N para la clonación de pET30 Xa/LIC: 5'-GGT ATT GAG GGT CGC ATGAATTATCCGGCAGAACC-3' (SEC. ID. nº 45). Cebador con terminal C para la clonación de pET 30 Xa/LIC: 5'-AGA GGA GAG TTA GAG CCTTAGATGTAATCAAAGCGTG-3' (SEC. ID. nº 46).

Se utilizó el ciclador térmico Eppendorf Mastercycler™ Gradient 5331 para todas las reacciones RCP. En 50 μ l se añadieron 0,5 μ g de plantilla (ADN genómico), 1,0 μ M de cada cebador, 0,4 mM de cada dNTP, 3,5 U de Expand High Fidelity Polymerase (Roche), 1 x tampón Expand con Mg, y DMSO al 5% (concentración final). El programa del termociclador para la RCP utilizado fue el siguiente: inicio en caliente a 96°C (5 minutos), 94°C - 30 segundos, 40-60°C - 1 minuto 45 segundos, 72°C - 2 minutos 15 segundos; 30 repeticiones. La etapa de polimerización final duró 7 minutos y las muestras se almacenaron entonces a 4°C.

Clonación

Para identificar los clones apropiados se utilizaron los procedimientos de clonación y de identificación de clones positivos detallados anteriormente en el Ejemplo 1.

Expresión génica y ensayos de actividad

Se subclonó el ADN plásmido (comprobado mediante análisis de las secuencias) en el hospedador de

expresión BL21(DE3) (Novagen). Se desarrollaron los cultivos en medio LB con 30 mg/l de kanamicina, se aislaron los plásmidos utilizando el kit miniprep de Qiagen y se analizaron mediante enzimas de restricción para confirmar la identidad.

Se realizaron experimentos de inducción con el hospedador de expresión BL21(DE3), los montajes se desarrollaron en medio LB que contenía 50 mg/l de kanamicina a 37°C. Se provocó la expresión de la proteína utilizando IPTG 0,1 mM una vez que la D.O.₆₀₀ del cultivo alcanzó aproximadamente 0,6. Las células se cultivaron durante 4 horas a 30°C y se recolectaron por centrifugación. Se lisaron a continuación las células en 5 ml/g de reactivo BugBuster™ (Novagen) en peso de células en húmedo que contenía 5 µl/ml de mezcla de inhibidor de proteasa serie III (Calbiochem) y 1 µl/ml de benzonasa nucleasa (Novagen) y las proteínas recombinantes activadas por His se purificaron utilizando los cartuchos His-Bind descritos anteriormente en el Ejemplo 1. Se desalaron las proteínas purificadas en una columna PD-10 (G25 Sephadex, Amersham Biosciences) y se eluyeron en tampón Tris-Cl 100 mM, pH 8,0. Se analizaron las proteínas por SDS-PAGE en geles con gradiente 4 al 15% para comprobar las concentraciones de proteína soluble al P.M. predicho de la proteína de fusión recombinante.

Mutagénesis

Algunos miembros del polipéptido clase 4.1.99.- (triptofanasa y β-tirosinasa) llevarán a cabo la reacción de la beta-liasa con aspartato o aminoácidos similares sin ninguna modificación. Sin embargo, puede ser necesario que algunos miembros de la clase experimenten mutagénesis para permitir la utilización de los sustratos y/o la creación del producto. Por otra parte, en algunos casos los polipéptidos que pueden realizar la conversión pueden optimizarse más por mutagénesis.

Se realizó la mutagénesis dirigida basándose en el análisis de estructura 3D de polipéptidos unidos a PLP. Dos ejemplos para cambiar la especificidad del sustrato de los polipéptidos se presentan a continuación.

Mutagénesis de la triptofanasa Ejemplo 1

El protocolo de la mutagénesis proporcionado a continuación introdujo dos mutaciones puntuales en la secuencia de aminoácidos. La primera mutación puntual cambió la arginina (R) en la posición 103 por treonina (T) y la segunda mutación puntual cambió la valina (V) en la posición 299 por arginina (R) (sistema de numeración para la proteína madura de *E. coli*). Se realizaron experimentos de mutagénesis por los laboratorios ATG (Eden Prairie, MN). Se introdujeron mutaciones sucesivamente por RCP de fragmentos génicos y el reensamblado de los fragmentos se llevó a cabo por RCP también. Cebadores para convertir arginina (R)103 a treonina (T): 5'-CCAGGGCACCGGCGCAGAGCAAATCTATATT-3' (SEC ID nº: 47) y 5'-TGCGCCGGTGCCCTGGTGTAGTCGGAATGGT-3' (SEC ID nº: 48).

Cebadores para convertir valina (V)299 en arginina (R): 5'-TCCTGCACGCGGCAAAGGGTTCTGCACTCGGT-3' (SEC ID nº: 49) y 5'-CTTTGCCGCGTGCAAGGCTTCCCGACA-3' (SEC ID nº: 50).

Los mutantes se identificaron por análisis con las enzimas de restricción Xba I/HindIII y SphI y se verificaron por secuenciado.

Mutagénesis de tirosina fenol liasa (β-tirosinasa) Ejemplo 2

Se realizaron dos mutaciones puntuales a la secuencia de aminoácidos de tirosina fenol liasa. Estas mutaciones convirtieron la arginina (R) en la posición 100 en treonina (T) y valina (V) en la posición 283 en arginina (R) (en la secuencia de la proteína madura de *C. freundii*).

Los cebadores para la conversión de R100T fueron: 5'-AGGGGACCGGCGCAGAAAACCTGTTATCG-3' (SEC ID nº: 51) y 5'-AGGGGACCGGCGCAGAAAACCTGTTATCG-3' (SEC ID nº: 52). Los cebadores para la conversión de V283R fueron: 5'-GTTAGTCCGCGTCTACGAAGGGATGCCAT-3' (SEC ID nº: 53) y 5'-GTAGACGCGGACTAACTTTGGCAGAAG-3' (SEC ID nº: 54).

Se utilizaron los procedimientos descritos anteriormente, y los clones se identificaron por digestión con KpnI/SacI y digestión con BstX I. Las secuencias fueron verificadas por secuenciado de la terminación de la cadena didesoxi. La proteína recombinante se produjo tal como se describió anteriormente para las enzimas naturales.

La mezcla de reacción constaba de Tris-Cl 50 mM pH 8,3, MgCl₂ 2 mM, fuente de carbono de C3 200 mM, sal sódica del alfa-cetoglutarato 5 mM, piridoxal fosfato 0,05 mM, agua desaireada para conseguir un volumen final de 0,5 ml tras la adición de las enzimas, fosfato potásico 3 mM, pH 7,5, 25 µg de ProA aldolasa de *C. testosteroni* recombinante en bruto preparada como en el Ejemplo 4, 500 µg de L-aspartato aminotransferasa (AspC) en bruto como se preparó en el Ejemplo 1 y triptófano sólido para dar una concentración > 60 mM (saturado; algo no disuelto a lo largo de la reacción). La mezcla de reacción se incubó a 30°C durante 30 minutos con mezclado. Se suministraron, serina, alanina y aspartato como fuentes 3 carbono C3. Los ensayos se realizaron con y sin enzimas con PLP secundarias (purificadas) capaces de llevar a cabo reacciones de eliminación en beta y de beta-liasa

(triptofanasa (TNA), triptofanasa doble mutante, β -tirosinasa (TPL)). Los resultados se presentan en la Tabla 4:

Tabla 4: Producción de monatina utilizando fuentes de carbono C3 alternativas

Fuente de carbono C3	Enzima con PLP adicional	Actividad relativa
ninguna	ninguna	0%
piruvato	ninguna	100%
serina	ninguna	3%
serina	11 μ g de TNA natural (1 U)	5,1%
serina	80 μ g de TNA doble mutante	4,6%
alanina	ninguna	32%
alanina	11 μ g de TNA natural	41,7%
alanina	80 μ g de TNA mutante	43,9%
aspartato	110 μ g de TNA natural (10 U)	7,7%
aspartato	5 U de TLP natural (en bruto)	5,1%
aspartato	80 μ g de TNA mutante	3,3%

La monatina producida a partir de alanina y serina como fuentes carbono C3 fue verificada por análisis LC/MS/MS modo *daughter scan* y fue idéntica a la monatina caracterizada producida en el Ejemplo 6. La alanina fue la mejor alternativa probada y fue transaminada por la enzima AspC. La cantidad de monatina producida aumentó mediante la adición de la triptofanasa, que puede realizar la transaminación como actividad secundaria. La cantidad de monatina producida con la serina como fuente de carbono, se duplicó casi con la adición de las enzimas triptofanasa, aún cuando solamente una quinta parte de la cantidad de triptofanasa en comparación con la aminotransferasa. AspC puede tener alguna cantidad de actividad de beta-eliminación sola. Los resultados con aspartato indican que la actividad de triptofanasa en aspartato no aumenta con las mismas mutaciones dirigidas que se sugirieron anteriormente para la β -tirosinasa. Es de esperar que la β -tirosinasa mutante presenta una mayor actividad para la producción de monatina.

EJEMPLO 9

Síntesis química de monatina

La adición de alanina al ácido indol-3-pirúvico produce monatina y esta reacción puede llevarse a cabo por síntesis con un reactivo de Grignard u organolitio.

Por ejemplo, a la 3-cloro- o 3-bromo-alanina que se ha bloqueado de manera apropiada en los grupos carboxilo y amino, se añade magnesio en condiciones anhidras. Se añade a continuación indol-3-piruvato (bloqueado de manera apropiada) para formar el producto acoplado seguido de eliminación de los grupos protectores para formar monatina. Los grupos protectores que son particularmente útiles incluyen THP (éter tetrahidropiranílico) que se acopla y se elimina fácilmente.

EJEMPLO 10

Detección de monatina y PM

Este ejemplo describe los procedimientos utilizados para detectar la presencia de monatina o de su precursor el 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico.

Análisis LC/MS

Los análisis de las mezclas para la forma alfa-cetoácido de monatina (precursor de monatina, PM) y de la monatina procedente de las reacciones bioquímicas *in vitro* o *in vivo* se realizaron utilizando un instrumento de espectrometría de masas en tándem con cromatografía líquida (LC/MS/MS) Micromass de Waters que incluye un cromatógrafo para líquidos 2690 de Waters con un monitor de absorbancia 996 Photo-Diode Array (PDA) de Waters colocado en serie entre el cromatógrafo y un espectrómetro de masas triple cuadrupolo Quattro Ultima de Micromass. Se realizaron separaciones por LC utilizando una columna de cromatografía en fase inversa Supelco Discovery C₁₈, 2,1 mm x 150 mm o una columna de cromatografía en fase inversa Xterra MS C₈, 2,1 mm x 250 mm, a temperatura ambiente. La fase móvil de la LC constaba de A) agua conteniendo 0,05% (v/v) de ácido trifluoroacético y B) metanol conteniendo 0,05% (v/v) de ácido trifluoroacético.

La elución en gradiente fue lineal desde el 5% B hasta el 35% B, 0 a 9 min, lineal desde 35% B hasta el 90% B, 9 a 16 min, isocrática al 90% B, 16 a 20 min, lineal desde el 90% B hasta el 5% B, 20 a 22 min, con un periodo de reequilibrado de 10 min entre las series. El caudal fue de 0,25 ml/min y la absorbancia de PDA se controló desde 200 nm hasta 400 nm. Todos los parámetros de la ESI-MS se optimizaron y seleccionaron basándose en la generación de iones moleculares protonados ([M+H]⁺) de los analitos de interés, y la producción

de iones de fragmentos característicos.

Se utilizaron los parámetros instrumentales siguientes para el análisis por LC/MS de monatina: capilar: 3,5 kV; cono: 40 V; Hex 1: 20 V; abertura: 0 V; Hex 2: 0 V; temperatura de la fuente: 100°C; temperatura de desolvatación: 350°C; gas de desolvatación: 500 l/h; cono de gas: 50 l/h; resolución baja de masas (Q1): 15,0; resolución alta de masas (Q1): 15,0; energía iónica: 0,2; entrada: 50 V; energía de colisión: 2; salida: 50 V; baja resolución de masas (Q2): 15; alta resolución de masas (Q2): 15; energía iónica (Q2): 3,5; multiplicador: 650. Las incertidumbres para las relaciones de masa/carga indicadas (m/z) y las masas moleculares son $\pm 0,01\%$. La detección inicial en las mezclas de la forma alfa-cetoácido de monatina (PM) y de monatina se llevó a cabo por seguimiento por LC/MS con recolección de los espectros de masas para la zona m/z 150-400. Los cromatogramas iónicos seleccionados para los iones moleculares protonados ($[M+H]^+$ = 292 para PM, $[M+H]^+$ = 293 para monatina) permitieron la identificación directa de estos analitos en las mezclas.

Análisis MS/MS

Se realizaron experimentos con ión derivado por LC/MS/MS en monatina de la forma siguiente. El análisis del ión derivado implica la transmisión del ión original (por ejemplo, m/z = 293 para monatina) de interés desde el primer analizador de masas (Q1) en la celda de colisión del espectrómetro de masas, donde se introduce el argón y se disocia químicamente el original en fragmentos de iones (derivados). Estos fragmentos de iones se detectan a continuación con el segundo analizador de masas (Q2) y pueden utilizarse para corroborar la asignación estructural del original.

Se utilizaron los siguientes parámetros instrumentales para el análisis LC/MS/MS de monatina: capilar: 3,5 kV; cono: 40 V; Hex 1: 20 V; abertura: 0 V; Hex 2: 0 V; temperatura de la fuente: 100°C; temperatura de desolvatación: 350°C; gas de desolvatación: 500 l/h; cono de gas: 50 l/h; resolución baja de masas (Q1): 13,0; resolución alta de masas (Q1): 13,0; energía iónica: 0,2; entrada: -5 V; energía de colisión: 1,4; salida: 1 V; baja resolución de masas (Q2): 15; alta resolución de masas (Q2): 15; energía iónica (Q2): 3,5; multiplicador: 650.

Determinación por alta resolución de monatina

Se realizaron análisis de alta resolución (< 5 min/muestra) de las mezclas para monatina procedente de las reacciones *in vitro* o *in vivo* utilizando la instrumentación descrita anteriormente, y los mismos parámetros que los descritos para LC/MS/MS. Se hicieron separaciones por LC utilizando la cromatografía de C_8 (2,1 mm x 50 mm) Xterra MS de Waters a temperatura ambiente con elución isocrática en MeOH acuoso al 15%, ácido acético al 0,25% a un caudal de 0,3 ml/min. La detección de monatina en las muestras se realizó utilizando espectrometría de masas en tándem con seguimiento de la reacción seleccionado (SRM). Esto implicaba el seguimiento del ión original provocado por colisiones específicas ($[M+H]^+$ = 293,1) a transiciones del ión derivado (por ejemplo, el fragmento iónico a m/z = 168,1, asignado provisionalmente como ión 3-buta-1,3-dienil-1H-indol carbonio) para maximizar la sensibilidad, selectividad y el rendimiento para la detección de monatina. Los datos de absorbancia del PDA se recogieron en paralelo para la verificación adicional de la identidad de la monatina.

EJEMPLO 11

Producción de monatina en bacterias

Este ejemplo describe los procedimientos utilizados para producir monatina en células de *E. coli*. Un experto en la materia apreciará que pueden utilizarse procedimientos similares para producir monatina en otras células bacterianas. Además, los pueden utilizarse vectores que contienen otros genes en la serie de reacciones de la síntesis de monatina (figura 2).

El medio con Trp-1+glucosa, medio mínimo que ha sido utilizado para la producción de triptófano en células de *E. coli* (Zeman *et al.*, *Folia Microbiol.* 35:200-4, 1990), se preparó de la forma siguiente. A 700 ml de agua nanopura se añadieron los siguientes reactivos: 2 g de $(NH_4)_2SO_4$, 13,6 g de KH_2PO_4 , 0,2 g de $MgSO_4 \cdot 7H_2O$, 0,01 g de $CaCl_2 \cdot 2H_2O$ y 0,5 mg de $FeSO_4 \cdot 7H_2O$. Se ajustó el pH a 7,0, se aumentó el volumen a 850 ml y se esterilizó el medio en autoclave. Se preparó una solución de glucosa al 50% por separado y se filtró estéril. Se agregaron cuarenta ml al medio básico (850 ml) para un volumen final de 1 l.

Se preparó una solución de 10 g/l de L-triptófano en fosfato sódico 0,1 M pH 7 y se filtró estéril. Un décimo del volumen se agregó por lo general a los cultivos como se especifica a continuación. Una solución de piruvato sódico al 10% se preparó también y se filtró estéril. Se utilizó por lo general una alícuota de 10 ml por litro de cultivo. Se prepararon soluciones madre de ampicilina (100 mg/ml), kanamicina (25 mg/ml) e IPTG (840 mM), se filtraron estériles y se almacenaron a -20°C antes de su uso. Se utilizó Tween 20 (polioxietileno 20-monolaurato de sorbitán) a una concentración final del 0,2% (vol/vol). Se utilizó ampicilina a concentraciones no letales, por lo general 1 a 10 µg/ml de concentración final.

Se prepararon placas recientes de *E. coli*, BL21(DE3)::C. *testosteroni* proA/pET 30 Xa/LIC (descritas en el

Ejemplo 4) en medio LB que contenía 50 µg/ml de kanamicina. Los cultivos de toda la noche (5 ml) de una sola colonia se inocularon y se cultivaron a 30°C en medio LB con kanamicina. Por lo general se utilizaron 1 a 50 inóculos para la inducción en el medio con trp-1+glucosa. Se agregó antibiótico reciente hasta una concentración final de 50 mg/l. En matraces oscilantes se cultivaron a 37°C antes de la inducción.

Se tomaron muestras de las células cada hora hasta que se obtuvo una D.O.₆₀₀ entre 0,35 y 0,8. Se indujeron a continuación las células con IPTG 0,1 mM, y se redujo la temperatura a 34°C. Se recolectaron muestras (1 ml) antes de la inducción (punto de tiempo cero) y se centrifugaron a 5000 x g. Se congeló el sobrenadante a -20°C para el análisis LC/MS. Cuatro horas después de la inducción, se recolectó otra muestra de 1 ml y se centrifugó para separar el caldo del cultivo del sedimento celular. Se añadieron triptófano, piruvato sódico, ampicilina y Tween tal como se describió anteriormente.

Se cultivaron las células durante 48 horas tras la inducción y se tomó otra muestra de 1 ml y se preparó como anteriormente. A las 48 horas, se añadió otra alícuota de triptófano y piruvato. El volumen del cultivo completo se centrifugó después de aproximadamente 70 horas de cultivo (después de la inducción), durante 20 minutos a 4°C y 3500 rpm. Se decantó el sobrenadante y tanto el caldo de cultivo como las células se congelaron a -80°C. Las fracciones de caldo de cultivo se filtraron y analizaron por LC/MS. Las alturas y las áreas de los picos $[M+H]^+ = 293$ se controlaron como se describe en el Ejemplo 10. Se sustrajo el nivel de fondo del medio. Los datos se normalizaron también para el cultivo celular representando la altura del pico $[M+H]^+ = 293$ dividido por la densidad óptica del cultivo a 600 nm.

Se produjeron niveles superiores de monatina cuando el piruvato, la ampicilina y Tween se añadieron 4 horas después de la inducción en lugar de en la inducción. Otros aditivos tales como PLP, fosfato adicional o MgCl₂ adicional no aumentaron la producción de monatina. Los títulos mayores de monatina se obtuvieron cuando se utilizaba triptófano en lugar de indol-3-piruvato, y cuando se añadía triptófano después de la inducción en lugar de en la inoculación, o en la inducción. Antes de la inducción y 4 horas después de la inducción (en el momento de la adición del sustrato), no existía por lo general ningún nivel detectable de monatina en el caldo de fermentación o en los extractos celulares. Se realizaron controles negativos utilizando células con vector pET30a solo, así como cultivos donde no se había añadido triptófano ni piruvato. Una ecografía por MS del original demostró que el compuesto con $(m+1)/z = 293$ no procedía de moléculas mayores, y las ecografías de los derivados realizadas (como en el Ejemplo 10) fueron similares a la monatina preparada *in vitro*.

Se estudió el efecto de Tween utilizando concentraciones de 0, 0,2% (vol/vol) y 0,6% final de Tween-20. La mayor cantidad de monatina producida por los matraces oscilantes fue en Tween al 0,2%. La concentración de ampicilina varió entre 0 y 10 µg/ml. La cantidad de monatina en el caldo de cultivo celular aumentó rápidamente (2,5 veces) entre 0 y 1 µg/ml y aumentó 1,3 veces cuando la concentración de ampicilina aumentaba desde 1 hasta 10 µg/ml.

En la figura 10 se muestra un experimento a lo largo del tiempo que presenta resultados típicos. La cantidad de monatina segregada en el caldo de cultivo celular aumentó, aun cuando los valores están normalizados para el crecimiento celular. Utilizando el coeficiente de extinción molar del triptófano la cantidad de monatina en el caldo de cultivo se estimó que era inferior a 10 µg/ml. El mismo experimento se repitió con las células que contenían el vector sin la inserción *proA*. Muchas de las cifras eran negativas, lo que se indica por la altura del pico a $m/z = 293$ que era inferior en estos cultivos que en el medio solo (figura 10). Las cifras eran habitualmente menores cuando faltaban el triptófano y el piruvato, lo que demuestra que la producción de monatina es un resultado de una reacción enzimática catalizada por la enzima aldolasa.

La producción *in vivo* de monatina en las células bacterianas se repitió en experimentos en matraz oscilante de 800 ml y en fermentadores. Una muestra de 250 ml de monatina (en caldo de cultivo sin células) se purificó por cromatografía de intercambio aniónico y cromatografía líquida de preparación en fase inversa. Esta muestra se evaporó y se sometió a análisis de espectroscopia de masas de alta resolución (descrito en el Ejemplo 6). La MS de alta resolución indicó que el metabolito que se produce es la monatina.

Los ensayos *in vitro* indican que la aminotransferasa necesita estar presente a concentraciones más altas que la aldolasa (véase el Ejemplo 6), por consiguiente la aspartato aminotransferasa de *E. coli* se sobreexpresó en combinación con el gen para aldolasa para aumentar la cantidad de monatina producida. Se diseñaron cebadores para introducir *proA* de *C. testosteroni* en un operón con *aspC*/pET30 Xa/LIC, de la manera siguiente: cebador 5': ACTCGGATCCGAAGGAGATATACATATGTACGAAGTGGGACT (SEC. ID. n° 67) y cebador en 3': CGGCTGTGACCGTTAGTCAATATATTTTCAGGC (SEC. ID. n° 68). El cebador en 5' contiene una secuencia BamHI, el cebador en 3' contiene una secuencia Sall para clonación. Se realizó la RCP tal como se describe en el Ejemplo 4, y se purificó en gel. El montaje *aspC*/pET30 Xa/LIC se digirió con BamHI y Sall dado que era el producto de la RCP. Se purificaron las enzimas utilizando una columna de centrifugación de Qiagen. El producto *proA* de la RCP se ligó al vector utilizando el kit de ligadura rápida de ADN de Roche (Indianapolis, IN) según las instrucciones del fabricante. Se hicieron transformaciones químicas utilizando Novablues Singles (Novagen) tal como se describe en el Ejemplo 1. Se cultivaron las colonias en medio LB que contenía 50 mg/l de kanamicina y se purificó el ADN plásmido utilizando el kit spin mimiprep de Qiagen. Se cribaron los clones por análisis con enzimas de restricción y se confirmó la secuencia mediante Seqwright (Houston, TX). Se subclonaron los montajes en BLR(DE3),

BLR(DE3)pLysS, BL21(DE3) y BL21(DE3)pLysS (Novagen). El montaje *proA*/pET30 Xa/LIC se transformó también en BL21(DE3)pLysS.

Las comparaciones iniciales de las muestras del matraz oscilante BLR(DE3) en las condiciones normales descritas anteriormente, demostraron que la adición del segundo gen (*aspC*) mejoró la cantidad de monatina producida en siete veces. Para activar el crecimiento, se utilizaron cepas hospedadoras procedentes de BL21(DE3). Los clones de *proA* y dos clones del operón del gen se indujeron en medio de Trp-1, los hospedadores pLysS tenían cloranfenicol (34 mg/l) añadido al medio también. Los experimentos en matraz en agitación se realizaron con y sin adición de Tween-20 al 0,2% y 1 mg/l de ampicilina. Se calculó la cantidad de monatina en el caldo de cultivo utilizando monatina *in vitro* purificada producida como patrón. Se realizaron análisis SRM tal como se describió en el Ejemplo 10. Se tomaron muestras de células a las cero, 4 horas, 24 horas, 48 horas, 72 horas y 96 horas de cultivo.

Los resultados se presentan en la Tabla 5 para las cantidades máximas producidas en los caldos de cultivo. En la mayoría de los casos los dos montajes génicos proporcionaron valores mayores que el montaje de *proA* solo. Las cepas de pLysS, que deberían tener envolturas celulares más permeables, tenían concentraciones mayores de monatina segregada, aun cuando estas cepas por lo general se desarrollan a velocidad más lenta. Las adiciones de Tween y ampicilina fueron beneficiosas.

Tabla 5: Cantidad de monatina producida por bacterias *E. coli*

Montaje	Hospedador	Tween + Amp	µg/ml de monatina	Tiempo
<i>proA</i>	BL21(DE3)	-	0,41	72 h
<i>proA</i>	BL21(DE3)	+	1,58	48 h
<i>proA</i>	BL21(DE3)pLysS	-	1,04	48 h
<i>proA</i>	BL21(DE3)pLysS	+	1,60	48 h
<i>aspC:proA</i>	BL21(DE3)	-	0,09	48 h
<i>aspC:proA</i>	BL21(DE3)	+	0,58	48 h
<i>aspC:proA</i>	BL21(DE3)pLysS	-	1,39	48 h
<i>aspC:proA</i>	BL21(DE3)pLysS	+	6,68	48 h

EJEMPLO 12

Producción de monatina en levadura

Este ejemplo describe los procedimientos utilizados para producir monatina en células eucarióticas. Un experto en la materia apreciará que pueden utilizarse procedimientos similares para producir monatina en cualquier célula de interés. Además, pueden utilizarse otros genes (por ejemplo, los listados en la figura 2) además de, o alternativamente a los descritos en este ejemplo.

El pESC Yeast Epitope Tagging Vector System (Stratagene, La Jolla, CA) se utilizó para clonar y expresar los genes *aspC* de *E. coli* y *proA* de *C. testosteroni* en *Saccharomyces cerevisiae*. Los vectores pESC contienen los activadores tanto GAL1 como el GAL10 en cadenas opuestas, con dos secuencias de clonación múltiples distintas, que permiten la expresión de dos genes al mismo tiempo. El vector pESC-His contiene también el gen *His3* para la complementación de la auxotrofia de histidina en el hospedador (YPH500). Los activadores GAL1 y GAL10 son reprimidos por la glucosa e inducidos por la galactosa; se utiliza una secuencia Kozak para la expresión óptima en la levadura. Los plásmidos pESC son vectores lanzadera, que permiten que se haga el montaje inicial en *E. coli* (con el gen *bla* para selección); sin embargo, ninguna de las secuencias de fijación del ribosoma bacteriano están presentes en las múltiples secuencias de clonación.

Se diseñaron los cebadores siguientes para la clonación en pESC-His (las secuencias de restricción están subrayadas, la secuencia Kozak está en negrita): *aspC* (BamHI/SalI), GAL1:

5'-CGCGGATCCATAATGGTTGAGAACATTACCG-3' (SEC ID n°: 69) y
 5'-ACGCGTCGACTTACAGCACTGCCACAATCG-3' (SEC ID n°: 70). *proA* (EcoRI/NotI), GAL10:
 5'-CCGGAATTCATAATGGTCTGAAGTGGGAGTTGT-3' (SEC ID n°: 71) y
 5'-GAATGCGGCCGCTTAGTCAATATATTCAGGCC-3' (SEC ID n°: 72).

El segundo codón para ambas proteínas maduras se cambió desde un aminoácido aromático a la valina debido a la introducción de la secuencia Kozak. Los genes de interés se ampliaron utilizando ADN pET30 Xa/LIC miniprep procedente de los clones descritos en los Ejemplos 1 y el Ejemplo 4 como plantilla. Se llevó a cabo la RCP utilizando el termociclador de gradiente del ciclador Eppendorf Master y el protocolo siguiente para una reacción de 50 µl: 1,0 µl de plantilla, 1,0 µM de cada cebador, 0,4 mM de cada dNTP, 3,5 U de Expand High Fidelity Polymerase (Roche, Indianapolis, IN) y 1 x tampón ExpandTM con Mg. El programa del termociclador utilizado consistía en un arranque en caliente a 94°C durante 5 minutos, seguido de 29 repeticiones en las etapas siguientes: 94°C durante 30 segundos, 50°C durante 1 minuto 45 segundos y 72°C durante 2 minutos 15 segundos. Después de las 29 repeticiones se mantuvo la muestra a 72°C durante 10 minutos y a continuación sea almacenó a 4°C. Se purificaron

los productos de la RCP por separación en un gel de TAE-agarosa al 1% seguido de recuperación utilizando un kit de extracción en gel QIA quick (Qiagen, Valencia, CA).

El ADN del vector pESC-His (2,7 µg) se digirió con *Bam*HI/*Sa*II y se purificó en gel como anteriormente. El producto *aspC* de la RCP se digirió con *Bam*HI/*Sa*II y se purificó con una columna de purificación por RCP de QIAquick. Se realizaron ligaduras con el kit de ligadura rápida de ADN de Roche siguiendo los protocolos del fabricante. Se electroporaron las ligaduras desaladas en 40 µl de células competentes DH10B de Electromax (Invitrogen) en una cubeta disponible BioRad de 0,2 cm utilizando un Gene Pulser II de BioRad con controlador de pulsaciones plus, según las instrucciones del fabricante. Después de 1 hora de recuperación en 1 ml de medio SOC, los transformados se colocaron en placas en medio LB que contenía 100 µg/ml de ampicilina. Las preparaciones del ADN plásmido para los clones se realizaron utilizando los kit Spin Miniprep de QIAprep. El ADN plásmido se identificó mediante enzimas de restricción, y se secuenció (Seqwright) para verificación utilizando los cebadores diseñados para el vector.

El clon *aspC*/pESC-His se digirió con *Eco*RI y *Not*I, ya que era el producto *proA* de la RCP. El ADN se purificó como anteriormente y se ligó como anteriormente. Los dos montajes génicos se transformaron en células DH10B y se cribaron mediante enzimas de restricción y secuenciado de ADN.

El montaje se transformó en la cepa YPH500 de *S. cerevisiae* utilizando el kit S.c. EasyComp™ Transformation (Invitrogen). Las reacciones de transformación se colocaron en medio mínimo SC-His (Manual de pYES2 de Invitrogen) que contenía glucosa al 2%. En cada una de las colonias de levadura se identificó la presencia de los genes *proA* y *aspC* por RCP de las colonias utilizando los cebadores de RCP anteriores. Las células sedimentadas (2 µl) se pusieron suspensión en 20 µl de tampón Y-Lysis (Zymo Research) que contenía 1 µl de zimolasa y se calentó a 37°C durante 10 minutos. Se utilizaron a continuación 4 µl de esta suspensión en una reacción RCP de 50 µl utilizando la mezcla de reacción RCP y el programa descrito anteriormente.

Se cultivaron durante la noche cultivos de cinco ml en SC-His + glucosa a 30°C y 225 rpm. Se ajustaron las células gradualmente al cultivo en rafinosa con objeto de minimizar el periodo de inactividad antes de la inducción con galactosa. Después de aproximadamente 12 horas de cultivo, se tomaron mediciones de absorbancia a 600 nm, y se sedimentó un volumen de células apropiado y se volvió a poner en suspensión para dar una D.O. de 0,4 en el medio SC-His reciente. Se utilizaron las siguientes fuentes de carbono sucesivamente: rafinosa al 1% + glucosa al 1%, glucosa al 0,5% + rafinosa al 1,5%, rafinosa al 2% y por último rafinosa al 1% + galactosa al 2% para la inducción.

Después de aproximadamente 16 horas de cultivo en el medio de inducción, los cultivos de 50 ml se dividieron en cultivos de 25 ml por duplicado, y se añadió lo siguiente a solamente uno de los duplicados: (concentraciones finales) 1 g/l de L-triptófano, fosfato sódico 5 mM pH 7,1, 1 g/l de piruvato sódico, MgCl₂ 1 mM. Se conservaron como referencias negativas las muestras de los caldos de cultivo y de los sedimentos celulares procedentes del medio sin inducción, y de los cultivos de 16 horas antes de la adición de sustratos para la serie de reacciones de monatina. Además, los montajes que contienen un gen *aspC* funcional, (y un gen *proA* truncado) se utilizaron como otra referencia negativa. Se dejaron cultivar las células durante un total de 69 horas después de la inducción. Ocasionalmente las células de levadura se indujeron a una D.O. inferior, y solamente se cultivaron durante 4 horas antes de la adición de triptófano y piruvato. Sin embargo, los sustratos de monatina parecen inhibir el crecimiento y la adición a D.O. mayor fue más eficaz.

Los sedimentos celulares procedentes de los cultivos se lisaron con 5 ml de YeastBuster™ más 50 µl THP(Novagen) por gramo (de peso húmedo) de las células siguiendo los protocolos del fabricante, con adición de inhibidores de plaza y benzonasa nucleasa como se describe en los ejemplos anteriores. El caldo de cultivo y los extractos celulares se filtraron y analizaron por SRM se tal como se describe en el Ejemplo 10. Utilizando este procedimiento, no se detectó monatina en ambas muestras, lo que indica que las células podrían no segregar monatina en estas condiciones. La fuerza protónica del motivo puede ser insuficiente en estas condiciones o los transportadores de aminoácidos generales pueden saturarse con triptófano. La expresión de la proteína no estaba a un nivel que permita la detección de cambios utilizando SDS-PAGE.

La monatina fue detectable (aproximadamente 60 µg/ml) temporalmente en extractos celulares del cultivo con dos genes funcionales, cuando se agregaron triptófano y piruvato al medio. No se detectó monatina en ninguno de los extractos celulares con referencia negativa. Los ensayos *in vitro* para monatina se realizaron por duplicado con 4,4 mg/ml de proteína completa (aproximadamente el doble de lo que se utiliza por lo general para los extractos celulares de *E. coli*) utilizando el ensayo optimizado descrito en el Ejemplo 6. Se realizaron otros ensayos con adición de 32 µg/ml de ProA aldolasa de *C. testosteroni* o 400 µg/ml de AspC aminotransferasa, para determinar qué enzima era limitativa en el extracto celular. Se realizaron controles negativos sin adición de enzima, o con adición de solamente AspC aminotransferasa (la condensación aldólica puede ocurrir en alguna medida sin enzima). Se realizaron controles positivos con enzimas parcialmente puras (30 a 40%), utilizando 16 µg/ml y 400 µg/ml de aminotransferasa.

Se analizaron los resultados *in vitro* por SRM. Los análisis de los extractos celulares demostraron que el triptófano se transportaba eficazmente en las células cuando se añadía al medio después de la inducción, dando

como resultado concentraciones de triptófano de dos órdenes de magnitud mayores que aquellas a las que no se añadió triptófano adicional. Los resultados de los análisis de monatina *in vitro* se presentan en la Tabla 6 (las cifras están indicadas en ng/ml).

5

Tabla 6: Producción de monatina con extractos de células de levadura

	Montaje aspC	+ aldolasa	+ AspC	Montaje de dos genes	+ aldolasa	+ AspC
Reprimido (medio de glucosa)	0	888,3	173,5	0	465,2	829
Inducido (24 h)	0	2832,8	642,4	0	1375,6	9146,6
Inducido (69 h)	0	4937,3	340,3	71,9	1652,8	23693,5
69 h + sus.	0	556,9	659,1	21,9	755,6	16688,2
+ referencia (enzimas purificadas)	21853			21853		
Control sin enzimas	0		254,3	0		254,3

10

Se obtuvieron resultados positivos con los extractos celulares de dos genes completos con y sin sustrato añadido de medio de cultivo. Estos resultados, en comparación con las referencias positivas, indican que las enzimas se expresaron a concentraciones próximas al 1% de la proteína total en levadura. La cantidad de monatina producida cuando el extracto celular del montaje aspC (con *ProA* truncado) se ensayó con aldolasa era significativamente mayor que cuando los extractos se ensayaban solos, e indica que la AspC aminotransferasa recombinante comprende aproximadamente del 1 al 2% de la proteína total de la levadura. Los extractos celulares de cultivos no inducidos tenían una cantidad mayor de actividad cuando se ensayaban con aldolasa debido a la presencia de aminotransferasas naturales en las células. Cuando se ensaya con AspC aminotransferasa, la actividad de los extractos de las células no inducidas aumentó hasta la cantidad de monatina producida por la referencia negativa AspC (aprox. 200 ng/ml). En cambio, la actividad observada cuando se ensaya el extracto celular del montaje de dos genes aumenta más cuando la aminotransferasa está enriquecida que cuando se añade la aldolasa. Ya que ambos genes deberían expresarse al mismo nivel, esto indica que la cantidad de monatina producida se maximiza cuando la concentración de aminotransferasa es mayor que la de aldolasa, de acuerdo con los resultados mostrados en el Ejemplo 6.

15

20

25

La adición de piruvato y triptófano no solamente inhibe el crecimiento celular, sino que aparentemente inhibe la expresión de la proteína también. La adición del plásmido pESC-Trp puede utilizarse para corregir la auxotrofia del triptófano de las células hospedadoras YPH500, para proporcionar un medio de suministrar triptófano con efectos menores sobre el crecimiento, la expresión y la secreción.

EJEMPLO 13

30

Mejora de los procesos enzimáticos utilizando reacciones acopladas

35

En teoría, si no se producen reacciones secundarias o degradación de sustratos o productos intermedios, la cantidad máxima de producto formada a partir de la reacción enzimática ilustrada en la figura 1 es directamente proporcional a las constantes de equilibrio de cada reacción, y las concentraciones de triptófano y piruvato. El triptófano no es un sustrato muy soluble, y las concentraciones de piruvato mayores de 200 mM parecen tener un efecto negativo en el rendimiento (véase el Ejemplo 6).

40

45

Lo ideal es que la concentración de monatina esté maximizada con respecto a los sustratos, con objeto de disminuir el coste de separación. Las separaciones físicas pueden realizarse de modo que la monatina se retire de la mezcla de reacción, evitando que ocurran reacciones inversas. Las materias primas y los catalizadores pueden regenerarse a continuación. Debido a la similitud de tamaño, carga e hidrofobia de la monatina para varios de los reactivos y compuestos intermedios, las separaciones físicas serán difíciles a menos que exista una gran cantidad de afinidad por la monatina (tal como una técnica de cromatografía por afinidad). Sin embargo, las reacciones de monatina pueden acoplarse a otras reacciones de modo que el equilibrio del sistema se desplace hacia la producción de monatina. Los siguientes son ejemplos de procedimientos para mejorar el rendimiento de monatina obtenida a partir de triptófano o indol-3-piruvato.

Reacciones acopladas que utilizan oxaloacetato descarboxilasa (CE 4.1.1.3)

50

La figura 11 es una ilustración de la reacción. La triptófano oxidasa y la catalasa se utilizan para conducir la reacción en la dirección de la producción de indol-3-piruvato. La catalasa se utiliza en exceso de modo que el peróxido de hidrógeno no está disponible para reaccionar en la dirección inversa o para dañar las enzimas o compuestos intermedios. El oxígeno se regenera durante la reacción con catalasa. Alternativamente, puede utilizarse como sustrato indol-3-piruvato.

55

El aspartato se utiliza como donante de amino para la aminación de PM, y se utiliza una aspartato

aminotransferasa. Lo ideal es que se utilice una aminotransferasa que tiene una baja especificidad por la reacción de triptófano/indol-3-piruvato en comparación con la PM para la reacción de monatina de modo que el aspartato no se utiliza para reaminar el indol-3-piruvato. La oxaloacetato descarboxilasa (procedente de *Pseudomonas sp.*) puede añadirse para convertir el oxaloacetato en piruvato y dióxido de carbono. Ya que el CO₂ es volátil, no está disponible para la reacción con las enzimas, disminuyendo o incluso evitando las reacciones inversas. El piruvato producido en esta etapa puede utilizarse también en la reacción de condensación aldólica. Otras enzimas de descarboxilasa pueden utilizarse, y se conocen homólogos que existen en *Actinobacillus actinomycetemcomitans*, *Aquifex aeolicus*, *Archaeoglobus fulgidus*, *Azotobacter vinelandii*, *Bacteroides fragilis*, varias especies de *Bordetella*, *Campylobacter jejuni*, *Chlorobium tepidum*, *Chloroflexus aurantiacus*, *Enterococcus faecalis*, *Fusobacterium nucleatum*, *Klebsiella pneumoniae*, *Legionella pneumophila*, *Magnetococcus MC-1*, *Mannheimia haemolytica*, *Methylobacillus flagellatus* KT, *Pasteurella multocida* Pm70, *Petrotoga mitherma*, *Porphyromonas gingivalis*, varias especies de *Pseudomonas*, varias especies de *Pyrococcus*, *Rhodococcus*, varias especies de *Salmonella*, varias especies de *Streptococcus*, *Thermochromatium tepidum*, *Thermotoga maritima*, *Treponema pallidum* y varias especies de *Vibrio*.

Se realizaron ensayos con triptófano aminotransferasa con la aspartato aminotransferasa (AspC) de *E. coli*, la tirosina aminotransferasa (TyrB) de *E. coli*, la aminotransferasa de amplio sustrato (BSAT) de *L. major*, y las dos glutamato-oxaloacetato aminotransferasas porcinas disponibles en el mercado como se describe en el Ejemplo 1. Tanto el oxaloacetato como el alfa-cetoglutarato se probaron como receptor de amino. Se comparó la relación de actividad utilizando monatina (Ejemplo 7) frente a la actividad utilizando triptófano, para determinar que enzima tenía la mayor especificidad para la reacción de monatina aminotransferasa. Estos resultados indican que la enzima con mayor especificidad para la reacción de monatina frente a la reacción de triptófano es la glutamato-oxaloacetato aminotransferasa tipo II-A porcina, GOAT (Sigma G7005). Esta especificidad era independiente de la que se utilizó con el receptor de amino. Por consiguiente, esta enzima se utilizó en las reacciones acopladas con oxaloacetato descarboxilasa.

Una reacción típica que parte del indol-3-piruvato incluía (concentraciones finales) Tris-Cl 50 mM pH 7,3, indol-3-piruvato 6 mM, piruvato sódico 6 mM, aspartato 6 mM, PLP 0,05 mM, fosfato potásico 3 mM, MgCl₂ 3 mM, 25 µg/ml de aminotransferasa, 50 µg/ml de ProA aldolasa de *C. testosteronei* y 3 Unidades/ml de descarboxilasa (Sigma O4878). Se dejaron proceder las reacciones durante 1 hora a 26°C. En algunos casos se omitió la descarboxilasa o se sustituyó el aspartato con alfa-cetoglutarato (como referencias negativas). Las enzimas aminotransferasas descritas anteriormente se probaron también en lugar de la GOAT para confirmar los experimentos de especificidad iniciales. Se filtraron las muestras y se analizaron por LC/MS como se describe en el Ejemplo 10. Los resultados demuestran que la enzima GOAT produjo la mayor cantidad de monatina por mg de proteína con la mínima cantidad de triptófano producido como subproducto, además, había un beneficio de 2 a 3 veces del que tiene la enzima descarboxilasa añadida. La enzima AspC de *E. coli* produjo grandes cantidades de monatina en comparación con otras aminotransferasas.

La producción de monatina aumentó por: 1) agregar periódicamente adiciones 2 mM de indol piruvato, piruvato y aspartato (cada media hora a hora), 2) llevar a cabo las reacciones en un medio anaeróbico o con tampones desgasificados, 3) dejar que procedan las reacciones durante la noche y 4) utilizar descarboxilasa recién preparada que no haya sido congelada y descongelada muchas veces. La descarboxilasa fue inhibida por concentraciones de piruvato superiores a 12 mM. A concentraciones de indol-3-piruvato superiores a 4 mM, las reacciones secundarias con indol-3-piruvato se aceleraron. La cantidad de indol-3-piruvato utilizada en la reacción pudo aumentarse si la cantidad de aldolasa aumentaba también. Se encontraron grandes concentraciones de fosfato (50 mM) y aspartato (50 mM) que eran inhibitorias para la enzima descarboxilasa. La cantidad de enzima descarboxilasa añadida pudo reducirse hasta 0,5 U/ml sin disminuir la producción de monatina en una reacción de una hora. La cantidad de monatina producida aumentaba cuando la temperatura aumentaba desde 26°C hasta 30°C y desde 30°C hasta 37°C; sin embargo, a 37°C las reacciones secundarias del indol-3-piruvato se aceleraron también. La cantidad de monatina producida aumentó al aumentar el pH desde 7 a 7,3 y era relativamente estable desde pH 7,3 a 8,3.

Una reacción típica que parte de triptófano incluía (concentraciones finales) Tris-Cl 50 mM pH 7,3, triptófano 20 mM, aspartato 6 mM, piruvato sódico 6 mM, PLP 0,05 mM, fosfato potásico 3 mM, MgCl₂ 3 mM, 25 µg/ml de aminotransferasa, 50 µg/ml de ProA aldolasa de *C. testosteronei* y 4 unidades/ml de descarboxilasa, 5 a 200 mU/ml de L-aminoácido oxidasa (Sigma A-2805), 168 U/ml de catalasa (Sigma C-3515) y 0,008 mg de FAD. Las reacciones se llevaron a cabo durante 30 minutos a 30°C. Se observó mejoría con la adición de descarboxilasa. La cantidad mayor de monatina se produjo cuando se utilizaron 50 mU/ml de oxidasa. Las mejoras fueron similares a las observadas cuando se utilizó como sustrato indol-3-piruvato. Además, la cantidad de monatina producida aumentó cuando 1) la concentración de triptófano era baja (es decir, inferior a la K_m de la enzima aminotransferasa y por consiguiente incapaz de competir con PM en la zona activa) y 2) la relación de oxidasa a aldolasa y aminotransferasa se mantuvo a una concentración tal que el indol-3-piruvato no se pudo acumular.

Ya sea partiendo de indol-3-piruvato o de triptófano, la cantidad de monatina producida en los ensayos con periodos de incubación de 1 a 2 horas aumentó cuando se utilizaban 2 a 4 veces las cantidades de todas las enzimas a la vez que se mantenían la misma relación enzimática. Utilizando uno de los dos sustratos, se consiguieron concentraciones de aproximadamente 1 mg/ml de monatina. La cantidad de triptófano producido si se

parte de indol-piruvato fue por lo general inferior al 20% de la cantidad de producto, lo que demuestra la utilidad de utilizar reacciones acopladas. Con más optimización y control de las concentraciones de los productos intermedios y de las reacciones secundarias, la productividad y el rendimiento pueden mejorarse en gran medida.

Reacciones acopladas que utilizan lisina épsilon aminotransferasa (CE 2.6.1.36)

La lisina epsilon aminotransferasa (L-lisina-6-transaminasa) se encuentra en varios organismos, incluyendo *Rhodococcus*, *Mycobacterium*, *Streptomyces*, *Nocardia*, *Flavobacterium*, *Candida utilis* y *Streptomyces*. Es utilizada por organismos como primera etapa en la producción de algunos antibióticos de beta-lactama (Rius y Demain, *J. Microbiol. Biotech.*, 7:95-100, 1997). Esta enzima convierte la lisina en L-2-aminoadipato 6-semialdehído (alisina) por una transaminación mediada por PLP del C-6 de la lisina, utilizando alfa-cetoglutarato como receptor de amino. La lisina es inestable y espontáneamente experimenta una deshidratación intramolecular para formar 1-piperideína 6-carboxilato, molécula cíclica. Esta inhibe con eficacia cualquier reacción inversa que ocurra. El esquema de reacción se representa en la figura 12. Puede utilizarse también una enzima alternativa, lisina-piruvato 6-transaminasa (CE 2.6.1.71).

Una reacción típica contenía en 1 ml: Tris-Cl 50 mM pH 7,3, indol-3-piruvato 20 mM, PLP 0,05 mM, fosfato potásico 6 mM, pH 8, piruvato sódico 2 a 50 mM, MgCl₂ 1,5 mM, lisina 50 mM, 100 µg de aminotransferasa (lisina épsilon aminotransferasa LAT-101, BioCatalytics Pasadena, CA) y 2000 µg de ProA aldolasa de *C. testosteroni*. La cantidad de monatina producida aumentó al aumentar las concentraciones de piruvato. La cantidad máxima utilizando estas condiciones de reacción (a piruvato 50 mM) fue 10 veces menos que la que se observó con reacciones acopladas que utilizan oxaloacetato descarboxilasa (aproximadamente 0,1 mg/ml).

Un pico con $[M+H]^+ = 293$ eluyó en el tiempo esperado para monatina y el espectro de masas contenía varios de los mismos fragmentos observados con otros procesos enzimáticos. Un segundo pico con la relación masa a carga (293) correcta eluyó ligeramente antes de que lo que se observa por lo general para la S,S monatina producida en el Ejemplo 6, y puede indicar la presencia de otro isómero de monatina. Esta enzima produjo muy poco triptófano. Sin embargo, existe probablemente alguna actividad sobre el piruvato (que produce alanina como subproducto). Además, se conoce que la enzima es inestable. Pueden introducirse mejoras realizando experimentos de evolución dirigida para aumentar la estabilidad, reducir la actividad con piruvato y aumentar la actividad con PM. Estas reacciones pueden acoplarse también a L-aminoácido oxidasa/catalasa como se describió anteriormente.

Otras reacciones acopladas

Otra reacción de acoplamiento que puede mejorar el rendimiento de monatina a partir del triptófano o del indol-piruvato, se muestra en la figura 13. La formiato deshidrogenasa (CE 1.2.1.2 o 1.2.1.43) es una enzima frecuente. Algunas formiato deshidrogenasas requieren NADH mientras que otras pueden utilizar NADPH. La glutamato deshidrogenasa catalizó la interconversión entre el precursor de monatina y monatina en los ejemplos anteriores, utilizando tampones a base de amonio. La presencia del formiato amónico y de la formiato deshidrogenasa es un sistema eficaz para la regeneración de cofactores, y la producción de dióxido de carbono es una manera eficaz de aumentar la velocidad de las reacciones inversas (Bommarius *et al.*, *Biocatalysis*, 10:37, 1994 y Galkin *et al.*, *Appl. Environ. Microbiol.*, 63:4651-6, 1997). Además, pueden disolverse grandes cantidades de formiato amónico en el tampón de reacción. El rendimiento de monatina producida por las reacciones con glutamato deshidrogenasa (o aminaciones reductoras similares) puede mejorarse por adición de formiato deshidrogenasa y formiato amónico.

Pueden utilizarse otros procedimientos para conducir el equilibrio hacia la producción de monatina. Por ejemplo, si se utiliza aminopropano como donante de aminoácidos en la conversión de PM a monatina con una omega-aminoácido aminotransferasa (CE 2.6.1.18) tal como las descritas por las patentes US nº 5.360.724 y nº 5.300.437, uno de los productos resultantes sería acetona, un producto más volátil que el sustrato amino-propano. La temperatura puede elevarse periódicamente durante periodos cortos para emitir la acetona, aliviando de este modo el equilibrio. La acetona tiene un punto de ebullición de 47°C, una temperatura improbable para degradar los compuestos intermedios si se utiliza durante breves periodos. La mayoría de las aminotransferasas que tiene actividad sobre la alfa-cetoglutarato también tienen actividad sobre el precursor de monatina. Asimismo, si se utiliza un glioxilato/acidoaromático aminotransferasa (CE 2.6.1.60) con glicina como donante de amino, se produce glioxalato que es relativamente estable y tiene un punto de ebullición muy reducido en comparación con la glicina.

EJEMPLO 14

Expresión recombinante

Con secuencias de ADNc enzimático y de aminoácidos disponibles al público, y las enzimas y secuencias dadas a conocer en la presente memoria, tal como las SEC. ID. nº 11 y nº 12, así como las variantes, polimorfismos, mutantes, fragmentos y fusiones de los mismos, se permite la expresión y purificación de cualquier proteína, tal como una enzima, mediante técnicas de laboratorio habituales. Un experto en la materia apreciará que

las enzimas y fragmentos de las mismas pueden producirse de manera recombinante en cualquier célula u organismo de interés, y purificarse antes de su utilización, por ejemplo antes de la producción de la SEC. ID. nº 2 y derivados de las mismas.

Los procedimientos para producir proteínas recombinantes son bien conocidos en la técnica. Por consiguiente, el alcance de la presente invención incluye la expresión recombinante de cualquier proteína o fragmento de la misma, tal como una enzima. Por ejemplo, véase la patente US nº 5.342.764 de Johnson *et al.*, la patente US nº 5.846.819 de Pausch *et al.*, la patente US nº 5.876.969 de Fleer *et al.* y Sambrook *et al.*, (*Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, Nueva York., 1989, Cap. 17).

En resumen, las secuencias de ADNc parcial, completo o variante, que codifican una proteína o péptido, pueden estar ligadas en un vector de expresión, tal como un vector de expresión bacteriano o eucariótico. Pueden producirse proteínas y/o péptidos colocando un activador corriente arriba de la secuencia de ADNc. Los ejemplos de activadores comprenden de manera no limitativa *lac*, *trp*, *tac*, *trc*, regiones principales del operador y del activador del fago lambda, la región de control de la proteína de la capa fd, los activadores precoz y tardío de SV40, los activadores procedentes de poliovirus, adenovirus, retrovirus, baculovirus y virus de simio, el activador para la 3-fosfoglicerato cinasa, los activadores de la levadura ácida fosfatasa, el activador de los factores emparejadores de la levadura alfa y combinaciones de los mismos.

Los vectores adecuados para la producción de proteínas naturales intactas incluyen pKC30 (Shimatake y Rosenberg, 1981, *Nature* 292:128), pKK177-3 (Amann y Brosius, 1985, *Gene* 40:183) y pET-3 (Studier y Moffatt, 1986, *J. Mol. Biol.*, 189:113). Una secuencia de ADN puede transferirse a otros vehículos de clonación, tales como otros plásmidos, bacteriófagos, cósmidos, virus animales y cromosomas artificiales de levadura (YAC) (Burke *et al.*, 1987, *Science*, 236:806-12). Estos vectores pueden introducirse en una variedad de hospedadores incluyendo las células somáticas, y organismos sencillos o complejos, tales como bacterias, hongos (Timberlake y Marshall, 1989, *Science*, 244:1313-7), invertebrados, plantas (Gasser y Fraley, 1989, *Science*, 244:1293) y mamíferos (Pursel *et al.*, 1989, *Science*, 244:1281-8), que resultan transgénicos mediante la introducción del ADNc heterólogo.

Para la expresión en células de mamífero, una secuencia de ADNc puede ligarse a activadores heterólogos, tales como el virus SV40 de simio, activador en el vector pSV2 (Mulligan y Berg, 1981, *Proc. Natl. Acad. Sci. USA.*, 78:2072-6) e introducirse en las células, tal como las células COS-1 (Gluzman, 1981, *Cell*, 23:175-82), para conseguir la expresión temporal o duradera. La integración estable del montaje génico híbrido puede mantenerse en células de mamífero por selección bioquímica, tal como neomicina (Southern y Berg, 1982, *J. Mol. Appl. Genet.*, 1:327-41) y ácido micofenólico (Mulligan y Berg, 1981, *Proc. Natl. Acad. Sci. USA.*, 78:2072-6).

La transferencia de ADN en células eucarióticas, tales como las del hombre o de otros mamíferos, es una técnica convencional. Los vectores se introducen en las células efectoras como ADN puro (transfección), por ejemplo por precipitación con fosfato cálcico (Graham y Vander Eb., 1973, *Virology*, 52:466), fosfato de estroncio (Brash *et al.*, 1987, *Mol. Cell. Biol.*, 7:2013), electroporación (Neumann *et al.*, 1982, *EMBO J.*, 1:841), lipofección (Felgner *et al.*, 1987, *Proc. Natl. Acad. Sci. USA.*, 84:7413), DEAE dextrano (McCuthan *et al.*, 1968, *J. Natl. Cancer Inst.*, 41:351), microinyección (Mueller *et al.*, 1978, *Cell*, 15:579), fusión del protoplasto (Schafner, 1980, *Proc. Natl. Acad. Sci. USA.*, 77:2163-7) o escopetas de perdigones (Klein *et al.*, 1987, *Nature* 327:70). Alternativamente, el ADNc puede introducirse por infección con vectores víricos, por ejemplo retrovirus (Bernstein *et al.*, 1985, *Gen Engng.*, 7:235) tales como adenovirus (Ahmad *et al.*, 1986, *J. Virol.*, 57:267) o Herpes (Spaete *et al.*, 1982, *Cell*, 30:295).

LISTADO DE SECUENCIAS

<110> Cargill, Incorporated

<120> Polipéptidos y rutas biosintéticas

<130> 6682-64349

<150> 60/374.831

<151> 2002-04-23

<160> 74

<170> PatentIn versión 3.1

<210> 1

<211> 1170

<212> ADN

<213> *Sinorhizobium meliloti*

<400> 1

```

atgttcgacg ccctcgcccc ccaagccgac gatcccttgc ttttcctgat cggcctgttc      60
aggaaggatg agcgccccgg aaaggatcgat ctcggcgtag gagtctatcg cgacgagacc      120
ggacgcacgc cgatcttccg ggccgtcaag gcggcggaag agcggcttct cgaaacacag      180
gacagcaagg cctatatcgg ccccggaagg gacctcgtct ttctcgatcg gctctgggaa      240
ctcgtcggcg gcgacacgat cgagcggagc catgttgcgg gcgtccagac gcccggcggc      300
tccggcgcgc tccgtttggc ggccgacctc atcgcccgca tgggcggccg aggcattctg      360
ctcgggctgc cgagctggcc gaaccacgcg ccgatcttca aggcggcccg gctcgatatc      420
gccacctacg acttcttcga cattccgtcg cagtcggtca tcttcgataa tctggtgagc      480
gcgctggaag gcgcgcgacg cggcgatcgc gtgctgctgc atgcaagctg ccacaacccg      540
accggcgcgg tcctgagcga agcacaatgg atggagatcg ccgcgctggt ggccgagcgc      600
ggcctgctgc cgctcgtcga tctcgcttat caggggttcg gcccgggcct cgaccaggat      660
gtcgcggggc tccggcatct tctcggcgtg gtcccgaag cgctcgtcgc ggtttcctgc      720
tcgaagtcct tcgggcttta tcgcgagcgc gcgggcgcga tcttcgcgcg gaccagctcg      780
actgcctcgg cggacagggt gcgctcaaac ctgcggggcc tcgcacgcac cagctattcc      840
atgccgcggg atcacggcgc agccgtcgtg cggacgatcc ttgacgaccc ggaactcagg      900
cgcgactgga cggaggagct cgagacgatg cggctcagga tgacgggcct ccggcggtcg      960
cttgccgagg gactccgcac ccgctggcag agcctcggcg cagtcgccga tcaggagggc     1020
atgttctcca tgctcgcgct ttccgaagcg gaggttatgc ggctcaggac cgagcacggc     1080
atctatatgc cggcatccgg ccgcatcaac atcgccgggc tgaagacggc ggaagccgcc     1140
gagattgccg gcaagtccac cagtctctga                                     1170

```

5

<210> 2

<211> 389

<212> PRT

<213> Sinorhizobium meliloti

10

<400> 2

```

Met Phe Asp Ala Leu Ala Arg Gln Ala Asp Asp Pro Leu Leu Phe Leu
 1              5              10              15

Ile Gly Leu Phe Arg Lys Asp Glu Arg Pro Gly Lys Val Asp Leu Gly
 20              25              30

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

Val Lys Ala Ala Glu Lys Arg Leu Leu Glu Thr Gln Asp Ser Lys Ala
 50              55              60

Tyr Ile Gly Pro Glu Gly Asp Leu Val Phe Leu Asp Arg Leu Trp Glu
 65              70              75              80

```

Leu Val Gly Gly Asp Thr Ile Glu Arg Ser His Val Ala Gly Val Gln
85 90 95

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

Arg Met Gly Gly Arg Gly Ile Trp Leu Gly Leu Pro Ser Trp Pro Asn
115 120 125

His Ala Pro Ile Phe Lys Ala Ala Gly Leu Asp Ile Ala Thr Tyr Asp
130 135 140

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

Ala Leu Glu Gly Ala Ala Ser Gly Asp Ala Val Leu Leu His Ala Ser
165 170 175

Cys His Asn Pro Thr Gly Gly Val Leu Ser Glu Ala Gln Trp Met Glu
180 185 190

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

Ala Tyr Gln Gly Phe Gly Arg Gly Leu Asp Gln Asp Val Ala Gly Leu
210 215 220

Arg His Leu Leu Gly Val Val Pro Glu Ala Leu Val Ala Val Ser Cys
225 230 235 240

Ser Lys Ser Phe Gly Leu Tyr Arg Glu Arg Ala Gly Ala Ile Phe Ala
245 250 255

Arg Thr Ser Ser Thr Ala Ser Ala Asp Arg Val Arg Ser Asn Leu Ala
260 265 270

Gly Leu Ala Arg Thr Ser Tyr Ser Met Pro Pro Asp His Gly Ala Ala
275 280 285

Val Val Arg Thr Ile Leu Asp Asp Pro Glu Leu Arg Arg Asp Trp Thr
290 295 300

Glu Glu Leu Glu Thr Met Arg Leu Arg Met Thr Gly Leu Arg Arg Ser
305 310 315 320

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

Asp Gln Glu Gly Met Phe Ser Met Leu Pro Leu Ser Glu Ala Glu Val
340 345 350

Met Arg Leu Arg Thr Glu His Gly Ile Tyr Met Pro Ala Ser Gly Arg
355 360 365

Ile Asn Ile Ala Gly Leu Lys Thr Ala Glu Ala Ala Glu Ile Ala Gly
370 375 380

Lys Phe Thr Ser Leu
385

<210> 3

<211> 1260

<212> ADN

<213> Rhodobacter sphaeroides

<400> 3

```

atgcgctcta cgacggctcc tgggccgagt ggggcatgta tgacgatctc aaggtcgcga      60
aaggatgacg aaggaatgct gaccgccttg aagccgcagc ccgcgacaa gatcctgcaa      120
ctgatccaga tgttcgcgga ggatgcgcgc gcggacaaga tcgatctggg cgtgggcgctc      180
tacaaggacc cgaccgggct caccgccgtc atgcggggcg tgaaggcggc cgagaagcgg      240
ctctggggagg tcgagaccac caagacctac accggccttg ccgacgagcc ggccataaat      300
gccgcgatgg cgaagctgat cctcgcgggc gcggtcccgc ccgaccgggt ggccctcggtc      360
gccacccccg gcggcacggg cgcggtgctg caggcgctcg agctgatccg catggcctcg      420
cccaggcca ccgtctggat ctcaaccgc acctggccga accatctgtc gatcgtgaaa      480
tatctcggca tcccgatgcg ggaataccgc tatttcgacg ccgagaccgg cgcgctcgat      540
gccgagggca tgatggagga tctggcccag gtgaaggcgg gcgacgtggt gctgctgcac      600
ggctgctgcc acaaccgcac cggcgccaac ccgaaccggg tgcagtggct ggccatctgc      660
gagagcctgg cccggacagg cgcggtgccg ctgatcgacc tcgcctatca gggcttcggc      720
gacgggctcg agatggatgc ggcgggcagc cggtctcttg ccaccagact gcccgagggtg      780
ctgatcgagg cctcctgctc gaagaacttc ggcatctacc gcgagcgac ggccatcctg      840
atcgccatcg gcgaggcggc gggccggggc acggtgcagg ccaacctcaa ctctctgaac      900
cggcagaact actccttccc gccggaccat ggcgcgcggc tcgtgacct gatcctcgag      960
gacgagacgc tgagcgccga ctggaaggcg gaactcgagg aggtgcggct caacatgctg     1020
aactgcgcc gccagcttgc cgatgcgctg caggccgaga ccggctcgaa ccgcttcggc     1080
ttcgtggccg agcatcgagg catgttctcg cgcctcggga tcacgcccgc cgaggtggag     1140
cggtgcgga ccgagcacgg ggtctacatg gtgggcgatt cgcggtgaa catcgcgggg     1200
ctgaaccgga cgaccgtgcc ggtgctggcg cgcgcggtgg ccaagggtgct gcgcggtga     1260

```

5

<210> 4

<211> 419

<212> PRT

<213> Rhodobacter sphaeroides

10

<400> 4

```

Met Arg Ser Thr Thr Ala Pro Gly Pro Ser Gly Ala Cys Met Thr Ile
1           5           10          15

Ser Arg Ser Arg Lys Asp Asp Glu Gly Met Leu Thr Ala Leu Lys Pro
20          25          30

Gln Pro Ala Asp Lys Ile Leu Gln Leu Ile Gln Met Phe Arg Glu Asp
35          40          45

Ala Arg Ala Asp Lys Ile Asp Leu Gly Val Gly Val Tyr Lys Asp Pro
50          55          60

Thr Gly Leu Thr Pro Val Met Arg Ala Val Lys Ala Ala Glu Lys Arg
65          70          75          80

```

Leu	Trp	Glu	Val	Glu	Thr	Thr	Lys	Thr	Tyr	Thr	Gly	Leu	Ala	Asp	Glu	
				85												95
Pro	Ala	Tyr	Asn	Ala	Ala	Met	Ala	Lys	Leu	Ile	Leu	Ala	Gly	Ala	Val	
			100					105					110			
Pro	Ala	Asp	Arg	Val	Ala	Ser	Val	Ala	Thr	Pro	Gly	Gly	Thr	Gly	Ala	
		115					120					125				
Val	Arg	Gln	Ala	Leu	Glu	Leu	Ile	Arg	Met	Ala	Ser	Pro	Glu	Ala	Thr	
	130					135						140				
Val	Trp	Ile	Ser	Asn	Pro	Thr	Trp	Pro	Asn	His	Leu	Ser	Ile	Val	Lys	
145					150					155					160	
Tyr	Leu	Gly	Ile	Pro	Met	Arg	Glu	Tyr	Arg	Tyr	Phe	Asp	Ala	Glu	Thr	
				165					170						175	
Gly	Ala	Val	Asp	Ala	Glu	Gly	Met	Met	Glu	Asp	Leu	Ala	Gln	Val	Lys	
			180					185						190		
Ala	Gly	Asp	Val	Val	Leu	Leu	His	Gly	Cys	Cys	His	Asn	Pro	Thr	Gly	
		195					200					205				
Ala	Asn	Pro	Asn	Pro	Val	Gln	Trp	Leu	Ala	Ile	Cys	Glu	Ser	Leu	Ala	
	210					215						220				
Arg	Thr	Gly	Ala	Val	Pro	Leu	Ile	Asp	Leu	Ala	Tyr	Gln	Gly	Phe	Gly	
225					230					235					240	
Asp	Gly	Leu	Glu	Met	Asp	Ala	Ala	Ala	Thr	Arg	Leu	Leu	Ala	Thr	Arg	
				245					250					255		
Leu	Pro	Glu	Val	Leu	Ile	Ala	Ala	Ser	Cys	Ser	Lys	Asn	Phe	Gly	Ile	
			260					265					270			
Tyr	Arg	Glu	Arg	Thr	Gly	Ile	Leu	Ile	Ala	Ile	Gly	Glu	Ala	Ala	Gly	
		275					280					285				
Arg	Gly	Thr	Val	Gln	Ala	Asn	Leu	Asn	Phe	Leu	Asn	Arg	Gln	Asn	Tyr	
		290				295					300					
Ser	Phe	Pro	Pro	Asp	His	Gly	Ala	Arg	Leu	Val	Thr	Met	Ile	Leu	Glu	
305					310					315					320	
Asp	Glu	Thr	Leu	Ser	Ala	Asp	Trp	Lys	Ala	Glu	Leu	Glu	Glu	Val	Arg	
				325					330					335		
Leu	Asn	Met	Leu	Thr	Leu	Arg	Arg	Gln	Leu	Ala	Asp	Ala	Leu	Gln	Ala	
			340					345					350			
Glu	Thr	Gly	Ser	Asn	Arg	Phe	Gly	Phe	Val	Ala	Glu	His	Arg	Gly	Met	
		355					360					365				
Phe	Ser	Arg	Leu	Gly	Ile	Thr	Pro	Ala	Glu	Val	Glu	Arg	Leu	Arg	Thr	
		370				375					380					
Glu	His	Gly	Val	Tyr	Met	Val	Gly	Asp	Ser	Arg	Leu	Asn	Ile	Ala	Gly	
385					390					395					400	
Leu	Asn	Arg	Thr	Thr	Val	Pro	Val	Leu	Ala	Arg	Ala	Val	Ala	Lys	Val	
				405					410					415		

Leu Arg Gly

<210> 5

$\langle 211 \rangle$ 1260

<212> ADN

<213> Rhodobacter sphaeroides

<400> 5

```

atgcgctcta cgacggctcc tgggccaggt ggggcatgta tgacgatctc aaggtcgcga      60
aaggatgacg aaggaatgct gaccgccctg aagccgcagc ccgcggacaa gatcctgcaa      120
ctgatccaga tgttccgcga ggatgcgcgc gcggacaaga tcgatctggg cgtgggcgtc      180
tacaaggacc cgacggggct caccgccgtc atgcgggcgc tgaaggccgc cgagaagcgg      240
ctctgggagg tcgagaccac caagacctac accggccttg ccggcgagcc cgcctacaat      300
gccgcgatgg cgaagctgat cctcgcaggc gcggtcccg ccgaccgggt ggccctcggtc      360
gccacccccg gcggcacggg cgcggtgcgt caggcgctcg agctgatccg catggcctcg      420
cccgaggcca ctgtctggat ctgaaacccg acctggccga accatctgtc gatcgtgaaa      480
tatctcggca tcccgatgcg ggaataccgc tatttcgacg ccgagaccgg cgcgctcgat      540
gccgaggggt tgatggagga tctggcccag gtgaaggcgg gcgacgtggt gctgctgcac      600
ggctgtgcc acaacccgac cggcgccaac ccgaaccggg tgcagtgggt ggccgtctgc      660
gagagcctgg cccggacagg cgcggtgccg ctgatcgacc tcgcctatca gggcttcggc      720
gacgggctcg agatggatgc ggccggcgacg cggcttctgg ccaccagact gcccgagggtg      780
ctgatcgagg cctcctgctc gaagaacttc ggcattctacc gcgagcgaac gggcattcctg      840
atcgccatcg gcgaggcggc gggccggggc acggtgcagg ccaacctcaa cttcctgaac      900
cggcagaact actccttccc gccggaccat ggcgcgcggc tcgtgaccat gatcctcgag      960
gacgagacgc tgagcgccga ctggaaggcg gaactcgagg aggtgcgggt caacatgctg     1020
acgctgcgcc gccagcttgc cgatgcgctg caggccgaga ccggctcgaa ccgcttcggc     1080
ttcgtggccg agcatcgcgg catgttctcg cgcctcgga tcacgcccgc cgagggtggag     1140
cggctgcgga ccgagcacgg ggtctacatg gtgggcgatt cgcggctgaa catcgcgggg     1200
ctgaaccgga cgaccgtgcc ggtgctggcg cgcgcggtgg ccaaggtgct gcgcggctga     1260

```

5

<210> 6

<211> 419

<212> PRT

<213> Rhodobacter sphaeroides

10

<400> 6

```

Met Arg Ser Thr Thr Ala Pro Gly Pro Ser Gly Ala Cys Met Thr Ile
1           5           10           15

Ser Arg Ser Arg Lys Asp Asp Glu Gly Met Leu Thr Ala Leu Lys Pro
          20           25           30

Gln Pro Ala Asp Lys Ile Leu Gln Leu Ile Gln Met Phe Arg Glu Asp
          35           40           45

Ala Arg Ala Asp Lys Ile Asp Leu Gly Val Gly Val Tyr Lys Asp Pro
          50           55           60

```

Thr Gly Leu Thr Pro Val Met Arg Ala Val Lys Ala Ala Glu Lys Arg
65 70 75 80

Leu Trp Glu Val Glu Thr Thr Lys Thr Tyr Thr Gly Leu Ala Gly Glu
85 90 95

Pro Ala Tyr Asn Ala Ala Met Ala Lys Leu Ile Leu Ala Gly Ala Val
100 105 110

Pro Ala Asp Arg Val Ala Ser Val Ala Thr Pro Gly Gly Thr Gly Ala
115 120 125

Val Arg Gln Ala Leu Glu Leu Ile Arg Met Ala Ser Pro Glu Ala Thr
130 135 140

Val Trp Ile Ser Asn Pro Thr Trp Pro Asn His Leu Ser Ile Val Lys
145 150 155 160

Tyr Leu Gly Ile Pro Met Arg Glu Tyr Arg Tyr Phe Asp Ala Glu Thr
165 170 175

Gly Ala Val Asp Ala Glu Gly Leu Met Glu Asp Leu Ala Gln Val Lys
180 185 190

Ala Gly Asp Val Val Leu Leu His Gly Cys Cys His Asn Pro Thr Gly
195 200 205

Ala Asn Pro Asn Pro Val Gln Trp Leu Ala Val Cys Glu Ser Leu Ala
 210 215 220

Arg Thr Gly Ala Val Pro Leu Ile Asp Leu Ala Tyr Gln Gly Phe Gly
 225 230 235 240

Asp Gly Leu Glu Met Asp Ala Ala Ala Thr Arg Leu Leu Ala Thr Arg
 245 250 255

Leu Pro Glu Val Leu Ile Ala Ala Ser Cys Ser Lys Asn Phe Gly Ile
 260 265 270

Tyr Arg Glu Arg Thr Gly Ile Leu Ile Ala Ile Gly Glu Ala Ala Gly
 275 280 285

Arg Gly Thr Val Gln Ala Asn Leu Asn Phe Leu Asn Arg Gln Asn Tyr
 290 295 300

Ser Phe Pro Pro Asp His Gly Ala Arg Leu Val Thr Met Ile Leu Glu
 305 310 315 320

Asp Glu Thr Leu Ser Ala Asp Trp Lys Ala Glu Leu Glu Glu Val Arg
 325 330 335

Leu Asn Met Leu Thr Leu Arg Arg Gln Leu Ala Asp Ala Leu Gln Ala
 340 345 350

Glu Thr Gly Ser Asn Arg Phe Gly Phe Val Ala Glu His Arg Gly Met
 355 360 365

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

Glu His Gly Val Tyr Met Val Gly Asp Ser Arg Leu Asn Ile Ala Gly
 385 390 395 400

Leu Asn Arg Thr Thr Val Pro Val Leu Ala Arg Ala Val Ala Lys Val
 405 410 415

Leu Arg Gly

<210> 7

<211> 1239

<212> ADN

<213> Leishmania major

<400> 7

atgtccatgc aggcggccat gaccacggcg gagcgctggc agaagattca ggcacaagct 60

cccgatgtca ttttcgatct cgcaaaacgc gccgcgctg ccaagggccc caaggccaac 120

```

ctcgctcattg gtgcctaccg cgacgagcag ggcggtccct atccggtacg cgtgggtccgc 180
aaggctgagc agcttctctt ggacatgaat ctgcactacg agtacctacc catctcgggc 240
taccagccct tcacgatga ggcggtaaag attatctacg gcaataccgt cgagctggag 300
aacctggttg cgggtgcagac gctgagcggg accggtgctg tctctctcgg ggcgaagctg 360
ctgactcgcg tcttcgacgc tgagacgacg cccatctacc tttccgaccc cacgtggccc 420
aaccactacg gcgtcgtgaa ggctgctggc tggagaaca tctgcacgta cgcctactac 480
gaccccaaga cggtcagcct gaatttcgag ggcataaga aagacattct ggcggcgccg 540
gacgggtccg tgttcattct gcaccagtgc gcgcacaacc ccaccggcgt ggaccgctcg 600
caggagcagt ggaacgagat cgcgtcactg atgctggcca agcaccatca ggtgttcttc 660
gactccgcct accaaggcta tgcgagcggc agcctcgaca cggacgcgta tgctgcccgc 720
ctgtttgccc gccgcggcat cgaggtactg ctggcgagct cgttctccaa gaacatgggc 780
ttgtacagcg agcgtgcagg cacgctgtcg ctgctcctca aggacaagac gaagcgcgcg 840
gatgtaaaga gcgtgatgga ttcgctgac cgtgaggagt acacgtgccc ccagcccac 900
ggtgcccgc tagcccacct aatcctgagc aacaacgaac tgcgaaagga gtgggaggca 960
gagctatcag ccatggcaga gcgcacccgt acgatgcgcc gcaccgtgta cgacgagctg 1020
ctgcgcctgc agacgcccgg gagctgggaa catgtcatta accagattgg catgttttcc 1080
ttcctcgggc tgtcaaaggc gcagtgcgaa tactgccaaa accacaacat cttcatcaca 1140
gtgtcggggc gcgctaacat ggcaggtctg acgcatgaga cggcgctgat gctagcacag 1200
acgatcaacg atgctgtgcg caatgtgaat cgtgagtga 1239

```

5

<210> 8
 <211> 412
 <212> PRT
 <213> Leishmania major

10

<400> 8

```

Met Ser Met Gln Ala Ala Met Thr Thr Ala Glu Arg Trp Gln Lys Ile
1          5          10          15

Gln Ala Gln Ala Pro Asp Val Ile Phe Asp Leu Ala Lys Arg Ala Ala
20          25          30

Ala Ala Lys Gly Pro Lys Ala Asn Leu Val Ile Gly Ala Tyr Arg Asp
35          40          45

Glu Gln Gly Arg Pro Tyr Pro Leu Arg Val Val Arg Lys Ala Glu Gln
50          55          60

Leu Leu Leu Asp Met Asn Leu Asp Tyr Glu Tyr Leu Pro Ile Ser Gly
65          70          75          80

```

Tyr Gln Pro Phe Ile Asp Glu Ala Val Lys Ile Ile Tyr Gly Asn Thr
 85 90 95
 Val Glu Leu Glu Asn Leu Val Ala Val Gln Thr Leu Ser Gly Thr Gly
 100 105 110
 Ala Val Ser Leu Gly Ala Lys Leu Leu Thr Arg Val Phe Asp Ala Glu
 115 120 125
 Thr Thr Pro Ile Tyr Leu Ser Asp Pro Thr Trp Pro Asn His Tyr Gly
 130 135 140
 Val Val Lys Ala Ala Gly Trp Lys Asn Ile Cys Thr Tyr Ala Tyr Tyr
 145 150 155 160
 Asp Pro Lys Thr Val Ser Leu Asn Phe Glu Gly Met Lys Lys Asp Ile
 165 170 175
 Leu Ala Ala Pro Asp Gly Ser Val Phe Ile Leu His Gln Cys Ala His
 180 185 190
 Asn Pro Thr Gly Val Asp Pro Ser Gln Glu Gln Trp Asn Glu Ile Ala
 195 200 205
 Ser Leu Met Leu Ala Lys His His Gln Val Phe Phe Asp Ser Ala Tyr
 210 215 220
 Gln Gly Tyr Ala Ser Gly Ser Leu Asp Thr Asp Ala Tyr Ala Ala Arg
 225 230 235 240
 Leu Phe Ala Arg Arg Gly Ile Glu Val Leu Leu Ala Gln Ser Phe Ser
 245 250 255
 Lys Asn Met Gly Leu Tyr Ser Glu Arg Ala Gly Thr Leu Ser Leu Leu
 260 265 270
 Leu Lys Asp Lys Thr Lys Arg Ala Asp Val Lys Ser Val Met Asp Ser
 275 280 285
 Leu Ile Arg Glu Glu Tyr Thr Cys Pro Pro Ala His Gly Ala Arg Leu
 290 295 300
 Ala His Leu Ile Leu Ser Asn Asn Glu Leu Arg Lys Glu Trp Glu Ala
 305 310 315 320
 Glu Leu Ser Ala Met Ala Glu Arg Ile Arg Thr Met Arg Arg Thr Val
 325 330 335
 Tyr Asp Glu Leu Leu Arg Leu Gln Thr Pro Gly Ser Trp Glu His Val
 340 345 350
 Ile Asn Gln Ile Gly Met Phe Ser Phe Leu Gly Leu Ser Lys Ala Gln
 355 360 365
 Cys Glu Tyr Cys Gln Asn His Asn Ile Phe Ile Thr Val Ser Gly Arg
 370 375 380
 Ala Asn Met Ala Gly Leu Thr His Glu Thr Ala Leu Met Leu Ala Gln
 385 390 395 400
 Thr Ile Asn Asp Ala Val Arg Asn Val Asn Arg Glu
 405 410

<210> 9

<211> 1182
 <212> ADN
 <213> Bacillus subtilis

5

<400> 9

```

atggaacatt tgctgaatcc gaaagcaaga gagatcgaaa tttcaggaat acgcaaattc      60
tcgaatcttg tagcccaaca cgaagacgtc atttcactta caatcggcca gcctgatttt      120
ttcacaccgc atcatgtgaa agctgccgca aaaaaagcca ttgatgaaaa cgtgacgtca      180
tatactccga atgccggcta cctggagctg agacaagctg tgcagcttta tatgaagaaa      240
aaagcggatt tcaactatga tgctgaatct gaaattatca tcacaacagg cgcaagccaa      300
gccattgatg ctgcattccg gacgatttta tctcccggtg atgaagtcac tatgccaggg      360
cctattttatc cgggctatga acctattatc aattttgtgcg gggccaagcc tgtcattggt      420
gatactacgt cacacggctt taagcttacc gcccggtga ttgaagatgc tctgacaccc      480
aacaccaagt gtgtcgtgct tccttatccg tcaaacccta ccggcgtgac tttatctgaa      540
gaagaactga aaagcatcgc agctctctta aaaggcagaa atgtcttcgt attgtctgat      600
gaaatataca gtgaattaac atatgacaga ccgcattact ccacgcgaac ctatttgcg      660
gatcaaacga ttgtcattaa cgggttgta aaatcacaca gcatgaccgg ttggagaatt      720
ggatttttat ttgcaccgaa agacattgca aagcacattt taaagggttca tcaatacaat      780
gtgtcgtgcg cctcatccat ttctcaaaaa gccgcgcttg aagctgtcac aaacggcttt      840
gacgatgcat tgattatgag agaacaatac aaaaaacgtc tggactatgt ttatgaccgt      900
cttggtttcca tgggacttga cgtagttaaa ccgtccggtg cgttttataat ctcccttct      960
attaaatcat ttggaatgac ttcatattgat tttagtatgg ctctttttgga agacgtggc      1020
gtggcactcg tgccgggag ctcgtttctca acatattggtg aaggatatgt aaggctgtct      1080
tttgcacgct caatggacac gctgagagaa ggcctagacc gtttagaatt atttgtatta      1140
aaaaaacgtg aagcaatgca gacgataaac aacggcggtt aa                        1182

```

10

<210> 10
 <211> 393
 <212> PRT
 <213> Bacillus subtilis

15

<400> 10

```

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

Ile Arg Lys Phe Ser Asn Leu Val Ala Gln His Glu Asp Val Ile Ser
20          25          30

Leu Thr Ile Gly Gln Pro Asp Phe Phe Thr Pro His His Val Lys Ala
35          40          45

Ala Ala Lys Lys Ala Ile Asp Glu Asn Val Thr Ser Tyr Thr Pro Asn
50          55          60

Ala Gly Tyr Leu Glu Leu Arg Gln Ala Val Gln Leu Tyr Met Lys Lys
65          70          75          80

Lys Ala Asp Phe Asn Tyr Asp Ala Glu Ser Glu Ile Ile Ile Thr Thr

```

85 90 95
 Gly Ala Ser Gln Ala Ile Asp Ala Ala Phe Arg Thr Ile Leu Ser Pro
 100 105 110
 Gly Asp Glu Val Ile Met Pro Gly Pro Ile Tyr Pro Gly Tyr Glu Pro
 115 120 125
 Ile Ile Asn Leu Cys Gly Ala Lys Pro Val Ile Val Asp Thr Thr Ser
 130 135 140
 His Gly Phe Lys Leu Thr Ala Arg Leu Ile Glu Asp Ala Leu Thr Pro
 145 150 155 160
 Asn Thr Lys Cys Val Val Leu Pro Tyr Pro Ser Asn Pro Thr Gly Val
 165 170 175
 Thr Leu Ser Glu Glu Glu Leu Lys Ser Ile Ala Ala Leu Leu Lys Gly
 180 185 190
 Arg Asn Val Phe Val Leu Ser Asp Glu Ile Tyr Ser Glu Leu Thr Tyr
 195 200 205
 Asp Arg Pro His Tyr Ser Ile Ala Thr Tyr Leu Arg Asp Gln Thr Ile
 210 215 220
 Val Ile Asn Gly Leu Ser Lys Ser His Ser Met Thr Gly Trp Arg Ile
 225 230 235 240
 Gly Phe Leu Phe Ala Pro Lys Asp Ile Ala Lys His Ile Leu Lys Val
 245 250 255
 His Gln Tyr Asn Val Ser Cys Ala Ser Ser Ile Ser Gln Lys Ala Ala
 260 265 270
 Leu Glu Ala Val Thr Asn Gly Phe Asp Asp Ala Leu Ile Met Arg Glu
 275 280 285
 Gln Tyr Lys Lys Arg Leu Asp Tyr Val Tyr Asp Arg Leu Val Ser Met
 290 295 300
 Gly Leu Asp Val Val Lys Pro Ser Gly Ala Phe Tyr Ile Phe Pro Ser
 305 310 315 320
 Ile Lys Ser Phe Gly Met Thr Ser Phe Asp Phe Ser Met Ala Leu Leu
 325 330 335
 Glu Asp Ala Gly Val Ala Leu Val Pro Gly Ser Ser Phe Ser Thr Tyr
 340 345 350
 Gly Glu Gly Tyr Val Arg Leu Ser Phe Ala Cys Ser Met Asp Thr Leu
 355 360 365
 Arg Glu Gly Leu Asp Arg Leu Glu Leu Phe Val Leu Lys Lys Arg Glu
 370 375 380
 Ala Met Gln Thr Ile Asn Asn Gly Val
 385 390

5

<210> 11
 <211> 1176
 <212> ADN
 <213> Lactobacillus amylovorus

10

<400> 11

```

atgccagaat tagctaata tttaggatta agcaaaaaga tcaactgatgt aaaagcttca      60
ggaattagaa tctttgataa caaagtttca gctattcctg gcattatcaa attgactttg     120
ggtgaaccag atatgaatac tcctgagcat gttaagcaag cggctattaa gaatattgca     180
gataatgatt cacactatgc tccacaaaag ggaaagcttg aattaagaaa agctatcagt     240
aaatatattga aaaagattac tggaattgaa tatgatccag aaacagaaat cgtagtaaca     300
gttggtgcaa ctgaagcaat taacgctacc ttgtttgcta ttactaatcc ggtgacaag      360
gttgcaattc ctacgccagt cttttctcta tattggcccg tggctacact tgctgatgcc     420
gattatgttt tgatgaatac tgcagaagat ggttttaagt taacacctaa gaagttagaa     480
gaaactatca aagaaaatcc aacaattaaa gcagtaattt tgaattatcc aactaaccce     540
actggtgttg aatatagcga agatgaaatt aaagctttgg ctaaggtaat taaagataat     600
catctgtacg taattaccga tgaaatttac agtactttga cttacggtgt aaaacacttt     660
tcaattgccca gcttaattcc agaagagca atttatatct ctggttttatc taaatcacat     720
gcgatgactg gttatcgttt aggctatgtt gccggacctg caaaaattat ggcagaaatt     780
ggtaaagttc atggccttat ggtgacgact acgacggatt catcacaagc tgccgcaatt     840
gaagcacttg aacacggact tgatgaccct gagaaatata gggaagttaa tgaaaagcgt     900
cgtgactatg ttttaaagga attagccgag atagagatgc aagcagttaa gccagaaggt     960
gcattttata tctttgctaa aattccagct aagtatggca aagacgatat gaaatttgcc    1020
ttggatttag cttttaaaga aaaagtgggt atcactccag gtatgtcatt tggctctggt    1080
ggtgaaggtc atattagatt atcttatgca tcaagtgatg aaaacttgca tgaggcaatg    1140
aagcgaatga agaaagtttt acaaggaggac gaataa                               1176

```

5

<210> 12
 <211> 391
 <212> PRT
 <213> Lactobacillus amylovorus

10

<400> 12

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

 Val Lys Ala Ser Gly Ile Arg Ile Phe Asp Asn Lys Val Ser Ala Ile
 20 25 30

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

 Glu His Val Lys Gln Ala Ala Ile Lys Asn Ile Ala Asp Asn Asp Ser
 50 55 60
 His Tyr Ala Pro Gln Lys Gly Lys Leu Glu Leu Arg Lys Ala Ile Ser
 65 70 75 80

 Lys Tyr Leu Lys Lys Ile Thr Gly Ile Glu Tyr Asp Pro Glu Thr Glu
 85 90 95

 Ile Val Val Thr Val Gly Ala Thr Glu Ala Ile Asn Ala Thr Leu Phe
 100 105 110

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

Ser Leu Tyr Trp Pro Val Ala Thr Leu Ala Asp Ala Asp Tyr Val Leu
130 135 140

Met Asn Thr Ala Glu Asp Gly Phe Lys Leu Thr Pro Lys Lys Leu Glu
145 150 155 160

Glu Thr Ile Lys Glu Asn Pro Thr Ile Lys Ala Val Ile Leu Asn Tyr
165 170 175

Pro Thr Asn Pro Thr Gly Val Glu Tyr Ser Glu Asp Glu Ile Lys Ala
180 185 190

Leu Ala Lys Val Ile Lys Asp Asn His Leu Tyr Val Ile Thr Asp Glu
195 200 205

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

Leu Ile Pro Glu Arg Ala Ile Tyr Ile Ser Gly Leu Ser Lys Ser His
225 230 235 240

Ala Met Thr Gly Tyr Arg Leu Gly Tyr Val Ala Gly Pro Ala Lys Ile
245 250 255

Met Ala Glu Ile Gly Lys Val His Gly Leu Met Val Thr Thr Thr Thr
260 265 270

Asp Ser Ser Gln Ala Ala Ala Ile Glu Ala Leu Glu His Gly Leu Asp
275 280 285

Asp Pro Glu Lys Tyr Arg Glu Val Tyr Glu Lys Arg Arg Asp Tyr Val
290 295 300

Leu Lys Glu Leu Ala Glu Ile Glu Met Gln Ala Val Lys Pro Glu Gly
305 310 315 320

Ala Phe Tyr Ile Phe Ala Lys Ile Pro Ala Lys Tyr Gly Lys Asp Asp
325 330 335

Met Lys Phe Ala Leu Asp Leu Ala Phe Lys Glu Lys Val Gly Ile Thr
340 345 350

Pro Gly Ser Ala Phe Gly Pro Gly Gly Glu Gly His Ile Arg Leu Ser
355 360 365

Tyr Ala Ser Ser Asp Glu Asn Leu His Glu Ala Met Lys Arg Met Lys
370 375 380

Lys Val Leu Gln Glu Asp Glu
385 390

5

<210> 13
<211> 1413
<212> ADN
<213> R. sphaeroides

10

<400> 13

atgcgcgagc ctcttgccct cgagatcgac ccggggccacg gcggcccgcgt gttcctcgcc	60
atcgccgagg cgatcacccct cgacatcacc cgcgggcggc tgaggcccg agcgagactg	120
cccggcacac gcgcgctggc gcgggcgctc ggcgtgcac gcaacacggt ggatgccgcc	180

```

tatacaggagt tgctgaccca gggctggctg caggccgagc ccgcgcgggg caccttcgtg 240
gcgacaggatc tgccgcaggg gatgctgggtg cacaggcccc cgcccgcgcc ggtcgagccg 300
gtcgcgatgc gcgcggggct cgccttctcc gatggcgcgc cggaccccgga gctggtgccc 360
gacaaggcgc tggcgcgggc ctttcgccgg gcgctcctgt cgcgcgcctt ccgcgcggga 420
gcggattacg gcgatgcccc cggcacctcc tcgctgcggg aggcgctggc agcctatctc 480
gcctcggacc ggggcgtggt cgcggatcct gcgcgctgc tgatcgcgcg ggcagccag 540
atggcgctgt tcctggtagc ccgggcggcg ctggcgccgg gagaggcgat cgcggtcgag 600
gagccgggct atccgctggc ctgggaggcg ttccgcgcag cgggagcgga ggtgcgcggc 660
gtgccggtgg atggcggcgg cctcaggatc gacgcgctcg aggcgcgct ggcccgggat 720
ccgcgaatcc gggcggtcta tgtcacgccc catcaccagt atccgacgac cgtcaccatg 780
ggcgcggcgc ggcggttgca gcttctggaa ctggcagagc gccaccggct cgcgctgac 840
gaggacgact acgaccacga ataccgcttc gagggccgct cgggtgctgcc gctggctgcc 900
cgcgcgccgg aaggtctgcc gctgatctat gtgggctcgc tgtcgaaact gctctcgccc 960
ggtatccggc tgggatacgc gctggcgccc gagcgctgc tgaccgcat ggccgcggcg 1020
cgcgcgccca tcgaccggca gggcgacgcg ccgctcgagg cggcgctggc cgagctgac 1080
cgcgacggcg atctgggccc tcatgcccgc aaggcgcgca gggctctacc ggcgcgggcg 1140
gatctgctgg cggagcgtct cacggcgag ctggccgggc gcgccgcctt cgatctgccg 1200
gccgggggcc tcgcgctgtg gctgcgctgc gcggcgctct cggccgagac ctgggcccga 1260
gccgcagggc aggcggggct cgcctgctg ccgggcacgc gcttcgcgct ggagagccc 1320
gcgcgcgagg ccttcgggct gggctatgcg gcgctggacg aggggcagat cgcgcggcg 1380
gtggagatcc tcgcccggag cttccccggc tga 1413

```

<210> 14

<211> 470

<212> PRT

<213> R. sphaeroides

<400> 14

```

Met Arg Glu Pro Leu Ala Leu Glu Ile Asp Pro Gly His Gly Gly Pro
1           5           10          15

```

```

Leu Phe Leu Ala Ile Ala Glu Ala Ile Thr Leu Asp Ile Thr Arg Gly
20           25           30

```

```

Arg Leu Arg Pro Gly Ala Arg Leu Pro Gly Thr Arg Ala Leu Ala Arg
35           40           45

```

Ala Leu Gly Val His Arg Asn Thr Val Asp Ala Ala Tyr Gln Glu Leu
 50 55 60

Leu Thr Gln Gly Trp Leu Gln Ala Glu Pro Ala Arg Gly Thr Phe Val
 65 70 75 80

Ala Gln Asp Leu Pro Gln Gly Met Leu Val His Arg Pro Ala Pro Ala
 85 90 95

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

Ala Pro Asp Pro Glu Leu Val Pro Asp Lys Ala Leu Ala Arg Ala Phe
 115 120 125

Arg Arg Ala Leu Leu Ser Pro Ala Phe Arg Ala Gly Ala Asp Tyr Gly
 130 135 140

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

Ala Ser Asp Arg Gly Val Val Ala Asp Pro Ala Arg Leu Leu Ile Ala
 165 170 175

Arg Gly Ser Gln Met Ala Leu Phe Leu Val Ala Arg Ala Ala Leu Ala
 180 185 190

Pro Gly Glu Ala Ile Ala Val Glu Glu Pro Gly Tyr Pro Leu Ala Trp
 195 200 205

Glu Ala Phe Arg Ala Ala Gly Ala Glu Val Arg Gly Val Pro Val Asp
 210 215 220

Gly Gly Gly Leu Arg Ile Asp Ala Leu Glu Ala Ala Leu Ala Arg Asp
 225 230 235 240

Pro Arg Ile Arg Ala Val Tyr Val Thr Pro His His Gln Tyr Pro Thr
 245 250 255

Thr Val Thr Met Gly Ala Ala Arg Arg Leu Gln Leu Leu Glu Leu Ala
 260 265 270

Glu Arg His Arg Leu Ala Leu Ile Glu Asp Asp Tyr Asp His Glu Tyr
 275 280 285

Arg Phe Glu Gly Arg Pro Val Leu Pro Leu Ala Ala Arg Ala Pro Glu
 290 295 300

Gly Leu Pro Leu Ile Tyr Val Gly Ser Leu Ser Lys Leu Leu Ser Pro
 305 310 315 320

Gly Ile Arg Leu Gly Tyr Ala Leu Ala Pro Glu Arg Leu Leu Thr Arg
325 330 335

Met Ala Ala Ala Arg Ala Ala Ile Asp Arg Gln Gly Asp Ala Pro Leu
340 345 350

Glu Ala Ala Leu Ala Glu Leu Ile Arg Asp Gly Asp Leu Gly Arg His
355 360 365

Ala Arg Lys Ala Arg Arg Val Tyr Arg Ala Arg Arg Asp Leu Leu Ala
370 375 380

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

Ala Gly Gly Leu Ala Leu Trp Leu Arg Cys Ala Gly Val Ser Ala Glu
405 410 415

Thr Trp Ala Glu Ala Ala Gly Gln Ala Gly Leu Ala Leu Leu Pro Gly
420 425 430

Thr Arg Phe Ala Leu Glu Ser Pro Ala Pro Gln Ala Phe Arg Leu Gly
435 440 445

Tyr Ala Ala Leu Asp Glu Gly Gln Ile Ala Arg Ala Val Glu Ile Leu
450 455 460

Ala Arg Ser Phe Pro Gly
465 470

<210> 15
<211> 35
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador

<400> 15

ggtattgagg gtcgcatgaa ggttttagtc aatgg

35

<210> 16
<211> 37
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador

<400> 16

agaggagagt tagagcctta tgaaatgcta gcagcct

37

<210> 17
<211> 35
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador

<400> 17

	ggtattgagg gtcgcatggt cgacgccctc gcccg	35
5	<210> 18 <211> 37 <212> ADN <213> Secuencia artificial	
10	<220> <223> cebador <400> 18	
15	agaggagagt tagagcctca gagactggtg aacttgc	37
20	<210> 19 <211> 35 <212> ADN <213> Secuencia artificial <220> <223> cebador <400> 19	
25	ggtattgagg gtcgcatgga acattgctg aatcc	35
30	<210> 20 <211> 37 <212> ADN <213> Secuencia artificial <220> <223> cebador <400> 20	
35	agaggagagt tagagcctta aacgccgttg ttatcg	37
40	<210> 21 <211> 35 <212> ADN <213> Secuencia artificial <220> <223> cebador <400> 21	
45	ggtattgagg gtcgcatgcg cgagcctctt gccct	35
50	<210> 22 <211> 37 <212> ADN <213> Secuencia artificial <220> <223> cebador <400> 22	
55	agaggagagt tagagcctca gccggggaag ctccggg	37
60	<210> 23 <211> 35 <212> ADN <213> Secuencia artificial	
65		

	<220>		
	<223> cebador		
5	<400> 23		
	ggtattgagg gtcgcatgtc cacgcaggcg gccat		35
10	<210> 24		
	<211> 37		
	<212> ADN		
	<213> Secuencia artificial		
15	<220>		
	<223> cebador		
	<400> 24		
	agaggagagt tagagcctca ctcacgattc acattgc		37
20	<210> 25		
	<211> 35		
	<212> ADN		
	<213> Secuencia artificial		
25	<220>		
	<223> cebador		
	<400> 25		
30	ggtattgagg gtcgcatgcc agaattagct aatga		35
35	<210> 26		
	<211> 37		
	<212> ADN		
	<213> Secuencia artificial		
40	<220>		
	<223> cebador		
	<400> 26		
	agaggagagt tagagcctta ttcgtcctct tgtaaaa		37
45	<210> 27		
	<211> 35		
	<212> ADN		
	<213> Secuencia artificial		
50	<220>		
	<223> cebador		
	<400> 27		
55	ggtattgagg gtcgcatgcg ctctacgacg gctcc		35
60	<210> 28		
	<211> 37		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> cebador		
	<400> 28		
65	agaggagagt tagagcctca gccgcgacg accttgg		37

5	<210> 29	
	<211> 35	
	<212> ADN	
	<213> Secuencia artificial	
10	<220>	
	<223> cebador	
	<400> 29	
	ggtattgagg gtcgcatggt tgagaacatt accgc	35
15	<210> 30	
	<211> 37	
	<212> ADN	
	<213> Secuencia artificial	
20	<220>	
	<223> cebador	
	<400> 30	
	agaggagagt tagagcctta cagcactgcc acaatcg	37
25	<210> 31	
	<211> 1194	
	<212> ADN	
	<213> E. coli	
30	<400> 31	
	gtgtttcaaa aagttgacgc ctacgctggc gacccgattc ttacgcttat ggagcgtttt	60
	aaagaagacc ctgcgacgca caaagtgaat ttaagtatcg gtctgtacta caacgaagac	120
	ggaattattc cacaactgca agccgtggcg gaggcggaag cgcgcctgaa tgcgcagcct	180
35	catggcgctt cgctttatctt accgatggaa gggcttaact gctatcgcca tgccattgcg	240
	ccgctgctgt ttggtgcgga ccatccggta ctgaaacaac agcgcgtagc aaccattcaa	300
	acccttgggc gctccggggc attgaaagtg ggcgcggatt tcctgaaacg ctacttcccg	360
	gaatcaggcg tctgggtcag cgatcctacc tgggaaaacc gcgtagcaat attcgccggg	420
40	gctggattcg aagttagtac ttacccctgg tatgacgaag cgactaacgg cgtgcgcttt	480
	aatgacctgt tggcgacgct gaaaacatta cctgcccgca gtatttgtgt gctgcatcca	540
	tgttgccaca acccaacggg tgccgatctc actaatgac agtgggatgc ggtgattgaa	600
	attctcaaag cccgcgagct tattccattc ctcgatattg cctatcaagg atttggtgcc	660
45	ggtatggaag aggatgccta cgctattcgc gccattgcca gcgctggatt acccgctctg	720
	gtgagcaatt cgttctcgaa aattttctcc ctttacggcg agcgcgctcg cggaactttct	780
	gttatgtgtg aagatgccga agccgctggc cgcgtactgg ggcaattgaa agcaacagtt	840
	cgccgcaact actccagccc gccgaatttt ggtgcgcagg tggtagctgc agtgctgaat	900
50	gacgaggcat tgaaagccag ctggctggcg gaagtagaag agatgcgtac tcgcattctg	960
	gcaatgcgtc aggaattggt gaaggtatta agcacagaga tgccagaacg caatttcgat	1020
	tatctgctta atcagcgcg catgttcagt tataccggtt taagtgcgcg tcaggttgac	1080
	cgactacgtg aagaatttgg tgtctatctc atcgccagcg gtcgcatgtg tgcgcgagg	1140
55	ttaaatacgg caaatgtaca acgtgtggca aaggcgtttg ctgcggtgat gtaa	1194
35	<210> 32	
	<211> 397	

<212> PRT

<213> E. coli

<400> 32

Val Phe Gln Lys Val Asp Ala Tyr Ala Gly Asp Pro Ile Leu Thr Leu
1 5 10 15

Met Glu Arg Phe Lys Glu Asp Pro Arg Ser Asp Lys Val Asn Leu Ser
20 25 30

Ile Gly Leu Tyr Tyr Asn Glu Asp Gly Ile Ile Pro Gln Leu Gln Ala
35 40 45

Val Ala Glu Ala Glu Ala Arg Leu Asn Ala Gln Pro His Gly Ala Ser
50 55 60

Leu Tyr Leu Pro Met Glu Gly Leu Asn Cys Tyr Arg His Ala Ile Ala
65 70 75 80

Pro Leu Leu Phe Gly Ala Asp His Pro Val Leu Lys Gln Gln Arg Val
85 90 95

Ala Thr Ile Gln Thr Leu Gly Gly Ser Gly Ala Leu Lys Val Gly Ala
100 105 110

Asp Phe Leu Lys Arg Tyr Phe Pro Glu Ser Gly Val Trp Val Ser Asp
115 120 125

Pro Thr Trp Glu Asn Arg Val Ala Ile Phe Ala Gly Ala Gly Phe Glu
130 135 140

Val Ser Thr Tyr Pro Trp Tyr Asp Glu Ala Thr Asn Gly Val Arg Phe
145 150 155 160

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

Leu Leu His Pro Cys Cys His Asn Pro Thr Gly Ala Asp Leu Thr Asn
180 185 190

Asp Gln Trp Asp Ala Val Ile Glu Ile Leu Lys Ala Arg Glu Leu Ile
195 200 205

Pro Phe Leu Asp Ile Ala Tyr Gln Gly Phe Gly Ala Gly Met Glu Glu
210 215 220

Asp Ala Tyr Ala Ile Arg Ala Ile Ala Ser Ala Gly Leu Pro Ala Leu
225 230 235 240

Val Ser Asn Ser Phe Ser Lys Ile Phe Ser Leu Tyr Gly Glu Arg Val
245 250 255

Gly Gly Leu Ser Val Met Cys Glu Asp Ala Glu Ala Ala Gly Arg Val
260 265 270

Leu Gly Gln Leu Lys Ala Thr Val Arg Arg Asn Tyr Ser Ser Pro Pro
275 280 285

<210> 33

<211> 35

<212> ADN

<213> Secuencia artificial

	<220>		
	<223> cebador		
5	<400> 33 35		
	ggtattgagg gtcgctgtt tcaaaaagtt gacgc		35
10	<210> 34		
	<211> 37		
	<212> ADN		
	<213> Secuencia artificial		
15	<220>		
	<223> cebador		
	<400> 34 37		
	agaggagagt tagagcctta catcaccgca gcaaacg		37
20	<210> 35		
	<211> 35		
	<212> ADN		
	<213> Secuencia artificial		
25	<220>		
	<223> cebador		
	<400> 35		
30	ggtattgagg gtcgcatgga gtccaaagtc gttga		35
35	<210> 36		
	<211> 37		
	<212> ADN		
	<213> Secuencia artificial		
40	<220>		
	<223> cebador		
	<400> 36		
	agaggagagt tagagcctta cacttgaaa acagcct		37
45	<210> 37		
	<211> 35		
	<212> ADN		
	<213> Secuencia artificial		
50	<220>		
	<223> cebador		
	<400> 37		
55	ggtattgagg gtcgcatgaa aaactggaaa acaag		35
60	<210> 38		
	<211> 37		
	<212> ADN		
	<213> Secuencia artificial		
65	<220>		
	<223> cebador		
	<400> 38		
	agaggagagt tagagcctta cagcttagcg ccttcta		37

5	<210> 39 <211> 35 <212> ADN <213> Secuencia artificial	
	<220> <223> cebador	
10	<400> 39 ggtattgagg gtcgcatgcg aggggcatta ttaa	35
15	<210> 40 <211> 36 <212> ADN <213> Secuencia artificial	
20	<220> <223> cebador	
	<400> 40 agaggagagt tagagcctca gcccttgagc gcgaag	36
25	<210> 41 <211> 1416 <212> ADN <213> E. coli	
30	<400> 41	

```

atggaaaact ttaaaccatct ccctgaaccg ttccgcattc gtgttattga gccagtaaaa    60
cgtaccacccc gcgcttatcg tgaagaggca attattaaat ccggtatgaa cccgttcctg    120
ctggatagcg aagatgtttt tatcgattta ctgaccgaca gcggcaccgg gccggtgacg    180
cagagcatgc aggctgcgat gatgcgcggc gacgaagcct acagcggcag tcgtagctac    240
tatgcgcttag ccgagtcagt gaaaaatata ttggttatc aatacaccat tccgactcac    300
cagggccgtg gcgcagagca aatctatat tccggtactga ttaaaaaacg cgagcaggaa    360
aaaggcctgg atcgagcaa aatggtggcg ttctctaact atttctttga taccacgcag    420
ggccatagcc agatcaacgg ctgtaccgtg cgtaacgtct atatcaaaga agccttcgat    480
acgggcgtgc gttacgactt taaaggcaac ttgaccttg agggattaga acgcggtatt    540
gaagaagtgt gtccgaataa cgtgccgtat atcgttgcaa ccatcaccag taactctgca    600
ggtggtcagc cggtttctact ggcaaaacta aaagcgatgt acagcatcgc gaagaaatac    660
gatattccgg tggtaatgga ctccgcgcgc ttgctgaaa acgcctatct catcaagcag    720
cgtgaagcag aatacaaaga ctggaccatc gagcagatca ccgcggaac ctacaaatat    780
gccgatatgc tggcgatgtc cgccaagaaa gatgcgatgg tgccgatggg cggcctgctg    840
tgcatgaaag acgacagctt ctttgatgtg tacaccgagt gcagaaccct ttgcgtggtg    900
caggaaggct tcccgaacata tggcggcctg gaaggcggcg cgatggagcg tctggcggta    960
ggtctgtatg acggcatgaa tctcgactgg ctggcttata gtatcgcgca ggtacagtat   1020
ctggctgatg gtctggaaga gattggcggt gtctgccagc aggcggggcg tcacgcggca   1080
ttcgttgatg ccggtaaact gttgccgat atcccggcag accagttccc ggcacaggcg   1140
ctggcctgcg agctgtataa agtcgccgtt atccgtgcgg tagaaattgg ctctttcctg   1200
ttaggcccgcg atccgaaaac cggtaaacaa ctgccatgcc cggctgaact gctgcgttta   1260
accattccgc gcgcaacata tactcaaaca catatggact tcattattga agcctttaa   1320
catgtgaaag agaacgcggc gaatattaaa ggattaacct ttacgtacga accgaaagta   1380
ttgcgtcact tcaccgcaa acttaaagaa gtttaa                                1416

```

<210> 42

<211> 1371

<212> ADN

<213> Citrobacter freundii

<400> 42

```

atgaattatc cggcagaacc cttccgtatt aaaagcgttg aaactgtatc tatgatcccg    60
cgtgatgaac gcctcaagaa aatgcaggaa gcgggttaca atactttcct gttaaattcg    120
aaagatattt atattgacct gctgacagac agtggcacta acgcaatgag cgacaagcag    180

```

tgggcccggaa tgatgatggg tgatgaagcg tacgcgggca gcgaaaactt ctatcatctg 240
 gaaagaaccg tgcaggaact gtttggcttt aaacatattg ttccgactca ccaggggctg 300
 ggcgcagaaa acctgttatc gcagctggct attaaacctg ggcaatatgt tgccgggaat 360
 atgtatttca ctaccacccg ttatcaccag gaaaaaatg gtgcggtggt tgctgatatc 420
 gttcgtgacg aagcgacga tgccggtctg aatattgctg ttaaagggtga tatcgatctt 480
 aaaaaattac aaaagctgat tgatgaaaaa ggcgcagaga atattgcgta tatctgcctg 540
 gcggtgacgg ttaacctcgc gggcgcccaa ccggtctcga tggctaacat gcgtgcggtg 600
 cgtgaactga cagaagcgca tggcattaaa gtgttctacg acgctaccg ctgcgtagaa 660
 aacgcctact ttatcaaaga gcaagagcag ggctttgaga acaagagcat cgccgagatc 720
 gtgcatgaga tgttcagcta cgcgcagcgt tgtaccatga gtggtaaaaa agactgtctg 780
 gtgaacatcg gcggcttctt gtgcatgaac gatgacgaaa tgttctcttc tgccaaagag 840
 ttagtcgtgg tctacgaagg gatgccatct tacggcggcc tggccggacg tgatatggaa 900
 gcgatggcga ttggcctgcg tgaagccatg cagtacgaat atattgagca ccgctgaag 960
 caggttcgtt acctgggcca taagctgaaa gccgctggcg taccgattgt tgaaccggta 1020
 ggcggtcacg cggatttctt cgatgcgcgt cgcttctgcg agcatctgac gcaagatgag 1080
 ttcccggcac aaagtctggc tgccagcatc tatgtggaaa ccggcgtgcg cagtatggag 1140
 cgcggaatta tctctgcggg ccgtaataac gtgaccggtg aacaccacag accgaaactg 1200
 gaaaccgtgc gtctgactat tccacgtcgt gtttatacct acgcacatat ggatgttgtg 1260
 gctgacggta ttattaaact ttaccagcac aaagaagata ttccggggct gaagtttatt 1320
 tacgagccga agcagttgcg tttctttact gcacgctttg attacatcta a 1371

<210> 43
 <211> 35
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador

<400> 43

ggtattgagg gtcgcatgga aaactttaa catct

<210> 44
 <211> 37
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador

<400> 44

agaggagagt tagagcctta aactcttta agtttg

<210> 45
 <211> 35
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador

<400> 45

35

37

	ggtattgagg gtcgcatgaa ttatccggca gaacc	35
5	<210> 46 <211> 37 <212> ADN <213> Secuencia artificial	
10	<220> <223> cebador	
	<400> 46	
15	agaggagagt tagagcctta gatgtaatca aagcgtg	37
20	<210> 47 <211> 31 <212> ADN <213> Secuencia artificial	
	<220> <223> cebador	
25	<400> 47	
	ccagggcacc ggcgcagagc aaatctatat t	31
30	<210> 48 <211> 30 <212> ADN <213> Secuencia artificial	
35	<220> <223> cebador	
	<400> 48	
40	tgcgccggtg ccctggtgag tcggaatggt	30
45	<210> 49 <211> 32 <212> ADN <213> Secuencia artificial	
	<220> <223> cebador	
	<400> 49	
50	tctgcacgc ggcaaagggt tctgcactcg gt	32
55	<210> 50 <211> 30 <212> ADN <213> Secuencia artificial	
	<220> <223> cebador	
60	<400> 50	
	cttgccgcg tgcaggaagg ctcccgaca	30
65	<210> 51 <211> 29 <212> ADN <213> Secuencia artificial	

	<220> <223> cebador	
5	<400> 51 aggggaccgg cgcagaaaac ctgttatcg	29
10	<210> 52 <211> 32 <212> ADN <213> Secuencia artificial	
15	<220> <223> cebador <400> 52	
20	tctgcgccgg tcccctggtg agtcggaaca at	32
25	<210> 53 <211> 29 <212> ADN <213> Secuencia artificial	
30	<220> <223> cebador <400> 53 gtagtccgc gtctacgaag ggatgccat	29
35	<210> 54 <211> 29 <212> ADN <213> Secuencia artificial	
40	<220> <223> cebador <400> 54 gtagacgagg actaactctt tggcagaag	29
45	<210> 55 <211> 35 <212> ADN <213> Secuencia artificial	
50	<220> <223> cebador <400> 55	
55	ggtattgagg gtcgcatgta cgaactggga gttgt	35
60	<210> 56 <211> 37 <212> ADN <213> Secuencia artificial	
65	<220> <223> cebador <400> 56 agaggagagt tagagcctta gtcaatatat ttcaggc	37

5	<210> 57 <211> 35 <212> ADN <213> Secuencia artificial	
10	<220> <223> cebador <400> 57 ggtattgagg gtcgcatgtc cggcatcggt gtcca	35
15	<210> 58 <211> 37 <212> ADN <213> Secuencia artificial	
20	<220> <223> cebador <400> 58 agaggagagt tagagcctca gacatatttc agtccca	37
25	<210> 59 <211> 35 <212> ADN <213> Secuencia artificial	
30	<220> <223> cebador <400> 59 ggtattgagg gtcgcatgcg actgaacaac ctcgg	35
35	<210> 60 <211> 37 <212> ADN <213> Secuencia artificial	
40	<220> <223> cebador <400> 60 agaggagagt tagagcctca gttctccacg tattcca	37
45	<210> 61 <211> 35 <212> ADN <213> Secuencia artificial	
50	<220> <223> cebador <400> 61 ggtattgagg gtcgcatgag cgtggtcac cggaa	35
55	<210> 62 <211> 37 <212> ADN <213> Secuencia artificial	
60	<220>	
65		

	<223> cebador	
	<400> 62	
5	agaggagagt tagagcctca atcgatatat ttcagtc	37
	<210> 63	
	<211> 35	
10	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> cebador	
15	<400> 63	
	ggtattgagg gtcgcatgag cctggtaaat atgaa	35
	<210> 64	
20	<211> 37	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
25	<223> cebador	
	<400> 64	
30	agaggagagt tagagcctta tgactttaac gcgttga	37
	<210> 65	
	<211> 684	
	<212> ADN	
35	<213> C. testosteroni	
	<400> 65	
	atgtacgaac tgggagttgt ctaccgcaat atccagcgcg ccgaccgcgc tgctgctgac	60
	ggcctggcgc ccctgggctc cgccaccgtg cacgaggcca tgggcccgcgt cggctctgctc	120
	aagccctata tgcgccccat ctatgccggc aagcaggtct cgggcaccgc cgtcacgggtg	180
	ctgctgcagc cggcgacaaa ctggatgatg catgtggctg ccgagcagat tcagcccggc	240
	gacatcgtgg tcgcagccgt caccgcagag tgcaccgacg gctacttcgg cgatctgctg	300
	gccaccagct tccaggcgcg cggcgcacgt gcgctgatca tcgatgccgg cgtgcgcgac	360
	gtgaagacgc tgcaggagat ggactttccg gtctggagca aggccatctc ttccaagggc	420
	acgatcaagg ccaccctggg ctcggtcaac atccccatcg tctgcgccgg catgctggtc	480
	acgcccgggtg acgtgatcgt ggccgacgac gacggcgtgg tctgcgtgcc cgccgcgcgt	540
	gccgtggaag tgctggccgc cgcccagaag cgtgaaagct tcgaaggcga aaagcgcgcc	600
	aagctggcct cgggcatcct cggcctggat atgtacaaga tgcgcgagcc cctggaaaag	660
	gccggcctga aatatattga ctaa	684
40	<210> 66	
	<211> 227	
	<212> PRT	
	<213> C. testosteroni	
45	<400> 66	

Met Tyr Glu Leu Gly Val Val Tyr Arg Asn Ile Gln Arg Ala Asp Arg
 1 5 10 15

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

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

Ala Gly Lys Gln Val Ser Gly Thr Ala Val Thr Val Leu Leu Gln Pro
 50 55 60

Gly Asp Asn Trp Met Met His Val Ala Ala Glu Gln Ile Gln Pro Gly
 65 70 75 80

Asp Ile Val Val Ala Ala Val Thr Ala Glu Cys Thr Asp Gly Tyr Phe
 85 90 95

Gly Asp Leu Leu Ala Thr Ser Phe Gln Ala Arg Gly Ala Arg Ala Leu
 100 105 110

Ile Ile Asp Ala Gly Val Arg Asp Val Lys Thr Leu Gln Glu Met Asp
 115 120 125

Phe Pro Val Trp Ser Lys Ala Ile Ser Ser Lys Gly Thr Ile Lys Ala
 130 135 140

Thr Leu Gly Ser Val Asn Ile Pro Ile Val Cys Ala Gly Met Leu Val
 145 150 155 160

Thr Pro Gly Asp Val Ile Val Ala Asp Asp Asp Gly Val Val Cys Val
 165 170 175

Pro Ala Ala Arg Ala Val Glu Val Leu Ala Ala Ala Gln Lys Arg Glu
 180 185 190

Ser Phe Glu Gly Glu Lys Arg Ala Lys Leu Ala Ser Gly Ile Leu Gly
 195 200 205

Leu Asp Met Tyr Lys Met Arg Glu Pro Leu Glu Lys Ala Gly Leu Lys
 210 215 220

Tyr Ile Asp
 225

5

<210> 67
 <211> 42
 <212> ADN
 <213> Secuencia artificial

10

<220>
 <223> cebador

<400> 67

15

actcggatcc gaaggagata tacatatgta cgaactggga ct

42

<210> 68
 <211> 33
 <212> ADN
 <213> Secuencia artificial

20

<220>
 <223> cebador

	<400> 68	
5	cggtgtcga ccgtagtca atatattca ggc	33
	<210> 69	
	<211> 31	
	<212> ADN	
10	<213> Secuencia artificial	
	<220>	
	<223> cebador	
15	<400> 69	
	cgcggtacca taatggtga gaacattacc g	31
	<210> 70	
20	<211> 30	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
25	<223> cebador	
	<400> 70	
	acgcgtcgac ttacagcact gccacaatcg	30
30	<210> 71	
	<211> 32	
	<212> ADN	
	<213> Secuencia artificial	
35	<220>	
	<223> cebador	
	<400> 71	
40	ccggaattca taatggtcga actgggagtt gt	32
	<210> 72	
	<211> 33	
45	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> cebador	
50	<400> 72	
	gaatgcggcc gcttagtcaa tatattcag gcc	33
55	<210> 73	
	<211> 15	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
60	<223> cebador	
	<400> 73	
	ggtattgagg gtcgc	15
65	<210> 74	
	<211> 17	

5	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> cebador	
	<400> 74	
	agaggagagt tagagcc	17

REIVINDICACIONES

1. Procedimiento de producción de monatina, que comprende convertir el triptófano y/o el ácido indol-3-láctico en monatina, en el que el procedimiento comprende:

- (a) poner en contacto el triptófano o el ácido indol-3-láctico con un primer polipéptido, en el que, cuando el triptófano está en contacto con el primer polipéptido, el primer polipéptido se selecciona de entre triptófano aminotransferasa (CE 2.6.1.27), triptófano deshidrogenasa (CE 1.4.1.19), tirosina aminotransferasa (aromática) (CE 2.6.1.5), triptofan-fenilpiruvato transaminasa (CE 2.6.1.28), sustrato múltiple aminotransferasa (CE 2.6.1.-), aspartato aminotransferasa (CE 2.6.1.1), triptófano oxidasa, L-aminoácido oxidasa (CE 1.4.3.2), D-aminoácido deshidrogenasa (CE 1.4.99.1), D-aminoácido oxidasa (CE 1.4.3.3), D-triptófano aminotransferasa, D-alanina aminotransferasa (CE 2.6.1.21), glutamato deshidrogenasa (CE 1.4.1.2-4), fenilalanina deshidrogenasa (CE 1.4.1.20), o una combinación de las mismas, en el que, cuando el ácido indol-3-láctico se pone en contacto con el primer polipéptido, el primer polipéptido se selecciona de entre indol-lactato deshidrogenasa (CE 1.1.1.110), R-4 hidroxifenil-lactato deshidrogenasa (CE 1.1.1.222), 3-(4)-hidroxifenilpiruvato reductasa (CE 1.1.1.237), lactato deshidrogenasa (CE 1.1.1.27, 1.1.1.28, 1.1.2.3), (3-imidazol-5-il) lactato deshidrogenasa (CE 1.1.1.111), y lactato oxidasa (CE 1.1.3.-), o una combinación de las mismas, en el que el primer polipéptido convierte el triptófano o el ácido indol-3-láctico en indol-3-piruvato;
- (b) poner en contacto el indol-3-piruvato con un segundo polipéptido y una fuente de carbono C3, en el que el segundo polipéptido es una enzima CE 4.1.2 o CE 4.1.3., en el que el segundo polipéptido convierte el indol-3-piruvato y la fuente de carbono C3 en 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico; y
- (c) poner en contacto el 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico con un tercer polipéptido, en el que el tercer polipéptido se selecciona de entre tirosina aminotransferasa (aromática) (CE 2.6.1.5), triptófano deshidrogenasa (CE 1.4.1.19), triptófano-fenilpiruvato transaminasa (CE 2.6.1.28), triptófano aminotransferasa (CE 2.6.1.27), aspartato aminotransferasa (CE 2.6.1.1), D-alanina aminotransferasa (CE 2.6.1.21), D-triptófano aminotransferasa, glutamato deshidrogenasa (CE 1.4.1.2-4), fenilalanina deshidrogenasa (CE 1.4.1.20), sustrato múltiple aminotransferasa (CE 2.6.1.-), D-aminoácido deshidrogenasa (CE 1.4.99.1), una lisina épsilon aminotransferasa, o una combinación de las mismas, en el que el tercer polipéptido convierte el 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico en monatina.

2. Procedimiento según la reivindicación 1, en el que, cuando el primer polipéptido que convierte el triptófano en indol-3-piruvato es una aminoácido oxidasa (CE 1.4.3.-), se utiliza además una catalasa (CE 1.11.1.6).

3. Procedimiento según la reivindicación 2, en el que la aminoácido oxidasa es una triptófano oxidasa.

4. Procedimiento según la reivindicación 1, en el que la enzima CE 4.1.3 es la 4-hidroxi-2-oxoglutarato glioxilato-liasa (CE 4.1.3.16), la 4-hidroxi-4-metil-2-oxoglutarato piruvato-liasa (CE 4.1.3.17) o una combinación de las mismas.

5. Procedimiento según la reivindicación 1, en el que el segundo polipéptido es una secuencia de aminoácidos seleccionada de entre el grupo constituido por:

- (a) SEC. ID. nº 66, nº de registro en Genbank: CAC46344, nº de registro en Genbank: CAB 14127.1, nº de registro en Genbank: AAC74920.1 o nº de registro en Genbank: CAC47463.1;
- (b) una secuencia de aminoácidos que comprende por lo menos 90% de identidad de la secuencia con la SEC. ID. nº 66, nº de registro en Genbank: CAC46344, nº de registro en Genbank: CAB14127.1, nº de registro en Genbank: AAC74920.1 o nº de registro en Genbank: CAC47463.1; y
- (c) secuencias de aminoácidos que se diferencian de la SEC. ID. nº 66, nº de registro en Genbank: CAC46344, nº de registro en Genbank: CAB 14127.1, nº de registro en Genbank: AAC74920.1 o nº de registro en Genbank: CAC47463.1 en menos de 50 sustituciones de aminoácidos conservadoras, en el que la secuencia de aminoácidos convierte el indol-3-piruvato y el piruvato en 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico.

6. Procedimiento según la reivindicación 1, en el que cuando el tercer polipéptido es una aspartato aminotransferasa, el 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutarico se pone en contacto con el aspartato para producir oxaloacetato y monatina.

7. Procedimiento según la reivindicación 6, que comprende además poner en contacto oxaloacetato con una descarboxilasa.

8. Procedimiento según la reivindicación 1, en el que cuando el tercer polipéptido es una lisina épsilon aminotransferasa, el 2-hidroxiácido 2-(indol-3-ilmetil)-cetoglutarico se pone en contacto además con la lisina.

9. Procedimiento según la reivindicación 1, en el que el primer y/o el tercer polipéptido son seleccionados de entre el grupo constituido por:

- (a) una secuencia de aminoácidos representada en las SEC. ID. nº: 2, 4, 6, 8, 10, 12, 14, 32, nº de registro en Genbank: NP_388848.1, nº de registro en Genbank: ZP00005082.1 o nº de registro en Genbank: ACC74014.1;
- (b) una secuencia de aminoácidos que comprende por lo menos 90% de identidad de secuencia con la secuencia representada en las SEC. ID. nº 2, 4, 6, 8, 10, 12, 14, 32, nº de registro en Genbank: NP_388848.1, nº de registro en Genbank: ZP00005082.1 o nº de registro en Genbank: ACC74014.1; y
- (c) secuencias de aminoácidos que se diferencian de las SEC. ID. nº 2, 4, 6, 8, 10, 12, 14, 32, nº de registro en Genbank: NP_388848.1, nº de registro en Genbank: ZP00005082.1 o nº de registro en Genbank: ACC74014.1 en menos de 50 sustituciones de aminoácidos conservadoras, en el que la secuencia de aminoácidos convierte el triptófano en indol-3-piruvato.

10. Procedimiento según la reivindicación 1, en el que el triptófano es producido a partir de indol utilizando el polipéptido de CE 4.1.99.1.

11. Procedimiento según la reivindicación 1, en el que la fuente de carbono C3 se selecciona de entre piruvato, fosfoenolpiruvato, alanina, serina, cisteína o una combinación de los mismos.

12. Procedimiento según la reivindicación 11, en el que el piruvato es producido por:

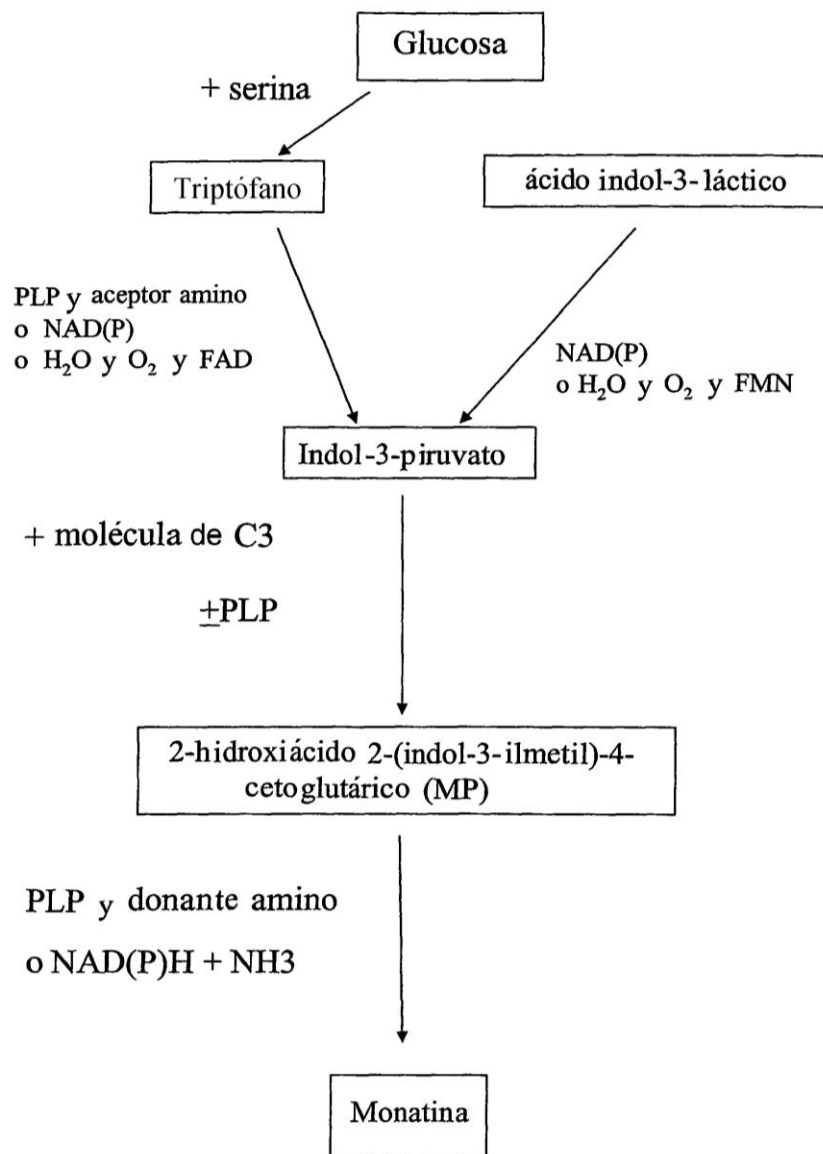
- (a) la alanina y un polipéptido que puede realizar la transaminación de la alanina;
- (b) la serina y un polipéptido que puede realizar la beta-eliminación de la serina;
- (c) la cisteína y un polipéptido que puede realizar la beta-eliminación de la cisteína;
- (d) el aspartato y un polipéptido que puede realizar las reacciones de beta-liasa con un aminoácido dicarboxílico;
- (e) el lactato y la lactato-oxidasas (CE 1.1.3.-) o una lactato-deshidrogenasa (CE 1.1.1.27, 1.1.1.28, 1.1.1.2.3); o
- (f) una combinación de los mismos.

13. Procedimiento según la reivindicación 1, en el que el procedimiento comprende además reducir una cantidad de peróxido de hidrógeno producido cuando el triptófano se convierte en indol-3-piruvato poniendo en contacto el peróxido de hidrógeno con una catalasa.

14. Procedimiento según la reivindicación 1, en el que se produce por lo menos 10 g/l de monatina.

15. Procedimiento según la reivindicación 1, en el que el procedimiento comprende además la detección del 2-hidroxiácido 2-(indol-3-ilmetil)-4-cetoglutárico por espectrometría de masas.

16. Procedimiento de producción de monatina, que comprende convertir el indol-3-piruvato en monatina en presencia de piruvato y un primer y un segundo polipéptidos, en el que el primer polipéptido es una aldolasa, y en el que el segundo polipéptido se selecciona de entre tirosina aminotransferasa (aromática) (CE 2.6.1.5), triptófano deshidrogenasa (CE 1.4.1.19), triptófano-fenilpiruvato transaminasa (CE 2.6.1.28), triptófano aminotransferasa (CE 2.6.1.27), aspartato aminotransferasa (CE 2.6.1.1), D-alanina aminotransferasa (CE 2.6.1.21), D-triptófano aminotransferasa, glutamato deshidrogenasa (CE 1.4.1.2-4), fenilalanina deshidrogenasa (CE 1.4.1.20), sustrato múltiple aminotransferasa (CE 2.6.1.-), D-aminoácido deshidrogenasa (CE 1.4.99.1) o una combinación de las mismas.

**FIG. 1**

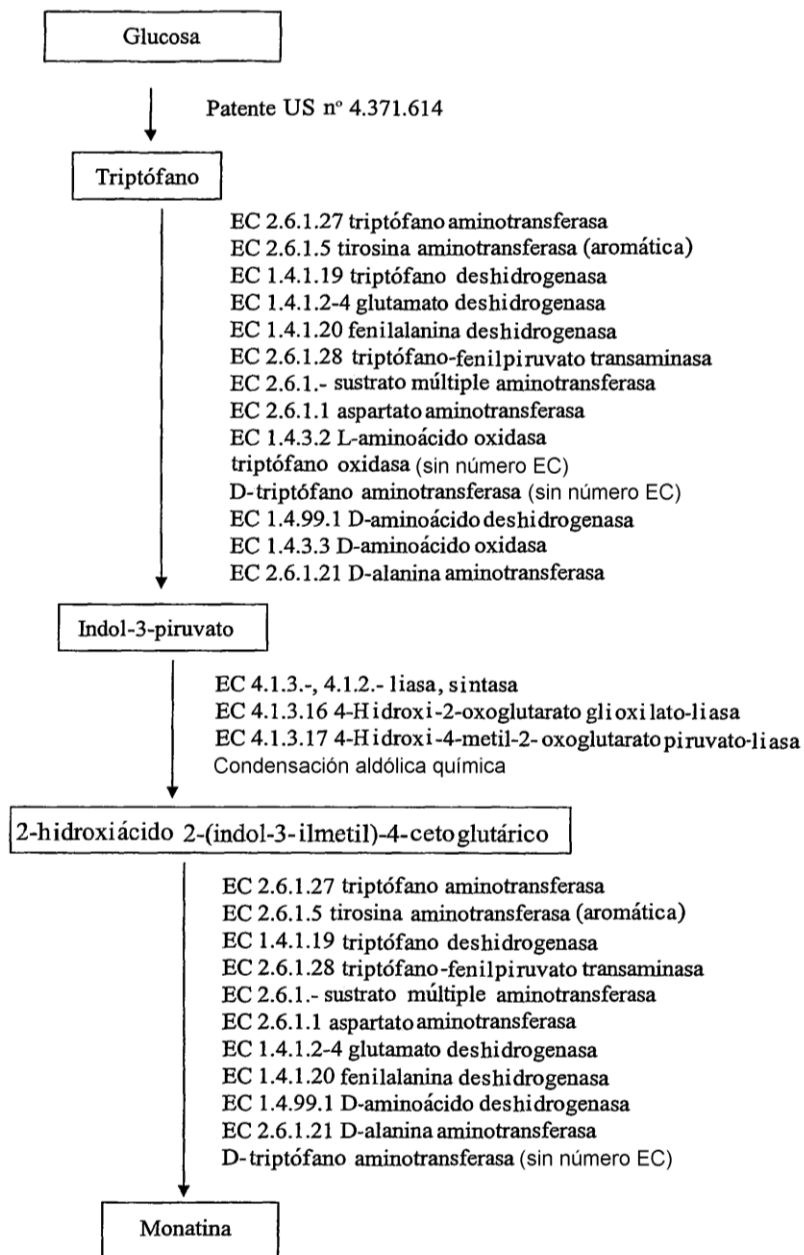
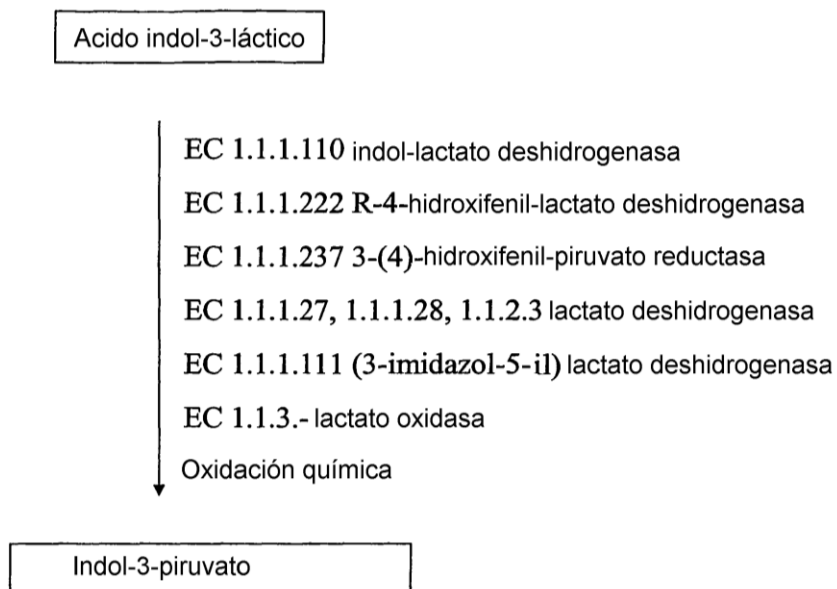


FIG. 2

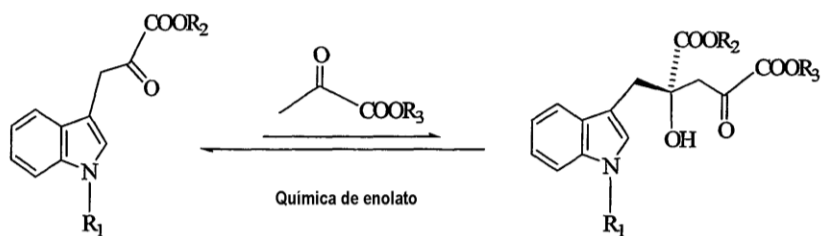
FIG. 3



EC 1.1.1.110 indol-lactato deshidrogenasa
 EC 1.1.1.222 R-4-hidroxifenil-lactato deshidrogenasa
 EC 1.1.1.237 3-(4)-hidroxifenil-piruvato reductasa
 EC 1.1.1.27, 1.1.1.28, 1.1.2.3 lactato deshidrogenasa
 EC 1.1.1.111 (3-imidazol-5-il) lactato deshidrogenasa
 EC 1.1.3.- lactato oxidasa
 Oxidación química

Indol-3-piruvato

FIG. 4



R_1 = Boc, Cbz, etc.

R_2 y R_3 = Alquilo, Amino, etc.

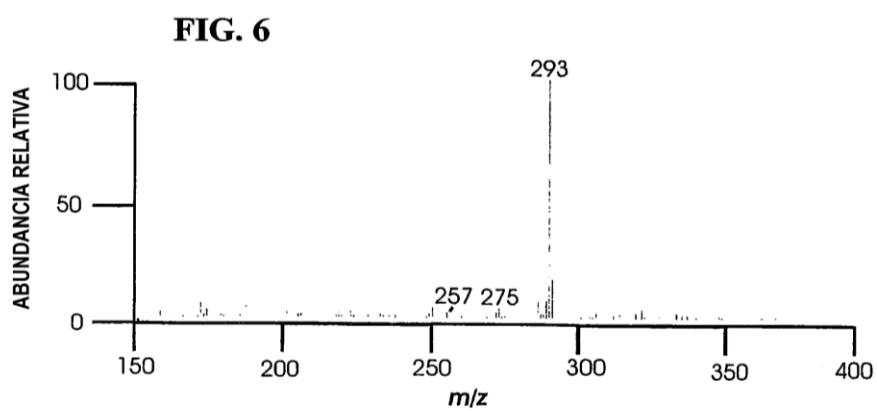
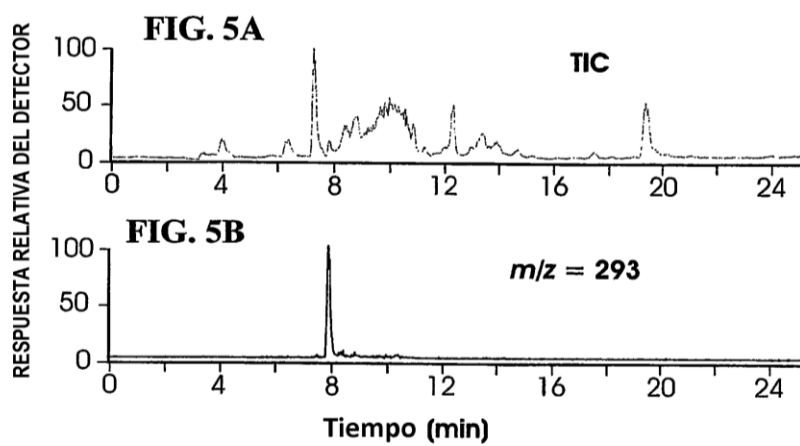


FIG. 7A

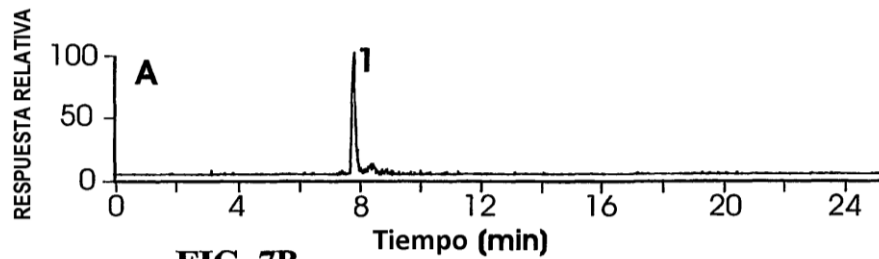


FIG. 7B

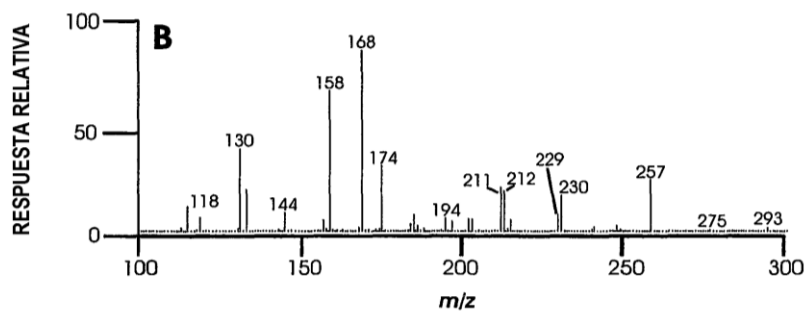
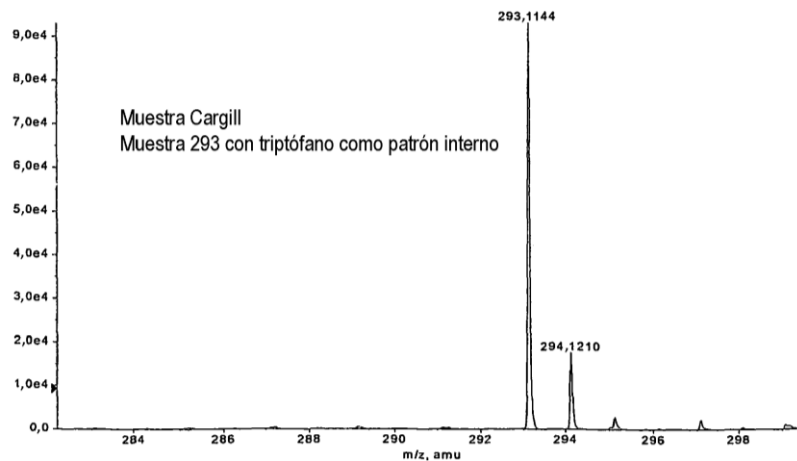


FIG. 8

+TOF MS: 124 MCA scans from Sample 7 (293jld_with Tryptophan) of car0709a.wif...
a=3,56970806537750280e-004, t0=-1,24626879588351580e+002 R; Max. 9,3e4 recuentos



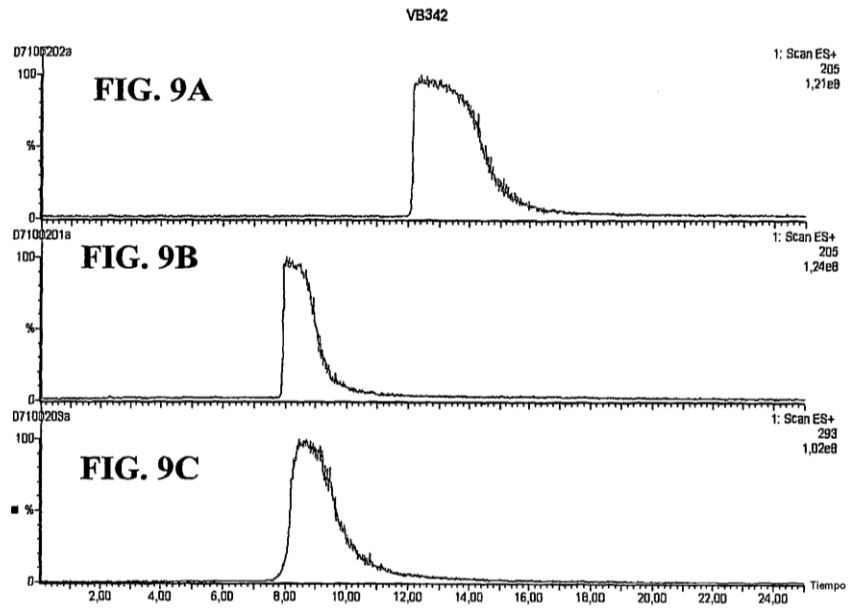


FIG. 10

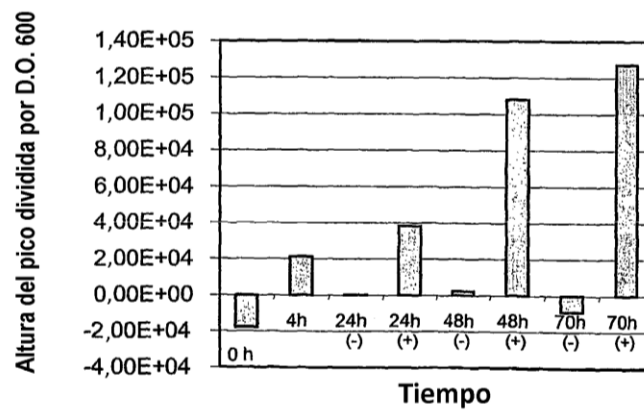


FIG. 11

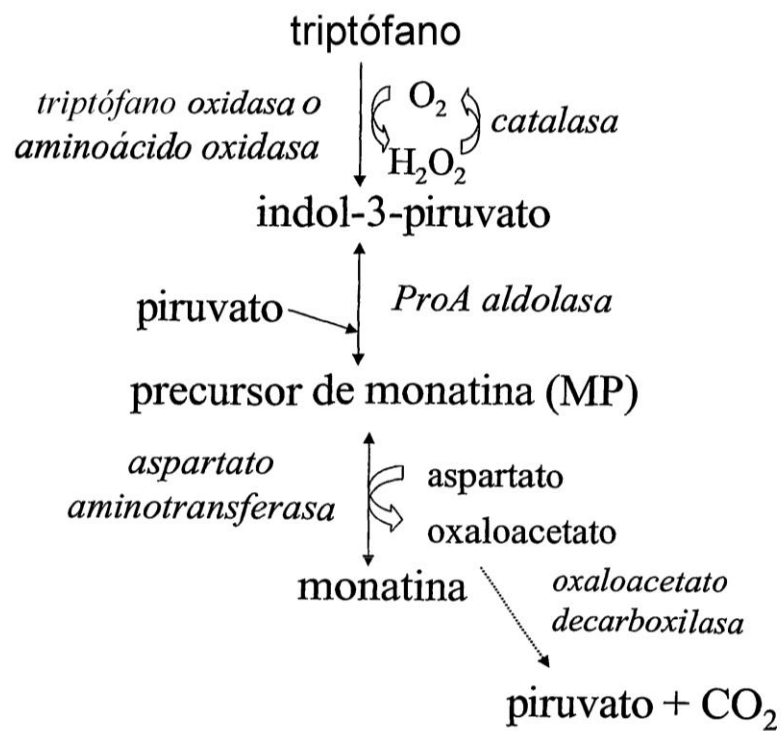


FIG. 12

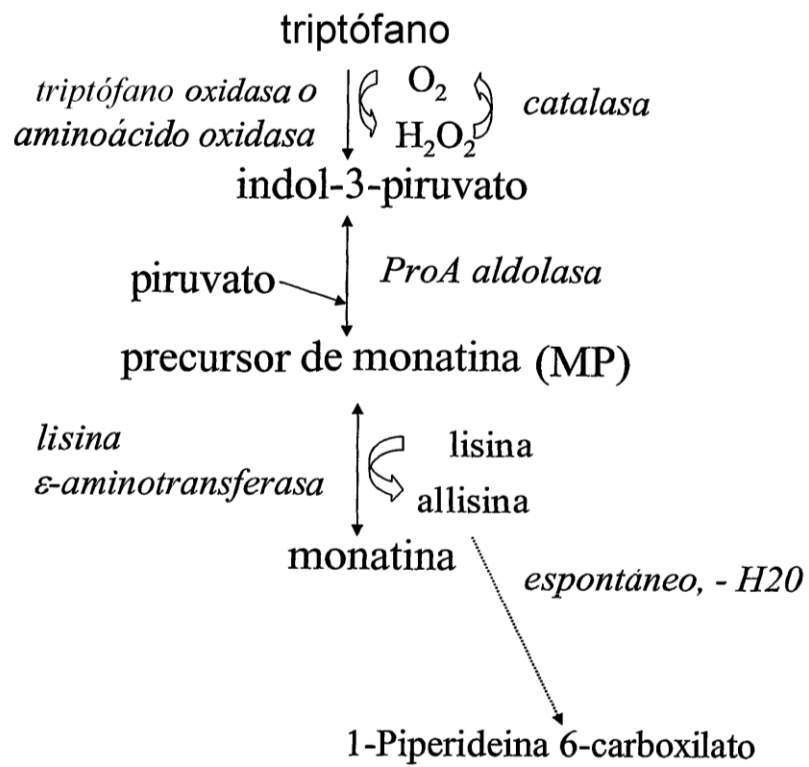


FIG. 13