

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 564 176**

51 Int. Cl.:

C12P 7/42 (2006.01)

C12N 9/78 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **31.10.2008 E 08843908 (8)**

97 Fecha y número de publicación de la concesión europea: **02.12.2015 EP 2217714**

54 Título: **Secuestración de formaldehído para estabilizar la actividad específica de nitrilasa cuando se convierte glicolonitrilo en ácido glicólico**

30 Prioridad:

31.10.2007 US 931069

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

18.03.2016

73 Titular/es:

**E. I. DU PONT DE NEMOURS AND COMPANY
(100.0%)
Chestnut Run Plaza, 974 Center Road, P.O. Box
2915
Wilmington, DE 19805, US**

72 Inventor/es:

**DICOSIMO, ROBERT;
PANOVA, ANNA;
CHENAULT, HENRY KEITH y
ARTHUR, SAMUEL DAVID**

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 564 176 T3

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

DESCRIPCIÓN

Secuestración de formaldehído para estabilizar la actividad específica de nitrilasa cuando se convierte glicolonitrilo en ácido glicólico

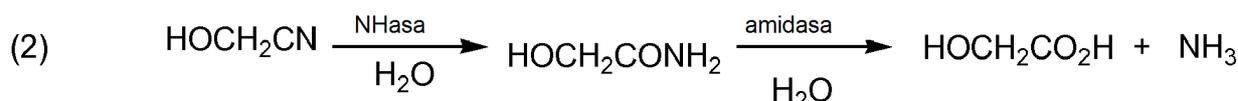
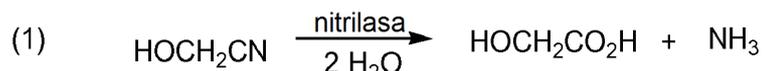
Campo de la invención

- 5 Esta invención se refiere al campo de la síntesis de ácidos orgánicos, a la biología molecular y a la microbiología. Más específicamente, se proporciona un procedimiento para proteger y/o incrementar la actividad específica de un catalizador enzimático que comprende un polipéptido que tiene actividad de nitrilasa cuando se convierte glicolonitrilo en ácido glicólico llevando a cabo la reacción en presencia de un protector de amina.

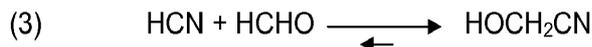
Antecedentes de la invención

- 10 El ácido glicólico (HOCH₂COOH; Número de registro en *Chemical Abstracts* (CAS) 79-14-1) es el miembro más sencillo de la familia de los α-hidroxi-ácidos de los ácidos carboxílicos. Sus propiedades lo hacen ideal para un amplio espectro de aplicaciones de consumo e industriales, incluyendo el uso en la rehabilitación de pozos de agua, la industria del cuero, la industria del petróleo y gas, y la industria de lavandería y textil, como monómero en la preparación de poli(ácido glicólico) (PGA) y como componente en productos para el cuidado personal. El ácido glicólico es también un ingrediente principal para agentes de limpieza en una variedad de industrias (agentes de limpieza para equipos de tratamiento en las industrias láctea y alimentaria, agentes de limpieza para el hogar e instituciones, agentes de limpieza industriales [para equipo de transporte, mampostería, placas de circuitos impresos, calderas de acero inoxidable y equipos de procedimientos, torres de refrigeración/intercambiadores de calor] y el procesamiento de metales [para el decapado de metales, abrillantamiento de cobre, ataque químico, electrochapado y electropulido]). Recientemente, se ha informado de que el poli(ácido glicólico) es útil como material barrera para gases (es decir, presenta altas características de barrera para el oxígeno) para el envasado de alimentos y bebidas carbónicas (documento WO 2005/106005 A1). Sin embargo, la síntesis química tradicional del ácido glicólico produce una cantidad significativa de impurezas que se deben eliminar antes de su uso en la preparación de poli(ácido glicólico) para los materiales barrera para gases. Por tanto sería muy bien recibida en la industria una nueva tecnología para producir comercialmente ácido glicólico, especialmente una que produzca ácido glicólico de alta pureza y a bajo coste.

- Los catalizadores microbianos pueden hidrolizar un nitrilo (por ejemplo, glicolonitrilo) directamente a los correspondientes ácidos carboxílicos (por ejemplo, ácido glicólico) usando una nitrilasa (EC 3.5.5.7), donde no hay producción intermedia de la amida correspondiente (Ecuación 1), o por una combinación de las enzimas nitrilo-hidratasa (EC 4.2.1.84) y amidasa (EC 3.5.1.4), donde una nitrilo-hidratasa (NHasa) convierte inicialmente un nitrilo en una amida, y luego la amida es convertida posteriormente por la amidasa en el correspondiente ácido carboxílico (Ecuación 2):



- 35 La síntesis enzimática de ácido glicólico requiere típicamente una forma sustancialmente pura de glicolonitrilo. Se han descrito previamente métodos para sintetizar glicolonitrilo por reacción de soluciones acuosas de formaldehído y cianuro de hidrógeno (Patentes de EE.UU. 2.175.805; 2.890.238; y 5.187.301; Ecuación 3).



- 40 Sin embargo, estos métodos dan como resultado típicamente un producto de reacción de glicolonitrilo acuoso que requiere una purificación significativa (por ejemplo, purificación por destilación) puesto que muchas de las impurezas y/o subproductos de la reacción (incluyendo el exceso de formaldehído reactivo) pueden interferir con la conversión enzimática de glicolonitrilo en ácido glicólico, incluyendo la supresión de la actividad del catalizador (es decir, disminución de la actividad específica). En particular, es bien sabido que el formaldehído puede crear modificaciones indeseables en las proteínas por reacción con los grupos amino de residuos de aminoácidos N-terminales y las cadenas laterales de residuos de arginina, cisteína, histidina, y lisina (Metz et al., *J. Biol. Chem.*, 279 (8): 6235-6243 (2004)). La supresión de la actividad del catalizador disminuye la productividad global del catalizador (es decir, gramos totales de ácido glicólico formados por gramo de catalizador), añadiendo un costo significativo al procedimiento global que puede hacer la producción enzimática económicamente no viable en comparación con la síntesis química. Como tales, se necesitan condiciones de reacción que puedan ayudar a proteger la actividad enzimática contra las impurezas indeseables que describen una disminución de la actividad del catalizador.

Se ha descrito un método de producir glicolonitrilo de alta pureza sometiendo formaldehído a un tratamiento térmico antes de la reacción de síntesis del glicolonitrilo (solicitudes de patente US 2006/0160196 y US 2006/0247467; Ecuación 3). Sin embargo, el glicolonitrilo puede disociarse reversiblemente en formaldehído y cianuro de hidrógeno. Por tanto, sigue habiendo necesidad de proteger un catalizador que tiene actividad de nitrilasa contra los efectos indeseables del formaldehído libre.

Se conocen varios métodos para preparar α -hidroxi-ácidos utilizando el correspondiente α -hidroxi-nitrilo como el material de partida y un microorganismo con el catalizador. Ejemplos de α -hidroxi-ácidos producidos incluyen: ácido glicólico, ácido láctico, ácido 2-hidroxiisobutírico, ácido 2-hidroxi-2-fenil propiónico, ácido mandélico, 2-hidroxi-3,3-dimetil-4-butirolactona, y ácido 4-metiltiobutírico. Estos productos se sintetizan utilizando microorganismos, tales como aquellos pertenecientes al género *Nocardia*, *Bacillus*, *Brevibacterium*, *Aureobacterium*, *Pseudomonas*, *Caseobacter*, *Alcaligenes*, *Acinetobacter*, *Enterobacter*, *Arthrobacter*, *Escherichia*, *Micrococcus*, *Streptomyces*, *Flavobacterium*, *Aeromonas*, *Mycoplana*, *Cellulomonas*, *Erwinia*, *Candida*, *Bacteridium*, *Aspergillus*, *Penicillium*, *Cochliobolus*, *Fusarium*, *Rhodopseudomonas*, *Rhodococcus*, *Corynebacterium*, *Microbacterium*, *Obsumbacterium* and *Gordona*. (JP-A-4-99495, JP-A-4-99496 y JP-A-4-218385 correspondiente a la patente de EE.UU. 5, 223,416; JP-A-4-99497 correspondiente a la patente de EE.UU. 5,234,826; JP-A-5-95795 correspondiente a la patente de EE.UU. 5,296,373; JP-A-5-21987; JP-A-5-192189 correspondiente a la patente de EE.UU. 5,326,702; JP-A-6-237789 correspondiente a la patente europea EP-A-0610048; JP-A-6-284899 correspondiente a la patente europea EP-A-0610049; JP-A-7-213296 correspondiente a la patente de EE.UU. 5,508,181).

Sin embargo, los métodos más conocidos para preparar los α -hidroxi-ácidos a partir de los correspondientes α -hidroxi-nitrilos, como se ha mencionado anteriormente, no producen ni acumulan un producto a una concentración suficientemente alta para satisfacer las necesidades comerciales. Esto es frecuentemente el resultado de la inactivación enzimática previa en el periodo de reacción. La patente de EE.UU. 5.756.306 muestra que "Cuando un α -hidroxi-nitrilo se hidroliza o hidrata enzimáticamente utilizando nitrilasa o nitrilo hidratasa para producir un α -hidroxi-ácido o una α -hidroxi-amida, se presenta un problema en el que la enzima se inactiva en un corto periodo de tiempo. Por lo tanto, es difícil obtener el α -hidroxi-ácido o la α -hidroxi-amida en alta concentración y alto rendimiento." (col. 1, líneas 49-54). Mantener la concentración de aldehído (formado por la disociación del α -hidroxi-nitrilo a aldehído y cianuro de hidrógeno) y/o la concentración de α -hidroxi-nitrilo en la mezcla de reacción dentro de un intervalo específico es un método para evitar este problema.

La patente de EE.UU. 5.508.181 menciona dificultades adicionales que se refieren a la rápida inactivación enzimática. Específicamente, la patente de EE.UU. 5.508.181 menciona que los compuestos de α -hidroxi-nitrilo se disocian parcialmente en los aldehídos correspondientes, de acuerdo con el equilibrio de disociación. Se informó de estos aldehídos inactivaban la enzima en un corto periodo de tiempo uniéndose a la proteína, lo que hace que sea difícil obtener el α -hidroxi-ácido o la α -hidroxi-amida en una alta concentración con una alta productividad a partir de α -hidroxi-nitrilos (col. 2, líneas 16-29). Como solución para evitar la inactivación de la enzima debido a la acumulación de aldehídos, se añadieron a la mezcla de reacción iones fosfato o hipofosfito. La patente de EE.UU. 5.326.702 informa del uso de iones sulfito, disulfito o ditionito para secuestrar el aldehído y prevenir la inactivación de la enzima, pero concluye que la concentración de α -hidroxi-ácido producido y acumulada incluso usando tales aditivos, como se ha descrito anteriormente, no es muy alta.

La patente de EE.UU. 6.037.155 enseña que la baja acumulación de producto α -hidroxi-ácido está relacionada con la inactivación enzimática en un corto periodo de tiempo, debido a la acumulación del aldehído disociado. Los inventores de dicha patente sugieren que la actividad enzimática es inhibida en presencia de cianuro de hidrógeno (Asano et al., *Agricultural Biological Chemistry*, Vol. 46, pp. 1165-1174 (1982)) generado en la disociación parcial del α -hidroxi-nitrilo en agua junto con el aldehído o cetona correspondiente (Mowry, David T., *Chemical Reviews*, vol. 42, pp. 189-283 (1948)). Los inventores resuelven el problema de la inactivación enzimática inducida por el aldehído usando microorganismos cuya actividad enzimática podría ser mejorada por adición de una sustancia de cianuro a la mezcla de reacción. La adición de una sustancia de cianuro limitaría la disociación del α -hidroxi-nitrilo a aldehído y cianuro de hidrógeno.

El documento WO 99/30781 describe un método y composición de tratamiento para neutralizar el formaldehído. La composición de tratamiento contiene polietilenimina, urea y un ácido que tiene un pH de menos de 6.

Con consideración específica a la producción de ácido glicólico, se conoce que el glicolonitrilo se disocia reversiblemente a cianuro de hidrógeno y formaldehído, cualquiera de los cuales puede estar implicado en la reducción de la actividad del catalizador. La patente de EE.UU. 3.940.316 describe un procedimiento para preparar un ácido orgánico a partir del nitrilo correspondiente utilizando una bacteria con actividad "nitrilásica", y menciona al glicolonitrilo como un sustrato. En particular, esta patente describe el uso de *Bacillus*, *Bacteridium*, *Micrococcus*, y *Brevibacterium* para este propósito. Aunque se describe que tiene actividad nitrilásica, *Brevibacterium* R312 es la única cepa utilizada en todos los ejemplos de la patente de EE.UU. 3.940.316. *Brevibacterium* R312 se conoce que tiene actividad de nitrilo hidratasa y amidasa, pero no actividad de nitrilasa (Tourneix et al., *Antonie van Leeuwenhoek*, 52:173:182 (1986)).

Un método para preparar ácido láctico, ácido glicólico y ácido 2-hidroxiisobutírico utilizando un microorganismo que pertenece a *Corynebacterium spp.* Se describe en la Patente Japonesa abierta a inspección pública No. Sho 61-

56086. La patente japonesa JP 09028390 describe un método para la fabricación de ácido glicólico a partir de glicolonitrilo por la acción de *Rhodococcus* o *Gordona* hidrolasa. La selectividad del ácido glicólico se describe a casi el 100%, sin la formación de la amida del ácido glicólico. La patente de EE.UU. 6.037.155 describe ejemplos de métodos para producir α -hidroxi-ácidos a partir de α -hidroxi-nitrilos, incluyendo el ácido glicólico. Esta descripción admite que no todos los catalizadores microbianos pueden producir altas concentraciones de ácido glicólico, debido a los problemas anteriormente mencionados e informa que deben llevarse a cabo estudios de selección para encontrar los microorganismos industrialmente ventajosos. La patente de EE.UU. 6.037.155 identifica específicamente los microorganismos *Variovorax spp.* y *Arthrobacter spp.*, que son resistentes al efecto supresor del α -hidroxi-nitrilo o α -hidroxi-ácido, tienen actividad duradera y pueden producir el producto deseado a alta concentración.
- 10 *Acidovorax facilis* 72W (ATCC 57746) se caracteriza por la actividad de nitrilasa alifática (EC 3.5.5.7), así como una combinación de las actividades nitrolo hidratasa (EC 4.2.1.84) y amidasa (EC 3.5.1.4). El gen que codifica la nitrilasa de *A. facilis* 72W (atcc 55746) ha sido clonado y expresado recombinantemente (documento WO 01/75077 correspondiente a la patente de EE.UU. 6.870.038) y Chauhan et al., Appl. Microbiol Biotechnol, 61:118-122 (2003)).
- 15 El documento WO 2006/069114 proporciona un procedimiento para producir ácido glicólico a partir de formaldehído y cianuro de hidrógeno. Más específicamente, el formaldehído tratado térmicamente y el cianuro de hidrógeno se hacen reaccionar para producir glicolonitrilo que tiene bajas concentraciones de impurezas. El glicolonitrilo se convierte posteriormente en una solución acuosa de glicolato de amonio usando un catalizador enzimático que tiene actividad de nitrilasa derivado de *Acidovorax facilis* 72W (ATCC 57746).
- 20 La nitrilasa de *A. facilis* 72W convierte α -hidroxinitrilos en los correspondientes ácidos α -hidroxicarboxílicos en alto rendimiento (patente de EE.UU. 6.383.786), incluyendo ácido glicólico (patente de EE.UU. 6.416.980). Un procedimiento mejorado para producir ácido glicólico a partir de glicolonitrilo utilizando mutantes derivados de la nitrilasa de *A. facilis* 72W se describe en los documentos WO2006/068110 y WO2006/069114 (correspondientes a la patente de EE.UU. 7.198.927 y solicitud de patente de EE.UU. 2006/0247467, respectivamente). En la solicitud patente de EE.UU. 2006/0247467 asignada, se describen varios medios para mejorar la producción de ácido glicólico, incluyendo (1) el uso de mutantes de *A. facilis* 72W que tienen una mejor actividad, (2) la adición de al menos un agente estabilizante (por ejemplo, tiosulfato de potasio, ditionito de sodio, exceso de HCN), (3) efectuar la reacción bajo condiciones libres de oxígeno, (4) controlar el caudal de alimentación del glicolonitrilo, y (5) el uso de glicolonitrilo de alta pureza. A pesar de que muchos de los medios mejoraron la productividad del ácido glicólico, con el tiempo se observó generalmente una disminución en la actividad enzimática. Esta disminución en la actividad se atribuyó típicamente, al menos en parte, a la presencia de formaldehído (aunque a niveles bajos) en la mezcla de reacción. Un procedimiento para proteger la actividad específica de un catalizador enzimático que tiene actividad de nitrilasa cuando se convierte glicolonitrilo en ácido glicólico en presencia de formaldehído mejoraría significativamente la economía de la síntesis del ácido glicólico.
- 35 El problema a resolver es proporcionar un procedimiento para estabilizar y/o incrementar la actividad específica de un catalizador enzimático que tiene actividad de nitrilasa cuando se convierte glicolonitrilo en ácido glicólico en presencia de formaldehído.

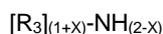
Sumario de la invención

40 El presente problema ha sido resuelto proporcionando un procedimiento para mejorar la actividad específica de un catalizador enzimático cuando se convierte enzimáticamente glicolonitrilo en ácido glicólico, comprendiendo dicho procedimiento:

(a) proporcionar una serie de componentes de reacción que comprenden:

- (i) una solución acuosa de glicolonitrilo que comprende al menos 0,01 ppm de formaldehído;
- (ii) un catalizador enzimático que comprende un polipéptido que tiene actividad de nitrilasa, en el que dicho polipéptido comprende un resto distintivo catalítico de la SEQ ID NO.: 1; en donde dicho catalizador enzimático comprende una actividad específica para hidrolizar glicolonitrilo a ácido glicólico; y
- (iii) un cantidad eficaz de al menos un protector de amina seleccionada del grupo que consiste en:

a) un compuesto de la fórmula



50 en donde X es 0 o 1, y R_3 es independientemente un grupo hidrocarbilo sustituido o grupo hidrocarbilo C1 a C20, en el que R_3 comprende opcionalmente uno o más enlaces éter; con la condición de que (i) el grupo sustituido no sea un grupo ciano capaz de reaccionar con el catalizador enzimático y (ii) R_3 no sea un grupo carbonilo,

55 b) un polímero poliaminado que comprende un número eficaz de grupos amino libres en donde dichos grupos amino libres son grupos amino primarios y/o grupos amino secundarios; y

- c) un material con funcionalidad amina que comprende un número eficaz de grupos amino libres en donde dichos grupos amino libres son grupos amino primarios y/o grupos amino secundarios;

en donde el protector de amina no se produce naturalmente por dicho catalizador enzimático;

- 5 (b) combinar la serie de componentes de reacción bajo condiciones de reacción acuosas adecuadas en las que se produce el ácido glicólico; en donde dicha actividad específica se incrementa por la adición de dicho al menos un protector de amina; y

- (c) recuperar el ácido glicólico o la sal del mismo producido en (b);

en donde el protector de amina es insoluble bajo dichas condiciones de reacción acuosa adecuadas y,

en donde el catalizador enzimático no está inmovilizado en una matriz que comprende dicho protector de amina.

- 10 También se proporciona en la presente memoria un procedimiento para estabilizar y/o incrementar la actividad específica de un catalizador nitrilasa que comprende un polipéptido que tiene actividad de nitrilasa cuando se convierte enzimáticamente glicolonitrilo en ácido glicólico, comprendiendo dicho procedimiento:

- (a) proporcionar una serie de componentes de reacción que comprenden:

- (i) una solución acuosa de glicolonitrilo que comprende al menos 0,01 ppm de formaldehído;

- 15 (ii) un catalizador enzimático que comprende un polipéptido que tiene actividad de nitrilasa, en el que dicho polipéptido comprende un resto distintivo catalítico de la SEQ ID NO.: 1; en donde dicho catalizador enzimático comprende una actividad específica para hidrolizar glicolonitrilo a ácido glicólico; y

- (iii) un cantidad eficaz de al menos un protector de amina seleccionada del grupo que consiste en:

- a) un compuesto de la fórmula



en donde X es 0 o 1, y R₃ es independientemente un grupo hidrocarbilo sustituido o grupo hidrocarbilo C1 a C20, en el que R₃ comprende opcionalmente uno o más enlaces éter; con la condición de que (i) el grupo sustituido no sea preferiblemente un grupo ciano capaz de reaccionar con el catalizador enzimático y (ii) R₃ no sea un grupo carbonilo,

- 25 b) un polímero poliaminado que comprende un número eficaz de grupos amino libres; y

- c) un material con funcionalidad amina que comprende un número eficaz de grupos amino libres;

en donde el protector de amina no se produce naturalmente por dicho catalizador enzimático;

- 30 (b) combinar la serie de componentes de reacción bajo condiciones de reacción acuosas adecuadas en las que se produce el ácido glicólico; en donde dicha actividad específica se incrementa por la adición de dicho al menos un protector de amina; y

- (c) recuperar el ácido glicólico o la sal del mismo producido en (b).

Breve descripción de la figura, listado de secuencias y los depósitos biológicos

La invención puede ser entendida más completamente a partir de la lista de secuencias, los depósitos biológicos, la figura, y la descripción detallada que juntos forman esta solicitud.

35 Figura

La Figura 1, paneles A-G, es un alineamiento por el programa CLUSTAL W (versión 1.83 utilizando parámetros por defecto) de varias secuencias de nitrilasa. La secuencia distintiva del catalizador conservada que rodea al residuo de cisteína del catalizador está resaltada con sombreado gris. Están subrayados los aminoácidos que representan la tríada catalítica (Glu₄₈, Lys₁₃₀ y Cys₁₆₄; con la numeración basada en la secuencia de aminoácidos de la SEQ ID NO: 4).

40 Listado de secuencias

Las siguientes descripciones de secuencias y listados de secuencias adjuntas cumplen con las normas que rigen las descripciones de secuencias de nucleótidos y/o aminoácidos en las solicitudes de patentes como se establece en el 37 C.F.R §1.821-1.825. Las descripciones de secuencias contienen el código de una letra para los caracteres de las secuencias de nucleótidos y los códigos de tres letras para los aminoácidos como se define de conformidad con las normas de la IUPAC-IYUB descritas en *Nucleic Acids Research* 13: 3021-3030 (1985) y en el *Biochemical*

ES 2 564 176 T3

Journal 219 (Nº 2): 345-373 (1984). Los símbolos y formatos utilizados para los datos de secuencias de nucleótidos y aminoácidos cumplen con las normas establecidas en 37 C.F.R §1.822.

La SEQ ID NO: 1 es la secuencia de aminoácidos del resto distintivo catalítico que abarca el residuo de cisteína esencial de las enzimas nitrilasas (Fórmula 1).

- 5 La SEQ ID NO: 2 es la secuencia de aminoácidos de un resto distintivo catalítico preferido que abarca el residuo de cisteína esencial de las enzimas nitrilasas (Fórmula 2).

La SEQ ID NO: 3 es la secuencia de nucleótidos de la secuencia codificadora de nitrilasa de *Acidovorax facilis* 72W que comprende un cambio en el codón de inicio de TTG a ATG para facilitar la expresión recombinante en *E. coli*.

La SEQ ID NO: 4 es la secuencia de aminoácidos deducida de la nitrilasa de *Acidovorax facilis* 72W (ATCC 55746).

- 10 La SEQ ID NO: 5 es la secuencia de aminoácidos de la nitrilasa de *Alcaligenes faecalis* JM3 (GENBANK® BAA02684.1).

La SEQ ID NO: 6 es la secuencia de aminoácidos de la nitrilasa de *Rhodococcus rhodochrous* J1 (GENBANK® Q03217).

- 15 La SEQ ID NO: 7 es la secuencia de aminoácidos de la nitrilasa de *Rhodococcus rhodochrous* K22 (GENBANK® Q02068).

La SEQ ID NO: 8 es la secuencia de aminoácidos de la nitrilasa de *Nocardia* sp. C-14-1 (GENBANK® AAX18182.1).

La SEQ ID NO: 9 es la secuencia de aminoácidos de la nitrilasa de *Bordetella bronchiseptica* RB50 (GENBANK® NP_887662.1).

- 20 La SEQ ID NO: 10 es la secuencia de aminoácidos de la nitrilasa de *Arabidopsis thaliana* (GENBANK® AAB60275.1 y AAA19627.1).

La SEQ ID NO: 11 es la secuencia de aminoácidos de la nitrilasa de *Synechococcus elongatus* PCC 7942 (GENBANK® YP_399857.1).

La SEQ ID NO: 12 es la secuencia de aminoácidos de la nitrilasa de *Synechococcus elongatus* PCC 6301 (GENBANK® YP_171411.1).

- 25 La SEQ ID NO: 13 es la secuencia de aminoácidos de la nitrilasa de *Synechocystis* sp. PCC 6803 (GENBANK® NP_442646.1).

La SEQ ID NO: 14 es la secuencia de aminoácidos de la nitrilasa de *Entomophila Pseudomonas* L48 (GENBANK® YP_609048.1).

- 30 La SEQ ID NO: 15 es la secuencia de aminoácidos de la nitrilasa de *Zymomonas mobilis* (GENBANK® YP_162942.1).

La SEQ ID NO: 16 es la secuencia de aminoácidos de la nitrilasa de *Bacillus* sp. OxB-1 (GENBANK® BAA90460.1).

La SEQ ID NO: 17 es la secuencia de aminoácidos de la nitrilasa de *Comamonas testosteroni* (GENBANK® AAA82085.1).

- 35 La SEQ ID NO: 18 es la secuencia de aminoácidos de la nitrilasa de *Synechococcus* sp. CC9605 (GENBANK® YP_381420.1).

La SEQ ID NO: 19 es la secuencia de aminoácidos de la nitrilasa de *Pseudomonas fluorescens* Pf-5 (GENBANK® YP_260015.1).

La SEQ ID NO: 20 es la secuencia de aminoácidos de la nitrilasa de *Nocardia farcinica* IFM 10152 (GENBANK® YP_119480.1).

- 40 La SEQ ID NO: 21 es la secuencia de aminoácidos de la nitrilasa de *Alcaligenes faecalis* 1650 (GENBANK® AAY06506.1).

La SEQ ID NO: 22 es la secuencia de aminoácidos de la nitrilasa de *Pseudomonas syringae* pv. *syringae* B728a (GENBANK® AAY35081.1).

- 45 La SEQ ID NO: 23 es la secuencia de aminoácidos de la nitrilasa de *Bradyrhizobium* sp. BTail (GENBANK® ZP_00859948.1).

La SEQ ID NO: 24 es la secuencia de aminoácidos de la nitrilasa de *Rhodococcus rhodochrous* NCIMB 11216 (GENBANK® CAC88237).

ES 2 564 176 T3

La SEQ ID NO: 25 es la secuencia de aminoácidos de la nitrilasa de *Rhodococcus rhodochrous* ATCC™ 39484.

La SEQ ID NO: 26 es la secuencia de nucleótidos de una nitrilasa mutante de *A. facilis* 72W que comprende un cambio de codón que dio como resultado una única sustitución de aminoácido en el residuo de la posición 201 (L201Q; Leu → Gln).

- 5 La SEQ ID NO: 27 es la secuencia de aminoácidos deducida de la nitrilasa mutante (SEQ ID NO: 26) que comprende una única sustitución de aminoácido en el residuo de la posición 201 (Leu201 → Gln) de la nitrilasa de *A. facilis* 72W.

- 10 La SEQ ID NO: 28 es la secuencia de nucleótidos de una nitrilasa mutante de *A. facilis* 72W que comprende un cambio de codón que dio como resultado una única sustitución de aminoácido en el residuo de la posición 201 (L201A; Leu → Ala).

La SEQ ID NO: 29 es la secuencia de aminoácidos deducida de la nitrilasa mutante (SEQ ID NO: 28) que comprende una única sustitución de aminoácido en el residuo de la posición 201 (Leu201 → Ala) de la nitrilasa de *A. facilis* 72W.

- 15 La SEQ ID NO: 30 es la secuencia de nucleótidos de una nitrilasa mutante de *A. facilis* 72W que comprende un cambio de codón que dio como resultado una única sustitución de aminoácido en el residuo de la posición 201 (L201C; Leu → Cys).

La SEQ ID NO: 31 es la secuencia de aminoácidos deducida de la nitrilasa mutante (SEQ ID NO: 30) que comprende una única sustitución de aminoácido en el residuo de la posición 201 (Leu201 → Cys) de la nitrilasa de *A. facilis* 72W.

- 20 La SEQ ID NO: 32 es la secuencia de nucleótidos de una nitrilasa mutante de *A. facilis* 72W que comprende un cambio de codón que dio como resultado una única sustitución de aminoácido en el residuo de la posición 201 (L201T; Leu → Thr).

- 25 La SEQ ID NO: 33 es la secuencia de aminoácidos deducida de la nitrilasa mutante (SEQ ID NO: 32) que comprende una única sustitución de aminoácido en el residuo de la posición 201 (Leu201 → Thr) de la nitrilasa de *A. facilis* 72W.

La SEQ ID NO: 34 es la secuencia de nucleótidos de una nitrilasa mutante de *A. facilis* 72W que comprende un cambio de codón que dio como resultado una única sustitución de aminoácido en el residuo de la posición 201 (L201 G; Leu → Gly).

- 30 La SEQ ID NO: 35 es la secuencia de aminoácidos deducida de la nitrilasa mutante (SEQ ID NO: 34) que comprende una única sustitución de aminoácido en el residuo de la posición 201 (Leu201 → Gly) de la nitrilasa de *A. facilis* 72W.

La SEQ ID NO: 36 es la secuencia de nucleótidos de una nitrilasa mutante de *A. facilis* 72W que comprende un cambio de codón que dio como resultado una única sustitución de aminoácido en el residuo de la posición 201 (L201 H; Leu → His).

- 35 La SEQ ID NO: 37 es la secuencia de aminoácidos deducida de la nitrilasa mutante (SEQ ID NO: 36) que comprende una única sustitución de aminoácido en el residuo de la posición 201 (Leu201 → His) de la nitrilasa de *A. facilis* 72W.

- 40 La SEQ ID NO: 38 es la secuencia de nucleótidos de una nitrilasa mutante de *A. facilis* 72W que comprende un cambio de codón que dio como resultado una única sustitución de aminoácido en el residuo de la posición 201 (L201 K; Leu → Lys).

La SEQ ID NO: 39 es la secuencia de aminoácidos deducida de la nitrilasa mutante (SEQ ID NO: 38) que comprende una única sustitución de aminoácido en el residuo de la posición 201 (Leu201 → Lys) de la nitrilasa de *A. facilis* 72W.

- 45 La SEQ ID NO: 40 es la secuencia de nucleótidos de una nitrilasa mutante de *A. facilis* 72W que comprende un cambio de codón que dio como resultado una única sustitución de aminoácido en el residuo de la posición 201 (L201 N; Leu → Asn).

La SEQ ID NO: 41 es la secuencia de aminoácidos deducida de la nitrilasa mutante (SEQ ID NO: 40) que comprende una única sustitución de aminoácido en el residuo de la posición 201 (Leu201 → Asn) de la nitrilasa de *A. facilis* 72W.

- 50 La SEQ ID NO: 42 es la secuencia de nucleótidos de una nitrilasa mutante de *A. facilis* 72W que comprende un cambio de codón que dio como resultado una única sustitución de aminoácido en el residuo de la posición 201 (L201S; Leu → Ser).

La SEQ ID NO: 43 es la secuencia de aminoácidos deducida de la nitrilasa mutante (SEQ ID NO: 42) que comprende una única sustitución de aminoácido en el residuo de la posición 201 (Leu201 → Ser) de la nitrilasa de *A. facilis* 72W.

5 La SEQ ID NO: 44 es la secuencia de nucleótidos de una nitrilasa mutante de *A. facilis* 72W que comprende un cambio de codón que dió como resultado una única sustitución de aminoácido en el residuo de la posición 168 (F168K; Phe → Lys).

La SEQ ID NO: 45 es la secuencia de aminoácidos deducida de la nitrilasa mutante (SEQ ID NO: 44) que comprende una única sustitución de aminoácido en el residuo de la posición 168 (Phe168 → Lys) de la nitrilasa de *A. facilis* 72W.

10 La SEQ ID NO: 46 es la secuencia de nucleótidos de una nitrilasa mutante de *A. facilis* 72W que comprende un cambio de codón que dió como resultado una única sustitución de aminoácido en el residuo de la posición 168 (F168M; Phe → Met).

15 La SEQ ID NO: 47 es la secuencia de aminoácidos deducida de la nitrilasa mutante (SEQ ID NO: 46) que comprende una única sustitución de aminoácido en el residuo de la posición 168 (Phe168 → Met) de la nitrilasa de *A. facilis* 72W.

La SEQ ID NO: 48 es la secuencia de nucleótidos de una nitrilasa mutante de *A. facilis* 72W que comprende un cambio de codón que dió como resultado una única sustitución de aminoácido en el residuo de la posición 168 (F168T; Phe → Thr).

20 La SEQ ID NO: 49 es la secuencia de aminoácidos deducida de la nitrilasa mutante (SEQ ID NO: 48) que comprende una única sustitución de aminoácido en el residuo de la posición 168 (Phe168 → Thr) de la nitrilasa de *A. facilis* 72W.

La SEQ ID NO: 50 es la secuencia de nucleótidos de una nitrilasa mutante de *A. facilis* 72W que comprende un cambio de codón que dió como resultado una única sustitución de aminoácido en el residuo de la posición 168 (F168V; Phe → Val).

25 La SEQ ID NO: 51 es la secuencia de aminoácidos deducida de la nitrilasa mutante (SEQ ID NO: 50) que comprende una única sustitución de aminoácido en el residuo de la posición 168 (Phe168 → Val) de la nitrilasa de *A. facilis* 72W.

30 La SEQ ID NO: 52 es la secuencia de nucleótidos de una nitrilasa mutante de *A. facilis* 72W que comprende un cambio de codón que dió como resultado una única sustitución de aminoácidos en el residuo de la posición 168 (T210A; Thr → Ala).

La SEQ ID NO: 53 es la secuencia de aminoácidos deducida de la nitrilasa mutante (SEQ ID NO: 52) que comprende una única sustitución de aminoácido en el residuo de la posición 210 (Thr210 → Ala) de la nitrilasa de *A. facilis* 72W.

35 La SEQ ID NO: 54 es la secuencia de nucleótidos de una nitrilasa mutante de *A. facilis* 72W que comprende un cambio de codón que dió como resultado una única sustitución de aminoácido en el residuo de la posición 168 (T210C; Thr → Cys).

La SEQ ID NO: 55 es la secuencia de aminoácidos deducida de la nitrilasa mutante (SEQ ID NO: 54) que comprende una única sustitución de aminoácido en el residuo de la posición 210 (Thr210 → Cys) de la nitrilasa de *A. facilis* 72W.

40 La SEQ ID NO: 56 es la secuencia de nucleótidos de la nitrilasa de *A. facilis* 72W expresada en la cepa *i* SS1001 de *E. coli* (ATCC PTA-1177).

La SEQ ID NO: 57 es la secuencia de aminoácidos deducida de la nitrilasa mutante de *A. facilis* 72W expresada en la cepa SS1001 de *E. coli* (ATCC PTA-1177).

Depósitos biológicos

45 Los siguientes depósitos biológicos se han hecho en los términos del Tratado de Budapest sobre Reconocimiento internacional del Depósito de microorganismos para los fines del procedimiento en materia de patentes:

Referencia para identificación del depositante	Designación internacional del depósito	Fecha del depósito
<i>Acidovorax facilis</i> 72W	ATCC 55746	8 de marzo 1996
<i>E. coli</i> SS1001	ATCC PTA-1177	11 de enero 2000

Como se usa en la presente memoria, "ATCC" se refiere al Organismo internacional para el depósito de microorganismos *American Type Culture Collection* domiciliado en ATCC, 10801 University Blvd., Manassas, VA 20110-2209, EE.UU. La "Designación del depósito Internacional" es el número de acceso al cultivo depositado en ATCC.

- 5 Los depósitos mencionados se mantendrán en el depósito internacional indicado durante al menos treinta (30) años y se pondrán a disposición del público tras la concesión de una patente que los describe. La disponibilidad de un depósito no constituye una licencia para practicar la presente invención en derogación de los derechos de patente concedidos por la acción gubernamental.

Descripción detallada de la invención

- 10 Se proporciona un procedimiento para estabilizar y/o incrementar la actividad específica de un catalizador enzimático que tiene actividad de nitrilasa cuando se convierte glicolonitrilo en ácido glicólico bajo condiciones de reacción acuosa adecuadas, en donde las condiciones de reacción incluyen un cantidad eficaz de al menos un protector de amina.

Definiciones:

- 15 En esta descripción, se utilizan una serie de términos y abreviaturas. Salvo que se especifique lo contrario se aplican las siguientes definiciones.

Tal como se utiliza en la presente memoria, el término "que comprende" significa la presencia de las características, números enteros, etapas o componentes indicados a los que se refieren las reivindicaciones, pero que no excluye la presencia o adición de una o más de otras características, números enteros, etapas, componentes o sus grupos.

- 20 Tal como se usa en la presente memoria, el término "aproximadamente" que modifica la cantidad de un ingrediente o reaccionante de la invención empleado se refiere a la variación en la cantidad numérica que puede ocurrir, por ejemplo, por los procedimientos típicos de medición y manipulación de líquidos utilizados para preparar concentrados o soluciones de uso en el mundo real; por errores involuntarios en estos procedimientos; por diferencias en la fabricación, la fuente o la pureza de los ingredientes empleados para preparar las composiciones o llevar a cabo los métodos; y similares.

- 25 El término "aproximadamente" incluye también cantidades que difieren debido a las diferentes condiciones de equilibrio para una composición que resulta de una mezcla inicial particular. Sean o no modificadas por el término "aproximadamente", las reivindicaciones incluyen equivalentes a las cantidades. En una realización, el término "aproximadamente" significa dentro de 10% del valor numérico descrito, preferiblemente dentro de 5% del valor numérico descrito.

- 30 Tal como se utiliza en la presente memoria, el término "protector de amina", "protector de amina añadido exógenamente", y "protector" se usan intercambiamente para referirse a un compuesto incluido en la mezcla de reacción que no se produce naturalmente, aislado, o naturalmente una parte del catalizador nitrilasa, en donde el protector de amina comprende una cantidad eficaz de al menos un grupo amino primario (R-NH₂) y/o un grupo amino secundario (R-NH-R₂) en donde R y R₁ pueden ser el mismo o diferente, y en donde ni R ni R₂ son grupos carbonilo (es decir, el grupo amino primario o secundario no forma parte de un grupo amida). En una realización preferida, el protector de amina comprende un número eficaz de grupos amino primarios. En una realización, el protector de amina se selecciona del grupo que consiste en polímeros poliaminados y materiales con funcionalidad amina que incluyen materiales de soporte con funcionalidad amina y polímeros con funcionalidad amina.

- 35 Tal como se utiliza en la presente memoria, el término "cantidad eficaz de un protector de amina" o "cantidad eficaz de al menos un protector de amina" se referirá a una cantidad del protector de amina que mejora la estabilidad del catalizador nitrilasa, la actividad específica, y/o la productividad catalítica cuando se convierte glicolonitrilo en ácido glicólico (o la sal de amonio del mismo). En una realización, la "cantidad eficaz de un protector de amina" se refiere a la cantidad de protector de amina presente en la mezcla de reacción que da como resultado un incremento medido en la actividad específica del catalizador cuando se compara con la actividad específica del mismo catalizador en ausencia del protector de amina bajo condiciones de reacción idénticas. Se sabe que una solución acuosa de glicolonitrilo contiene típicamente formaldehído libre e impurezas derivadas de formaldehído producidas durante la síntesis de glicolonitrilo o generadas como resultado de la disociación de glicolonitrilo. Se entiende que el protector de amina tendrá una "cantidad eficaz de grupos amino primarios y/o secundarios libres" capaces de reaccionar con el formaldehído o impurezas derivadas de formaldehído. Como condición, el protector de amina no incluye compuestos producidos de forma natural por o naturalmente presentes en el catalizador nitrilasa (por ejemplo, compuestos naturales producidos por o dentro de un catalizador de células enteras o aquellos encontrados dentro de un extracto de células enteras).

- 40 Tal como se utiliza en la presente memoria, el término "con funcionalidad amina" o "material con funcionalidad amina" se refiere a un compuesto o material que ha sido modificado químicamente para tener uno o más grupos amino primarios y/o secundarios. En una realización el material con funcionalidad amina se selecciona del grupo que consiste en materiales de soporte con funcionalidad amina y polímeros con funcionalidad amina. Son bien conocidos en la técnica medios para compuestos funcionalizados químicamente para contener uno o más grupos

amino (véase, por ejemplo, Bickerstaff, G. F., *Immobilization of Enzymes and Cells, Methods in Biotechnology*, Volúmen 1. 1997. Humana Press, Totowa, NJ, and Wong, S. S., *Chemistry of Protein Conjugation and Crosslinking*, 1991. CRC Press, Boca Raton, FL).

5 Tal como se utiliza en la presente memoria, "materiales de soporte con funcionalidad amina" incluyen, aunque sin limitación, materiales inorgánicos tales como perlas de vidrio y vidrio de poro controlado (por ejemplo, perlas de vidrio aminopropilsilado), sílice, magnetita, y alúmina así como materiales de soporte utilizados con frecuencia, resinas de intercambio aniónico débilmente básicas, así como ω -aminoheptilagarosa, ω -aminododecilagarosa, y ω -aminoetilagarosa. En una realización, el material con funcionalidad amina comprende polietilenimina (es decir, "un material con funcionalidad polietilenimina").

10 Tal como se utiliza en la presente memoria, "polímeros con funcionalidad amina" se refiere a polímeros que han sido modificados para incluir uno o más grupos amino primarios o secundarios. Típicamente, los polímeros no funcionalizados no contienen normalmente un número eficaz de grupos amino libres.

15 Tal como se utiliza en la presente memoria, el término "grupo amino libre" se refiere a un grupo amino primario (R-NH₂) y/o un grupo amino secundario (R-NH-R₂) en un protector de amina que es capaz de reaccionar con formaldehído.

20 Tal como se utiliza en la presente memoria, el término "polímero poliaminado" se refiere a polímeros que comprenden uno o más grupos amino libres. Los polímeros poliaminados se preparan a partir de al menos un monómero que comprende al menos un grupo amino primario y/o al menos un grupo amino secundario. En una realización, los polímeros poliaminados incluyen copolímeros que comprenden al menos un grupo amino libre. Los polímeros poliaminados pueden variar en tamaño desde oligómeros (típicamente menos de 1.000 Daltons) a polímeros con mayor peso molecular que oscila de 1.000 Daltons a 10.000.000 Daltons, preferiblemente de 1.000 Daltons a 2.000.000 Daltons, más preferiblemente de 1.000 Daltons a 1.000.000 Daltons, y más preferiblemente de 10.000 Daltons a 1.000.000 Daltons.

25 Tal como se utiliza en la presente memoria, el término "glicolonitrilo" se abrevia como "GLN" y es sinónimo de hidroxiacetonitrilo, 2-hidroxiacetonitrilo, hidroximetilnitrilo y todos los demás sinónimos del compuesto de número de registro en el *Chemical Abstracts Service* (CAS) 107-16-4.

30 Tal como se utiliza en la presente memoria, el término "ácido glicólico" se abrevia como "GLA" y es sinónimo de ácido hidroxiacético, ácido hidroxietanoico y todos los demás sinónimos del compuesto de número de registro en el CAS 79-14-1. El ácido glicólico producido por los presentes procedimientos puede estar en forma del ácido carboxílico protonizado y/o la sal de amonio correspondiente.

Tal como se utiliza en la presente memoria, el término "glicolato de amonio" se abrevia "NH₄GLA".

Tal como se utiliza en la presente memoria, el término "glicolamida" es la amida derivada de la reacción de amoníaco con ácido glicólico y se refiere a todos los otros sinónimos de compuesto que tienen el número de registro CAS 598-42-5.

35 Tal como se utiliza en la presente memoria, el término "glicólido" se refiere al compuesto de número de registro en el CAS 502-97-6.

40 Tal como se utiliza en la presente memoria, el término "formaldehído" se abrevia como "FA" y es sinónimo de aldehído fórmico, aldehído metílico, oxometano, y todos los demás sinónimos del compuesto de número de registro en el CAS 50-00-0. El formaldehído comercialmente disponible se compone típicamente de una mezcla de formaldehído monómero ("formaldehído libre") y diversos oligómeros de formaldehído junto con algo de metanol (típicamente alrededor de 1% en peso a alrededor de 15% en peso).

Tal como se utiliza en la presente memoria, el término "cianuro de hidrógeno" es sinónimo de ácido prúsico, ácido cianhídrico y todos los demás sinónimos del compuesto de número de registro en el CAS 200-821-6.

45 Tal como se utiliza en la presente memoria, el término "recuperación" significa aislamiento, purificación o transferencia del producto formado por el presente procedimiento. Los métodos para aislar y purificar el o los productos a partir de la mezcla de reacción son bien conocidos en la técnica y pueden incluir, aunque sin limitación, precipitación selectiva, cristalización, filtración, extracción con disolventes reactivos, intercambio iónico, electrodiálisis, polimerización, destilación, descomposición térmica, alcoholisis, cromatografía en columna y sus combinaciones. En una realización, el término "recuperación" también puede incluir transferencia de la mezcla de productos (típicamente después de filtrar el catalizador enzimático) a otra reacción para crear uno o más productos. En una realización preferida, se utiliza intercambio iónico para recuperar el ácido glicólico.

55 Tal como se usa en la presente memoria, los términos "catalizador enzimático", "catalizador nitrilasa" o "catalizador de células microbianas" se refiere a un catalizador que se caracteriza por una actividad de nitrilasa (es decir, comprende al menos un polipéptido que tiene actividad de nitrilasa) para la conversión de glicolonitrilo en ácido glicólico y amoníaco. Una enzima nitrilasa convierte directamente un nitrilo (preferiblemente, un nitrilo alifático) en el ácido carboxílico correspondiente, sin formación de la amida correspondiente como producto intermedio (véase la

ecuación 1). Las nitrilasas comparten varios dominios distintivos conservados conocidos en la técnica, incluyendo un dominio distintivo en la presente invención denominado "secuencia distintiva catalítica" o "secuencia distintiva". Esta región comprende un residuo de cisteína esencial (por ejemplo, Cys₁₆₄ de la SEQ ID NO: 4). Como tales, los polipéptidos que tienen actividad de nitrilasa se pueden identificar por la existencia de la secuencia distintiva de dominio catalítico (SEQ ID NO: 1). En una realización preferida, la secuencia distintiva es la SEQ ID NO: 2. El catalizador enzimático puede estar en forma de una célula microbiana entera, célula(s) microbiana(s) permeabilizada(s), uno o más componentes celulares de un extracto celular microbiano, enzima parcialmente purificada, o enzima purificada. El catalizador enzimático puede ser libre (no inmovilizado) o inmovilizado en o sobre un soporte soluble o insoluble. Tal como se usa en la presente memoria, "catalizador enzimático reciclado" se refiere a un catalizador enzimático que se reutiliza como catalizador enzimático en reacciones por lotes. En una realización preferida, el catalizador enzimático es una célula hospedante microbiana transformada que expresa recombinantemente al menos una enzima que tiene actividad de nitrilasa; en donde dicha enzima comprende el resto distintivo catalítico de la SEQ ID NO: 1 o SEQ ID NO: 2.

Tal como se utiliza en la presente memoria, los términos "*Acidovorax facilis*" y "*A. facilis*" se usan intercambiamente y se refieren a *Acidovorax facilis* 72W depositado en la *American Type Culture Collection* (una autoridad internacional de depósito), que tiene el número de acceso 55746 ("ATCC 55746"). Los nitrilasas mutantes derivadas de *A. facilis* 72W caracterizadas por su mejor actividad de nitrilasa en convertir glicolonitrilo en ácido glicólico han sido descritas previamente (véase la patente de EE.UU. 7.198.927 compartida con la sociedad titular de la presente). Ejemplos de estas nitrilasas mutantes derivadas de *A. facilis* 72W se proporcionan en las SEQ ID NO: 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53 y 55.

Tal como se utiliza en la presente memoria, los términos "*Escherichia coli*" y "*E. coli*". se usan intercambiamente. Varias cepas de *E. coli* adecuadas para la expresión recombinante se describen en la presente memoria, incluyendo, aunque sin limitación, *E. coli* MG1655 que tiene el número de depósito internacional ATCC 47076, *E. coli* FM5 que tiene el número de depósito internacional ATCC 53911, *E. coli* W3110 que tiene el número de depósito internacional ATCC 27325, *E. coli* MC4100 que tiene el número de depósito internacional ATCC 35695 y *E. coli* W1485 que tiene el número de depósito internacional ATCC 12435. En una realización las cepas de *Escherichia coli* adecuadas incluyen *E. coli* FM5 (ATCC 53911) y *E. coli* MG1655 (ATCC 47076).

Tal como se utiliza en la presente memoria, los términos "*E. coli* SS1001" o "SS1001" se refieren a una cepa de *E. coli* transformada que expresa la nitrilasa de *Acidovorax facilis* 72W que tiene el número de acceso en ATCC PTA-1177 (véase la patente de EE.UU. 6.870.038; incorporada en la presente memoria en su totalidad como referencia). La nitrilasa de *E. coli* SS1001 expresada recombinantemente (SEQ ID NO: 57) contiene 2 cambios de secuencia menores en comparación con la secuencia de nitrilasa de 72W de tipo natural (SEQ ID NO: 4). El codón de iniciación estaba cambiado de GTG a ATG para facilitar la expresión recombinante y se introdujo un artefacto durante la clonación que dio como resultado un cambio de un único aminoácido cerca del extremo C (Pro367 [CCA] → Ser [TCA]).

Tal como se utiliza en la presente memoria, los términos "mezcla de reacción de glicolonitrilo acuosa adecuada", "mezcla de reacción acuosa adecuada", y "condiciones de reacción adecuadas" se refieren a los materiales (incluyendo al menos un protector de amina) y condiciones de reacción bajo las cuales se lleva a cabo el presente procedimiento. Los componentes de la mezcla de reacción acuosa adecuada se proporcionan en la presente memoria y los expertos en la técnica apreciarán la gama de variaciones de componentes adecuada para este procedimiento.

Tal como se utiliza en la presente memoria, los términos "solución de glicolato de amonio acuosa", "solución acuosa que comprende glicolato de amonio" y "solución acuosa de glicolato de amonio" se utilizarán para describir una solución acuosa que comprende glicolato de amonio producido por la hidrólisis enzimática de glicolonitrilo bajo condiciones típicas de reacción enzimática (es decir, un intervalo de pH de aproximadamente 6 a aproximadamente 8). La solución acuosa de glicolato de amonio comprende glicolato de amonio a una concentración de al menos aproximadamente 0,1 por ciento en peso a aproximadamente 99% en peso de glicolato de amonio. En otra realización, la solución acuosa de glicolato de amonio se compone de al menos aproximadamente 10% en peso a aproximadamente 75% en peso glicolato de amonio. En una realización adicional, la solución acuosa de glicolato de amonio se compone de al menos aproximadamente 20% en peso a aproximadamente 50% en peso de glicolato de amonio. El pH de la solución acuosa de glicolato de amonio puede ser de aproximadamente 2 a aproximadamente 12, preferiblemente 5 a aproximadamente 10, más preferiblemente 6 a aproximadamente 8. El pH se puede ajustar según sea necesario antes de iniciar las etapas del procedimiento relacionadas con la recuperación del ácido glicólico (en forma de ácido o sal) de la solución acuosa de glicolato de amonio.

Tal como se usa en la presente memoria, los términos "productividad del catalizador" y "productividad del catalizador enzimático" se refieren a la cantidad total de producto producida por gramo de catalizador. En el presente procedimiento, el catalizador comprende una enzima nitrilasa (EC 3.5.5.7) y el producto formado es ácido glicólico y/o glicolato de amonio (dependiendo del pH de la reacción). En general, los métodos presentes se realizan en condiciones esencialmente de pH neutro de modo que el ácido glicólico producido esté predominantemente en forma de la sal correspondiente del ácido glicólico (es decir, glicolato de amonio). Generalmente, en reacciones por lotes con reciclaje del catalizador, la actividad del catalizador disminuye con cada reacción de reciclaje (inactivación enzimática). Como se muestra en el Ejemplo 6, y en una realización, la actividad

específica del catalizador se incrementa durante múltiples reacciones de reciclaje, que dan como resultado un incremento en la productividad del catalizador.

Los términos "actividad específica de catalizador", "actividad específica" y "actividad específica de nitrilasa" se refieren a la actividad enzimática por unidad de masa (por ejemplo, miligramos) de proteína, peso de células secas o peso de perlas (catalizador inmovilizado) cuando se convierte glicolonitrilo en ácido glicólico (o el glicolato de amonio correspondiente). Se midieron comparaciones de la actividad de nitrilasa proporcionales al peso de células secas o al peso de perlas. Dado que el mismo catalizador y el sistema de expresión se usaron para comparar la actividad específica entre reacciones con o sin (es decir, el control) el protector de amina, se midieron comparaciones y mejoras/incrementos documentados de la actividad específica relativa al peso de células secas (pcs) o al peso de perlas (pp). Como se usa en la presente memoria, el término "mejor actividad específica del catalizador" y "mejor actividad específica" se refiere al incremento de la actividad específica observada en relación con la actividad específica del control correspondiente. La mejora de la actividad específica se mide en condiciones de reacción idénticas en las que la única diferencia es la inclusión y omisión del protector de amina. En una realización, la mejora de la actividad específica se mide durante múltiples reacciones de reciclaje. En otra realización, la mejora o incremento de la actividad específica de catalizador utilizando el presente procedimiento se incrementa al menos un 1%, preferiblemente al menos un 5%, más preferiblemente al menos un 10%, incluso más preferiblemente al menos un 25%.

Tal como se usa en la presente memoria, el término "una unidad de actividad enzimática" o "una unidad de actividad de nitrilasa" o "U" se define como la cantidad de actividad enzimática requerida para la producción de 1 μ mol de producto ácido glicólico por minuto (U de GLA/g en peso de células secas o en peso de perlas) a una temperatura especificada (por ejemplo, 25°C).

Tal como se utiliza en la presente memoria, los términos "organismo recombinante", "célula hospedante transformada", "célula hospedante", "transformante", "organismo transgénico", y "hospedante microbiano transformado" se refieren a un organismo hospedante que ha sido transformado con DNA heterólogo o extraño. Los organismos recombinantes de la presente invención expresan secuencias codificadoras extrañas o genes que codifican la enzima nitrilasa activa. "Transformación" se refiere a la transferencia de un fragmento de DNA al organismo hospedante. El fragmento de DNA transferido puede ser incorporado cromosómica o extracromosómicamente (es decir, a través de un vector) en el organismo hospedante. Tal como se utiliza en la presente memoria, el término "casete de transformación" se refiere a un fragmento específico de DNA que contiene un conjunto de elementos genéticos convenientemente dispuestos para su inserción en una célula hospedante, por lo general como parte de un plásmido. Tal como se utiliza en la presente memoria, el término "casete de expresión" se refiere a un fragmento específico de DNA que contiene un conjunto de elementos genéticos convenientemente dispuestos para su inserción en una célula hospedante, por lo general como parte de un plásmido que también permite una mayor expresión génica en el hospedante.

Tal como se utiliza en la presente memoria, los términos "fragmento de ácido nucleico" y "molécula de ácido nucleico" se refieren a una molécula de DNA que puede codificar un gen completo, la secuencia codificadora, y/o secuencias reguladoras que preceden (5', aguas arriba) o siguen (3', aguas abajo) a la secuencia codificadora. En un aspecto, las presentes moléculas de ácidos nucleicos codifican polipéptidos que tienen actividad de nitrilasa.

Tal como se utiliza en la presente memoria, el término "gen" se refiere a una molécula de ácido nucleico que expresa una proteína específica. Como se usa en la presente memoria, puede o no puede incluir secuencias reguladoras que preceden (secuencias 5' no codificadoras) y que siguen (secuencias 3' no codificadoras) a la secuencia codificadora. "Gen quimérico" se refiere a cualquier gen que no es un gen natural, que comprende secuencias reguladoras y codificadoras que no se encuentran juntas en la naturaleza. Por consiguiente, un gen quimérico puede comprender secuencias reguladoras y secuencias codificadoras que se derivan de diferentes fuentes o secuencias reguladoras y secuencias codificadoras derivadas de la misma fuente, pero dispuestas de una manera diferente a la encontrada en la naturaleza. "Gen endógeno" se refiere a un gen natural en su localización natural en el genoma de un organismo. Un gen "extraño" se refiere a un gen que no se encuentra normalmente en el organismo hospedante, pero que se introduce en el organismo hospedante por transferencia de genes. Los genes extraños pueden comprender genes naturales insertados en un organismo no natural o genes quiméricos. Un "transgén" es un gen que ha sido introducido en el genoma por un procedimiento de transformación.

Tal como se utiliza en la presente memoria, el término "secuencia codificadora" se refiere a una secuencia de DNA que codifica una secuencia específica de aminoácidos. Tal como se usa en la presente memoria, "secuencias reguladoras adecuadas" se refieren a secuencias de nucleótidos localizadas aguas arriba (secuencias no codificadoras en 5'), dentro de, o aguas abajo (secuencias no codificadoras en 3') de una secuencia codificadora, y que influyen en la transcripción, procesamiento o estabilidad del RNA, o la traducción de la secuencia codificadora asociada. Las secuencias reguladoras pueden incluir promotores, secuencias situadas antes de las secuencias que se traducen, intrones, secuencias de reconocimiento de poliadenilación, sitios de procesamiento del RNA, sitios de unión a efectores, y estructuras de tallo-bucle.

"Promotor" se refiere a una secuencia de DNA capaz de controlar la expresión de una secuencia codificadora o RNA funcional. En general, una secuencia codificadora está situada en 3' respecto a una secuencia promotora. Los promotores pueden derivarse en su totalidad de un gen natural o estar compuestos de diferentes elementos

derivados de diferentes promotores encontrados en la naturaleza, o incluso comprender segmentos de DNA sintéticos. Los promotores que hacen que un gen se exprese en la mayoría de tipos de células, en la mayoría de veces, o bajo la mayoría de condiciones ambientales se denominan comúnmente "promotores constitutivos". Los promotores que hacen que un gen se exprese sólo en presencia de un compuesto particular o condición ambiental se denominan comúnmente "promotores inducibles". Dado que en la mayoría de los casos los límites exactos de las secuencias reguladoras no han sido completamente definidos, los fragmentos de DNA de diferentes longitudes pueden tener idéntica actividad promotora.

Tal como se utiliza en la presente memoria, el término "unida operativamente" se refiere a la asociación de secuencias de ácido nucleico con una única molécula de ácido nucleico de modo que la función de una secuencia se ve afectada por la otra. Por ejemplo, un promotor está unido operativamente a una secuencia codificadora cuando es capaz de afectar a la expresión de esa secuencia codificadora (es decir, que la secuencia codificadora está bajo el control transcripcional del promotor). Las secuencias codificadoras pueden unirse operativamente a las secuencias reguladoras en orientación con sentido o antisentido.

Tal como se usa en la presente memoria, el término "secuencias no codificadoras en 3'" se refiere a secuencias de DNA localizadas aguas abajo de una secuencia codificadora e incluyen secuencias de reconocimiento de la poliadenilación (normalmente limitadas a eucariotas) y otras secuencias que codifican señales reguladoras capaces de afectar el procesamiento del mRNA o a la expresión génica. La señal de poliadenilación (normalmente limitada a los eucariotas) se caracteriza normalmente por afectar a la adición de tramos de poli(ácido adenílico) al extremo 3' del precursor del mRNA.

Los expertos en la técnica conocen bien el "sesgo de codones" exhibido por una célula hospedante específica en el uso de codones de nucleótidos para especificar un aminoácido dado. Por lo tanto, cuando se sintetiza un gen para mejorar la expresión en una célula hospedante, es deseable diseñar el gen de modo que su uso de codones refleje el sesgo de codones preferido por la célula hospedante. Un estudio de genes derivados de la célula hospedante cuando está disponible la información de la secuencia puede determinar su sesgo de codones. La optimización de codones es bien conocida en la técnica y se ha descrito para varios sistemas, incluyendo, aunque sin limitación, levadura (Outchkourov et al., *Protein Expr. Purif.* 24 (1): 18-24 (2002)) y *E. coli* (Feng et al., *Biochemistry*, 39 (50):15399-15409 (2000)).

Catalizadores nitrilasa.

Todas las nitrilasas (EC 3.5.5.7) comparten una triada catalítica conservada (Glu, Lys y Cys) (Chauhan et al., *Appl. Microbiol. Biotechnol.* 61:118-122 (2003); Pace, H. and Brenner, C., *Genome Biol.* 2(1):reviews 0001.1-0001.9 (2001)). Todas las nitrilasas conocidas tienen una cisteína nucleófila en el sitio de la actividad enzimática (Cowan et al., *Extremophiles*, 2:207-216 (1998); Pace, H. and Brenner, C., *supra*; y Chauhan et al., *supra*) y todas son susceptibles de inactivación por reactivos de tiol (concentraciones 1,0 mM de cloruro de cobre, nitrato de plata, acetato mercúrico o cloruro férrico producían cada uno disminuciones importantes en la actividad de la enzima nitrilasa de *A. facilis* 72W). Los residuos de cisteína son también capaces de ser oxidados irreversiblemente a ácido sulfínico, dando como resultado una pérdida de la actividad enzimática. A pesar de la sensibilidad de las enzimas nitrilasas a diversos mecanismos inactivantes, las células de *A. facilis* 72W inmovilizadas son robustas, y capaces de retener mucha de su actividad de nitrilasa después de numerosas reacciones de reciclaje (Patentes de EE.UU. 6.870.038; 7.148.051 y 7.198.927; y Chauhan et al., *supra*). Los catalizadores nitrilasas derivados de la nitrilasa de *A. facilis* 72W han demostrado también catalizar la conversión de α -hidroxinitrilos (es decir, gliconitrilos) a ácidos α -hidroxicarboxílicos (es decir, ácido glicólico) (véanse las patentes de EE.UU. 6.383.786; 6.416.980 y 7.198.927).

Se han descrito comparaciones de secuencias de la nitrilasa de *A. facilis* 72W con otras nitrilasas bacterianas (Patente de EE.UU. 6.870.038; Chauhan et al., *supra*). La nitrilasa de 72W tiene diversos dominios distintivos conservados incluyendo una región de 16 aminoácidos cerca del extremo amino (residuos de aminoácidos 40-55 de la SEQ ID NO: 4) y una región catalítica de 12 aminoácidos (residuos de aminoácidos 160-171 de la SEQ ID NO: 4) que contiene el residuo de cisteína esencial. Este residuo de cisteína esencial (Cys₁₆₄ de la SEQ ID NO: 4), junto con el ácido glutámico conservado (Glu₄₈ de la SEQ ID NO:4) y los residuos de lisina (Lys₁₃₀ de la SEQ ID NO:4), forman el resto de la triada catalítica encontrado en todas las nitrilasas (Pace, H., y Brenner, C., *supra*).

Las regiones que rodean cada uno de los residuos de la triada catalítica están altamente conservadas, especialmente la región que rodea al residuo catalítico de cisteína. El residuo catalítico de cisteína esencial está localizado en una región altamente conservada denominada el "resto distintivo catalítico" o "resto distintivo". Por tanto, el presente procedimiento es útil para proteger la actividad enzimática de cualquier nitrilasa que contenga el resto distintivo catalítico definido por la Fórmula 1 (los residuos de aminoácidos en letras negritas indican los residuos de aminoácidos estrictamente conservados, los residuos de aminoácidos en letra cursiva son los que presentan mínima variabilidad [es decir, una variación mínima de 3 o menos residuos de aminoácidos], el residuo catalítico de cisteína está subrayado):

Fórmula 1 (SEQ ID NO: 1).

Gly-Xaa₁-Xaa₂-Xaa₃-Cys-Trp-Glu-Xaa₄-Xaa₅-Xaa₆-Xaa₆-Xaa₇-Xaa₈

en donde

Xaa₁ = Ala o Gly;

Xaa₂ = Leu, Val o Ala;

Xaa₃ = Ala, Asn, Ile, Cys, Val o Gln;

Xaa₄ = His o Asn;

5 Xaa₅ = Leu, Tyr, Phe, Ala, Met, Lys, Val, Thr o Arg;

Xaa₆ = Asn, Gln, Met, Leu o Ser;

Xaa₇ = Pro o Thr; y

Xaa₈ = Leu o Val.

10 En una realización preferida, el resto distintivo de nitrilasa de Fórmula 1 es Xaa₁ = Ala o Gly; Xaa₂ = Leu; Xaa₃ = Ala, Asn, Ile, Cys, Val o Gln; Xaa₄ = His; Xaa₅ = Leu, Tyr, Phe, Ala, Met, Lys, Val, Thr o Arg; Xaa₆ = Ser, Gln, Asn o Met; Xaa₇ = Pro; y Xaa₈ = Leu; dando como resultado el resto distintivo catalítico representado por la siguiente fórmula:

Gly-Xaa₁-Leu-Xaa₃-Cys-Trp-Glu-His-Xaa₅-Xaa₆-Pro-Leu (SEQ ID NO: 2)

15 Ejemplos de nitrilasas, que incluyen las secuencias y la posición en la secuencia de resto distintivo catalítico correspondiente se proporcionan en la Tabla 1.

Tabla 1. Región de cisteína catalítica conservada - Restos distintivos catalíticos

Fuente de nitrilasa	Número de acceso en GenBank®	Secuencia de aminoácidos SEQ ID NO:	Secuencia del resto distintivo (posiciones de los residuos de aminoácidos)
<i>Acidovorax facilis</i> 72W	ABD98457.1	4	GGLNCWEHFQPL (160-171)
<i>Alcaligenes faecalis</i> JM3	BAA02684.1	5	GALCCWEHLSPL (159-170)
<i>Rhodococcus rhodochrous</i> J1	Q03217	6	GALNCWEHFQTL (161-172)
<i>Rhodococcus rhodochrous</i> K22	Q02068	7	GGLNCWEHFQPL (166-177)
<i>Nocardia sp.</i> C-14-1	AAX18182.1	8	GGLNCWEHFQPL (154-165)
<i>Bordetella bronchiseptica</i> RB50	NP_887662.1	9	GAVVWENYMP (161-172)
<i>Arabidopsis thaliana</i>	AAB60275.1 AAA19627.1	10	GAAICWENRMPL (175-186)
<i>Synechococcus elongatus</i> PCC 7942	YP_399857.1	11	GALACWEHYNPL (157-168)
<i>Synechococcus elongatus</i> PCC 6301	YP_171411.1	12	GALACWEHYNPL (157-168)
<i>Synechocystis sp.</i> PCC 6803	NP_442646.1	13	GALACWEHYNPL (165-176)
<i>Pseudomonas entomophila</i> L48	YP_6090481.1	14	GAAVCWENYMP (161-172)
<i>Zymomonas mobilis</i>	YP_162942.1	15	GAAICWENYMPV (161-172)
<i>Bacillus sp.</i> OxB-1	BAA90460.1	16	GGLQCWEHFLPL (158-169)
<i>Comamonas testosteroni</i>	AAA82085.1	17	GGLQCWEHALPL (159-170)
<i>Synechococcus sp.</i> CC9605	YP_381420.1	18	GALACWEHYNPL (156-167)
<i>Pseudomonas fluorescens</i> Pf-5	YP_260015.1	19	GAVICWENMMPL (161-172)
<i>Nocardia farcinica</i> IFM 10152	YP_119480.1	20	GALCCWEHLQPL (159-170)
<i>Alcaligenes faecalis</i> 1650	AAY06506.1	21	GALCCWEHLSPL (159-170)
<i>Pseudomonas syringae</i> pv. <i>syringae</i> B728a	AAY35081.1	22	GALCCWEHLQPL (157-168)
<i>Bradyrhizobium sp.</i> BTail	ZP_00859948.1	23	GALCCWEHLQPL (163-174)

Fuente de nitrilasa	Número de acceso en GenBank®	Secuencia de aminoácidos SEQ ID NO:	Secuencia del resto distintivo (posiciones de los residuos de aminoácidos)
<i>Rhodococcus rhodochrous</i> NCIMB 11216	CAC88237	24	GALNCWEHFQTL (161-172)
<i>Rhodococcus rhodochrous</i> ATCC 39484™	N/A	25	GALNCWEHFQTL (161-172)

En una realización, el catalizador nitrilasa que comprende un polipéptido que tiene actividad de nitrilasa, aislado de un género seleccionado del grupo que consiste en *Acidovorax*, *Rhodococcus*, *Nocardia*, *Bacillus* y *Alcaligenes*. En una realización, el catalizador nitrilasa comprende un polipéptido que tiene actividad de nitrilasa aislado de un género seleccionado del grupo que consiste en *Acidovorax* y *Rhodococcus*.

- 5 En otra realización, el polipéptido que tiene actividad de nitrilasa procede de *Acidovorax facilis* 72W (ATCC 55746) o un polipéptido (que tiene actividad de nitrilasa) que es sustancialmente similar a la nitrilasa de *Acidovorax facilis* 72W (SEQ ID NO: 4) o la enzima derivada de *A. facilis* 72W representada por la SEQ ID NO: 51.

En una realización, el catalizador nitrilasa es una célula microbiana hospedante transformada para expresar al menos un polipéptido que tiene actividad de nitrilasa. En una realización, la célula hospedante transformada se selecciona del grupo que consiste en: *Comamonas* sp., *Corynebacterium* sp., *Brevibacterium* sp., *Rhodococcus* sp., *Azotobacter* sp., *Citrobacter* sp., *Enterobacter* sp., *Clostridium* sp., *Klebsiella* sp., *Salmonella* sp., *Lactobacillus* sp., *Aspergillus* sp., *Saccharomyces* sp., *Yarrowia* sp., *Zygosaccharomyces* sp., *Pichia* sp., *Kluyveromyces* sp., *Candida* sp., *Hansenula* sp., *Dunaliella* sp., *Debaryomyces* sp., *Mucor* sp., *Torulopsis* sp., *Methylobacteria* sp., *Bacillus* sp., *Escherichia* sp., *Pseudomonas* sp., *Rhizobium* sp., y *Streptomyces* sp. En una realización preferida, la célula microbiana hospedante se selecciona del grupo que consiste en *Bacillus* sp., *Pseudomonas* sp. y *Escherichia* sp. En una realización preferida, el catalizador es una célula hospedante de *Escherichia coli* que expresa recombinantemente uno o más de los polipéptidos que tienen actividad de nitrilasa.

En otra realización, el catalizador nitrilasa que comprende un polipéptido que tiene actividad de nitrilasa, en donde dicho polipéptido que tiene actividad de nitrilasa tiene al menos 60% de identidad con la SEQ ID NO: 51, preferiblemente al menos 70% de identidad con la SEQ ID NO: 51, incluso más preferiblemente al menos 80% de identidad con la SEQ ID NO: 51, e incluso más preferiblemente al menos 90% de identidad con la SEQ ID NO: 51, y lo más preferiblemente al menos 95% de identidad con la SEQ ID NO: 51.

En la presente memoria se describen ejemplos de trabajo de diversos catalizadores que tienen actividad de nitrilasa derivados de diversas fuentes, incluyendo un catalizador derivado de nitrilasa de *A. facilis* 72W. Se han descrito en la técnica diversos mutantes derivados de la enzima nitrilasa de *Acidovorax facilis* 72W (Patentes de EE.UU. 7.148.051 y 7.198.927).

En una realización, el polipéptido que tiene actividad de nitrilasa se selecciona del grupo que consiste en las SEQ ID NO: 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55 y 57. En otra realización, el polipéptido que tiene actividad de nitrilasa se selecciona del grupo que consiste en las SEQ ID NO: 4, 24, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55 y 57. En otra realización, el polipéptido que tiene actividad de nitrilasa se selecciona del grupo que consiste en las SEQ ID NO: 4, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55 y 57. En otra realización, el polipéptido que tiene actividad de nitrilasa se selecciona del grupo que consiste en las SEQ ID NO: 4, 24, 25 y 51. En otra realización, el catalizador nitrilasa comprende el polipéptido de la SEQ ID NO SEQ ID NO: 51.

35 Protectores de amina

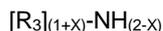
El protector de amina es cualquier compuesto añadido exógenamente a la mezcla de reacción acuosa (por ejemplo, no producida de forma natural por, encontrado de forma natural con, o aislado de forma natural a partir del catalizador enzimático) que comprende al menos un grupo amino primario (R-NH₂) y/o un grupo amino secundario (R-NH-R₂) capaz de reaccionar con formaldehído, en donde R y R₁ pueden ser el mismo o diferente, y donde ni R ni R₂ son grupos carbonilo (es decir, no son un grupo amida). Se entiende también que el presente catalizador enzimático hidroliza grupos nitrilo, y por lo tanto, el protector de amina no incluye preferiblemente un grupo nitrilo capaz de ser hidrolizado por un catalizador nitrilasa. En una realización preferida, el protector de amina es un compuesto o componente añadido exógenamente en la mezcla de reacción que comprende al menos un grupo amina primario.

En una realización, de la presente invención, una cantidad eficaz de al menos un protector de amina se añade a la mezcla de reacción. Como se usa en la presente memoria, "una cantidad eficaz de un protector de amina" es la cantidad de protector de amina necesaria para dar como resultado una mejora detectable en la actividad específica, estabilidad, o productividad catalítica del catalizador cuando se convierte enzimáticamente gliconitrilo en ácido glicólico. Está dentro de la experiencia del experto en la técnica ajustar la cantidad de protector de amina en la mezcla de reacción acuosa con el fin de lograr el efecto deseado. En un aspecto, la cantidad de protector de amina añadida a la

mezcla de reacción comprende un exceso molar de los grupos amino libres con relación a la cantidad de formaldehído presente en la mezcla de reacción. En un aspecto, la mezcla de reacción comprende al menos 0,01 ppm de formaldehído. En otro aspecto, la cantidad de protector de amina en la mezcla de reacción comprende una relación molar de grupos amino libres (primario, secundario, o una combinación de los mismos) con relación a la cantidad de formaldehído libre (HCHO) presente en la mezcla de reacción que varía de 0,001:1,0 a 1,0:0,001, preferiblemente al menos 0,5:1,0.

En un aspecto, el protector de amina se selecciona del grupo que consiste en:

- a) un compuesto de la fórmula



10 en donde X es 0 o 1, y R₃ es independientemente un grupo hidrocarbilo sustituido o grupo hidrocarbilo C1 a C20, en el que R₃ comprende opcionalmente uno o más enlaces éter; con la condición de que (i) el grupo sustituido no sea un grupo ciano capaz de reaccionar con el catalizador nitrilasa y (ii) R₃ no sea un grupo carbonilo,

15 b) un polímero poliaminado que comprende un número eficaz de grupos amino libres en donde dichos grupos amino libres son grupos amino primarios y/o grupos amino secundarios; y

- c) un material con funcionalidad amina que comprende un número eficaz de grupos amino libres en donde dichos grupos amino libres son grupos amino primarios y/o grupos amino secundarios;

en donde el protector de amina es insoluble en condiciones de reacción acuosa aceptables y,

en donde el protector de amina no se produce de forma natural por el catalizador enzimático.

20 El grupo hidrocarbilo de R₃ puede ser lineal, ramificado, cíclico, policíclico, o aromático. Como se usa en la presente memoria, por los términos "hidrocarbilo", "grupo hidrocarbilo", y "resto hidrocarbilo" se entiende una cadena lineal, ramificada o cíclica compuesta de átomos de carbono unidos por enlaces carbono a carbono simples, dobles o triples y/o por enlaces éter, y sustituidos de acuerdo con átomos de hidrógeno. Tales grupos hidrocarbilo pueden ser alifáticos o aromáticos. Ejemplos de grupos hidrocarbilo incluyen metilo, etilo, propilo, isopropilo, butilo, isobutilo, t-butilo, ciclopropilo, ciclobutilo, pentilo, ciclopentilo, metilciclopentilo, hexilo, ciclohexilo, bencilo, y fenilo. En una realización preferida, el resto hidrocarbilo es una cadena lineal, ramificada o cíclica compuesta de átomos de carbono unidos por enlaces carbono a carbono simples y/o por enlaces éter, y sustituido de acuerdo con átomos de hidrógeno. Como se usa en la presente memoria, "hidrocarbilo sustituido" significa un grupo que está sustituido y contiene uno o más grupos sustituyentes que no provocan el sustrato, catalizador, o producto a ser inestable o inadecuado para el uso o la reacción deseada. Los grupos sustituyentes que son generalmente útiles incluyen éter, éster, halo, amino (incluyendo grupos amino primarios, secundarios o terciarios), hidroxilo, vinilideno o vinilideno sustituido, sililo o sililo sustituido, nitro, nitroso, sulfonilo, sulfonilo, sal de metal alcalino de ácido sulfónico, boranilo o boranilo sustituido, y tioéter.

35 En una realización, el protector de amina es un polímero poliaminado que comprende un número eficaz de grupos amino libres. En otra realización, los polímeros poliaminados incluyen copolímeros que comprenden un número eficaz de grupos amino libres. Los polímeros poliaminados pueden variar en tamaño desde oligómeros (típicamente menos de 1.000 Daltons) a polímeros de peso molecular mayor que varían de 1.000 Daltons a 10.000.000 Daltons, preferiblemente de 1.000 Daltons a 2.000.000 Daltons, más preferiblemente de 1.000 Daltons a 1.000.000 Daltons, y más preferiblemente de 10.000 Daltons a 1.000.000 Daltons.

40 En una realización, el polímero poliaminado se selecciona del grupo que consiste en 2-amino-2-desoxi-(1→4)-β-D-glucopirranano (es decir, quitosano), polímeros de polialquilenamino que tienen restos alqueno de 2 a 12 átomos de carbono, polietilenimina, polialilamina, poli(alcohol vinílico)/copolímeros de polivinilamina, D-polilisina, L-polilisina, mezclas de D/L polilisina, polietilenimina reticulada con glutaraldehído, y mezclas de los mismos. En una realización preferida, el polímero poliaminado es polímero de polietilenimina (lineal o ramificado) y copolímeros que comprenden polietilenimina. En otra realización preferida, el polímero poliaminado es polietilenimina reticulada con glutaraldehído.

45 También se describe en esta memoria, el catalizador enzimático está inmovilizado en o sobre una partícula o perla que comprende polietilenimina reticulada con glutaraldehído.

El protector de amina, como se utiliza en la invención, es insoluble en la mezcla de reacción acuosa. Sin embargo, los protectores de amina solubles también se describen en la presente memoria y pueden estar separados de la mezcla de producto posterior utilizando cualquier número de metodologías de separación bien conocidas en la técnica (por ejemplo, técnicas de intercambio iónico o cromatográficas). Dependiendo de la naturaleza química del protector de amina soluble, un experto en la técnica puede separar el protector de amina del producto de ácido glicólico (o glicolato de amonio) ajustando las condiciones de procedimiento posteriores a la reacción para precipitar o adsorber selectivamente uno o más componentes de la mezcla de producto. Para algunas aplicaciones, puede que no sea necesario eliminar el protector de amina soluble de la mezcla de producto, especialmente para aplicaciones en las que no se requiera ácido glicólico de alta pureza y/o glicolato de amonio.

En una realización preferida, el protector de amina es sustancialmente insoluble en la mezcla de reacción acuosa. El uso de un protector de amina insoluble facilita la separación simple/aislamiento (es decir, filtración, centrifugación, etc.) de la mezcla de producto acuosa resultante. En otra realización preferida, el protector de amina es un polímero poliaminado que es insoluble en la mezcla de reacción acuosa.

5 En otra realización, el protector de amina es un material con funcionalidad amina insoluble y puede incluir materiales de soporte con funcionalidad amina y polímeros con funcionalidad amina. Los métodos para modificar químicamente los materiales insolubles que incluyen al menos un grupo amino primario son bien conocidos en la técnica (por ejemplo, véase Bickerstaff, G. F., *Immobilization of Enzymes and Cells, Methods in Biotechnology*, Volumen 1. 1997. Humana Press, Totowa, NJ, y Wong, S. S., *Chemistry of Protein Conjugation and Crosslinking*, 1991. CRC Press, Boca Raton, FL).

10 Los materiales con funcionalidad amina que comprenden un número eficaz de grupos amino libres (o funcionalizados para contener un número eficaz de grupos amino libres) útiles en el presente procedimiento pueden incluir, pero sin limitación, polisacáridos que han sido funcionalizados químicamente para contener uno o más grupos amino (por ejemplo, glicanos con funcionalidad amina, agarosa, carragenina, alginato, dextrano, y celulosa), metacrilatos, poliuretanos, poliésteres, nylons, alúmina, sílice, poliestireno, poli(alcoholes de vinilo), magnetita, y vidrio de poro controlado. Se pueden utilizar los polímeros de polialquilenammina (que tienen típicamente restos alquilenos de 2 a 12 átomos de carbono) y materiales de soporte con funcionalidad polialquilenammina, como polímeros de polietilenimina solubles o insolubles, materiales de soporte con funcionalidad polietilenamina, polietilenimina, polímeros de polietilenimina (lineales o ramificados), polietilenimina reticulada con glutaraldehído, resinas de intercambio iónico de polietilenimina, resinas de intercambio aniónico débilmente básicas (por ejemplo, resinas de intercambio aniónico débilmente básicas AMBERLITE® IRA-95, IRA-96, IRA-67, y IRA-92 (disponibles en Rohm and Haas, Philadelphia, PA), y DIAION® WA20 o WA21J (disponibles en Mitsubishi Chemical, Tokyo, Japan), celulosa-polietilenimina, sílice-polietilenimina, polialilamina, poli(alcohol vinílico)/copolímeros de polivinilamina, polilisina (D, L, o mezclas D/L), ω -aminohexilagarosa, ω -aminododecilagarosa, ω -aminoetilagarosa, perlas de vidrio aminopropilsilado, y combinaciones de los mismos. En una realización preferida, el protector de amina no incluye un grupo ciano capaz de ser hidrolizado por el catalizador nitrilasa.

En una realización, los derivados de celulosa de amino alquil-, amino(hidroxi)alquil-, aminoalquil-éter-, y amino(hidroxi)alquiléter, quitina y otros carbohidratos que se producen de forma natural se seleccionan preferiblemente del grupo que consiste en:

30 $H_2N-(CH_2)_n$ -[carbohidrato] donde $n=1-10$, incluyendo isómeros de alquilo,
 $H_2N-(CH_2)_m-CHOH-(CH_2)_n$ -[carbohidrato], donde $m=0-10$ y $n=0-10$,
 $H_2N-(CH_2)_n-O$ -[carbohidrato] donde $n=1-10$, y
 $H_2N-(CH_2)_m-CHOH-(CH_2)_n-O$ -[carbohidrato] donde $m=0-10$ y $n=0-10$.

35 En otra realización, los derivados de celulosa aminofenilos o aminobencilos, quitina u otros carbohidratos que se producen de forma natural se seleccionan preferiblemente del grupo que consiste en:

$H_2N-C_6H_4-(CH_2)_n$ -[carbohidrato],
 $H_2N-CH_2-C_6H_4-(CH_2)_n$ -[carbohidrato],
 $H_2N-C_6H_4-(CH_2)_n-O$ -[carbohidrato] donde $n=0-10$, y
 $H_2N-C_6H_4-(CH_2)_m-CHOH-(CH_2)_n-O$ -[carbohidrato],
40 donde $m=0-10$ y $n=0-10$, incluyendo sus isómeros *p*-, *o*- y *m*-amino, aminometilo y grupo alquilo del anillo bencénico de los mismos.

Los materiales con funcionalidad amina se pueden preparar adicionalmente como derivados mixtos de polisacárido polimérico en los que grupos funcionales amino primario, aminoalquilo (uno a diez carbonos por grupo alquilo), aminohidroxi)alquilo (uno a diez carbonos por grupo alquilo y uno a diez grupos hidroxilo por grupo alquilo), aminobenceno y/o aminoalquilbenceno (uno a diez carbonos por grupo alquilo) están unidos covalentemente a matrices tales como copolímeros epíclorhidrina de celulosa o quitina y en donde los grupos espaciadores de hidrocarburos pueden incluir alqueno, así como grupos alquilo.

Los derivados poliméricos que no son polisacáridos en los que los grupos funcionales de amino primario, aminoalquilo (uno a diez carbonos por grupo alquilo), aminohidroxi)alquilo (uno a diez carbonos por grupo alquilo y uno a diez grupos hidroxilo por grupo alquilo), aminobenceno y/o aminoalquilbenceno (uno a diez carbonos por grupo alquilo) están unidos covalentemente a uno cualquiera de una amplia variedad de polímeros sintéticos que incluyen poliestireno, copolímero de estireno-divinilbenceno, poli(alcohol vinílico) y derivados reticulados de los mismos, y en donde los grupos espaciadores de hidrocarburos pueden incluir alqueno así como grupos alquilo.

También descrito en la presente memoria, el catalizador enzimático está inmovilizado en una matriz que comprende al menos un protector de amina. En otro ejemplo, el catalizador enzimático está inmovilizado en una matriz insoluble que comprende un número eficaz de grupos amino libres. En otro ejemplo, se proporciona la matriz insoluble y un número eficaz de grupos amino primarios (es decir, perlas de carragenina reticuladas con glutaraldehído y polietilenimina).

- 5 También descrito en la presente memoria, en aquellas aplicaciones en las que un aditivo protector de amina insoluble no necesita ser eliminado del producto final, o en las que dicho aditivo se elimina fácilmente de la mezcla de producto (por ejemplo, mediante intercambio iónico o destilación), entonces se pueden utilizar protectores de amina solubles. Ejemplos de protectores de amina solubles pueden incluir, pero sin limitación, formas solubles de los polímeros insolubles anteriormente mencionados (por ejemplo, monómeros de bajo peso molecular), compuestos aromáticos, cicloalifáticos o alifáticos de C₁-C₁₂ que comprenden uno o más grupos amino primarios, 1-amino-2-etanol, 1-amino-2-propanol, tampón Tris, etilendiamina y protectores de amina solubles similares que comprenden al menos un grupo amino primario.

En una realización preferida, el protector de amina es polietilenimina. También descrito en la presente memoria, el catalizador nitrilasa está inmovilizado en una matriz insoluble reticulada con glutaraldehído y polietilenimina.

- 15 Un experto en la técnica puede ajustar la cantidad de protector de amina en la mezcla de reacción para lograr el efecto deseado (es decir, una mejora en la actividad de nitrilasa; también referido en la presente memoria como una "cantidad eficaz"). En una realización, la concentración de protector de amina es al menos 0,01 mg/mL, preferiblemente al menos 0,1 mg/L, más preferiblemente al menos 1 mg/mL, y lo más preferiblemente al menos 5 mg/mL.

Hidrólisis de glicolonitrilo a ácido glicólico usando un catalizador nitrilasa

- 20 La conversión enzimática de glicolonitrilo en ácido glicólico (en forma de ácido y/o de la sal de amonio correspondiente) se realizó poniendo en contacto un catalizador enzimático, (que comprende un polipéptido que tiene actividad de nitrilasa) con una mezcla de reacción acuosa adecuada que comprende glicolonitrilo utilizando una serie adecuada de condiciones de reacción enzimática (intervalo de pH, temperaturas, concentraciones, etc.) descritas a continuación. En una realización, las células microbianas recombinantes enteras se pueden usar como un catalizador enzimático sin ningún pretratamiento. En otra realización, el catalizador celular microbiano puede ser añadido directamente a una mezcla de reacción, o mantenido de forma separada de la mezcla de reacción a granel usando cartuchos de membrana de fibra hueca o membranas de ultrafiltración. En una realización adicional, las células microbianas se pueden inmovilizar en una matriz polimérica (por ejemplo, carragenina o partículas de gel de poli(acrilamida) (PAG)) o en soporte sólido insoluble (por ejemplo, celita) para facilitar la recuperación y la reutilización del catalizador enzimático (patente de EE.UU. 6.870.038; incorporada en la presente memoria como referencia). En aún una realización adicional, las enzimas purificadas o parcialmente purificadas también se pueden aislar de las células enteras y usar directamente como un catalizador, o las enzimas se pueden inmovilizar en una matriz polimérica o en un soporte insoluble. Los métodos para la inmovilización de células o para las enzimas aisladas han sido documentados ampliamente y son bien conocidos por los expertos en la técnica (*Methods in Biotechnology, Vol. 1: Immobilization of enzymes and Cells*; Gordon F. Bickerstaff, editor; Humana Press, Totowa, NJ, USA; 1997). Se ha documentado previamente la inmovilización del catalizador nitrilasa de *A. facilis* 72W (patente de EE.UU. 6.870.038).

- 40 La concentración de catalizador enzimático en la mezcla de reacción acuosa depende de la actividad específica del catalizador enzimático y se elige para obtener la velocidad de reacción deseada. El peso de células húmedas de las células microbianas utilizadas como catalizador en reacciones de hidrólisis varía típicamente desde 0,001 gramos hasta 0,250 gramos de células húmedas por mL de volumen de reacción total, preferiblemente desde 0,002 gramos hasta 0,050 gramos de células húmedas por mL.

- 45 La temperatura de la reacción de hidrólisis de glicolonitrilo se elige para controlar tanto la velocidad de reacción como la estabilidad de la actividad del catalizador enzimático. La temperatura de la reacción puede variar desde justo por encima del punto de congelación de la mezcla de reacción (aproximadamente 0°C) hasta aproximadamente 65°C, con un intervalo preferido de temperatura de reacción desde aproximadamente 5°C hasta aproximadamente 35°C. La suspensión de catalizador celular microbiano se puede preparar poniendo en suspensión las células en agua destilada, o en una solución acuosa de un tampón que mantenga el pH inicial de la reacción entre aproximadamente 5,0 y aproximadamente 10,0, preferiblemente entre aproximadamente 5,5 y 50 aproximadamente 8,0, más preferiblemente entre aproximadamente 5,5 y aproximadamente 7,7, y lo más preferiblemente desde aproximadamente 6,0 hasta aproximadamente 7,7. A medida que avanza la reacción, el pH de la mezcla de reacción puede cambiar debido a la formación de una sal de amonio del ácido carboxílico a partir de la funcionalidad nitrilo correspondiente. La reacción se puede realizar para completar la conversión de glicolonitrilo sin control de pH, o se puede añadir un ácido o una base adecuados durante el transcurso de la 55 reacción para mantener el pH deseado.

- 60 Se encontró que el glicolonitrilo era completamente miscible con agua en todas las proporciones a 25°C. En los casos en los que se eligen las condiciones de reacción, de tal manera que la solubilidad del sustrato (es decir, un α -hidroxinitrilo) también dependa de la temperatura de la solución y/o la concentración de sal (tampón o el producto sal de amonio del ácido glicólico, también conocido como glicolato de amonio) en la fase acuosa, la mezcla de reacción puede estar compuesta inicialmente de dos fases: una fase acuosa que contiene el catalizador enzimático

y α -hidroxinitrilo disuelto, y una fase orgánica (el α -hidroxinitrilo sin disolver). A medida que progresa la reacción, el α -hidroxinitrilo se disuelve en la fase acuosa, y finalmente se obtiene una mezcla de producto en una sola fase. La reacción también puede ser realizada añadiendo el α -hidroxinitrilo a la mezcla de reacción a una velocidad aproximadamente igual a la velocidad de reacción de la hidrólisis enzimática, manteniendo de este modo una

5

mezcla de reacción acuosa en una sola fase, y evitando el problema potencial de inhibición por el sustrato de la enzima a altas concentraciones de material de partida.

El ácido glicólico puede estar en la mezcla de productos como una mezcla del ácido carboxílico protonizado y/o su sal de amonio correspondiente (dependiendo del pH de la mezcla de productos; el pKa del ácido glicólico es aproximadamente 3,83) y adicionalmente puede estar presente como una sal del ácido carboxílico con cualquier

10

tampón que adicionalmente pueda estar presente en la mezcla de productos. Típicamente, el ácido glicólico producido está principalmente en forma de sal de amonio (el pH de la reacción de hidrólisis del glicolonitrilo está típicamente entre aproximadamente 5,5 y aproximadamente 7,7). El producto de ácido glicólico se puede aislar de la mezcla de reacción como el ácido carboxílico protonizado, o como una sal del ácido carboxílico, según se desee.

La concentración final de ácido glicólico en la mezcla de productos en la conversión completa de glicolonitrilo puede variar desde 0,001 M hasta el límite de solubilidad del producto ácido glicólico. En una realización, la concentración de ácido glicólico variará desde aproximadamente 0,10 M hasta aproximadamente 5,0 M. En otra realización, la concentración de ácido glicólico variará desde aproximadamente 0,2 M hasta aproximadamente 3,0 M.

15

El ácido glicólico puede ser recuperado en forma del ácido o la sal correspondiente usando una variedad de técnicas que incluyen, aunque sin limitación, intercambio iónico, electrodiálisis, extracción con disolvente reactivo, polimerización, descomposición térmica, alcoholisis y sus combinaciones (véase la publicación de solicitud de patente de EE.UU. copendiente No. 2006-0247467)

20

Expresión microbiana

El catalizador nitrilasa se puede producir en células hospedantes heterólogas, preferiblemente en hospedantes microbianos. Serán particularmente útiles en la presente invención las células que son fácilmente adaptables a los

métodos de fermentación a gran escala. Dichos organismos son bien conocidos en la técnica del biotratamiento industrial, cuyos ejemplos pueden encontrarse en "Recombinant Microbes for Industrial and Agricultural Applications", Murooka *et al.*, eds., Marcel Dekker, Inc., Nueva York, Nueva York (1994), e incluyen bacterias fermentativas así como levaduras y hongos filamentosos. Las células hospedantes pueden incluir, pero sin limitación, *Comamonas sp.*, *Corynebacterium sp.*, *Brevibacterium sp.*, *Rhodococcus sp.*, *Azotobacter sp.*, *Citrobacter sp.*, *Enterobacter sp.*, *Clostridium sp.*, *Klebsiella sp.*, *Salmonella sp.*, *Lactobacillus sp.*, *Aspergillus sp.*, *Saccharomyces sp.*, *Yarrowia sp.*, *Zygosaccharomyces sp.*, *Pichia sp.*, *Kluyveromyces sp.*, *Candida sp.*, *Hansenula sp.*, *Dunaliella sp.*, *Debaryomyces sp.*, *Mucor sp.*, *Torulopsis sp.*, *Methylobacteria sp.*, *Bacillus sp.*, *Escherichia sp.*, *Pseudomonas sp.*, *Rhizobium sp.*, y *Streptomyces sp.* Se prefiere particularmente *E. coli*. Ejemplos de células hospedantes de *E. coli* aecuadas, en las que se puede expresar un gen nitrilasa mutante incluyen, pero sin limitación, las células hospedantes especificadas en la presente memoria y MG1655 (ATCC 47076), FM5 (ATCC 53911), W3110 (ATCC 27325), MC4100 (ATCC 35695), W1485 (ATCC 12435), y sus derivados. En otro aspecto, las cepas hospedantes de *E. coli* preferidas son MG1655 (ATCC 47076) o FM5 (ATCC 53911),

25

30

35

La expresión heteróloga de la nitrilasa *A. facilis* 72W ha sido documentada previamente (Chauhan *et al.*, *supra*; patente de EE.UU. 6.870.038; patente de EE.UU. 7.148.051; y patente de EE.UU. 7.198.927). Chauhan *et al.* detallan una cepa de *E. coli* (*E. coli* SS1001 (ATCC PTA-1177)) que expresa la nitrilasa *A. facilis* 72W activa (SEQ ID NO: 57). La secuencia codificadora de la nitrilasa expresada recombinantemente (*E. coli* SS1001) contenía dos cambios de secuencia menores en comparación con la secuencia de nitrilasa de 72W de tipo natural (SEQ ID NOs: 3 y 4). El codón de iniciación estaba cambiado de GTG a ATG para facilitar la expresión recombinante y se introdujo un artefacto durante la clonación que dio como resultado un cambio de un único aminoácido cerca del extremo C (Pro367 [CCA] → Ser [TCA]).

40

45

La expresión recombinante en un hospedante industrialmente adecuado tiene diversas ventajas. En primer lugar, la caja de herramientas genética para muchos de los hospedantes de producción utilizados habitualmente está muy desarrollada en comparación con las herramientas genéticas disponibles para muchos de los microorganismos de los que se obtienen los genes de interés. La expresión recombinante en estos hospedantes es normalmente más rentable que la expresión en el hospedante nativo. Por ejemplo, se ha demostrado que las células de *A. facilis* 72W se cultivan en glicerol, un sustrato de carbono relativamente caro, cuando se cultivan por fermentación, y no se han cultivado con éxito utilizando glucosa de bajo coste. Por el contrario, los transformantes de *E. coli* se pueden cultivar en glucosa a la misma densidad celular que las células de *A. facilis* 72W en aproximadamente la mitad de tiempo, reduciendo significativamente los costes de producción de biocatalizador (patente de EE.UU. 6.870.038).

50

Los sistemas de expresión microbiana y vectores de expresión contienen secuencias reguladoras que dirigen la expresión de nivel alto de proteínas extrañas son bien conocidos por los expertos en la técnica. Estos se podrían utilizar para construir genes quiméricos para la producción de productos genéticos. Estos genes quiméricos podrían entonces ser introducidos en microorganismos apropiados a través de la transformación para proporcionar un alto nivel de expresión de la nitrilasa deseada.

Los genes quiméricos serán efectivos alterando las propiedades de una célula hospedante. Por ejemplo, introduciendo al menos una copia genes quiméricos que codifican las nitrilasas presentes bajo el control de los promotores apropiados en una célula hospedante da a la célula hospedante una capacidad mejorada para convertir glicolonitrilo en ácido glicólico. Los genes quiméricos comprenderán secuencias reguladoras adecuadas útiles para la conducción de la expresión genética de las secuencias de nitrilasa mutante presentes. Las secuencias reguladoras adecuadas pueden incluir, pero sin limitación, promotores, secuencias situadas antes de las secuencias que se traducen, y sitios de unión al ribosoma. Se prefiere que estas secuencias se deriven del organismo hospedante; sin embargo, el experto en la materia reconocerá que también se pueden utilizar secuencias reguladoras heterólogas.

Los genes quiméricos se pueden introducir en un hospedante apropiado mediante su clonación en un vector de expresión adecuado. Los vectores o casetes útiles para la transformación de células hospedantes adecuadas son bien conocidos en la técnica. Típicamente, el vector o casete contiene secuencias que dirigen la transcripción y traducción del gen relevante, un marcador seleccionable, y las secuencias que permiten la replicación autónoma o la integración cromosómica. Los vectores adecuados comprenden una región en 5' de la secuencia codificadora que alberga los controles de la iniciación transcripcional y una región en 3' del fragmento de DNA que controla la terminación transcripcional. Esto es lo más preferido cuando ambas regiones de control proceden de genes homólogos de la célula hospedante, aunque no es necesario que dichas regiones de control procedan de los genes nativos de la especie específica elegida como hospedante de producción.

En una realización, las secuencias reguladoras incluirán un promotor. Los promotores pueden ser constitutivos o inducibles. Los promotores inducibles pueden ser generalmente sensibles a un estímulo específico (por ejemplo, IPTG o lactosa que induce el promotor *lac*). Los promotores inducibles pueden ser sensibles a una variedad de estímulos, incluyendo productos químicos, ciclo de cultivo, cambios en la temperatura, cambios en el pH y cambios en la osmolaridad, por nombrar sólo unos pocos.

Los promotores o regiones de control de la iniciación, que son útiles para dirigir la expresión de la nitrilasa en la célula hospedante deseada son numerosos y familiares para los expertos en la técnica, incluyendo, pero sin limitación, *CYC1*, *HIS3*, *GAL1*, *GAL10*, *ADH1*, *PGK*, *PHO5*, *GAPDH*, *ADC1*, *TRP1*, *URA3*, *LEU2*, *ENO*, *TPI* (útil para la expresión en *Saccharomyces*); *AOX1* (útil para la expresión en *Pichia*); y *lac*, *trp*, *IP_L*, *IP_R*, *T7*, *tac*, *P_{BAD}*, *npr*, y *trc* (particularmente útil para la expresión en *Escherichia coli*). Ejemplos adicionales de promotores particularmente adecuados para dirigir la expresión en *E. coli* incluyen, pero sin limitación, el promotor del operón de triptófano *P_{trp}* de *E. coli*, el promotor del operón de lactosa *Plac* de *E. coli*, un promotor *P_{tac}* de *E. coli*, un promotor derecho del fago lambda *P_R*, un promotor izquierdo de fago lambda *P_L*, un promotor *T7*, y un promotor del gen *GAP* a partir de *Pichia pastoris*, o es al menos un promotor aislado del grupo de microorganismos seleccionados del grupo que consiste en *Comamonas*, *Corynebacterium*, *Brevibacterium*, *Rhodococcus*, *Azotobacter*, *Citrobacter*, *Enterobacter*, *Clostridium*, *Klebsiella*, *Salmonella*, *Lactobacillus*, *Aspergillus*, *Saccharomyces*, *Pichia*, *Zygosaccharomyces*, *Kluyveromyces*, *Candida*, *Hansenula*, *Dunaliella*, *Debaryomyces*, *Mucor*, *Torulopsis*, *Methylobacteria*, *Bacillus*, *Escherichia*, *Pseudomonas*, *Rhizobium*, y *Streptomyces*.

Las regiones de control de la terminación también pueden proceder de diversos genes nativos para los hospedantes preferidos. Opcionalmente, puede no ser necesario un sitio de terminación; sin embargo, es lo más preferido si está incluido.

Adicionalmente, el material genético insertado puede incluir un sitio de unión al ribosoma (RBS). El sitio de unión al ribosoma puede ser de un gen de fago lambda *Cil* o se selecciona del grupo que consiste en sitios de unión al ribosoma a partir de un gen de *Comamonas*, *Corynebacterium*, *Brevibacterium*, *Rhodococcus*, *Azotobacter*, *Citrobacter*, *Enterobacter*, *Clostridium*, *Klebsiella*, *Salmonella*, *Lactobacillus*, *Aspergillus*, *Saccharomyces*, *Zygosaccharomyces*, *Pichia*, *Kluyveromyces*, *Candida*, *Hansenula*, *Dunaliella*, *Debaryomyces*, *Mucor*, *Torulopsis*, *Methylobacteria*, *Bacillus*, *Escherichia*, *Pseudomonas*, *Rhizobium*, y *Streptomyces*.

Opcionalmente, los productos genéticos pueden ser preferiblemente un producto secretado del hospedante transformado. La secreción de las proteínas deseadas en el medio de cultivo simplifica los procedimientos de purificación y reduce los costes. Las secuencias señal de secreción son frecuentemente útiles facilitando el transporte activo proteínas expresables a través de membranas celulares. Un hospedante transformado capaz de secreción puede ser creado incorporando en el hospedante una secuencia de DNA que codifica para una señal de secreción. Los métodos para la elección de secuencias señal apropiadas son bien conocidos en la técnica (véase por ejemplo la patente EP 546049; documento WO 93/24631). El DNA de la señal de secreción se puede localizar entre el DNA que controla la expresión y la secuencia codificadora instantánea o secuencia codificadora, y en el marco de lectura con la última.

Producción industrial del catalizador microbiano

Cuando se desea la producción comercial del catalizador nitrilasa, se puede utilizar una variedad de metodologías de cultivo. Se pueden realizar operaciones de fermentación de modos por lotes, por lotes alimentados o continuo, siendo dichos métodos bien conocidos en la técnica (Thomas D. Brock en *Biotechnology: A Textbook of Industrial Microbiology*, Second edition (1989) Sinauer Associates, Inc., Sunderland, MA, (1989); Deshpande, Mukund V., *Appl. Biochem. Biotechnol.* 36(3): 227-234 (1992)).

Un método de cultivo por lotes clásico es un sistema cerrado donde la composición de los medios se ajusta al comienzo del cultivo y no está sometida a alteraciones artificiales durante el procedimiento de cultivo. Por tanto, al comienzo del procedimiento de cultivo el medio se inocula con el organismo u organismos deseados y se deja que tenga lugar el crecimiento o la actividad metabólica no añadiendo nada al sistema. Típicamente, sin embargo, un cultivo "por lotes" es por lote con respecto a la adición de la fuente de carbono y con frecuencia se realizan intentos de controlar factores tales como el pH y la concentración de oxígeno. En los sistemas por lotes, las composiciones de metabolitos y biomasa del sistema cambian constantemente hasta el momento en el que se termina el cultivo. Dentro de los cultivos por lotes, las células se moderan desde de una fase logarítmica estática hasta una fase logarítmica de alto crecimiento y finalmente hasta una fase estacionaria donde disminuye o se detiene la velocidad de crecimiento. Si no se tratan, las células en la fase estacionaria finalmente morirán. En algunos sistemas las células en la fase logarítmica son con frecuencia responsables de la mayor parte de la producción del producto final o producto intermedio. En otros sistemas se puede obtener una producción en fase estacionaria o post-exponencial.

Una variación en el sistema por lotes estándar es el sistema por lotes alimentado. Los procedimientos de cultivo por lotes alimentados son también adecuados en la presente invención y comprenden un sistema por lotes típico con la excepción de que el sustrato se añade en incrementos a medida que progresa el cultivo. Los sistemas por lotes alimentados son útiles cuando la represión de catabolitos es apta para inhibir el metabolismo de las células y cuando es deseable tener cantidades limitadas de sustrato en los medios. La medición de la concentración real del sustrato en los sistemas por lotes alimentados es difícil y por lo tanto se estima basándose en los cambios de factores medibles, tales como pH, oxígeno disuelto y presión parcial de los gases residuales, tales como CO₂. Los métodos de cultivo por lotes y por lotes alimentados son usuales y bien conocidos en la técnica y se pueden encontrar ejemplos en Brock (*supra*) y Deshpande (*supra*).

La producción comercial de los catalizadores nitrilasa también se puede realizar por un cultivo continuo. Los cultivos continuos son un sistema abierto en el que se añade continuamente a un biorreactor un medio de cultivo definido y se elimina simultáneamente durante el procedimiento una cantidad igual de medio acondicionado. Los cultivos continuos mantienen generalmente las células a una densidad en fase líquida alta y constante donde las células se desarrollan principalmente en fase logarítmica. Alternativamente, el cultivo continuo se puede realizar con células inmovilizadas a las que se añaden continuamente carbono y nutrientes y se eliminan continuamente de la masa celular productos valiosos, subproductos o productos residuales. La inmovilización celular se puede realizar utilizando una amplia gama de soportes sólidos compuestos de materiales naturales y/o sintéticos.

El cultivo continuo o semicontinuo permite la modulación de un factor o de cualquier número de factores que afecten al crecimiento celular o a la concentración celular final. Por ejemplo, un método mantendrá un nutriente limitante, tal como la fuente de carbono o el nivel de nitrógeno, en una tasa fija y permitirá que se moderen todos los demás parámetros. En otros sistemas, se pueden alterar continuamente una serie de factores que afecten al crecimiento mientras se mantenga constante la concentración celular, medida por la turbidez del medio. Los sistemas continuos intentan mantener las condiciones de crecimiento en estado estacionario y, por tanto, la pérdida de células debido a los medios que se retiran debe equilibrarse con la tasa de crecimiento celular en el cultivo. Los métodos para modular los nutrientes y los factores de crecimiento en los procedimientos de cultivo continuos, así como las técnicas para maximizar la velocidad de formación de células, son bien conocidos en la técnica de la microbiología industrial y una variedad de métodos han sido detallados por Brock (*supra*).

Los medios de fermentación en la presente invención deben contener sustratos de carbono adecuados. Los sustratos adecuados pueden incluir, aunque sin limitación, monosacáridos, tales como glucosa y fructosa, disacáridos, tales como lactosa o sacarosa, polisacáridos, tales como almidón o celulosa, o sus mezclas, y mezclas no purificadas de materias primas renovables, tales como líquido filtrado de suero de queso, licor de maíz fermentado, melazas de remolacha azucarera y malta de cebada. Por tanto, se considera que la fuente de carbono utilizada en la presente invención puede abarcar una amplia variedad de sustratos que contengan carbono y solamente estará limitada por la elección del organismo.

Además, cuando una cantidad, concentración u otro valor o parámetro se da en forma de un intervalo, un intervalo preferido o una lista de valores superiores preferibles y valores inferiores preferibles, debe entenderse que están descritos específicamente todos los intervalos formados a partir de una pareja de cualquier límite superior del intervalo o valor preferido y cualquier límite inferior del intervalo o valor preferido, independientemente de si los intervalos se describen por separado. Cuando en la presente memoria se enumera un intervalo de valores numéricos, a menos que se indique lo contrario, se pretende que el intervalo incluya sus extremos, y todos los números enteros y fracciones dentro del intervalo. No se pretende que el alcance de la invención esté limitado a los valores específicos citados cuando se define un intervalo.

Métodos generales

Los siguientes ejemplos se proporcionan para demostrar las realizaciones preferidas de la invención. Los expertos en la técnica deben apreciar que las técnicas descritas en los ejemplos siguientes representan técnicas descubiertas por el inventor para llevar bien a la práctica la invención, y por tanto se puede considerar que constituyen modos preferidos para su práctica.

Los materiales y métodos adecuados para el mantenimiento y crecimiento de cultivos bacterianos son muy conocidos en la técnica. Técnicas adecuadas para uso en los siguientes ejemplos se pueden encontrar descritos en *Manual of Methods for General Bacteriology* (1994) (Phillipp Gerhardt, R.G.E. Murray, Ralph N. Costilow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg and G. Briggs Phillips, eds.), *American Society for Microbiology*, Washington, DC.) o por Thomas D. Brock, en *Biotechnology: A Textbook of Industrial Microbiology*, (1989) Second Edition, (Sinauer Associates, Inc., Sunderland, MA). Los métodos para inmovilizar catalizadores enzimáticos se pueden encontrar en Bickerstaff, G. F., *supra*.

Los procedimientos requeridos para la preparación de DNA genómico, la amplificación por PCR, las modificaciones del DNA por endo- y exo-nucleasas para generar extremos deseados para la clonación del DNA, las ligamientos y la transformación bacteriana, son bien conocidos en la técnica. Técnicas de DNA recombinante y de clonación molecular estándares utilizadas en la presente invención son bien conocidas en la técnica y han sido descritas por Maniatis, *supra*; y por T. J. Silhavy, M. L. Bennan, and L. W. Enquist, *Experiments with Gene Fusions*, (1984) *Cold Spring Harbor Laboratory Press*, Cold Spring, NY; y por Ausubel, F.M. et al., *Current Protocols in Molecular Biology*, (1994-1998) John Wiley & Sons, Inc., New York.

Todos los reactivos y materiales se obtuvieron de Aldrich Chemicals (Milwaukee, WI), DIFCO Laboratories (Detroit, MI), GIBCO/BRL (Gaithersburg, MD) o Sigma/Aldrich Chemical Company (St. Louis, MO) salvo indicación contraria.

Las abreviaturas en la memoria descriptiva corresponden a unidades de medida, técnicas, propiedades o compuestos de la siguiente manera: "s" significa segundo(s), "min" significa minuto(s), "h" significa hora(s), "d" significa densidad en g/mL, "µL" significa microlitros, "mL" significa mililitros, "L" significa litros, "mM" significa milimolar, "M" significa molar, "pm" significa peso molecular medio, "mmol" significa milimol(es), "p" significa peso, "% en peso" significa porcentaje en peso, "g" significa gramos, "µg" significa microgramos, "HPLC" significa cromatografía de líquidos de alta resolución, "DO" significa densidad óptica a la longitud de onda designada, "pcs" significa peso de células secas, "U" significa unidades de actividad de nitrilasa; "EDTA" significa ácido etilendiamintetraacético, "DTT" significa ditioneitol, "DI" significa destilada y desionizada, y "NCIMB" significa Colección Nacional de Alimentos Industriales y Bacterias Marinas, Bucksburn, Aberdeen, Scotland, GB. Una U de actividad de nitrilasa corresponde a la hidrólisis de 1 µmol de glicolonitrilo/min.

Metodología analítica

Análisis por HPLC

A menos que se indique lo contrario, se utilizó el siguiente método de HPLC. Las mezclas de productos de reacción se analizaron por el siguiente método de HPLC. Se añadieron partes alícuotas (0,01 mL) de la mezcla de reacción a 1,50 mL de agua y se analizó por HPLC (columna HPX 87H, 30 cm x 7,8 mm; fase móvil H₂SO₄ 0,01 N; caudal 1,0 mL/min a 50°C; volumen de inyección 10 µL; detector infrarrojo (IR), tiempo de análisis 20 min). El método se calibró para glicolonitrilo en una serie de concentraciones utilizando glicolonitrilo comercialmente disponible adquirido a Aldrich.

35 Ejemplo 1

Dependencia de la Actividad Específica de Nitrilasa de *E. coli* MG1655/pSW138-168V inmovilizada en Polietilenimina añadida.

En un vial de cristal de 20 ml con una barra de agitación magnética se añaden 1,0 ml de 25,0 (peso de células secas)/ml de una suspensión de *E. coli* MG1655/pSW138-168V (SEQ ID NO: 51) en tampón KH₂PO₄ 0,1 M (pH 7,5) y bien 1,0 ml de tampón KH₂PO₄ 0,3 M (pH 7,5) o 1,0 ml de una solución de polietilenimina 10,2 mg/ml (BASF LUPASOL® PS, pm 750.000; BASF Aktiengesellschaft, Ludwigshafen, Alemania) en tampón KH₂PO₄ 0,3 M (pH 7,5) a 25°C. Con agitación, se añadieron 2,0 ml de una solución acuosa que contiene glicolonitrilo (1,0 M) y formaldehído (8 mM, 1 mM, o 0,04 mM) en agua destilada, desionizada a 25°C, y la mezcla de reacción se mantuvo a 25°C con un baño termostático. A los 5, 10, 15 y 30 minutos, se retiró una parte alícuota de 100 µl de la mezcla de reacción y se mezcló con 100 µl de agua, 10 µl de HCl 6,0 N y 200 µL de n-propanol 0,25 M en agua (patrón externo para HPLC), la mezcla se centrifugó y el líquido sobrenadante resultante se analizó por HPLC para determinar la velocidad de reacción inicial y la actividad específica del catalizador (U/g pcs) (Tabla 2).

Tabla 2. Dependencia de la actividad específica de biocatalizador en concentración de formaldehído con y sin polietilenimina añadida (PEI)

glicolonitrilo (mM)	formaldehído (mM)	PEI (mg/ml)	actividad específica (U/g pcs)
500	4,0	0	471
500	4,0	5,1	1484
500	0,50	0	1344

glicolonitrilo (mM)	formaldehído (mM)	PEI (mg/ml)	actividad específica (U/g pcs)
500	0,50	5,1	1726
500	0,02	0	1742
500	0,02	5,1	1629

Ejemplo 2

Preparación de copoli(vinilalcohol/vinilamina)

Se colocó una solución de 0,2g de dodecilbencenosulfonato de sodio (Sigma-Aldrich Catálogo nº 289957) y 0,2 g de fosfato dihidrógeno de sodio en 80 ml de agua desionizada en un matraz de cuatro cuellos de 250 ml con condensador y entrada de nitrógeno, termómetro, embudo de goteo y agitador magnético. El matraz fue barrido con nitrógeno y se agitó en un baño de agua a 72°C hasta que la temperatura de la solución fue de 65°C; después se añadió 0,1 g de iniciador VAZO®-64 (2,2'-azobisisobutironitrilo; pm: 164,2; Sigma-Aldrich catálogo nº 441090). Se colocó una solución de 40g de acetato de vinilo (Sigma-Aldrich catálogo nº V1503, filtrada a través de alúmina básica para eliminar el inhibidor), 4 g de N-vinilformamida (Sigma-Aldrich catálogo nº 447331, usada sin purificación adicional), y 0,3 g de VAZO® 64 en el embudo de goteo y se añadieron 5 ml de esta solución monomérica al matraz. La mezcla se agitó 20 min y se añadieron otros 5 ml de monómero. Se añadieron partes alícuotas de 5 ml de monómero cada 20 minutos hasta que se añadieron 20 ml (en 1 hora); después la mezcla se agitó a 70°C durante 1 hora. Después de esto se añadió el resto del monómero a un caudal de 5 ml cada 20 min. La adición de monómero se completó a las 4 horas, y después la mezcla se agitó a 70°C durante 3 horas y se dejó enfriar a temperatura ambiente (aproximadamente 22°C). Las perlas de polímero resultantes se lavaron varias veces con agua caliente y después se dejaron reposar en agua durante 72 horas. A continuación, las perlas se lavaron con éter dietílico y se dejaron reposar en éter durante 15 minutos para extraer monómeros de vinilo, seguido de filtración y secado a vacío con una capa de nitrógeno, después se secaron toda la noche a 70°C en un horno de vacío a 350 mm Hg con un barrido de nitrógeno para dar 39,0 g de copolímero de poli(vinilacetato-vinilformamida).

El copolímero de poli(vinilacetato-vinilformamida) (25,1 g) se agitó a reflujo con 200 ml de metanol que contenía 6 ml de HCl concentrado y 5 ml de agua durante 5 horas, obteniéndose una solución transparente que se dejó reposar a TA durante la noche. La solución se convirtió en un gel de caucho; éste se mezcló con varios cambios de acetona para precipitar la sal de HCL de polímero hidrolizada, que se secó bajo nitrógeno en horno de vacío a 75°C (rendimiento: 15 g). El análisis por ¹H NMR (DMSO-d₆) indicó aproximadamente 20% de acetatos dejados sin hidrolizar, por lo que el producto se agitó a reflujo con una mezcla de 150 ml de metanol y 5 ml de HCl conc. toda la noche para completar la hidrólisis. El metanol se separó por filtración y el producto de caucho, insoluble en metanol, se agitó con 100 ml de agua DI a 90°C; esta solución se filtró primero a través de una pantalla para eliminar una pequeña cantidad de gel y después a través de un prefiltro de celulosa Millipore (Millipore Corp., Bedford, MA) bajo presión. La solución filtrada se basificó a pH 9,0 con 10% de NaOH y después se dializó durante la noche frente a agua DI en un tubo de membrana de diálisis cortado de peso molecular MEMBRA-CEL® 3.5K (Viskase Co., Willowbrooke, IL). La liofilización produjo 8,7 g de copoli(alcohol vinílico/vinilamina).

Ejemplo 3

Dependencia de la Actividad Específica de Nitrilasa de *E. coli* MG1655/pSW138-168V inmovilizada en Protectores de Amina añadidos.

En un vial de cristal de 20 mL se añadieron 20 mg con agitación magnética a 400 mg de protector de amina (véase la Tabla 2), seguido de 1,825 ml de agua desionizada y 1,0 mL de tampón KH₂PO₄ 0,3 M (pH 7,5). El pH de la mezcla resultante se verificó y se reajustó a pH 7,5 con HCl 6 N. Cuando el pH de la mezcla resultante se estableció a pH 7,5, se añadió 1,0 mL de aproximadamente 25,0 mg (peso de células secas)/mL de suspensión de *E. coli* MG1655/pSW138-168V (SEQ ID NO: 51) en tampón KH₂PO₄ 0,1 M (pH 7,5), y después la reacción se inició por la adición de 0,175 mL (0,187 g) de glicolonitrilo acuoso (61% en peso de GLN en agua, 2,00 mmol de glicolonitrilo en solución que contiene también bien 0,80% en mol o 0,10% en mol de formaldehído con relación al glicolonitrilo), y la mezcla de reacción se mantuvo a 25°C con un baño de temperatura. A los 5, 10, 15, y 30 minutos, se retiró una parte alícuota de 100 µl de la mezcla de reacción y se mezcló con 100 µl de agua, 10 µl de HCl 6.0 N y 200 µl de n-propanol 0.25 M en agua (patrón externo para HPLC), la mezcla se centrifugó y el líquido sobrenadante resultante se analizó por HPLC para determinar la velocidad de reacción inicial y la actividad específica del catalizador (U/g pcs) (Tabla 3). Cada serie de experimentos enumerados en la Tabla 3 se realizaron con una suspensión celular recién preparada (suspensiones celulares A - G), en la que las diferencias en la actividad específica de las células en la reacción de control para cada suspensión celular se deben a diferencias en las concentraciones de peso de células secas de las suspensiones celulares individuales.

50

ES 2 564 176 T3

Tabla 3. Dependencia de la actividad específica de biocatalizador en concentración de formaldehído con y sin protector de amina añadido.

Suspensión celular	Glicolonitrilo (mM)	Formaldehído (mM)	Protector de amina	Protector de amina (mg/mL)	Actividad específica (U/g pcs)
A	500	4.0	PEI/sílice (Aldrich 24675-1)	100	491
A	500	4.0	PEI/celulosa (Sigma P6883)	100	1212
A	500	4.0	ninguno (control)	0	385
B	500	4.0	3-aminopropil-gel de sílice (Sigma A1409)	100	799
B	500	4.0	ninguno (control)	0	481
C	500	4.0	Vidrio de poro controlado de alquilamina de cadena larga (Sigma L8638)	100	1136
C	500	4.0	ninguno (control)	0	527
D	500	0.5	Vidrio de poro controlado de alquilamina de cadena larga (Sigma L8638)	100	1891
D	500	0.5	ninguno (control)	0	1325
E	500	4.0	Polialilamina (Aldrich 47914-4)	5	598
E	500	4.0	Polialilamina (Aldrich 47914-4)	10	648
E	500	4.0	PVOH/polivinilamina (Ejemplo 2)	70	1541
E	500	4.0	Poli-D-lisina (Sigma P0296)	11	502
E	500	4.0	ninguno (control)	0	464
F	500	0.5	PVOH/polivinilamina (Ejemplo 2)	70	1786
F	500	0.5	Poly-D-lysine (Sigma P0296)	11	1388
F	500	0.5	Ninguno (control)	0	1220
G	500	4.0	DIAION® WA20 (Mitsubishi Chemical)	100	1128
G	500	4.0	DIAION® WA21J (Mitsubishi Chemical)	100	1073
G	500	4.0	ninguno (control)	0	508
G	500	0.5	DIAION® WA20 (Mitsubishi Chemical)	100	1335
G	500	0.5	DIAION® WA21J (Mitsubishi Chemical)	100	1339
G	500	0.5	ninguno (control)	0	1182

Ejemplo 4

Preparación de perlas de carragenina/*E. coli* MG1655/pNM18-168V reticuladas con GA/PEI

- 5 Con agitación rápida, se añadieron lentamente 12 g de carragenina (FMC GP911, FMC Corp., Philadelphia, PA) a 228 g de agua destilada desionizada, a 50°C, la mezcla resultante se calentó hasta 80°C hasta que la carragenina

se disolvió completamente, y la solución resultante se enfrió con agitación hasta 52°C. En un vaso separado equipado con una barra de agitación, se añadieron 83,2 g de células congeladas de *E. coli* MG1655/pNM18-168V (SEQ ID NO: 51) (25,2% en pcs) a 84,8 g de Na₂HPO₄ 0,35 M (pH 7,3) a aproximadamente 25°C y se mezcló hasta que las células se pusieron en suspensión, y a continuación se añadió una solución de desoxirribonucleasa I (10 µL de 12.500 U/mL de DNasa (Sigma-Aldrich)/100 mL de suspensión celular). La suspensión celular se filtró consecutivamente por un elemento filtrante NUPRO® TF (Swagelok Company, Solon, OH) de 230 micrómetros y 140 micrómetros y se calentó con agitación hasta 50°C. A la solución de carragenina a 52°C, se añadieron con agitación, 160,0 g de una suspensión celular de *E. coli* MG1655/pNM18-168V a 50°C, y la suspensión de células/carragenina resultante se bombeó a través de una aguja de calibre 20 calentada eléctricamente a 47°C y se añadió gota a gota con agitación a KHCO₃ 0,25 M (pH = 7,3) a aproximadamente 37-38°C; el caudal a través de la aguja se ajustó a 5-8 mL/min. Se dejaron endurecer las perlas resultantes en este mismo tampón durante 1 hora a temperatura ambiente con agitación, y se conservaron en bicarbonato de potasio 0,25 M (pH 7,3).

La reticulación química de una porción de las perlas inmovilizadas de células/carragenina se realizó por adición de 1,0 g de glutaraldehído (GA) al 25% en agua (Sigma M 752-07) a 20 g de perlas en suspensión en 48 mL de bicarbonato de potasio 0,25 M (pH 7,3) y agitación durante 1 hora a temperatura ambiente. A la suspensión de perlas se añadieron luego 4,0 g de polietilenimina al 12,5% en peso (PEI, BASF LUPASOL® PS) en agua, y la suspensión de perlas se agitó durante 18 horas más a temperatura ambiente. Las perlas reticuladas con GA/PEI se recuperaron de la suspensión, se agitaron dos veces durante 15 minutos en 48 mL de bicarbonato de potasio 0,25 M (pH 7,3), y a continuación se almacenaron en bicarbonato de amonio 1,0 M (pH 7,3) a 5°C. Antes de su uso las perlas se lavaron dos veces durante 15 minutos con 180 mL de glicolato de amonio 0,1 M (pH 7,0) a temperatura ambiente.

Ejemplo 5

Dependencia de la actividad específica del biocatalizador en la reticulación con glutaraldehído/polietilenamina de transformantes de *E. coli* MG1655/psW138-168V inmovilizados en carragenina que expresa la nitrilasa de *A. facilis* 72W

En un procedimiento típico, un recipiente de reacción con camisa, de 50 mL, equipado con agitación superior y control de temperatura se cargó con 4,0 g de perlas de *E. Coli* MG1655/pSW138-168V/carragenina reticuladas con GA/PEI (preparadas utilizando el procedimiento como se describe en el Ejemplo 4) que contenía 5% (pcs) de transformante que expresa el mutante F168V de nitrilasa de *A. facilis* 72W (SEQ ID NO: 51). Al recipiente se añadieron después 10,85 ml de agua destilada y 3,0 ml de glicolato de amonio acuoso (4,0 M, pH 7,0). La mezcla se agitó a 25°C mientras se añadieron simultáneamente 1,75 ml de glicolonitrilo (GLN) al 60,8% en peso en agua (1,876 g, GLN 20,0 mmol, formaldehído 0,160 mmol; estabilizado con ácido glicólico al 0,7% en peso) y se añadieron 0,40 mL de hidróxido de amonio acuoso (NH₃ al 1,875% en peso) (pH final 7,5). Cuatro muestras de reacción de 0,50 mL se retiraron en tiempos predeterminados después de la adición de GLN y se analizaron mediante HPLC para determinar la velocidad de reacción inicial y la actividad específica del catalizador (µmol de ácido glicólico/min/g en pcs de biocatalizador).

Tabla 4. Dependencia de la actividad específica del biocatalizador inmovilizado en reticulación con glutaraldehído/polietilenimina

Glicolonitrilo (mM)	Formaldehído (mM)	GA/PEI reticulado	Actividad específica (U/g de perlas)	Actividad específica (U/g pcs)
1000	8,0	no	36,2	723
1000	8,0	sí	68,2	1364
1000	1,0	no	58,4	1169
1000	1,0	sí	99,7	1994
1000	0,04	no	68,8	1375
1000	0,04	sí	90,1	1801

Ejemplo 6

Comparación de la actividad específica del biocatalizador para transformantes de *E. coli* MG1655/pSW138-168V inmovilizada en carragenina y reticulado con glutaraldehído/polietilenamina y no reticulado que expresa la nitrilasa de *A. facilis* 72W en reacciones por lotes consecutivas con reciclaje del catalizador

Un recipiente de reacción con camisa, de 50 mL, equipado con agitación superior y control de temperatura se cargó con 8 g de perlas de *E. coli* MG1655/pSW138-168V/carragenina reticuladas con GA/PEI (preparadas utilizando el procedimiento como se ha descrito en el Ejemplo 4) que contenía 5% (pcs) de transformante que expresa el mutante F168V de nitrilasa de *A. facilis* 72W (SEQ ID NO: 51). Al recipiente se añadieron después 14,78 ml de agua destilada y 6,0 ml de glicolato de amonio acuoso (4,0 M, pH 7,0) y el recipiente de reacción se purgó con

ES 2 564 176 T3

5 nitrógeno. La mezcla se agitó a 25°C mientras se utilizaban bombas de jeringa programables para añadir simultáneamente 1,07 mL de glicolonitrilo (GLN) al 60% en peso en agua (GLN 12,0 mmol, formaldehído 0,084 mmol; Fluka (redestilada, estabilizada con ácido glicólico al 0,5 % en peso, disponible en Sigma-Aldrich) y 0,150 mL de hidróxido de amonio acuoso (1,875% en peso); cada 2 h se añadió simultáneamente un volumen equivalente de soluciones de GLN e hidróxido de amonio para mantener la concentración de GLN a ≤ 400 mM y el pH en un intervalo de 6,5-7,5. Se retiraron cuatro muestras de reacción de 0,050 mL en momentos predeterminados después de la primera adición de GLN y se analizaron por HPLC para determinar la velocidad de reacción inicial y la actividad específica del catalizador (μmol de ácido glicólico/min/g en pcs de biocatalizador). Al completarse la reacción, se produjo una conversión del 100% de GLN en ácido glicólico (en forma de sal de amonio) con un rendimiento $> 99\%$ y la concentración de glicolato de amonio producida por el GLN añadido fue aproximadamente 2,5 M (glicolato de amonio total 3,1 M cuando se incluye tampón de glicolato de amonio inicial en un volumen de reacción final de aproximadamente 38,0 mL).

10 Al final de la primera reacción, la mezcla acuosa de producto se decantó del catalizador (bajo nitrógeno), dejando aproximadamente 10,3 g de una mezcla de catalizador celular inmovilizado (8,0 g) y permaneciendo la mezcla del producto (aproximadamente 2,3 g). A continuación se añadió al recipiente de reacción 20,78 mL de agua destilada desionizada, y se realizó una segunda reacción a 25°C por adición de partes alícuotas de GLN acuoso e hidróxido de amonio como se acaba de describir. La actividad específica del biocatalizador recuperado en reacciones por lotes consecutivas con reciclaje del catalizador se recogen en la Tabla 5. Para comparar el efecto de la reticulación con glutaraldehído/polietilenimina de células inmovilizadas en carragenina en la actividad específica, se realizó una segunda serie de reacciones de reciclaje, como se describe anteriormente, excepto que las perlas de *E. Coli* MG1655/pSW138-168V/carragenina no se reticularon químicamente con glutaraldehído y polietilenimina (Tabla 5).

Tabla 5. Dependencia de la actividad específica del biocatalizador en la reticulación con glutaraldehído/PEI de catalizador microbiano inmovilizado en carragenina en reacciones por lotes consecutivas con reciclaje del catalizador

Actividad específica del biocatalizador (U/g pcs)		
Número de reacción	Reticulación con GA/PEI	Sin reticulación con GA/PEI
1	1312	215
2	871	385
3	897	270
4	816	
5	898	296
6	697	
7	663	
8	792	
9	732	
10	760	393
11	539	
12	495	373
13	396	
14	560	
15	324	443
16	426	
17	500	366
18	722	394
19	489	
20		
21		283
22	444	

Actividad específica del biocatalizador (U/g pcs)		
Número de reacción	Reticulación con GA/PEI	Sin reticulación con GA/PEI
23		325
24		
25	501	
26	508	285
27	487	224
28	349	
29	371	
30	345	224

Ejemplo 7

Dependencia de la Actividad Específica de Nitrilasa de *Rhodococcus sp.*(ATCC 39484™) inmovilizado en Polietilenimina añadida

5 Un cultivo liofilizado de *Rhodococcus sp.* ATCC 39484 (SEQ ID NO: 25) se puso en suspensión en 5 ml de caldo nutriente (Difco; catálogo nº 0003-01-6) en un tubo de ensayo cónico de 50-mL y reactivado después de la incubación a 30°C durante 3 días a 60 rpm, seguido de conservación a -80°C después de la adición de glicerol (10% p/p). El cultivo reactivado (1 mL) se inoculó en un matraz de agitación sin deflectores de 125-mL que contenía 50 mL de caldo Miller's Luria-Bertani (Mediatech, Inc.; 46-050-CM; Manassas, VA) y se cultivaron durante 16 h a 30°C y 200 rpm en un DO de 9,0 a 600 nm, después se añadió glicerol (10% p/p) y el inóculo se conservó a -80°C. El inóculo (2 mL) se transfirió a un matraz de agitación sin deflectores de 1-L que contenía 200 mL de un medio que consistía en 1 g de polipeptona (Becton-Dickinson; catálogo nº 11910; Franklin Lakes, NJ), 0,6 g de extracto de malta (Difco; 0186-02-4), 0,6 g de extracto de levadura (Difco; 0127-17), 2 g de glicerol, y 0,2 mL de isovaleronitrilo (T. Nagasawa, M. Kobayashi, H. Yamada, Archives of Microbiology, (1988) 150:89-94) a pH 7,0, seguido de crecimiento a 30°C con agitación a 200 rpm. A las 50 horas y 70 horas, el cultivo se alimentó con 0,1% (p/p) y 0,2% (p/p) de isovaleronitrilo, respectivamente. La pasta celular se recogió a las 96 horas por centrifugación a 13.000 x g (5°C) y se conservó a -80°C.

20 En un vial de cristal de 20 mL con agitación magnética se añadieron 1,0 mL de una suspensión de 73,0 mg (peso de células secas)/mL de *Rhodococcus sp.* (ATCC 39484™) en un tampón KH₂PO₄ 0,1 M (pH 7,5), y bien 1,0 ml de tampón KH₂PO₄ 0,3 M (pH 7,5) o 1,0 mL de 10,2 mg/mL de una solución de polietilenimina (BASF LUPASOL® PS, 750.000 mw) en un tampón KH₂PO₄ 0,3 M (pH 7,5) a 25°C. Se añadieron con agitación, 2,0 mL de una solución acuosa que contenía glicolonitrilo (1,0 M) y formaldehído (8 mM o 0,04 mM) en agua destilada, desionizada a 25°C, y la mezcla de reacción se mantuvo a 25°C. A los 5, 10, 15, y 30 minutos, se retiró una parte alícuota de 100 µl de la mezcla de reacción y se mezcló con 100 µl de agua, 10 µl de HCl 6.0 N y 200 µl de n-propanol 0,25 M en agua (patrón externo para HPLC), la mezcla se centrifugó y el líquido sobrenadante resultante se analizó por HPLC para determinar la velocidad de reacción inicial y la actividad específica del catalizador (U/g pcs) (Tabla 6).

Tabla 6. Dependencia de la actividad específica de la nitrilasa de *Rhodococcus sp.* (ATCC 39484™) en concentración de formaldehído con y sin polietilenimina (PEI) añadida.

Glicolonitrilo (mM)	Formaldehído (mM)	PEI (mg/mL)	Actividad específica (U/g pcs)
500	4.0	0	598
500	4.0	5.0	780

Ejemplo 8

30 Dependencia de la Actividad Específica de Nitrilasa de *Rhodococcus rhodochrous* (NCIMB 11216) no inmovilizado en Polietilenimina añadida.

35 El *Rhodococcus rhodochrous* (NCIMB 11216; SEQ ID NO: 24) se reactivó después de poner en suspensión el cultivo preservado en 100 µL de solución salina normal, seguido de veteado de la suspensión celular salina normal sobre placas que contenían agar nutriente (Difco catálogo nº 0001-01-8) complementado con extracto de levadura (2 g/L; Difco catálogo nº 0127-17) y cloruro de sodio (5 g/L). La secuencia de la nitrilasa del *R. rhodochrous* NCIMB 11216 ya ha sido descrita (Publicación de Solicitud de Patente de EE.UU. 2003/0157672 y Número de acceso en GENBANK® No. CAC88237; SEQ ID NO: 24) Las placas se incubaron a 30°C durante 48 horas y las colonias se inocularon en medio de E2 basal acuoso complementado con ácido adípico (0,8% en peso; fuente de carbono) y propionitrilo (0,2% en peso; fuente de nitrógeno) y se incubaron durante 48 horas a 30°C con agitación a 200 rpm.

ES 2 564 176 T3

5 El medio de E2 basal consiste en: KH_2PO_4 , 1.4 g/L; NaH_2PO_4 , 0,69 g/L; citrato de sodio, 0,1 g/L; $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 0,025 g/L; KCl, 0.5 g/L; NaCl, 1,0 g/L, $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0,5 g/L; $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 0,05 g/L; $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$, 0,01 g/L; $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$, 0,001 g/L; ZnCl_2 , 0,0005 g/L; $\text{NaMoO}_4 \cdot 2\text{H}_2\text{O}$, 0,0025 g/L; $\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$, 0,01 g/L; $\text{CuSO}_4 \cdot 2\text{H}_2\text{O}$, 0,005 g/L; Biotina, 0,0002 g/L; Ácido Fólico, 0,0002 g/L; Piridoxina•HCl, 0,001 g/L; Riboflavina, 0,0005 g/L; Tiamina•HCl, 0,00005 g/L; Ácido Nicotínico, 0,0005 g/L; Ácido Pantoténico, 0,0005 g/L; Vitamina B₁₂, 0,00001 g/L, Ácido p-Aminobenzoico, 0,0005 g/L. Después de 48 horas de cultivo, 1,59 g de pasta celular húmeda se recogió por centrifugación a 13.000 x g, y la pasta celular resultante se conservó a -80°C.

10 En un vial de cristal de 20 mL con agitación magnética se añadieron 1,0 mL de una suspensión de 79,0 mg (peso de células secas)/mL de *Rhodococcus sp.* (NCIMB 11216) en un tampón KH_2PO_4 0,1 M (pH 7,5), y bien 1,0 ml de tampón KH_2PO_4 0,3 M (pH 7,5) o 1,0 mL de 10,2 mg/mL de una solución de polietilenimina (BASF LUPASOL® PS, 750.000 mw) en un tampón KH_2PO_4 0,3 M (pH 7,5) a 25°C. Se añadieron con agitación, 2,0 mL de una solución acuosa que contenía glicolonitrilo (1,0 M) y formaldehído (8 mM o 0,04 mM) en agua destilada, desionizada a 25°C, y la mezcla de reacción se mantuvo a 25°C. A los 5, 10, 15, y 30 minutos, se retiró una parte alícuota de 100 µl de la mezcla de reacción y se mezcló con 100 µl de agua, 10 µl de HCl 6.0 N y 200 µl de n-propanol 0,25 M en agua (patrón externo para HPLC), la mezcla se centrifugó y el líquido sobrenadante resultante se analizó por HPLC para determinar la velocidad de reacción inicial y la actividad específica del catalizador (U/g pcs) (Tabla 7).

Glicolonitrilo (mM)	Formaldehído (mM)	PEI (mg/mL)	Actividad específica (U/g pcs)
500	4.0	0	37.9
500	4.0	5.0	45.4

Listado de Secuencias

- <110> E.I. duPont de Nemours and Company, Inc.
- <120> Secuestración de formaldehído para estabilizar la actividad específica de nitrilasa cuando se convierte glicolonitrilo en ácido glicólico
- 5 <130> CL3584PCT
- <160> 57
- <170> PatentIn versión 3.4
- <210> 1
- <211> 12
- 10 <212> PRT
- <213> Secuencia artificial
- <220>
- <223> Resto distintivo catalítico
- <220>
- 15 <221> CARACTERÍSTICA_NUEVA
- <222> (2)..(2)
- <223> Xaa = Ala o Gly
- <220>
- 20 <221> CARACTERÍSTICA_NUEVA
- <222> (3)..(3)
- <223> Xaa = Leu, Val, o Ala
- <220>
- 25 <221> CARACTERÍSTICA_NUEVA
- <222> (4)..(3)
- <223> Xaa = Ala, Asn, Ile, Cys, Val, o Gln;
- <220>
- <221> CARACTERÍSTICA_NUEVA
- <222> (4)..(4)
- <223> Xaa puede ser cualquier aminoácido de origen natural
- 30 <220>
- <221> CARACTERÍSTICA_NUEVA
- <222> (8)..(8)
- <223> Xaa = His o Asn;
- <220>
- 35 <221> CARACTERÍSTICA_NUEVA
- <222> (9) .. (9)
- <223> Xaa = Leu, Tyr, Phe, Ala, Met, Lys, Val, Thr, o Arg
- <220>
- 40 <221> CARACTERÍSTICA_NUEVA
- <222> (10)..(10)
- <223> Xaa = Asn, Gln, Met, Leu, o Ser;
- <220>
- 45 <221> CARACTERÍSTICA_NUEVA
- <222> (11)..(11)
- <223> Xaa = Pro o Thr
- <220>
- <221> CARACTERÍSTICA_NUEVA
- <222> (12)..(12)
- <223> Xaa = Leu o Val
- 50 <400> 1
- Gly Xaa Xaa Xaa Cys Trp Glu Xaa Xaa Xaa Xaa Xaa
- 1 5 10

- <210> 2
 <211> 12
 <212> PRT
 <213> Secuencia artificial
- 5 <220>
 <223> resto nº 2 de secuencia catalítico
- <220>
 <221> CARACTERÍSTICA_NUEVA
 <222> (2)..(2)
 10 <223> Xaa = Ala o Gly
- <220>
 <221> CARACTERÍSTICA_NUEVA
 <222> (4)..(4)
 <223> Xaa = Ala o Gly
- 15 <220>
 <221> CARACTERÍSTICA_NUEVA
 <222> (9) .. (9)
 <223> Xaa = Leu, Tyr, Phe, Ala, Met, Lys, Val, Thr, o Arg
- 20 <220>
 <221> CARACTERÍSTICA_NUEVA
 <222> (10)..(10)
 <223> Xaa = Asn, Gln, Met, o Ser
- <400> 2
 Gly Xaa Leu Xaa Cys Trp Glu His Xaa Xaa Pro Leu
 1 5 10
- 25 <210> 3
 <211> 1110
 <212> DNA
 <213> Acidovorax facilis 72W
- 30 <220>
 <221> CDS
 <222> (1)..(1110)
- <220>
 <221> CARACTERÍSTICA_NUEVA
 <222> (1)..(3)
 35 <223> codón de inicio TTG cambiado a ATG para facilitar la expresión recombinante
- <400> 3

ES 2 564 176 T3

atg gtt tcg tat aac agc aag ttc ctc gcg gca acc gtt cag gca gag	48
Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu	
1 5 10 15	
ccg gta tgg ctc gac gca gac gca acg atc gac aag tcg atc ggc atc	96
Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile	
20 25 30	
atc gaa gaa gct gcc caa aag ggc gcg agt ctg atc gct ttc ccg gaa	144
Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu	
35 40 45	
gta ttc att ccg ggc tac ccc tat tgg gcg tgg ctc ggc gac gtg aag	192
Val Phe Ile Pro Gly Tyr Pro Tyr Trp Ala Trp Leu Gly Asp Val Lys	
50 55 60	
tac agc cta agc ttt act tca cgc tat cac gag aat tcg ttg gag cta	240
Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu	
65 70 75 80	
ggt gac gac cgt atg cgt cgc ctc cag ctg gcc gcg cgc cgc aac aaa	288
Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys	
85 90 95	
atc gca ctc gtc atg ggc tat tcg gag cgg gaa gcc gga tcg cgc tat	336
Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr	
100 105 110	
ctg agc cag gtg ttc atc gac gag cgt ggc gag atc gtt gcc aat cgg	384
Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg	
115 120 125	
cgc aag ctg aag ccc aca cac gtt gag cgt acg atc tac ggc gaa ggc	432
Arg Lys Leu Lys Pro Thr His Val Glu Arg Thr Ile Tyr Gly Glu Gly	
130 135 140	
aac gga acc gat ttc ctc acg cac gac ttc gcg ttc gga cgc gtc ggt	480
Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly	
145 150 155 160	
gga ttg aac tgc tgg gaa cat ttc caa ccg ctc agc aag ttc atg atg	528
Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met	
165 170 175	
tac agc ctc ggt gag cag gtc cac gtt gca tcg tgg ccg gcg atg tcc	576
Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser	
180 185 190	
cct ctt cag ccg gat gtt ttc caa ctg agc atc gaa gcc aac gcg acg	624
Pro Leu Gln Pro Asp Val Phe Gln Leu Ser Ile Glu Ala Asn Ala Thr	
195 200 205	
gtc acc cgc tcg tac gca atc gaa ggc caa acc ttt gtg ctt tgc tcg	672
Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser	
210 215 220	

ES 2 564 176 T3

acg cag gtg atc gga cct agc gcg atc gaa acg ttc tgc ctc aac gac 720
 Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
 225 230 235 240

gaa cag cgc gca ctg ttg ccg caa gga tgt ggc tgg gcg cgc att tac 768
 Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
 245 250 255

ggc ccg gat gga agc gag ctt gcg aag cct ctg gcg gaa gat gct gag 816
 Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
 260 265 270

ggg atc ttg tac gca gag atc gat ctg gag cag att ctg ctg gcg aag 864
 Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
 275 280 285

gct gga gcc gat ccg gtc ggg cac tat tcg ccg cct gac gtg ctg tcg 912
 Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
 290 295 300

gtc cag ttc gac ccg cgc aat cat acg cca gtt cat cgc atc ggc att 960
 Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
 305 310 315 320

gac ggt cgc ttg gat gtg aat acc cgc agt cgc gtg gag aat ttc cga 1008
 Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
 325 330 335

ctg cga caa gcg gct gag cag gag cgt cag gca tcc aag cgg ctc gga 1056
 Leu Arg Gln Ala Ala Glu Gln Glu Arg Gln Ala Ser Lys Arg Leu Gly
 340 345 350

acg aaa ctc ttt gaa caa tcc ctt ctg gct gaa gaa ccg gtc cca gca 1104
 Thr Lys Leu Phe Glu Gln Ser Leu Leu Ala Glu Glu Pro Val Pro Ala
 355 360 365

aag tag 1110
 Lys

<210> 4

<211> 369

<212> PRT

5 <213> Acidovorax facilis 72W

<400> 4

Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
 1 5 10 15

Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
 35 40 45

Val Phe Ile Pro Gly Tyr Pro Tyr Trp Ala Trp Leu Gly Asp Val Lys

ES 2 564 176 T3

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

ES 2 564 176 T3

Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
305 310 315 320

Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
325 330 335

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

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

Lys

<210> 5

<211> 356

<212> PRT

5 <213> *Alcaligenes faecalis* LM3

<400> 5

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

Pro Asn Tyr Asp Leu Ala Thr Gly Val Asp Lys Thr Ile Glu Leu Ala
20 25 30

Arg Gln Ala Arg Asp Glu Gly Cys Asp Leu Ile Val Phe Gly Glu Thr
35 40 45

Trp Leu Pro Gly Tyr Pro Phe His Val Trp Leu Gly Ala Pro Ala Trp
50 55 60

Ser Leu Lys Tyr Ser Ala Arg Tyr Tyr Ala Asn Ser Leu Ser Leu Asp
65 70 75 80

Ser Ala Glu Phe Gln Arg Ile Ala Gln Ala Ala Arg Thr Leu Gly Ile
85 90 95

Phe Ile Ala Leu Gly Tyr Ser Glu Arg Ser Gly Gly Ser Leu Tyr Leu
100 105 110

Gly Gln Cys Leu Ile Asp Asp Lys Gly Gln Met Leu Trp Ser Arg Arg
115 120 125

ES 2 564 176 T3

Lys Leu Lys Pro Thr His Val Glu Arg Thr Val Phe Gly Glu Gly Tyr
 130 135 140

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

Leu Cys Cys Trp Glu His Leu Ser Pro Leu Ser Lys Tyr Ala Leu Tyr
 165 170 175

Ser Gln His Glu Ala Ile His Ile Ala Ala Trp Pro Ser Phe Ser Leu
 180 185 190

Tyr Ser Glu Gln Ala His Ala Leu Ser Ala Lys Val Asn Met Ala Ala
 195 200 205

Ser Gln Ile Tyr Ser Val Glu Gly Gln Cys Phe Thr Ile Ala Ala Ser
 210 215 220

Ser Val Val Thr Gln Glu Thr Leu Asp Met Leu Glu Val Gly Glu His
 225 230 235 240

Asn Ala Ser Leu Leu Lys Val Gly Gly Gly Ser Ser Met Ile Phe Ala
 245 250 255

Pro Asp Gly Arg Thr Leu Ala Pro Tyr Leu Pro His Asp Ala Glu Gly
 260 265 270

Leu Ile Ile Ala Asp Leu Asn Met Glu Glu Ile Ala Phe Ala Lys Ala
 275 280 285

Ile Asn Asp Pro Val Gly His Tyr Ser Lys Pro Glu Ala Thr Arg Leu
 290 295 300

Val Leu Asp Leu Gly His Arg Glu Pro Met Thr Arg Val His Ser Lys
 305 310 315 320

Ser Val Ile Gln Glu Glu Ala Pro Glu Pro His Val Gln Ser Thr Ala
 325 330 335

Ala Pro Val Ala Val Ser Gln Thr Gln Asp Ser Asp Thr Leu Leu Val
 340 345 350

Gln Glu Pro Ser
 355

<210> 6
 <211> 366

ES 2 564 176 T3

<212> PRT

<213> Rhodococcus rhodochrous J1

<400> 6

```

Met Val Glu Tyr Thr Asn Thr Phe Lys Val Ala Ala Val Gln Ala Gln
 1          5          10          15

Pro Val Trp Phe Asp Ala Ala Lys Thr Val Asp Lys Thr Val Ser Ile
 20          25          30

Ile Ala Glu Ala Ala Arg Asn Gly Cys Glu Leu Val Ala Phe Pro Glu
 35          40          45

Val Phe Ile Pro Gly Tyr Pro Tyr His Ile Trp Val Asp Ser Pro Leu
 50          55          60

Ala Gly Met Ala Lys Phe Ala Val Arg Tyr His Glu Asn Ser Leu Thr
 65          70          75          80

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

Asn Ile Ala Val Val Val Gly Ile Ser Glu Arg Asp Gly Gly Ser Leu
 100         105         110

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

Arg Arg Lys Leu Lys Pro Thr His Val Glu Arg Ser Val Tyr Gly Glu
 130         135         140

Gly Asn Gly Ser Asp Ile Ser Val Tyr Asp Met Pro Phe Ala Arg Leu
 145         150         155         160

Gly Ala Leu Asn Cys Trp Glu His Phe Gln Thr Leu Thr Lys Tyr Ala
 165         170         175

Met Tyr Ser Met His Glu Gln Val His Val Ala Ser Trp Pro Gly Met
 180         185         190

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

Thr Ala Thr Arg Met Tyr Ala Leu Glu Gly Gln Thr Phe Val Val Cys
 210         215         220

```

ES 2 564 176 T3

Thr Thr Gln Val Val Thr Pro Glu Ala His Glu Phe Phe Cys Asp Asn
225 230 235 240

Asp Glu Gln Arg Lys Leu Ile Gly Arg Gly Gly Gly Phe Ala Arg Ile
245 250 255

Ile Gly Pro Asp Gly Arg Asp Leu Ala Thr Pro Leu Ala Glu Asp Glu
260 265 270

Glu Gly Ile Leu Tyr Ala Asp Ile Asp Leu Ser Ala Ile Thr Leu Ala
275 280 285

Lys Gln Ala Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu
290 295 300

Ser Leu Asn Phe Asn Gln Arg His Thr Thr Pro Val Asn Thr Ala Ile
305 310 315 320

Ser Thr Ile His Ala Thr His Thr Leu Val Pro Gln Ser Gly Ala Leu
325 330 335

Asp Gly Val Arg Glu Leu Asn Gly Ala Asp Glu Gln Arg Ala Leu Pro
340 345 350

Ser Thr His Ser Asp Glu Thr Asp Arg Ala Thr Ala Ser Ile
355 360 365

<210> 7

<211> 383

<212> PRT

5 <213> Rhodococcus rhodochrous K22

<400> 7

Met Ser Ser Asn Pro Glu Leu Lys Tyr Thr Gly Lys Val Lys Val Ala
1 5 10 15

Thr Val Gln Ala Glu Pro Val Ile Leu Asp Ala Asp Ala Thr Ile Asp
20 25 30

Lys Ala Ile Gly Phe Ile Glu Glu Ala Ala Lys Asn Gly Ala Glu Phe
35 40 45

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

Ile Gly Asp Val Lys Trp Ala Val Ser Asp Phe Ile Pro Lys Tyr His

ES 2 564 176 T3

Pro Val Gln Tyr Ile Thr Ala Asp Gly Arg Thr Ser Leu Asn Ser Asn
 325 330 335

Ser Arg Val Glu Asn Tyr Arg Leu His Gln Leu Ala Asp Ile Glu Lys
 340 345 350

Tyr Glu Asn Ala Glu Ala Ala Thr Leu Pro Leu Asp Ala Pro Ala Pro
 355 360 365

Ala Pro Ala Pro Glu Gln Lys Ser Gly Arg Ala Lys Ala Glu Ala
 370 375 380

<210> 8

<211> 381

<212> PRT

5 <213> Norcardia sp. C-14-1

<400> 8

Met Lys Val Ala Thr Val Gln Ala Glu Pro Val Ile Leu Asp Ala Asp
 1 5 10 15

Ala Thr Ile Asp Lys Ala Ile Gly Tyr Ile Glu Glu Ala Ser Lys Asn
 20 25 30

Gly Ala Glu Phe Ile Ala Phe Pro Glu Val Trp Ile Pro Gly Tyr Pro
 35 40 45

Tyr Trp Ala Trp Ile Gly Asp Val Lys Trp Ala Val Ser Glu Phe Ile
 50 55 60

Pro Lys Tyr His Glu Asn Ser Leu Thr Leu Gly Asp Asp Arg Met Arg
 65 70 75 80

Arg Leu Gln Leu Ala Ala Arg Gln His Asn Ile Ala Met Val Val Gly
 85 90 95

Tyr Ser Glu Lys Asp Gly Ala Ser Arg Tyr Leu Ser Gln Val Phe Ile
 100 105 110

Asp Gln Asn Gly Asp Ile Val Ala Asn Arg Arg Lys Leu Lys Pro Thr
 115 120 125

His Val Glu Arg Thr Ile Tyr Gly Glu Gly Asn Gly Thr Asp Phe Leu
 130 135 140

ES 2 564 176 T3

Thr His Asp Phe Gly Phe Gly Arg Val Gly Gly Leu Asn Cys Trp Glu
145 150 155 160

His Phe Gln Pro Leu Ser Lys Tyr Met Met Tyr Ser Leu Asn Glu Gln
165 170 175

Ile His Val Ala Ser Trp Pro Ala Met Phe Ala Leu Thr Pro Asp Val
180 185 190

His Gln Leu Ser Val Glu Ala Asn Asp Thr Val Thr Arg Ser Tyr Ala
195 200 205

Ile Glu Gly Gln Thr Phe Val Leu Ala Ala Thr His Val Ile Gly Lys
210 215 220

Ala Thr Gln Asp Leu Phe Ala Gly Asp Asp Glu Ala Lys Arg Ala Leu
225 230 235 240

Leu Pro Leu Gly Gln Gly Trp Ala Arg Ile Tyr Gly Pro Asp Gly Lys
245 250 255

Ser Leu Ala Glu Pro Leu Ala Glu Asn Ala Glu Gly Leu Leu Tyr Ala
260 265 270

Glu Leu Asp Leu Glu Gln Ile Ile Val Ala Lys Ala Ala Ala Asp Pro
275 280 285

Ala Gly His Tyr Ser Arg Pro Asp Val Leu Ser Leu Lys Val Asp Thr
290 295 300

Arg Asn His Thr Pro Val Gln Tyr Val Thr Glu Asp Gly Gly Ser Ser
305 310 315 320

Leu Asn Ser Asn Ser Arg Val Glu Asn Tyr Arg Leu Arg Gln Leu Ala
325 330 335

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

Val Thr Thr Pro Glu Lys Gln Ser Gly Asp Val Asn Ala Asn Gly Asn
355 360 365

Ala Lys Val Asn Thr Asn Pro Ser Ala Lys Ala Lys Ala
370 375 380

<210> 9
<211> 310

ES 2 564 176 T3

<212> PRT

<213> Bordetella bronchiseptica RB50

<400> 9

Met Thr Thr His Arg Ile Ala Val Ile Gln Asp Gly Pro Val Pro Gly
 1 5 10 15

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

Lys Ala Gln Gly Ala Arg Leu Ala Leu Phe Pro Glu Ala Phe Val Gly
 35 40 45

Gly Tyr Pro Lys Gly Ala Asp Phe His Ile Phe Leu Gly Gly Arg Thr
 50 55 60

Pro Gln Gly Arg Ala Gln Tyr Gln Arg Tyr Ala Glu Thr Ala Ile Ala
 65 70 75 80

Val Pro Gly Pro Val Thr Glu Arg Ile Gly Gln Ile Ala Ala Glu Gln
 85 90 95

Asp Met Phe Ile Val Val Gly Val Ile Glu Arg Asp Gly Gly Thr Leu
 100 105 110

Tyr Cys Thr Ile Leu Phe Phe Ser Pro Glu Gly Glu Leu Leu Gly Lys
 115 120 125

His Arg Lys Leu Met Pro Thr Ala Leu Glu Arg Leu Leu Trp Gly Tyr
 130 135 140

Gly Asp Gly Ser Thr Phe Pro Val Tyr Asp Thr Pro Leu Gly Lys Leu
 145 150 155 160

Gly Ala Val Val Cys Trp Glu Asn Tyr Met Pro Leu Leu Arg Met Ala
 165 170 175

Met Tyr Gly Lys Gln Ile Gln Ile Tyr Cys Ala Pro Thr Ala Asp Asp
 180 185 190

Lys Pro Thr Trp Val Ser Thr Met Gln His Val Ala Leu Glu Gly Arg
 195 200 205

Cys Phe Val Leu Ser Ala Cys Gln His Leu Arg Gly Lys Asp Phe Pro
 210 215 220

ES 2 564 176 T3

Pro Glu Phe His Asn Ala Leu Asp Val Gln Pro Asp Thr Val Leu Met
225 230 235 240

Arg Gly Gly Ser Cys Ile Val Asp Pro Met Gly Gln Leu Leu Ala Gly
245 250 255

Pro Val Tyr Asp Glu Asp Ala Ile Leu Val Ala Asp Ile Asp Leu Asp
260 265 270

Ala Val Thr Arg Gly Lys Met Asp Phe Asp Val Val Gly His Tyr Ala
275 280 285

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

Val Thr Thr Leu Lys Pro
305 310

<210> 10

<211> 339

<212> PRT

5 <213> *Arabidopsis thaliana*

<400> 10

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

Ile Val Arg Ala Thr Ile Val Gln Ala Ser Thr Val Tyr Asn Asp Thr
20 25 30

Pro Ala Thr Leu Glu Lys Ala Asn Lys Phe Ile Val Glu Ala Ala Ser
35 40 45

Lys Gly Ser Glu Leu Val Val Phe Pro Glu Ala Phe Ile Gly Gly Tyr
50 55 60

Pro Arg Gly Phe Arg Phe Gly Leu Gly Val Gly Val His Asn Glu Glu
65 70 75 80

Gly Arg Asp Glu Phe Arg Lys Tyr His Ala Ser Ala Ile Lys Val Pro
85 90 95

Gly Pro Glu Val Glu Lys Leu Ala Glu Leu Ala Gly Lys Asn Asn Val
100 105 110

Tyr Leu Val Met Gly Ala Ile Glu Lys Asp Gly Tyr Thr Leu Tyr Cys

ES 2 564 176 T3

115 120 125

Thr Ala Leu Phe Phe Ser Pro Gln Gly Gln Phe Leu Gly Lys His Arg
 130 135 140

Lys Leu Met Pro Thr Ser Leu Glu Arg Cys Ile Trp Gly Gln Gly Asp
 145 150 155 160

Gly Ser Thr Ile Pro Val Tyr Asp Thr Pro Ile Gly Lys Leu Gly Ala
 165 170 175

Ala Ile Cys Trp Glu Asn Arg Met Pro Leu Tyr Arg Thr Ala Leu Tyr
 180 185 190

Ala Lys Gly Ile Glu Leu Tyr Cys Ala Pro Thr Ala Asp Gly Ser Lys
 195 200 205

Glu Trp Gln Ser Ser Met Leu His Ile Ala Ile Glu Gly Gly Cys Phe
 210 215 220

Val Leu Ser Ala Cys Gln Phe Cys Leu Arg Lys Asp Phe Pro Asp His
 225 230 235 240

Pro Asp Tyr Leu Phe Thr Asp Trp Tyr Asp Asp Lys Glu Pro Asp Ser
 245 250 255

Ile Val Ser Gln Gly Gly Ser Val Ile Ile Ser Pro Leu Gly Gln Val
 260 265 270

Leu Ala Gly Pro Asn Phe Glu Ser Glu Gly Leu Ile Thr Ala Asp Leu
 275 280 285

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

His Tyr Ser Arg Pro Asp Val Leu His Leu Thr Val Asn Glu His Pro
 305 310 315 320

Lys Lys Pro Val Thr Phe Ile Ser Lys Val Glu Lys Ala Glu Asp Asp
 325 330 335

Ser Asn Lys

<210> 11

<211> 334

<212> PRT

5 <213> Synechococcus elongatus PCC 7942

ES 2 564 176 T3

<400> 11

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

Phe Ser Leu Glu Gly Ser Val Ala Arg Val Leu Ala Ala Met Ala Glu
20 25 30

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

Pro Tyr Tyr Pro Tyr Phe Ser Phe Val Glu Pro Pro Val Leu Met Gly
50 55 60

Arg Ser His Leu Lys Leu Tyr Glu Gln Ala Phe Thr Met Thr Gly Pro
65 70 75 80

Glu Leu Gln Gln Ile Ala Arg Ala Ala Arg Gln His Arg Leu Phe Val
85 90 95

Leu Leu Gly Val Asn Glu Arg Asp Gly Gly Ser Leu Tyr Asn Thr Gln
100 105 110

Leu Leu Ile Ser Asp Gln Gly Asp Leu Leu Leu Lys Arg Arg Lys Ile
115 120 125

Thr Pro Thr Tyr His Glu Arg Met Val Trp Gly Gln Gly Gly Gly Ala
130 135 140

Gly Leu Thr Val Val Glu Thr Val Leu Gly Lys Val Gly Ala Leu Ala
145 150 155 160

Cys Trp Glu His Tyr Asn Pro Leu Ala Arg Phe Ser Leu Met Thr Gln
165 170 175

Gly Glu Glu Ile His Cys Ala Gln Phe Pro Gly Ser Leu Val Gly Pro
180 185 190

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

Ala Gly Cys Phe Val Leu Ser Ser Thr Ala Trp Leu Asp Pro Ala Asp
210 215 220

ES 2 564 176 T3

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

Gly Cys His Thr Ala Ile Ile Ser Pro Glu Gly Arg Tyr Leu Ala Gly
245 250 255

Pro Leu Pro Glu Gly Glu Gly Leu Ala Ile Ala Glu Leu Asp Lys Ser
260 265 270

Leu Ile Thr Lys Arg Lys Arg Met Met Asp Ser Val Gly His Tyr Ser
275 280 285

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

Val Gln Ala Ile Gly Ser Ala Ala Ala Leu Pro Glu Leu Pro Asn Leu
305 310 315 320

Glu Ala Ala Pro Ala Glu Thr Ala Glu Asp Tyr Leu His Ala
325 330

<210> 12

<211> 334

<212> PRT

5 <213> *Synechococcus elongatus* PCC 6301

<400> 12

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

Phe Ser Leu Glu Gly Ser Val Ala Arg Val Leu Ala Ala Met Ala Glu
20 25 30

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

Pro Tyr Tyr Pro Tyr Phe Ser Phe Val Glu Pro Pro Val Leu Met Gly
50 55 60

Arg Ser His Leu Lys Leu Tyr Glu Gln Ala Phe Thr Met Thr Gly Pro
65 70 75 80

Glu Leu Gln Gln Ile Ala Arg Ala Ala Arg Gln His Arg Leu Phe Val
85 90 95

Leu Leu Gly Val Asn Glu Arg Asp Gly Gly Ser Leu Tyr Asn Thr Gln
100 105 110

ES 2 564 176 T3

Leu Leu Ile Ser Asp Gln Gly Asp Leu Leu Leu Lys Arg Arg Lys Ile
 115 120 125

Thr Pro Thr Tyr His Glu Arg Met Val Trp Gly Gln Gly Gly Gly Ala
 130 135 140

Gly Leu Thr Val Val Glu Thr Val Leu Gly Lys Val Gly Ala Leu Ala
 145 150 155 160

Cys Trp Glu His Tyr Asn Pro Leu Ala Arg Phe Ser Leu Met Thr Gln
 165 170 175

Gly Glu Glu Ile His Cys Ala Gln Phe Pro Gly Ser Leu Val Gly Pro
 180 185 190

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

Ala Gly Cys Phe Val Leu Ser Ser Thr Ala Trp Leu Asp Pro Ala Asp
 210 215 220

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

Gly Cys His Thr Ala Ile Ile Ser Pro Glu Gly Arg Tyr Leu Ala Gly
 245 250 255

Pro Leu Pro Glu Gly Glu Gly Leu Ala Ile Ala Glu Leu Asp Lys Ser
 260 265 270

Leu Ile Thr Lys Arg Lys Arg Met Met Asp Ser Val Gly His Tyr Ser
 275 280 285

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

Val Gln Ala Ile Gly Ser Ala Ala Ala Leu Pro Glu Leu Pro Asn Leu
 305 310 315 320

Glu Ala Ala Pro Ala Glu Thr Ala Glu Asp Tyr Leu His Ala
 325 330

<210> 13

<211> 346

<212> PRT

5 <213> Synechocystis sp. PCC 6803

<400> 13

ES 2 564 176 T3

Met Leu Gly Lys Ile Met Leu Asn Tyr Thr Lys Asn Ile Arg Ala Ala
1 5 10 15

Ala Ala Gln Ile Ser Pro Val Leu Phe Ser Gln Gln Gly Thr Met Glu
20 25 30

Lys Val Leu Asp Ala Ile Ala Asn Ala Ala Lys Lys Gly Val Glu Leu
35 40 45

Ile Val Phe Pro Glu Thr Phe Val Pro Tyr Tyr Pro Tyr Phe Ser Phe
50 55 60

Val Glu Pro Pro Val Leu Met Gly Lys Ser His Leu Lys Leu Tyr Gln
65 70 75 80

Glu Ala Val Thr Val Pro Gly Lys Val Thr Gln Ala Ile Ala Gln Ala
85 90 95

Ala Lys Thr His Gly Met Val Val Val Leu Gly Val Asn Glu Arg Glu
100 105 110

Glu Gly Ser Leu Tyr Asn Thr Gln Leu Ile Phe Asp Ala Asp Gly Ala
115 120 125

Leu Val Leu Lys Arg Arg Lys Ile Thr Pro Thr Tyr His Glu Arg Met
130 135 140

Val Trp Gly Gln Gly Asp Gly Ala Gly Leu Arg Thr Val Asp Thr Thr
145 150 155 160

Val Gly Arg Leu Gly Ala Leu Ala Cys Trp Glu His Tyr Asn Pro Leu
165 170 175

Ala Arg Tyr Ala Leu Met Ala Gln His Glu Gln Ile His Cys Gly Gln
180 185 190

Phe Pro Gly Ser Met Val Gly Gln Ile Phe Ala Asp Gln Met Glu Val
195 200 205

Thr Met Arg His His Ala Leu Glu Ser Gly Cys Phe Val Ile Asn Ala
210 215 220

Thr Gly Trp Leu Thr Ala Glu Gln Lys Leu Gln Ile Thr Thr Asp Glu

ES 2 564 176 T3

Gly Ala Ser Leu Val Leu Gly Val Ile Glu Arg Ser Gly Asn Thr Leu
 100 105 110

Tyr Cys Thr Val Leu Phe Phe Glu Pro Glu Gly Gly Leu Val Ala Lys
 115 120 125

His Arg Lys Leu Met Pro Thr Gly Thr Glu Arg Leu Ile Trp Gly Lys
 130 135 140

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

Gly Ala Ala Val Cys Trp Glu Asn Tyr Met Pro Leu Leu Arg Thr Ala
 165 170 175

Met Tyr Ala Lys Gly Val Gln Leu Trp Cys Ala Pro Thr Val Asp Glu
 180 185 190

Arg Glu Leu Trp Gln Val Ser Met Arg His Val Ala Ala Glu Gly Arg
 195 200 205

Cys Phe Val Ile Ser Ala Cys Gln Val Gln Asp Ser Pro Ala Ala Leu
 210 215 220

Gly Met Glu Val Ala Asn Trp Pro Ala Glu Arg Pro Leu Ile Asn Gly
 225 230 235 240

Gly Ser Leu Ile Val Gly Pro Leu Gly Asp Val Leu Ala Gly Pro Leu
 245 250 255

Leu Gly Ala Arg Gly Leu Val Cys Ala Glu Val Asp Thr Asp Glu Leu
 260 265 270

Val Arg Ala Arg Tyr Asp Phe Asp Val Val Gly His Tyr Ala Arg Pro
 275 280 285

Asp Val Phe Glu Leu Ser Val Asp Glu Arg Pro Arg Pro Gly Val Arg
 290 295 300

Phe Ile Gly
 305

<210> 15

<211> 329

<212> PRT

5 <213> *Zymomonas mobilis* subsp. *mobilis* ZM4

<400> 15

ES 2 564 176 T3

Met Ser Cys His Arg Val Ala Val Ile Gln Ala Gly Thr Ser Leu Phe
 1 5 10 15

Asp Thr Glu Lys Thr Leu Asp Arg Met Glu Ala Leu Cys Arg Gln Ala
 20 25 30

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

Gly Tyr Pro Lys Gly Leu Asp Phe Gly Ala Arg Met Gly Thr Arg Thr
 50 55 60

Glu Ala Gly Arg Glu Asp Phe Leu Arg Tyr Trp Lys Ala Ala Ile Asp
 65 70 75 80

Val Pro Gly Lys Glu Thr Ala Arg Ile Gly Ser Phe Ala Ala Lys Met
 85 90 95

Lys Ala Tyr Leu Val Val Gly Val Ile Glu Arg Ser Glu Ala Thr Leu
 100 105 110

Tyr Cys Thr Ala Leu Phe Phe Ala Pro Asp Gly Thr Leu Ile Gly Lys
 115 120 125

His Arg Lys Leu Met Pro Thr Ala Thr Glu Arg Leu Val Trp Gly Gln
 130 135 140

Gly Asp Gly Ser Thr Ile Glu Ile Leu Asp Thr Ala Val Gly Lys Leu
 145 150 155 160

Gly Ala Ala Ile Cys Trp Glu Asn Tyr Met Pro Val Leu Arg Gln Val
 165 170 175

Met Tyr Ala Gly Gly Val Asn Ile Trp Cys Ala Pro Thr Val Asp Gln
 180 185 190

Arg Glu Ile Trp Gln Val Ser Met Arg His Ile Ala Tyr Glu Gly Arg
 195 200 205

Leu Phe Val Leu Ser Ala Cys Gln Tyr Met Thr Arg Ala Asp Ala Pro
 210 215 220

Ala Asp Tyr Asp Cys Ile Gln Gly Asn Asp Pro Glu Thr Glu Leu Ile
 225 230 235 240

ES 2 564 176 T3

Ala Gly Gly Ser Val Ile Ile Asp Pro Met Gly Asn Ile Leu Ala Gly
245 250 255

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

Asp Thr Ile Lys Ala Arg Tyr Asp Leu Asp Val Ser Gly His Tyr Gly
275 280 285

Arg Pro Asp Ile Phe Glu Ile Lys Val Asp Arg Gln Ser His Gln Val
290 295 300

Ile Thr Asp Gln Phe Ser Arg Asp Gln Ala Thr Glu Lys Lys Pro Val
305 310 315 320

Ser Asp Ser Glu Ile Ser Gln Leu Asp
325

<210> 16

<211> 339

<212> PRT

5 <213> Bacillus sp. OxB-1

<400> 16

Met Ser Asn Tyr Pro Lys Tyr Arg Val Ala Ala Val Gln Ala Ser Pro
1 5 10 15

Val Leu Leu Asp Leu Asp Ala Thr Ile Asp Lys Thr Cys Arg Leu Val
20 25 30

Asp Glu Ala Ala Ala Asn Gly Ala Lys Val Ile Ala Phe Pro Glu Ala
35 40 45

Phe Ile Pro Gly Tyr Pro Trp Trp Ile Trp Leu Gly Asn Ala Asp Tyr
50 55 60

Gly Met Lys Tyr Tyr Ile Gln Leu Tyr Lys Asn Ser Val Glu Ile Pro
65 70 75 80

Ser Leu Ala Val Gln Lys Leu Ser Ser Ala Gly Thr Asn Lys Val Tyr
85 90 95

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

ES 2 564 176 T3

Gln Leu Trp Phe Asp Pro Asn Gly Asp Leu Ile Gly Lys His Arg Lys
 115 120 125

Leu Lys Ala Thr Asn Ala Glu Lys Thr Ile Trp Gly Asp Gly Asp Gly
 130 135 140

Ser Met Met Pro Val Phe Glu Thr Glu Phe Gly Asn Leu Gly Gly Leu
 145 150 155 160

Gln Cys Trp Glu His Phe Leu Pro Leu Asn Val Ala Ala Met Ala Ser
 165 170 175

Met Asn Glu Gln Val His Val Ala Ser Trp Pro Ile Gly Met Pro Gln
 180 185 190

Glu Gly His Leu Phe Gly Pro Glu Gln Cys Val Thr Ala Thr Lys Tyr
 195 200 205

Tyr Ala Ile Ser Asn Gln Val Phe Cys Leu Leu Ser Ser Gln Ile Trp
 210 215 220

Thr Glu Glu Gln Arg Asp Lys Ile Cys Glu Thr Glu Glu Gln Arg Asn
 225 230 235 240

Phe Met Lys Val Gly His Gly Phe Ser Lys Ile Ile Ala Pro Asn Gly
 245 250 255

Met Glu Ile Gly Asn Lys Leu Ala His Asp Glu Glu Gly Ile Thr Tyr
 260 265 270

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

Ser Ala Gly His Tyr Ser Thr Pro Gly Phe Leu Ser Leu Ser Phe Asp
 290 295 300

Arg Thr Glu Lys Lys Pro Ile Lys His Ile Gly Glu Ser Ala Gln Glu
 305 310 315 320

Thr Val Thr Tyr Glu Glu Ile Gln Tyr Gly Asn Lys Ala Asn Val Lys
 325 330 335

Val His Ser

<210> 17
 <211> 354
 <212> PRT

ES 2 564 176 T3

<213> Comamonas testosterone

<400> 17

Met Lys Asn Tyr Pro Thr Val Lys Val Ala Ala Val Gln Ala Ala Pro
 1 5 10 15

Val Phe Met Asn Leu Glu Ala Thr Val Asp Lys Thr Cys Lys Leu Ile
 20 25 30

Ala Glu Ala Ala Ser Met Gly Ala Lys Val Ile Gly Phe Pro Glu Ala
 35 40 45

Phe Ile Pro Gly Tyr Pro Tyr Trp Ile Trp Thr Ser Asn Met Asp Phe
 50 55 60

Thr Gly Met Met Trp Ala Val Leu Phe Lys Asn Ala Ile Glu Ile Pro
 65 70 75 80

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

Tyr Val Cys Val Ser Val Ser Glu Lys Asp Asn Ala Ser Leu Tyr Leu
 100 105 110

Thr Gln Leu Trp Phe Asp Pro Asn Gly Asn Leu Ile Gly Lys His Arg
 115 120 125

Lys Phe Lys Pro Thr Ser Ser Glu Arg Ala Val Trp Gly Asp Gly Asp
 130 135 140

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

Leu Gln Cys Trp Glu His Ala Leu Pro Leu Asn Ile Ala Ala Met Gly
 165 170 175

Ser Leu Asn Glu Gln Val His Val Ala Ser Trp Pro Ala Phe Val Pro
 180 185 190

Lys Gly Ala Val Ser Ser Arg Val Ser Ser Ser Val Cys Ala Ser Thr
 195 200 205

Asn Ala Met His Gln Ile Ile Ser Gln Phe Tyr Ala Ile Ser Asn Gln
 210 215 220

ES 2 564 176 T3

Val Tyr Val Ile Met Ser Thr Asn Leu Val Gly Gln Asp Met Ile Asp
225 230 235 240

Met Ile Gly Lys Asp Glu Phe Ser Lys Asn Phe Leu Pro Leu Gly Ser
245 250 255

Gly Asn Thr Ala Ile Ile Ser Asn Thr Gly Glu Ile Leu Ala Ser Ile
260 265 270

Pro Gln Asp Ala Glu Gly Ile Ala Val Ala Glu Ile Asp Leu Asn Gln
275 280 285

Ile Ile Tyr Gly Lys Trp Leu Leu Asp Pro Ala Gly His Tyr Ser Thr
290 295 300

Pro Gly Phe Leu Ser Leu Thr Phe Asp Gln Ser Glu His Val Pro Val
305 310 315 320

Lys Lys Ile Gly Glu Gln Thr Asn His Phe Ile Ser Tyr Glu Asp Leu
325 330 335

His Glu Asp Lys Met Asp Met Leu Thr Ile Pro Pro Arg Arg Val Ala
340 345 350

Thr Ala

<210> 18

<211> 332

<212> PRT

5 <213> Synechococcus sp. CC9605

<400> 18

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

Ser Leu Asp Gly Ser Leu Gln Lys Val Leu Asp Ala Met Ala Glu Ala
20 25 30

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

Tyr Tyr Pro Tyr Phe Ser Phe Val Glu Pro Pro Val Leu Met Gly Arg
50 55 60

Ser His Leu Ala Leu Tyr Glu Gln Ala Val Val Val Pro Gly Pro Val

ES 2 564 176 T3

Asp Ala Leu Ser Ser Met Glu Ala Leu Asn His Val
 325 330

<210> 19

<211> 306

<212> PRT

5 <213> Pseudomonas fluorescens Pf-5

<400> 19

Met Pro Lys Ser Val Val Ala Ala Leu Gln Ile Gly Ala Leu Pro Glu
 1 5 10 15

Gly Lys Ala Ala Thr Leu Glu Gln Ile Leu Ser Tyr Glu Ala Ala Ile
 20 25 30

Ile Glu Ala Gly Ala Gln Leu Val Val Met Pro Glu Ala Leu Leu Gly
 35 40 45

Gly Tyr Pro Lys Gly Glu Gly Phe Gly Thr Gln Leu Gly Tyr Arg Leu
 50 55 60

Pro Glu Gly Arg Glu Ala Phe Ala Arg Tyr Phe Ala Asn Ala Ile Glu
 65 70 75 80

Val Pro Gly Val Glu Thr Asp Ala Leu Ala Ala Leu Ser Ala Arg Thr
 85 90 95

Gly Ala Asn Leu Val Leu Gly Val Ile Glu Arg Ser Gly Ser Thr Leu
 100 105 110

Tyr Cys Thr Ala Leu Tyr Phe Asp Pro Gln Gln Gly Leu Ser Gly Lys
 115 120 125

His Arg Lys Leu Met Pro Thr Gly Thr Glu Arg Leu Ile Trp Gly Lys
 130 135 140

Gly Asp Gly Ser Thr Leu Pro Val Leu Asp Thr Gln Val Gly Arg Val
 145 150 155 160

Gly Ala Val Ile Cys Trp Glu Asn Met Met Pro Leu Leu Arg Thr Ala
 165 170 175

Met Tyr Ala Gln Gly Ile Glu Val Trp Cys Ala Pro Thr Val Asp Glu
 180 185 190

ES 2 564 176 T3

Arg Glu Met Trp Gln Val Ser Met Arg His Ile Ala His Glu Gly Arg
 195 200 205

Cys Phe Val Val Ser Ala Cys Gln Val Gln Ala Ser Pro Glu Glu Leu
 210 215 220

Gly Leu Glu Ile Ala Asn Trp Pro Ala Gln Arg Pro Leu Ile Ala Gly
 225 230 235 240

Gly Ser Val Ile Val Gly Pro Met Gly Asp Val Leu Ala Gly Pro Leu
 245 250 255

Val Gly Arg Ala Gly Leu Ile Ser Ala Gln Ile Asp Thr Ala Asp Leu
 260 265 270

Val Arg Ala Arg Tyr Asp Tyr Asp Val Val Gly His Tyr Ala Arg Pro
 275 280 285

Asp Val Phe Glu Leu Thr Val Asp Gln Arg Pro Arg Pro Gly Val Arg
 290 295 300

Phe Thr
 305

<210> 20

<211> 336

<212> PRT

5 <213> *Nocardia farcinica* IFM 10152]

<400> 20

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

Val Trp Leu Asp Gly Ala Ala Thr Val Asp Lys Cys Val Ala Leu Ile
 20 25 30

Glu Glu Ala Ala Asp Asn Gly Ala Ala Leu Ile Ala Phe Pro Glu Thr
 35 40 45

Phe Val Pro Gly Tyr Pro Trp Trp Leu Trp Leu Asp Ser Pro Ala Trp
 50 55 60

Gly Met Gln Phe Val Ala Arg Tyr Phe Asp Asn Ser Leu Ala Leu Asp
 65 70 75 80

Gly Pro Leu Phe Ala Arg Leu Arg Glu Ala Ala Arg Arg Ser Ala Ile
 85 90 95

ES 2 564 176 T3

Thr Val Val Thr Gly His Ser Glu Arg Asp Gly Gly Ser Leu Tyr Met
 100 105 110

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

Lys Leu Lys Pro Thr His Val Glu Arg Thr Val Phe Gly Glu Ser Asp
 130 135 140

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

Leu Cys Cys Trp Glu His Leu Gln Pro Leu Thr Lys Tyr Ala Met Tyr
 165 170 175

Ser Gln His Glu Gln Ile His Val Ala Ala Trp Pro Ser Phe Ser Val
 180 185 190

Tyr Arg Gly Ala Ala Tyr Ala Leu Gly Pro Glu Val Asn Thr Gly Ala
 195 200 205

Ala Arg Gln Tyr Ala Val Glu Gly Gln Cys Phe Val Leu Ser Pro Cys
 210 215 220

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

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

Pro Asp Gly Arg Glu Leu Gly Thr Ala Leu Pro Glu Thr Glu Glu Gly
 260 265 270

Leu Val Tyr Ala Asp Leu Glu Ala Ser Ala Val Ala Val Ala Lys Ser
 275 280 285

Ala Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Gln Leu
 290 295 300

Leu Trp Asp Pro Arg Pro Arg Ser Val Val Arg Gln Val Ala Leu Ser
 305 310 315 320

Val Ala Ser Pro Ala Glu Ser Ala Asp Asp Ala Glu Pro Ala Val Arg
 325 330 335

<210> 21
 <211> 356
 <212> PRT

ES 2 564 176 T3

<213> Alcaligenes faecalis 1650

<400> 21

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

Pro Asn Tyr Asp Leu Ala Thr Gly Val Asp Lys Thr Ile Glu Leu Ala
20 25 30

Arg Gln Ala Arg Asp Glu Gly Cys Asp Leu Ile Val Phe Gly Glu Thr
35 40 45

Trp Leu Pro Gly Tyr Pro Phe His Val Trp Leu Gly Ala Pro Ala Trp
50 55 60

Ser Leu Lys Tyr Ser Ala Arg Tyr Tyr Ala Asn Ser Leu Ser Leu Asp
65 70 75 80

Ser Ala Glu Phe Gln Arg Ile Ala Gln Ala Ala Arg Thr Leu Gly Ile
85 90 95

Phe Ile Ala Leu Gly Tyr Ser Glu Arg Ser Gly Gly Ser Leu Tyr Leu
100 105 110

Gly Gln Cys Leu Ile Asp Asp Lys Gly Glu Met Leu Trp Ser Arg Arg
115 120 125

Lys Leu Lys Pro Thr His Val Glu Arg Thr Val Phe Gly Glu Gly Tyr
130 135 140

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

Leu Cys Cys Trp Glu His Leu Ser Pro Leu Ser Lys Tyr Ala Leu Tyr
165 170 175

Ser Gln His Glu Ala Ile His Ile Ala Ala Trp Pro Ser Phe Ser Leu
180 185 190

Tyr Ser Glu Gln Ala His Ala Leu Ser Ala Lys Val Asn Met Ala Ala
195 200 205

Ser Gln Ile Tyr Ser Val Glu Gly Gln Cys Phe Thr Ile Ala Ala Ser

ES 2 564 176 T3

210

215

220

Ser Val Val Thr Gln Glu Thr Leu Asp Met Leu Glu Val Gly Glu His
225 230 235 240

Asn Ala Pro Leu Leu Lys Val Gly Gly Gly Ser Ser Met Ile Phe Ala
245 250 255

Pro Asp Gly Arg Thr Leu Ala Pro Tyr Leu Pro His Asp Ala Glu Gly
260 265 270

Leu Ile Ile Ala Asp Leu Asn Met Glu Glu Ile Ala Phe Ala Lys Ala
275 280 285

Ile Asn Asp Pro Val Gly His Tyr Ser Lys Pro Glu Ala Thr Arg Leu
290 295 300

Val Leu Asp Leu Gly His Arg Asp Pro Met Thr Arg Val His Ser Lys
305 310 315 320

Ser Val Thr Arg Glu Glu Ala Pro Glu Gln Gly Val Gln Ser Lys Ile
325 330 335

Ala Ser Val Ala Ile Ser His Pro Gln Asp Ser Asp Thr Leu Leu Val
340 345 350

Gln Glu Pro Ser
355

<210> 22

<211> 336

<212> PRT

5 <213> *Pseudomonas syringae* pv. *syringae* B728a

<400> 22

Met Lys Glu Pro Leu Lys Val Ala Cys Val Gln Ala Ala Pro Val Phe
1 5 10 15

Leu Asp Leu Asp Ala Thr Val Asp Lys Thr Ile Thr Leu Met Glu Gln
20 25 30

Ala Ala Ala Ala Gly Ala Gly Leu Ile Ala Phe Pro Glu Thr Trp Ile
35 40 45

Pro Gly Tyr Pro Trp Phe Leu Trp Leu Asp Ala Pro Ala Trp Asn Met
50 55 60

ES 2 564 176 T3

Pro Leu Val Gln Arg Tyr His Gln Gln Ser Leu Val Leu Asp Ser Val
65 70 75 80

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

Val Leu Gly Tyr Ser Glu Arg Asn Lys Ala Ser Leu Tyr Ile Gly Gln
100 105 110

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

Lys Ala Thr His Val Glu Arg Thr Met Phe Gly Glu Gly Asp Gly Ala
130 135 140

Ser Leu Arg Thr Phe Glu Thr Pro Val Gly Val Leu Gly Ala Leu Cys
145 150 155 160

Cys Trp Glu His Leu Gln Pro Leu Ser Lys Tyr Ala Met Tyr Ala Gln
165 170 175

Asn Glu Gln Ile His Val Ala Ala Trp Pro Ser Phe Ser Leu Tyr Arg
180 185 190

Asn Ala Thr Ser Ala Leu Gly Pro Glu Val Asn Thr Ala Ala Ser Arg
195 200 205

Val Tyr Ala Ala Glu Gly Gln Cys Phe Val Leu Ala Pro Cys Ala Ile
210 215 220

Val Ser Pro Glu Met Ile Glu Met Leu Cys Asp Ser Asp Ala Lys Arg
225 230 235 240

Ser Leu Leu Gln Ala Gly Gly Gly His Ala Arg Ile Phe Gly Pro Asp
245 250 255

Gly Ser Asp Leu Ala Thr Pro Leu Gly Glu His Glu Glu Gly Leu Leu
260 265 270

Tyr Ala Thr Leu Asp Pro Ala Ala Leu Thr Leu Ala Lys Val Ala Ala
275 280 285

Asp Pro Ala Gly His Tyr Ser Arg Pro Asp Val Thr Arg Leu Met Phe
290 295 300

ES 2 564 176 T3

Asn Pro Asn Pro Thr Pro Cys Val Val Asp Leu Pro Asp Leu Pro Ile
305 310 315 320

Ser Ser Glu Ser Ile Glu Leu Leu Arg Pro Asp Ile Ala Leu Glu Val
325 330 335

<210> 23

<211> 345

<212> PRT

5 <213> Bradyrhizobium sp. BTAi1

<400> 23

Met Gly Leu Ala His Pro Lys Tyr Lys Val Ala Val Val Gln Ala Ala
1 5 10 15

Pro Ala Trp Leu Asp Leu Asp Ala Ser Ile Lys Lys Thr Ile Ala Leu
20 25 30

Ile Glu Glu Ala Ala Asp Lys Gly Ala Lys Leu Ile Ala Phe Pro Glu
35 40 45

Val Phe Ile Pro Gly Tyr Pro Trp His Ile Trp Met Asp Ser Pro Ala
50 55 60

Trp Cys Ile Gly Arg Gly Phe Val Gln Arg Tyr Phe Asp Asn Ser Leu
65 70 75 80

Ala Tyr Asp Ser Pro Gln Ala Glu Ala Leu Arg Ala Ala Val Arg Lys
85 90 95

Ala Gln Leu Thr Ala Val Leu Gly Leu Ser Glu Arg Asp Gly Gly Ser
100 105 110

Leu Tyr Ile Ala Gln Trp Leu Ile Gly Ala Asp Gly Glu Thr Ile Ala
115 120 125

Lys Arg Arg Lys Leu Arg Pro Thr His Ala Glu Arg Thr Val Tyr Gly
130 135 140

Glu Gly Asp Gly Ser Asp Leu Ala Val His Glu Arg Pro Asp Ile Gly
145 150 155 160

Arg Ile Gly Ala Leu Cys Cys Trp Glu His Leu Gln Pro Leu Ser Lys
165 170 175

Tyr Ala Met Tyr Ala Gln Asn Glu Gln Val His Val Ala Ala Trp Pro
180 185 190

ES 2 564 176 T3

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

Asn Asn Ala Ala Ser Arg Val Tyr Ala Val Glu Gly Ser Cys Phe Val
 210 215 220

Leu Ala Pro Cys Ala Thr Val Ser Gln Ala Met Ile Asp Glu Leu Cys
 225 230 235 240

Asp Arg Pro Asp Lys His Ala Leu Leu His Ala Gly Gly Gly His Ala
 245 250 255

Ala Ile Phe Gly Pro Asp Gly Ser Ala Leu Ala Ala Gln Leu Pro Pro
 260 265 270

Asp Gln Glu Gly Leu Leu Ile Ala Glu Ile Asp Leu Gly Met Ile Gly
 275 280 285

Ile Ala Lys Asn Ala Ala Asp Pro Ala Gly His Tyr Ser Arg Pro Asp
 290 295 300

Val Thr Arg Leu Leu Leu Asn Lys Lys Pro Leu Asn Arg Val Glu His
 305 310 315 320

Phe Ser Leu Pro Val Asp Ser Ala Ala Ala Leu Pro Gly Glu Ala
 325 330 335

Ala Val Ala Arg Pro Asp Gln Ser Ile
 340 345

<210> 24

<211> 366

<212> PRT

5 <213> Rhodococcus rhodochrous NCIMB 11216

<400> 24

Met Val Glu Tyr Thr Asn Thr Phe Lys Val Ala Ala Val Gln Ala Gln
 1 5 10 15

Pro Val Trp Phe Asp Ala Ala Lys Thr Val Asp Lys Thr Val Ser Ile
 20 25 30

Ile Ala Glu Ala Ala Arg Asn Gly Cys Glu Leu Val Ala Phe Pro Glu
 35 40 45

ES 2 564 176 T3

Val Phe Ile Pro Gly Tyr Pro Tyr His Ile Trp Val Asp Ser Pro Leu
50 55 60

Ala Gly Met Ala Lys Phe Ala Val Arg Tyr His Glu Asn Ser Leu Thr
65 70 75 80

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

Asn Ile Ala Val Val Val Gly Ile Ser Glu Arg Asp Gly Gly Ser Leu
100 105 110

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

Arg Arg Lys Leu Lys Pro Thr His Val Glu Arg Ser Val Tyr Gly Glu
130 135 140

Gly Asn Gly Ser Asp Ile Ser Val Tyr Asp Met Pro Phe Ala Arg Leu
145 150 155 160

Gly Ala Leu Asn Cys Trp Glu His Phe Gln Thr Leu Thr Lys Tyr Ala
165 170 175

Met Tyr Ser Met His Glu Gln Val His Val Ala Ser Trp Pro Gly Met
180 185 190

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

Thr Ala Thr Arg Met Tyr Ala Leu Glu Gly Gln Thr Phe Val Val Cys
210 215 220

Thr Thr Gln Val Val Thr Pro Glu Ala His Glu Phe Phe Cys Glu Asn
225 230 235 240

Glu Glu Gln Arg Lys Leu Ile Gly Arg Gly Gly Gly Phe Ala Arg Ile
245 250 255

Ile Gly Pro Asp Gly Arg Asp Leu Ala Thr Pro Leu Ala Glu Asp Glu
260 265 270

Glu Gly Ile Leu Tyr Ala Asp Ile Asp Leu Ser Ala Ile Thr Leu Ala
275 280 285

Lys Gln Ala Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu

ES 2 564 176 T3

290

295

300

Ser Leu Asn Phe Asn Gln Arg Arg Thr Thr Pro Val Asn Thr Pro Leu
305 310 315 320

Ser Thr Ile His Ala Thr His Thr Phe Val Pro Gln Phe Gly Ala Leu
325 330 335

Asp Gly Val Arg Glu Leu Asn Gly Ala Asp Glu Gln Arg Ala Leu Pro
340 345 350

Ser Thr His Ser Asp Glu Thr Asp Arg Ala Thr Ala Thr Leu
355 360 365

<210> 25

<211> 366

<212> PRT

5 <213> *Rhodococcus rhodochrous* ATCC 39484

<400> 25

Met Val Glu Tyr Thr Asn Thr Phe Lys Val Ala Ala Val Gln Ala Gln
1 5 10 15

Pro Val Trp Phe Asp Ala Ala Lys Thr Val Asp Lys Thr Val Ser Ile
20 25 30

Ile Ala Glu Ala Ala Arg Asn Gly Cys Glu Leu Val Ala Phe Pro Glu
35 40 45

Val Phe Ile Pro Gly Tyr Pro Tyr His Ile Trp Val Asp Ser Pro Leu
50 55 60

Ala Gly Met Ala Lys Phe Ala Val Arg Tyr His Glu Asn Ser Leu Thr
65 70 75 80

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

Asn Ile Ala Val Val Val Gly Ile Ser Glu Arg Asp Gly Gly Ser Leu
100 105 110

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

Arg Arg Lys Leu Lys Pro Thr His Val Glu Arg Ser Val Tyr Gly Glu
130 135 140

ES 2 564 176 T3

Gly Asn Gly Ser Asp Ile Ser Val Tyr Asp Met Pro Phe Ala Arg Leu
 145 150 155 160

Gly Ala Leu Asn Cys Trp Glu His Phe Gln Thr Leu Thr Lys Tyr Ala
 165 170 175

Met Tyr Ser Met His Glu Gln Val His Val Ala Ser Trp Pro Gly Met
 180 185 190

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

Thr Ala Thr Arg Met Tyr Ala Leu Glu Gly Gln Thr Phe Val Val Cys
 210 215 220

Thr Thr Gln Val Val Thr Pro Glu Ala His Glu Phe Phe Cys Glu Asn
 225 230 235 240

Glu Glu Gln Arg Met Leu Ile Gly Arg Gly Gly Gly Phe Ala Arg Ile
 245 250 255

Ile Gly Pro Asp Gly Arg Asp Leu Ala Thr Pro Leu Ala Glu Asp Glu
 260 265 270

Glu Gly Ile Leu Tyr Ala Asp Ile Asp Leu Ser Ala Ile Thr Leu Ala
 275 280 285

Lys Gln Ala Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu
 290 295 300

Ser Leu Asn Phe Asn Gln Arg Arg Thr Thr Pro Val Asn Thr Pro Leu
 305 310 315 320

Ser Thr Ile His Ala Thr His Thr Phe Val Pro Gln Phe Gly Ala Leu
 325 330 335

Asp Gly Val Arg Glu Leu Asn Gly Ala Asp Glu Gln Arg Ala Leu Pro
 340 345 350

Ser Thr His Ser Asp Glu Thr Asp Arg Ala Thr Ala Thr Leu
 355 360 365

<210> 26

<211> 1110

<212> DNA

5 <213> *Acidovorax facilis* 72W

<220>

<221> CDS

<222> (1)..(1110)

ES 2 564 176 T3

<400> 26

atg gtt tcg tat aac agc aag ttc ctc gcg gca acc gtt cag gca gag	48
Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu	
1 5 10 15	
ccg gta tgg ctc gac gca gac gca acg atc gac aag tcg atc ggc atc	96
Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile	
20 25 30	
atc gaa gaa gct gcc caa aag ggc gcg agt ctg atc gct ttc ccg gaa	144
Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu	
35 40 45	
gta ttc att ccg ggc tac ccc tat tgg gcg tgg ctc ggc gac gtg aag	192
Val Phe Ile Pro Gly Tyr Pro Tyr Trp Ala Trp Leu Gly Asp Val Lys	
50 55 60	
tac agc cta agc ttt act tca cgc tat cac gag aat tcg ttg gag cta	240
Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu	
65 70 75 80	
ggt gac gac cgt atg cgt cgc ctc cag ctg gcc gcg cgc cgc aac aaa	288
Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys	
85 90 95	
atc gca ctc gtc atg ggc tat tcg gag cgg gaa gcc gga tcg cgc tat	336
Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr	
100 105 110	
ctg agc cag gtg ttc atc gac gag cgt ggc gag atc gtt gcc aat cgg	384
Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg	
115 120 125	
cgc aag ctg aag ccc aca cac gtt gag cgt acg atc tac ggc gaa ggc	432
Arg Lys Leu Lys Pro Thr His Val Glu Arg Thr Ile Tyr Gly Glu Gly	
130 135 140	
aac gga acc gat ttc ctc acg cac gac ttc gcg ttc gga cgc gtc ggt	480
Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly	
145 150 155 160	
gga ttg aac tgc tgg gaa cat ttc caa ccg ctc agc aag ttc atg atg	528
Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met	
165 170 175	
tac agc ctc ggt gag cag gtc cac gtt gca tcg tgg ccg gcg atg tcc	576
Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser	
180 185 190	
cct ctt cag ccg gat gtt ttc caa cag agc atc gaa gcc aac gcg acg	624
Pro Leu Gln Pro Asp Val Phe Gln Gln Ser Ile Glu Ala Asn Ala Thr	
195 200 205	
gtc acc cgc tcg tac gca atc gaa ggc caa acc ttt gtg ctt tgc tcg	672
Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser	

ES 2 564 176 T3

210	215	220	
acg cag gtg atc gga cct agc gcg atc gaa acg ttc tgc ctc aac gac			720
Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp			
225	230	235	240
gaa cag cgc gca ctg ttg ccg caa gga tgt ggc tgg gcg cgc att tac			768
Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr			
	245	250	255
ggc ccg gat gga agc gag ctt gcg aag cct ctg gcg gaa gat gct gag			816
Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu			
	260	265	270
ggg atc ttg tac gca gag atc gat ctg gag cag att ctg ctg gcg aag			864
Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys			
	275	280	285
gct gga gcc gat ccg gtc ggg cac tat tcg cgg cct gac gtg ctg tcg			912
Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser			
	290	295	300
gtc cag ttc gac ccg cgc aat cat acg cca gtt cat cgc atc ggc att			960
Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile			
305	310	315	320
gac ggt cgc ttg gat gtg aat acc cgc agt cgc gtg gag aat ttc cga			1008
Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg			
	325	330	335
ctg cga caa gcg gct gag cag gag cgt cag gca tcc aag cgg ctc gga			1056
Leu Arg Gln Ala Ala Glu Gln Glu Arg Gln Ala Ser Lys Arg Leu Gly			
	340	345	350
acg aaa ctc ttt gaa caa tcc ctt ctg gct gaa gaa ccg gtc cca gca			1104
Thr Lys Leu Phe Glu Gln Ser Leu Leu Ala Glu Glu Pro Val Pro Ala			
	355	360	365
aag tag			1110
Lys			

<210> 27

<211> 369

<212> PRT

5 <213> Acidovorax facilis 72W

<400> 27

Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
1 5 10 15

Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
20 25 30

Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
35 40 45

ES 2 564 176 T3

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

Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
65 70 75 80

Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
85 90 95

Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
100 105 110

Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
115 120 125

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

Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
145 150 155 160

Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met
165 170 175

Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
180 185 190

Pro Leu Gln Pro Asp Val Phe Gln Gln Ser Ile Glu Ala Asn Ala Thr
195 200 205

Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
210 215 220

Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
225 230 235 240

Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
245 250 255

Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
260 265 270

Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
275 280 285

ES 2 564 176 T3

Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
 290 295 300

Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
 305 310 315 320

Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
 325 330 335

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

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

Lys

<210> 28

<211> 1110

<212> DNA

5 <213> *Acidovorax facilis* 72W

<220>

<221> CDS

<222> (1)..(1110)

<400> 28

atg gtt tcg tat aac agc aag ttc ctc gcg gca acc gtt cag gca gag 48
 Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
 1 5 10 15

ccg gta tgg ctc gac gca gac gca acg atc gac aag tcg atc ggc atc 96
 Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

atc gaa gaa gct gcc caa aag ggc gcg agt ctg atc gct ttc ccg gaa 144
 Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
 35 40 45

gta ttc att ccg ggc tac ccc tat tgg gcg tgg ctc ggc gac gtg aag 192
 Val Phe Ile Pro Gly Tyr Pro Tyr Trp Ala Trp Leu Gly Asp Val Lys
 50 55 60

tac agc cta agc ttt act tca cgc tat cac gag aat tcg ttg gag cta 240
 Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
 65 70 75 80

ggc gac gac cgt atg cgt cgc ctc cag ctg gcc gcg cgc cgc aac aaa 288
 Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
 85 90 95

10 atc gca ctc gtc atg ggc tat tcg gag cgg gaa gcc gga tcg cgc tat 336
 Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr

ES 2 564 176 T3

100					105					110						
ctg	agc	cag	gtg	ttc	atc	gac	gag	cgt	ggc	gag	atc	gtt	gcc	aat	cgg	384
Leu	Ser	Gln	Val	Phe	Ile	Asp	Glu	Arg	Gly	Glu	Ile	Val	Ala	Asn	Arg	
		115					120					125				
cgc	aag	ctg	aag	ccc	aca	cac	gtt	gag	cgt	acg	atc	tac	ggc	gaa	ggc	432
Arg	Lys	Leu	Lys	Pro	Thr	His	Val	Glu	Arg	Thr	Ile	Tyr	Gly	Glu	Gly	
	130					135					140					
aac	gga	acc	gat	ttc	ctc	acg	cac	gac	ttc	gcg	ttc	gga	cgc	gtc	ggt	480
Asn	Gly	Thr	Asp	Phe	Leu	Thr	His	Asp	Phe	Ala	Phe	Gly	Arg	Val	Gly	
145					150					155					160	
gga	ttg	aac	tgc	tgg	gaa	cat	ttc	caa	ccg	ctc	agc	aag	ttc	atg	atg	528
Gly	Leu	Asn	Cys	Trp	Glu	His	Phe	Gln	Pro	Leu	Ser	Lys	Phe	Met	Met	
			165					170						175		
tac	agc	ctc	ggt	gag	cag	gtc	cac	gtt	gca	tcg	tgg	ccg	gcg	atg	tcc	576
Tyr	Ser	Leu	Gly	Glu	Gln	Val	His	Val	Ala	Ser	Trp	Pro	Ala	Met	Ser	
		180					185						190			
cct	ctt	cag	ccg	gat	gtt	ttc	caa	gct	agc	atc	gaa	gcc	aac	gcg	acg	624
Pro	Leu	Gln	Pro	Asp	Val	Phe	Gln	Ala	Ser	Ile	Glu	Ala	Asn	Ala	Thr	
		195					200					205				
gtc	acc	cgc	tcg	tac	gca	atc	gaa	ggc	caa	acc	ttt	gtg	ctt	tgc	tcg	672
Val	Thr	Arg	Ser	Tyr	Ala	Ile	Glu	Gly	Gln	Thr	Phe	Val	Leu	Cys	Ser	
	210					215					220					
acg	cag	gtg	atc	gga	cct	agc	gcg	atc	gaa	acg	ttc	tgc	ctc	aac	gac	720
Thr	Gln	Val	Ile	Gly	Pro	Ser	Ala	Ile	Glu	Thr	Phe	Cys	Leu	Asn	Asp	
225					230					235					240	
gaa	cag	cgc	gca	ctg	ttg	ccg	caa	gga	tgt	ggc	tgg	gcg	cgc	att	tac	768
Glu	Gln	Arg	Ala	Leu	Leu	Pro	Gln	Gly	Cys	Gly	Trp	Ala	Arg	Ile	Tyr	
			245					250						255		
ggc	ccg	gat	gga	agc	gag	ctt	gcg	aag	cct	ctg	gcg	gaa	gat	gct	gag	816
Gly	Pro	Asp	Gly	Ser	Glu	Leu	Ala	Lys	Pro	Leu	Ala	Glu	Asp	Ala	Glu	
			260					265					270			
ggg	atc	ttg	tac	gca	gag	atc	gat	ctg	gag	cag	att	ctg	ctg	gcg	aag	864
Gly	Ile	Leu	Tyr	Ala	Glu	Ile	Asp	Leu	Glu	Gln	Ile	Leu	Leu	Ala	Lys	
		275					280						285			
gct	gga	gcc	gat	ccg	gtc	ggg	cac	tat	tcg	cgg	cct	gac	gtg	ctg	tcg	912
Ala	Gly	Ala	Asp	Pro	Val	Gly	His	Tyr	Ser	Arg	Pro	Asp	Val	Leu	Ser	
	290					295					300					
gtc	cag	ttc	gac	ccg	cgc	aat	cat	acg	cca	ggt	cat	cgc	atc	ggc	att	960
Val	Gln	Phe	Asp	Pro	Arg	Asn	His	Thr	Pro	Val	His	Arg	Ile	Gly	Ile	
305					310					315				320		
gac	ggt	cgc	ttg	gat	gtg	aat	acc	cgc	agt	cgc	gtg	gag	aat	ttc	cga	1008
Asp	Gly	Arg	Leu	Asp	Val	Asn	Thr	Arg	Ser	Arg	Val	Glu	Asn	Phe	Arg	
				325					330					335		
ctg	cga	caa	gcg	gct	gag	cag	gag	cgt	cag	gca	tcc	aag	cgg	ctc	gga	1056
Leu	Arg	Gln	Ala	Ala	Glu	Gln	Glu	Arg	Gln	Ala	Ser	Lys	Arg	Leu	Gly	
			340					345					350			

ES 2 564 176 T3

Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
 180 185 190

Pro Leu Gln Pro Asp Val Phe Gln Ala Ser Ile Glu Ala Asn Ala Thr
 195 200 205

Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
 210 215 220

Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
 225 230 235 240

Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
 245 250 255

Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
 260 265 270

Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
 275 280 285

Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
 290 295 300

Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
 305 310 315 320

Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
 325 330 335

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

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

Lys

<210> 30

<211> 1110

<212> DNA

5 <213> Acidovorax facilis 72W

<220>

<221> CDS

<222> (1)..(1110)

<400> 30

ES 2 564 176 T3

atg gtt tcg tat aac agc aag ttc ctc gcg gca acc gtt cag gca gag Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu 1 5 10 15	48
ccg gta tgg ctc gac gca gac gca acg atc gac aag tcg atc ggc atc Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile 20 25 30	96
atc gaa gaa gct gcc caa aag ggc gcg agt ctg atc gct ttc ccg gaa Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu 35 40 45	144
gta ttc att ccg ggc tac ccc tat tgg gcg tgg ctc ggc gac gtg aag Val Phe Ile Pro Gly Tyr Pro Tyr Trp Ala Trp Leu Gly Asp Val Lys 50 55 60	192
tac agc cta agc ttt act tca cgc tat cac gag aat tcg ttg gag cta Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu 65 70 75 80	240
ggt gac gac cgt atg cgt cgc ctc cag ctg gcc gcg cgc cgc aac aaa Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys 85 90 95	288
atc gca ctc gtc atg ggc tat tcg gag cgg gaa gcc gga tcg cgc tat Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr 100 105 110	336
ctg agc cag gtg ttc atc gac gag cgt ggc gag atc gtt gcc aat ccg Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg 115 120 125	384
cgc aag ctg aag ccc aca cac gtt gag cgt acg atc tac ggc gaa ggc Arg Lys Leu Lys Pro Thr His Val Glu Arg Thr Ile Tyr Gly Glu Gly 130 135 140	432
aac gga acc gat ttc ctc acg cac gac ttc gcg ttc gga cgc gtc ggt Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly 145 150 155 160	480
gga ttg aac tgc tgg gaa cat ttc caa ccg ctc agc aag ttc atg atg Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met 165 170 175	528
tac agc ctc ggt gag cag gtc cac gtt gca tcg tgg ccg gcg atg tcc Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser 180 185 190	576
cct ctt cag ccg gat gtt ttc caa tgt agc atc gaa gcc aac gcg acg Pro Leu Gln Pro Asp Val Phe Gln Cys Ser Ile Glu Ala Asn Ala Thr 195 200 205	624
gtc acc cgc tcg tac gca atc gaa ggc caa acc ttt gtg ctt tgc tcg Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser 210 215 220	672
acg cag gtg atc gga cct agc gcg atc gaa acg ttc tgc ctc aac gac Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp 225 230 235 240	720

ES 2 564 176 T3

gaa cag cgc gca ctg ttg ccg caa gga tgt ggc tgg gcg cgc att tac 768
 Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
 245 250 255

ggc ccg gat gga agc gag ctt gcg aag cct ctg gcg gaa gat gct gag 816
 Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
 260 265 270

ggg atc ttg tac gca gag atc gat ctg gag cag att ctg ctg gcg aag 864
 Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
 275 280 285

gct gga gcc gat ccg gtc ggg cac tat tcg cgg cct gac gtg ctg tcg 912
 Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
 290 295 300

gtc cag ttc gac ccg cgc aat cat acg cca gtt cat cgc atc ggc att 960
 Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
 305 310 315 320

gac ggt cgc ttg gat gtg aat acc cgc agt cgc gtg gag aat ttc cga 1008
 Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
 325 330 335

ctg cga caa gcg gct gag cag gag cgt cag gca tcc aag cgg ctc gga 1056
 Leu Arg Gln Ala Ala Glu Gln Glu Arg Gln Ala Ser Lys Arg Leu Gly
 340 345 350

acg aaa ctc ttt gaa caa tcc ctt ctg gct gaa gaa ccg gtc cca gca 1104
 Thr Lys Leu Phe Glu Gln Ser Leu Leu Ala Glu Glu Pro Val Pro Ala
 355 360 365

aag tag 1110
 Lys

<210> 31

<211> 369

<212> PRT

5 <213> *Acidovorax facilis* 72W

<400> 31

Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
 1 5 10 15

Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
 35 40 45

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

ES 2 564 176 T3

Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
65 70 75 80

Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
85 90 95

Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
100 105 110

Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
115 120 125

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

Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
145 150 155 160

Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met
165 170 175

Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
180 185 190

Pro Leu Gln Pro Asp Val Phe Gln Cys Ser Ile Glu Ala Asn Ala Thr
195 200 205

Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
210 215 220

Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
225 230 235 240

Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
245 250 255

Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
260 265 270

Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
275 280 285

Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
290 295 300

Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile

ES 2 564 176 T3

cgc aag ctg aag ccc aca cac gtt gag cgt acg atc tac ggc gaa ggc	432
Arg Lys Leu Lys Pro Thr His Val Glu Arg Thr Ile Tyr Gly Glu Gly	
130 135 140	
aac gga acc gat ttc ctc acg cac gac ttc gcg ttc gga cgc gtc ggt	480
Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly	
145 150 155 160	
gga ttg aac tgc tgg gaa cat ttc caa ccg ctc agc aag ttc atg atg	528
Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met	
165 170 175	
tac agc ctc ggt gag cag gtc cac gtt gca tcg tgg ccg gcg atg tcc	576
Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser	
180 185 190	
cct ctt cag ccg gat gtt ttc caa acc agc atc gaa gcc aac gcg acg	624
Pro Leu Gln Pro Asp Val Phe Gln Thr Ser Ile Glu Ala Asn Ala Thr	
195 200 205	
gtc acc cgc tcg tac gca atc gaa ggc caa acc ttt gtg ctt tgc tcg	672
Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser	
210 215 220	
acg cag gtg atc gga cct agc gcg atc gaa acg ttc tgc ctc aac gac	720
Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp	
225 230 235 240	
gaa cag cgc gca ctg ttg ccg caa gga tgt ggc tgg gcg cgc att tac	768
Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr	
245 250 255	
ggc ccg gat gga agc gag ctt gcg aag cct ctg gcg gaa gat gct gag	816
Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu	
260 265 270	
ggg atc ttg tac gca gag atc gat ctg gag cag att ctg ctg gcg aag	864
Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys	
275 280 285	
gct gga gcc gat ccg gtc ggg cac tat tcg cgg cct gac gtg ctg tcg	912
Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser	
290 295 300	
gtc cag ttc gac ccg cgc aat cat acg cca gtt cat cgc atc ggc att	960
Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile	
305 310 315 320	
gac ggt cgc ttg gat gtg aat acc cgc agt cgc gtg gag aat ttc cga	1008
Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg	
325 330 335	
ctg cga caa gcg gct gag cag gag cgt cag gca tcc aag cgg ctc gga	1056
Leu Arg Gln Ala Ala Glu Gln Glu Arg Gln Ala Ser Lys Arg Leu Gly	
340 345 350	
acg aaa ctc ttt gaa caa tcc ctt ctg gct gaa gaa ccg gtc cca gca	1104
Thr Lys Leu Phe Glu Gln Ser Leu Leu Ala Glu Glu Pro Val Pro Ala	
355 360 365	

ES 2 564 176 T3

1110

aag tag
Lys

<210> 33

<211> 369

<212> PRT

5 <213> Acidovorax facilis 72W

<400> 33

Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
1 5 10 15

Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
20 25 30

Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
35 40 45

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

Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
65 70 75 80

Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
85 90 95

Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
100 105 110

Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
115 120 125

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

Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
145 150 155 160

Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met
165 170 175

Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
180 185 190

Pro Leu Gln Pro Asp Val Phe Gln Thr Ser Ile Glu Ala Asn Ala Thr

ES 2 564 176 T3

195 200 205

Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
 210 215 220

Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
 225 230 235 240

Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
 245 250 255

Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
 260 265 270

Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
 275 280 285

Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
 290 295 300

Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
 305 310 315 320

Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
 325 330 335

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

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

Lys

<210> 34

<211> 1110

<212> DNA

5 <213> *Acidovorax facilis* 72W

<220>

<221> CDS

<222> (1)..(1110)

<400> 34

atg gtt tcg tat aac agc aag ttc ctc gcg gca acc gtt cag gca gag
 Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
 1 5 10 15

10

48

ES 2 564 176 T3

ccg gta tgg ctc gac gca gac gca acg atc gac aag tcg atc ggc atc 96
 Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

atc gaa gaa gct gcc caa aag ggc gcg agt ctg atc gct ttc ccg gaa 144
 Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
 35 40 45

gta ttc att ccg ggc tac ccc tat tgg gcg tgg ctc ggc gac gtg aag 192
 Val Phe Ile Pro Gly Tyr Pro Tyr Trp Ala Trp Leu Gly Asp Val Lys
 50 55 60

tac agc cta agc ttt act tca cgc tat cac gag aat tcg ttg gag cta 240
 Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
 65 70 75 80

ggt gac gac cgt atg cgt cgc ctc cag ctg gcc gcg cgc cgc aac aaa 288
 Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
 85 90 95

atc gca ctc gtc atg ggc tat tcg gag cgg gaa gcc gga tcg cgc tat 336
 Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
 100 105 110

ctg agc cag gtg ttc atc gac gag cgt ggc gag atc gtt gcc aat cgg 384
 Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
 115 120 125

cgc aag ctg aag ccc aca cac gtt gag cgt acg atc tac ggc gaa ggc 432
 Arg Lys Leu Lys Pro Thr His Val Glu Arg Thr Ile Tyr Gly Glu Gly
 130 135 140

aac gga acc gat ttc ctc acg cac gac ttc gcg ttc gga cgc gtc ggt 480
 Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
 145 150 155 160

gga ttg aac tgc tgg gaa cat ttc caa ccg ctc agc aag ttc atg atg 528
 Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met
 165 170 175

tac agc ctc ggt gag cag gtc cac gtt gca tcg tgg ccg gcg atg tcc 576
 Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
 180 185 190

cct ctt cag ccg gat gtt ttc caa gga agc atc gaa gcc aac gcg acg 624
 Pro Leu Gln Pro Asp Val Phe Gln Gly Ser Ile Glu Ala Asn Ala Thr
 195 200 205

gtc acc cgc tcg tac gca atc gaa ggc caa acc ttt gtg ctt tgc tcg 672
 Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
 210 215 220

acg cag gtg atc gga cct agc gcg atc gaa acg ttc tgc ctc aac gac 720
 Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
 225 230 235 240

gaa cag cgc gca ctg ttg ccg caa gga tgt ggc tgg gcg cgc att tac 768
 Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
 245 250 255

ES 2 564 176 T3

ggc ccg gat gga agc gag ctt gcg aag cct ctg gcg gaa gat gct gag 816
 Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
 260 265 270

ggg atc ttg tac gca gag atc gat ctg gag cag att ctg ctg gcg aag 864
 Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
 275 280 285

gct gga gcc gat ccg gtc ggg cac tat tcg cgg cct gac gtg ctg tcg 912
 Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
 290 295 300

gtc cag ttc gac ccg cgc aat cat acg cca gtt cat cgc atc ggc att 960
 Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
 305 310 315 320

gac ggt cgc ttg gat gtg aat acc cgc agt cgc gtg gag aat ttc cga 1008
 Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
 325 330 335

ctg cga caa gcg gct gag cag gag cgt cag gca tcc aag cgg ctc gga 1056
 Leu Arg Gln Ala Ala Glu Gln Glu Arg Gln Ala Ser Lys Arg Leu Gly
 340 345 350

acg aaa ctc ttt gaa caa tcc ctt ctg gct gaa gaa ccg gtc cca gca 1104
 Thr Lys Leu Phe Glu Gln Ser Leu Leu Ala Glu Glu Pro Val Pro Ala
 355 360 365

aag tag 1110
 Lys

<210> 35

<211> 369

<212> PRT

5 <213> *Acidovorax facilis* 72W

<400> 35

Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
 1 5 10 15

Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
 35 40 45

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

Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
 65 70 75 80

Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys

ES 2 564 176 T3

85 90 95

Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
100 105 110

Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
115 120 125

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

Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
145 150 155 160

Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met
165 170 175

Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
180 185 190

Pro Leu Gln Pro Asp Val Phe Gln Gly Ser Ile Glu Ala Asn Ala Thr
195 200 205

Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
210 215 220

Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
225 230 235 240

Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
245 250 255

Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
260 265 270

Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
275 280 285

Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
290 295 300

Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
305 310 315 320

Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
325 330 335

ES 2 564 176 T3

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

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

Lys

<210> 36

<211> 1110

<212> DNA

5 <213> *Acidovorax facilis* 72W

<220>

<221> CDS

<222> (1)..(1110)

<400> 36

atg gtt tcg tat aac agc aag ttc ctc gcg gca acc gtt cag gca gag 48
 Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
 1 5 10 15

ccg gta tgg ctc gac gca gac gca acg atc gac aag tcg atc ggc atc 96
 Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

atc gaa gaa gct gcc caa aag ggc gcg agt ctg atc gct ttc ccg gaa 144
 Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
 35 40 45

gta ttc att ccg ggc tac ccc tat tgg gcg tgg ctc ggc gac gtg aag 192
 Val Phe Ile Pro Gly Tyr Pro Tyr Trp Ala Trp Leu Gly Asp Val Lys
 50 55 60

tac agc cta agc ttt act tca cgc tat cac gag aat tcg ttg gag cta 240
 Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
 65 70 75 80

ggc gac gac cgt atg cgt cgc ctc cag ctg gcc gcg cgc cgc aac aaa 288
 Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
 85 90 95

atc gca ctc gtc atg ggc tat tcg gag cgg gaa gcc gga tcg cgc tat 336
 Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
 100 105 110

ctg agc cag gtg ttc atc gac gag cgt ggc gag atc gtt gcc aat cgg 384
 Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
 115 120 125

cgc aag ctg aag ccc aca cac gtt gag cgt acg atc tac ggc gaa ggc 432
 Arg Lys Leu Lys Pro Thr His Val Glu Arg Thr Ile Tyr Gly Glu Gly
 130 135 140

10

ES 2 564 176 T3

aac gga acc gat ttc ctc acg cac gac ttc gcg ttc gga cgc gtc ggt 480
 Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
 145 150 155 160

gga ttg aac tgc tgg gaa cat ttc caa ccg ctc agc aag ttc atg atg 528
 Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met
 165 170 175

tac agc ctc ggt gag cag gtc cac gtt gca tcg tgg ccg gcg atg tcc 576
 Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
 180 185 190

cct ctt cag ccg gat gtt ttc caa cac agc atc gaa gcc aac gcg acg 624
 Pro Leu Gln Pro Asp Val Phe Gln His Ser Ile Glu Ala Asn Ala Thr
 195 200 205

gtc acc cgc tcg tac gca atc gaa ggc caa acc ttt gtg ctt tgc tcg 672
 Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
 210 215 220

acg cag gtg atc gga cct agc gcg atc gaa acg ttc tgc ctc aac gac 720
 Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
 225 230 235 240

gaa cag cgc gca ctg ttg ccg caa gga tgt ggc tgg gcg cgc att tac 768
 Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
 245 250 255

ggc ccg gat gga agc gag ctt gcg aag cct ctg gcg gaa gat gct gag 816
 Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
 260 265 270

ggg atc ttg tac gca gag atc gat ctg gag cag att ctg ctg gcg aag 864
 Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
 275 280 285

gct gga gcc gat ccg gtc ggg cac tat tcg cgg cct gac gtg ctg tcg 912
 Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
 290 295 300

gtc cag ttc gac ccg cgc aat cat acg cca gtt cat cgc atc ggc att 960
 Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
 305 310 315 320

gac ggt cgc ttg gat gtg aat acc cgc agt cgc gtg gag aat ttc cga 1008
 Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
 325 330 335

ctg cga caa gcg gct gag cag gag cgt cag gca tcc aag cgg ctc gga 1056
 Leu Arg Gln Ala Ala Glu Gln Glu Arg Gln Ala Ser Lys Arg Leu Gly
 340 345 350

acg aaa ctc ttt gaa caa tcc ctt ctg gct gaa gaa ccg gtc cca gca 1104
 Thr Lys Leu Phe Glu Gln Ser Leu Leu Ala Glu Glu Pro Val Pro Ala
 355 360 365

aag tag 1110
 Lys

<210> 37
 <211> 369

ES 2 564 176 T3

<212> PRT

<213> *Acidovorax facilis* 72W

<400> 37

Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
 1 5 10 15

Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
 35 40 45

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

Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
 65 70 75 80

Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
 85 90 95

Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
 100 105 110

Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
 115 120 125

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

Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
 145 150 155 160

Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met
 165 170 175

Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
 180 185 190

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

Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
 210 215 220

ES 2 564 176 T3

Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
225 230 235 240

Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
245 250 255

Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
260 265 270

Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
275 280 285

Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
290 295 300

Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
305 310 315 320

Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
325 330 335

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

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

Lys

<210> 38

<211> 1110

<212> DNA

5 <213> Acidovorax facilis 72W

<220>

<221> CDS

<222> (1)..(1110)

<400> 38

atg gtt tcg tat aac agc aag ttc ctc gcg gca acc gtt cag gca gag 48
Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
1 5 10 15

ccg gta tgg ctc gac gca gac gca acg atc gac aag tcg atc ggc atc 96
Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
20 25 30

10

ES 2 564 176 T3

atc gaa gaa gct gcc caa aag ggc gcg agt ctg atc gct ttc ccg gaa 144
Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
35 40 45

gta ttc att ccg ggc tac ccc tat tgg gcg tgg ctc ggc gac gtg aag 192
Val Phe Ile Pro Gly Tyr Pro Tyr Trp Ala Trp Leu Gly Asp Val Lys
50 55 60

tac agc cta agc ttt act tca cgc tat cac gag aat tcg ttg gag cta 240
Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
65 70 75 80

ggc gac gac cgt atg cgt cgc ctc cag ctg gcc gcg cgc cgc aac aaa 288
Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
85 90 95

atc gca ctc gtc atg ggc tat tcg gag cgg gaa gcc gga tcg cgc tat 336
Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
100 105 110

ctg agc cag gtg ttc atc gac gag cgt ggc gag atc gtt gcc aat cgg 384
Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
115 120 125

cgc aag ctg aag ccc aca cac gtt gag cgt acg atc tac ggc gaa ggc 432
Arg Lys Leu Lys Pro Thr His Val Glu Arg Thr Ile Tyr Gly Glu Gly
130 135 140

aac gga acc gat ttc ctc acg cac gac ttc gcg ttc gga cgc gtc ggt 480
Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
145 150 155 160

gga ttg aac tgc tgg gaa cat ttc caa ccg ctc agc aag ttc atg atg 528
Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met
165 170 175

tac agc ctc ggt gag cag gtc cac gtt gca tcg tgg ccg gcg atg tcc 576
Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
180 185 190

cct ctt cag ccg gat gtt ttc caa aag agc atc gaa gcc aac gcg acg 624
Pro Leu Gln Pro Asp Val Phe Gln Lys Ser Ile Glu Ala Asn Ala Thr
195 200 205

gtc acc cgc tcg tac gca atc gaa ggc caa acc ttt gtg ctt tgc tcg 672
Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
210 215 220

acg cag gtg atc gga cct agc gcg atc gaa acg ttc tgc ctc aac gac 720
Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
225 230 235 240

gaa cag cgc gca ctg ttg ccg caa gga tgt ggc tgg gcg cgc att tac 768
Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
245 250 255

ggc ccg gat gga agc gag ctt gcg aag cct ctg gcg gaa gat gct gag 816
Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
260 265 270

ggg atc ttg tac gca gag atc gat ctg gag cag att ctg ctg gcg aag 864

ES 2 564 176 T3

Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
 275 280 285

gct gga gcc gat ccg gtc ggg cac tat tcg cgg cct gac gtg ctg tcg 912
 Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
 290 295 300

gtc cag ttc gac ccg cgc aat cat acg cca gtt cat cgc atc ggc att 960
 Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
 305 310 315 320

gac ggt cgc ttg gat gtg aat acc cgc agt cgc gtg gag aat ttc cga 1008
 Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
 325 330 335

ctg cga caa gcg gct gag cag gag cgt cag gca tcc aag cgg ctc gga 1056
 Leu Arg Gln Ala Ala Glu Gln Glu Arg Gln Ala Ser Lys Arg Leu Gly
 340 345 350

acg aaa ctc ttt gaa caa tcc ctt ctg gct gaa gaa ccg gtc cca gca 1104
 Thr Lys Leu Phe Glu Gln Ser Leu Leu Ala Glu Glu Pro Val Pro Ala
 355 360 365

aag tag 1110
 Lys

<210> 39

<211> 369

<212> PRT

5 <213> *Acidovorax facilis* 72W

<400> 39

Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
 1 5 10 15

Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
 35 40 45

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

Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
 65 70 75 80

Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
 85 90 95

Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
 100 105 110

ES 2 564 176 T3

Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
 115 120 125

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

Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
 145 150 155 160

Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met
 165 170 175

Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
 180 185 190

Pro Leu Gln Pro Asp Val Phe Gln Lys Ser Ile Glu Ala Asn Ala Thr
 195 200 205

Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
 210 215 220

Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
 225 230 235 240

Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
 245 250 255

Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
 260 265 270

Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
 275 280 285

Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
 290 295 300

Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
 305 310 315 320

Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
 325 330 335

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

ES 2 564 176 T3

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

Lys

<210> 40

<211> 1110

<212> DNA

5 <213> *Acidovorax facilis* 72W

<220>

<221> CDS

<222> (1)..(1110)

<400> 40

atg gtt tcg tat aac agc aag ttc ctc gcg gca acc gtt cag gca gag 48
 Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Thr Val Gln Ala Glu
 1 5 10 15

ccg gta tgg ctc gac gca gac gca acg atc gac aag tcg atc ggc atc 96
 Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

atc gaa gaa gct gcc caa aag ggc gcg agt ctg atc gct ttc ccg gaa 144
 Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
 35 40 45

gta ttc att ccg ggc tac ccc tat tgg gcg tgg ctc ggc gac gtg aag 192
 Val Phe Ile Pro Gly Tyr Pro Tyr Trp Ala Trp Leu Gly Asp Val Lys
 50 55 60

tac agc cta agc ttt act tca cgc tat cac gag aat tcg ttg gag cta 240
 Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
 65 70 75 80

ggc gac gac cgt atg cgt cgc ctc cag ctg gcc gcg cgc cgc aac aaa 288
 Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
 85 90 95

atc gca ctc gtc atg ggc tat tcg gag cgg gaa gcc gga tcg cgc tat 336
 Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
 100 105 110

ctg agc cag gtg ttc atc gac gag cgt ggc gag atc gtt gcc aat cgg 384
 Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
 115 120 125

cgc aag ctg aag ccc aca cac gtt gag cgt acg atc tac ggc gaa ggc 432
 Arg Lys Leu Lys Pro Thr His Val Glu Arg Thr Ile Tyr Gly Glu Gly
 130 135 140

aac gga acc gat ttc ctc acg cac gac ttc gcg ttc gga cgc gtc ggt 480
 Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
 145 150 155 160

10 gga ttg aac tgc tgg gaa cat ttc caa ccg ctc agc aag ttc atg atg 528

ES 2 564 176 T3

Gly	Leu	Asn	Cys	Trp	Glu	His	Phe	Gln	Pro	Leu	Ser	Lys	Phe	Met	Met	
				165					170					175		
tac	agc	ctc	ggt	gag	cag	gtc	cac	gtt	gca	tcg	tgg	ccg	gcg	atg	tcc	576
Tyr	Ser	Leu	Gly	Glu	Gln	Val	His	Val	Ala	Ser	Trp	Pro	Ala	Met	Ser	
			180					185					190			
cct	ctt	cag	ccg	gat	gtt	ttc	caa	aat	agc	atc	gaa	gcc	aac	gcg	acg	624
Pro	Leu	Gln	Pro	Asp	Val	Phe	Gln	Asn	Ser	Ile	Glu	Ala	Asn	Ala	Thr	
			195				200					205				
gtc	acc	cgc	tcg	tac	gca	atc	gaa	ggc	caa	acc	ttt	gtg	ctt	tgc	tcg	672
Val	Thr	Arg	Ser	Tyr	Ala	Ile	Glu	Gly	Gln	Thr	Phe	Val	Leu	Cys	Ser	
	210					215				220						
acg	cag	gtg	atc	gga	cct	agc	gcg	atc	gaa	acg	ttc	tgc	ctc	aac	gac	720
Thr	Gln	Val	Ile	Gly	Pro	Ser	Ala	Ile	Glu	Thr	Phe	Cys	Leu	Asn	Asp	
225					230					235					240	
gaa	cag	cgc	gca	ctg	ttg	ccg	caa	gga	tgt	ggc	tgg	gcg	cgc	att	tac	768
Glu	Gln	Arg	Ala	Leu	Leu	Pro	Gln	Gly	Cys	Gly	Trp	Ala	Arg	Ile	Tyr	
				245				250						255		
ggc	ccg	gat	gga	agc	gag	ctt	gcg	aag	cct	ctg	gcg	gaa	gat	gct	gag	816
Gly	Pro	Asp	Gly	Ser	Glu	Leu	Ala	Lys	Pro	Leu	Ala	Glu	Asp	Ala	Glu	
			260					265					270			
ggg	atc	ttg	tac	gca	gag	atc	gat	ctg	gag	cag	att	ctg	ctg	gcg	aag	864
Gly	Ile	Leu	Tyr	Ala	Glu	Ile	Asp	Leu	Glu	Gln	Ile	Leu	Leu	Ala	Lys	
		275					280					285				
gct	gga	gcc	gat	ccg	gtc	ggg	cac	tat	tcg	cgg	cct	gac	gtg	ctg	tcg	912
Ala	Gly	Ala	Asp	Pro	Val	Gly	His	Tyr	Ser	Arg	Pro	Asp	Val	Leu	Ser	
	290					295					300					
gtc	cag	ttc	gac	ccg	cgc	aat	cat	acg	cca	gtt	cat	cgc	atc	ggc	att	960
Val	Gln	Phe	Asp	Pro	Arg	Asn	His	Thr	Pro	Val	His	Arg	Ile	Gly	Ile	
305					310					315					320	
gac	ggt	cgc	ttg	gat	gtg	aat	acc	cgc	agt	cgc	gtg	gag	aat	ttc	cga	1008
Asp	Gly	Arg	Leu	Asp	Val	Asn	Thr	Arg	Ser	Arg	Val	Glu	Asn	Phe	Arg	
				325					330					335		
ctg	cga	caa	gcg	gct	gag	cag	gag	cgt	cag	gca	tcc	aag	cgg	ctc	gga	1056
Leu	Arg	Gln	Ala	Ala	Glu	Gln	Glu	Arg	Gln	Ala	Ser	Lys	Arg	Leu	Gly	
			340					345					350			
acg	aaa	ctc	ttt	gaa	caa	tcc	ctt	ctg	gct	gaa	gaa	ccg	gtc	cca	gca	1104
Thr	Lys	Leu	Phe	Glu	Gln	Ser	Leu	Leu	Ala	Glu	Glu	Pro	Val	Pro	Ala	
		355					360					365				
aag	tag															1110
Lys																

<210> 41

<211> 369

<212> PRT

5 <213> Acidovorax facilis 72W

<400> 41

ES 2 564 176 T3

Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
1 5 10 15

Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
20 25 30

Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
35 40 45

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

Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
65 70 75 80

Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
85 90 95

Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
100 105 110

Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
115 120 125

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

Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
145 150 155 160

Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met
165 170 175

Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
180 185 190

Pro Leu Gln Pro Asp Val Phe Gln Asn Ser Ile Glu Ala Asn Ala Thr
195 200 205

Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
210 215 220

Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
225 230 235 240

ES 2 564 176 T3

Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
 245 250 255

Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
 260 265 270

Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
 275 280 285

Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
 290 295 300

Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
 305 310 315 320

Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
 325 330 335

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

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

Lys

<210> 42

<211> 1110

<212> DNA

5 <213> Acidovorax facilis 72W

<220>

<221> CDS

<222> (1)..(1110)

<400> 42

atg gtt tcg tat aac agc aag ttc ctc gcg gca acc gtt cag gca gag 48
 Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
 1 5 10 15

ccg gta tgg ctc gac gca gac gca acg atc gac aag tcg atc ggc atc 96
 Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

atc gaa gaa gct gcc caa aag ggc gcg agt ctg atc gct ttc ccg gaa 144
 Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
 35 40 45

10 gta ttc att ccg ggc tac ccc tat tgg gcg tgg ctc ggc gac gtg aag 192

ES 2 564 176 T3

Val	Phe	Ile	Pro	Gly	Tyr	Pro	Tyr	Trp	Ala	Trp	Leu	Gly	Asp	Val	Lys	
	50					55					60					
tac	agc	cta	agc	ttt	act	tca	cgc	tat	cac	gag	aat	tcg	ttg	gag	cta	240
Tyr	Ser	Leu	Ser	Phe	Thr	Ser	Arg	Tyr	His	Glu	Asn	Ser	Leu	Glu	Leu	
65					70					75					80	
ggt	gac	gac	cgt	atg	cgt	cgc	ctc	cag	ctg	gcc	gcg	cgc	cgc	aac	aaa	288
Gly	Asp	Asp	Arg	Met	Arg	Arg	Leu	Gln	Leu	Ala	Ala	Arg	Arg	Asn	Lys	
				85					90					95		
atc	gca	ctc	gtc	atg	ggc	tat	tcg	gag	cgg	gaa	gcc	gga	tcg	cgc	tat	336
Ile	Ala	Leu	Val	Met	Gly	Tyr	Ser	Glu	Arg	Glu	Ala	Gly	Ser	Arg	Tyr	
			100					105					110			
ctg	agc	cag	gtg	ttc	atc	gac	gag	cgt	ggc	gag	atc	gtt	gcc	aat	cgg	384
Leu	Ser	Gln	Val	Phe	Ile	Asp	Glu	Arg	Gly	Glu	Ile	Val	Ala	Asn	Arg	
		115					120					125				
cgc	aag	ctg	aag	ccc	aca	cac	gtt	gag	cgt	acg	atc	tac	ggc	gaa	ggc	432
Arg	Lys	Leu	Lys	Pro	Thr	His	Val	Glu	Arg	Thr	Ile	Tyr	Gly	Glu	Gly	
	130					135					140					
aac	gga	acc	gat	ttc	ctc	acg	cac	gac	ttc	gcg	ttc	gga	cgc	gtc	ggt	480
Asn	Gly	Thr	Asp	Phe	Leu	Thr	His	Asp	Phe	Ala	Phe	Gly	Arg	Val	Gly	
145					150					155					160	
gga	ttg	aac	tgc	tgg	gaa	cat	ttc	caa	ccg	ctc	agc	aag	ttc	atg	atg	528
Gly	Leu	Asn	Cys	Trp	Glu	His	Phe	Gln	Pro	Leu	Ser	Lys	Phe	Met	Met	
			165					170						175		
tac	agc	ctc	ggt	gag	cag	gtc	cac	gtt	gca	tcg	tgg	ccg	gcg	atg	tcc	576
Tyr	Ser	Leu	Gly	Glu	Gln	Val	His	Val	Ala	Ser	Trp	Pro	Ala	Met	Ser	
			180					185					190			
cct	ctt	cag	ccg	gat	gtt	ttc	caa	tct	agc	atc	gaa	gcc	aac	gcg	acg	624
Pro	Leu	Gln	Pro	Asp	Val	Phe	Gln	Ser	Ser	Ile	Glu	Ala	Asn	Ala	Thr	
		195					200					205				
gtc	acc	cgc	tcg	tac	gca	atc	gaa	ggc	caa	acc	ttt	gtg	ctt	tgc	tcg	672
Val	Thr	Arg	Ser	Tyr	Ala	Ile	Glu	Gly	Gln	Thr	Phe	Val	Leu	Cys	Ser	
	210					215					220					
acg	cag	gtg	atc	gga	cct	agc	gcg	atc	gaa	acg	ttc	tgc	ctc	aac	gac	720
Thr	Gln	Val	Ile	Gly	Pro	Ser	Ala	Ile	Glu	Thr	Phe	Cys	Leu	Asn	Asp	
225					230					235				240		
gaa	cag	cgc	gca	ctg	ttg	ccg	caa	gga	tgt	ggc	tgg	gcg	cgc	att	tac	768
Glu	Gln	Arg	Ala	Leu	Leu	Pro	Gln	Gly	Cys	Gly	Trp	Ala	Arg	Ile	Tyr	
			245					250						255		
ggc	ccg	gat	gga	agc	gag	ctt	gcg	aag	cct	ctg	gcg	gaa	gat	gct	gag	816
Gly	Pro	Asp	Gly	Ser	Glu	Leu	Ala	Lys	Pro	Leu	Ala	Glu	Asp	Ala	Glu	
			260					265					270			
ggg	atc	ttg	tac	gca	gag	atc	gat	ctg	gag	cag	att	ctg	ctg	gcg	aag	864
Gly	Ile	Leu	Tyr	Ala	Glu	Ile	Asp	Leu	Glu	Gln	Ile	Leu	Leu	Ala	Lys	
		275					280					285				
gct	gga	gcc	gat	ccg	gtc	ggg	cac	tat	tcg	cgg	cct	gac	gtg	ctg	tcg	912
Ala	Gly	Ala	Asp	Pro	Val	Gly	His	Tyr	Ser	Arg	Pro	Asp	Val	Leu	Ser	

ES 2 564 176 T3

290	295	300	
gtc cag ttc gac ccg cgc aat cat acg cca gtt cat cgc atc ggc att			960
Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile			
305	310	315	320
gac ggt cgc ttg gat gtg aat acc cgc agt cgc gtg gag aat ttc cga			1008
Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg			
	325	330	335
ctg cga caa gcg gct gag cag gag cgt cag gca tcc aag cgg ctc gga			1056
Leu Arg Gln Ala Ala Glu Gln Glu Arg Gln Ala Ser Lys Arg Leu Gly			
	340	345	350
acg aaa ctc ttt gaa caa tcc ctt ctg gct gaa gaa ccg gtc cca gca			1104
Thr Lys Leu Phe Glu Gln Ser Leu Leu Ala Glu Glu Pro Val Pro Ala			
	355	360	365
aag tag			1110
Lys			
<210> 43			
<211> 369			
<212> PRT			
5 <213> <i>Acidovorax facilis</i> 72W			
<400> 43			
Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu			
1	5	10	15
Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile			
	20	25	30
Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu			
	35	40	45
Val Phe Ile Pro Gly Tyr Pro Tyr Trp Ala Trp Leu Gly Asp Val Lys			
	50	55	60
Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu			
65	70	75	80
Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys			
	85	90	95
Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr			
	100	105	110
Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg			
	115	120	125

ES 2 564 176 T3

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

Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
 145 150 155 160

Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met
 165 170 175

Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
 180 185 190

Pro Leu Gln Pro Asp Val Phe Gln Ser Ser Ile Glu Ala Asn Ala Thr
 195 200 205

Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
 210 215 220

Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
 225 230 235 240

Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
 245 250 255

Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
 260 265 270

Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
 275 280 285

Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
 290 295 300

Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
 305 310 315 320

Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
 325 330 335

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

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

Lys

ES 2 564 176 T3

<210> 44
 <211> 1110
 <212> DNA
 <213> Acidovorax facilis 72W

5 <220>
 <221> CDS
 <222> (1)..(1110)

<400> 44

atg gtt tcg tat aac agc aag ttc ctc gcg gca acc gtt cag gca gag	48
Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu	
1 5 10 15	
ccg gta tgg ctc gac gca gac gca acg atc gac aag tcg atc ggc atc	96
Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile	
20 25 30	
atc gaa gaa gct gcc caa aag ggc gcg agt ctg atc gct ttc ccg gaa	144
Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu	
35 40 45	
gta ttc att ccg ggc tac ccc tat tgg gcg tgg ctc ggc gac gtg aag	192
Val Phe Ile Pro Gly Tyr Pro Tyr Trp Ala Trp Leu Gly Asp Val Lys	
50 55 60	
tac agc cta agc ttt act tca cgc tat cac gag aat tcg ttg gag cta	240
Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu	
65 70 75 80	
ggt gac gac cgt atg cgt cgc ctc cag ctg gcc gcg cgc cgc aac aaa	288
Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys	
85 90 95	
atc gca ctc gtc atg ggc tat tcg gag cgg gaa gcc gga tcg cgc tat	336
Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr	
100 105 110	
ctg agc cag gtg ttc atc gac gag cgt ggc gag atc gtt gcc aat cgg	384
Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg	
115 120 125	
cgc aag ctg aag ccc aca cac gtt gag cgt acg atc tac ggc gaa ggc	432
Arg Lys Leu Lys Pro Thr His Val Glu Arg Thr Ile Tyr Gly Glu Gly	
130 135 140	
aac gga acc gat ttc ctc acg cac gac ttc gcg ttc gga cgc gtc ggt	480
Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly	
145 150 155 160	
gga ttg aac tgc tgg gaa cat aaa caa ccg ctc agc aag ttc atg atg	528
Gly Leu Asn Cys Trp Glu His Lys Gln Pro Leu Ser Lys Phe Met Met	
165 170 175	
tac agc ctc ggt gag cag gtc cac gtt gca tcg tgg ccg gcg atg tcc	576
Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser	

ES 2 564 176 T3

	180		185		190	
cct ctt cag ccg gat gtt ttc caa ctg agc atc gaa gcc aac gcg acg						624
Pro Leu Gln Pro Asp Val Phe Gln Leu Ser Ile Glu Ala Asn Ala Thr						
	195		200		205	
gtc acc cgc tcg tac gca atc gaa ggc caa acc ttt gtg ctt tgc tcg						672
Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser						
	210		215		220	
acg cag gtg atc gga cct agc gcg atc gaa acg ttc tgc ctc aac gac						720
Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp						
	225		230		235	240
gaa cag cgc gca ctg ttg ccg caa gga tgt ggc tgg gcg cgc att tac						768
Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr						
		245		250		255
ggc ccg gat gga agc gag ctt gcg aag cct ctg gcg gaa gat gct gag						816
Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu						
		260		265		270
ggg atc ttg tac gca gag atc gat ctg gag cag att ctg ctg gcg aag						864
Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys						
		275		280		285
gct gga gcc gat ccg gtc ggg cac tat tcg cgg cct gac gtg ctg tcg						912
Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser						
	290		295		300	
gtc cag ttc gac ccg cgc aat cat acg cca gtt cat cgc atc ggc att						960
Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile						
	305		310		315	320
gac ggt cgc ttg gat gtg aat acc cgc agt cgc gtg gag aat ttc cga						1008
Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg						
		325		330		335
ctg cga caa gcg gct gag cag gag cgt cag gca tcc aag cgg ctc gga						1056
Leu Arg Gln Ala Ala Glu Gln Glu Arg Gln Ala Ser Lys Arg Leu Gly						
		340		345		350
acg aaa ctc ttt gaa caa tcc ctt ctg gct gaa gaa ccg gtc cca gca						1104
Thr Lys Leu Phe Glu Gln Ser Leu Leu Ala Glu Glu Pro Val Pro Ala						
		355		360		365

aag tag 1110
Lys

<210> 45

<211> 369

<212> PRT

5 <213> Acidovorax facilis 72W

<400> 45

Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
1 5 10 15

ES 2 564 176 T3

Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
 35 40 45

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

Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
 65 70 75 80

Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
 85 90 95

Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
 100 105 110

Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
 115 120 125

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

Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
 145 150 155 160

Gly Leu Asn Cys Trp Glu His Lys Gln Pro Leu Ser Lys Phe Met Met
 165 170 175

Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
 180 185 190

Pro Leu Gln Pro Asp Val Phe Gln Leu Ser Ile Glu Ala Asn Ala Thr
 195 200 205

Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
 210 215 220

Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
 225 230 235 240

Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
 245 250 255

ES 2 564 176 T3

Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
 260 265 270

Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
 275 280 285

Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
 290 295 300

Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
 305 310 315 320

Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
 325 330 335

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

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

Lys

<210> 46

<211> 1110

<212> DNA

5 <213> *Acidovorax facilis* 72W

<220>

<221> CDS

<222> (1)..(1110)

<400> 46

atg gtt tcg tat aac agc aag ttc ctc gcg gca acc gtt cag gca gag 48
 Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
 1 5 10 15

ccg gta tgg ctc gac gca gac gca acg atc gac aag tcg atc ggc atc 96
 Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

atc gaa gaa gct gcc caa aag ggc gcg agt ctg atc gct ttc ccg gaa 144
 Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
 35 40 45

gta ttc att ccg ggc tac ccc tat tgg gcg tgg ctc ggc gac gtg aag 192
 Val Phe Ile Pro Gly Tyr Pro Tyr Trp Ala Trp Leu Gly Asp Val Lys
 50 55 60

tac agc cta agc ttt act tca cgc tat cac gag aat tcg ttg gag cta 240
 Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu

10

ES 2 564 176 T3

65					70						75				80	
ggt	gac	gac	cgt	atg	cgt	cgc	ctc	cag	ctg	gcc	gcg	cgc	cgc	aac	aaa	288
Gly	Asp	Asp	Arg	Met	Arg	Arg	Leu	Gln	Leu	Ala	Ala	Arg	Arg	Asn	Lys	
				85					90					95		
atc	gca	ctc	gtc	atg	ggc	tat	tcg	gag	cgg	gaa	gcc	gga	tcg	cgc	tat	336
Ile	Ala	Leu	Val	Met	Gly	Tyr	Ser	Glu	Arg	Glu	Ala	Gly	Ser	Arg	Tyr	
			100					105					110			
ctg	agc	cag	gtg	ttc	atc	gac	gag	cgt	ggc	gag	atc	ggt	gcc	aat	cgg	384
Leu	Ser	Gln	Val	Phe	Ile	Asp	Glu	Arg	Gly	Glu	Ile	Val	Ala	Asn	Arg	
		115					120					125				
cgc	aag	ctg	aag	ccc	aca	cac	ggt	gag	cgt	acg	atc	tac	ggc	gaa	ggc	432
Arg	Lys	Leu	Lys	Pro	Thr	His	Val	Glu	Arg	Thr	Ile	Tyr	Gly	Glu	Gly	
	130					135					140					
aac	gga	acc	gat	ttc	ctc	acg	cac	gac	ttc	gcg	ttc	gga	cgc	gtc	ggt	480
Asn	Gly	Thr	Asp	Phe	Leu	Thr	His	Asp	Phe	Ala	Phe	Gly	Arg	Val	Gly	
145					150					155					160	
gga	ttg	aac	tgc	tgg	gaa	cat	atg	caa	ccg	ctc	agc	aag	ttc	atg	atg	528
Gly	Leu	Asn	Cys	Trp	Glu	His	Met	Gln	Pro	Leu	Ser	Lys	Phe	Met	Met	
			165					170						175		
tac	agc	ctc	ggt	gag	cag	gtc	cac	ggt	gca	tcg	tgg	ccg	gcg	atg	tcc	576
Tyr	Ser	Leu	Gly	Glu	Gln	Val	His	Val	Ala	Ser	Trp	Pro	Ala	Met	Ser	
			180					185					190			
cct	ctt	cag	ccg	gat	gtt	ttc	caa	ctg	agc	atc	gaa	gcc	aac	gcg	acg	624
Pro	Leu	Gln	Pro	Asp	Val	Phe	Gln	Leu	Ser	Ile	Glu	Ala	Asn	Ala	Thr	
		195					200					205				
gtc	acc	cgc	tcg	tac	gca	atc	gaa	ggc	caa	acc	ttt	gtg	ctt	tgc	tcg	672
Val	Thr	Arg	Ser	Tyr	Ala	Ile	Glu	Gly	Gln	Thr	Phe	Val	Leu	Cys	Ser	
	210					215					220					
acg	cag	gtg	atc	gga	cct	agc	gcg	atc	gaa	acg	ttc	tgc	ctc	aac	gac	720
Thr	Gln	Val	Ile	Gly	Pro	Ser	Ala	Ile	Glu	Thr	Phe	Cys	Leu	Asn	Asp	
225					230					235					240	
gaa	cag	cgc	gca	ctg	ttg	ccg	caa	gga	tgt	ggc	tgg	gcg	cgc	att	tac	768
Glu	Gln	Arg	Ala	Leu	Leu	Pro	Gln	Gly	Cys	Gly	Trp	Ala	Arg	Ile	Tyr	
			245					250						255		
ggc	ccg	gat	gga	agc	gag	ctt	gcg	aag	cct	ctg	gcg	gaa	gat	gct	gag	816
Gly	Pro	Asp	Gly	Ser	Glu	Leu	Ala	Lys	Pro	Leu	Ala	Glu	Asp	Ala	Glu	
			260					265					270			
ggg	atc	ttg	tac	gca	gag	atc	gat	ctg	gag	cag	att	ctg	ctg	gcg	aag	864
Gly	Ile	Leu	Tyr	Ala	Glu	Ile	Asp	Leu	Glu	Gln	Ile	Leu	Leu	Ala	Lys	
		275					280					285				
gct	gga	gcc	gat	ccg	gtc	ggg	cac	tat	tcg	cgg	cct	gac	gtg	ctg	tcg	912
Ala	Gly	Ala	Asp	Pro	Val	Gly	His	Tyr	Ser	Arg	Pro	Asp	Val	Leu	Ser	
	290					295					300					
gtc	cag	ttc	gac	ccg	cgc	aat	cat	acg	cca	ggt	cat	cgc	atc	ggc	att	960
Val	Gln	Phe	Asp	Pro	Arg	Asn	His	Thr	Pro	Val	His	Arg	Ile	Gly	Ile	
305					310					315					320	

ES 2 564 176 T3

gac ggt cgc ttg gat gtg aat acc cgc agt cgc gtg gag aat ttc cga 1008
 Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
 325 330 335

ctg cga caa gcg gct gag cag gag cgt cag gca tcc aag cgg ctc gga 1056
 Leu Arg Gln Ala Ala Glu Gln Glu Arg Gln Ala Ser Lys Arg Leu Gly
 340 345 350

acg aaa ctc ttt gaa caa tcc ctt ctg gct gaa gaa ccg gtc cca gca 1104
 Thr Lys Leu Phe Glu Gln Ser Leu Leu Ala Glu Glu Pro Val Pro Ala
 355 360 365

aag tag 1110
 Lys

<210> 47

<211> 369

<212> PRT

5 <213> *Acidovorax facilis* 72W

<400> 47

Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
 1 5 10 15

Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
 35 40 45

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

Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
 65 70 75 80

Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
 85 90 95

Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
 100 105 110

Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
 115 120 125

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

ES 2 564 176 T3

Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
145 150 155 160

Gly Leu Asn Cys Trp Glu His Met Gln Pro Leu Ser Lys Phe Met Met
165 170 175

Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
180 185 190

Pro Leu Gln Pro Asp Val Phe Gln Leu Ser Ile Glu Ala Asn Ala Thr
195 200 205

Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
210 215 220

Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
225 230 235 240

Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
245 250 255

Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
260 265 270

Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
275 280 285

Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
290 295 300

Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
305 310 315 320

Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
325 330 335

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

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

Lys

<210> 48
<211> 1110
<212> DNA

ES 2 564 176 T3

<213> Acidovorax facilis 72W

<220>

<221> CDS

<222> (1)..(1110)

5 <400> 48

atg gtt tcg tat aac agc aag ttc ctc gcg gca acc gtt cag gca gag 48
Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
1 5 10 15

ccg gta tgg ctc gac gca gac gca acg atc gac aag tcg atc ggc atc 96
Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
20 25 30

atc gaa gaa gct gcc caa aag ggc gcg agt ctg atc gct ttc ccg gaa 144
Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
35 40 45

gta ttc att ccg ggc tac ccc tat tgg gcg tgg ctc ggc gac gtg aag 192
Val Phe Ile Pro Gly Tyr Pro Tyr Trp Ala Trp Leu Gly Asp Val Lys
50 55 60

tac agc cta agc ttt act tca cgc tat cac gag aat tcg ttg gag cta 240
Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
65 70 75 80

ggt gac gac cgt atg cgt cgc ctc cag ctg gcc gcg cgc cgc aac aaa 288
Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
85 90 95

atc gca ctc gtc atg ggc tat tcg gag cgg gaa gcc gga tcg cgc tat 336
Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
100 105 110

ctg agc cag gtg ttc atc gac gag cgt ggc gag atc gtt gcc aat cgg 384
Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
115 120 125

cgc aag ctg aag ccc aca cac gtt gag cgt acg atc tac ggc gaa ggc 432
Arg Lys Leu Lys Pro Thr His Val Glu Arg Thr Ile Tyr Gly Glu Gly
130 135 140

aac gga acc gat ttc ctc acg cac gac ttc gcg ttc gga cgc gtc ggt 480
Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
145 150 155 160

gga ttg aac tgc tgg gaa cat acc caa ccg ctc agc aag ttc atg atg 528
Gly Leu Asn Cys Trp Glu His Thr Gln Pro Leu Ser Lys Phe Met Met
165 170 175

tac agc ctc ggt gag cag gtc cac gtt gca tcg tgg ccg gcg atg tcc 576
Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
180 185 190

cct ctt cag ccg gat gtt ttc caa ctg agc atc gaa gcc aac gcg acg 624
Pro Leu Gln Pro Asp Val Phe Gln Leu Ser Ile Glu Ala Asn Ala Thr
195 200 205

ES 2 564 176 T3

gtc acc cgc tcg tac gca atc gaa ggc caa acc ttt gtg ctt tgc tcg 672
 Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
 210 215 220

acg cag gtg atc gga cct agc gcg atc gaa acg ttc tgc ctc aac gac 720
 Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
 225 230 235 240

gaa cag cgc gca ctg ttg ccg caa gga tgt ggc tgg gcg cgc att tac 768
 Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
 245 250 255

ggc ccg gat gga agc gag ctt gcg aag cct ctg gcg gaa gat gct gag 816
 Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
 260 265 270

ggg atc ttg tac gca gag atc gat ctg gag cag att ctg ctg gcg aag 864
 Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
 275 280 285

gct gga gcc gat ccg gtc ggg cac tat tcg cgg cct gac gtg ctg tcg 912
 Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
 290 295 300

gtc cag ttc gac ccg cgc aat cat acg cca gtt cat cgc atc ggc att 960
 Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
 305 310 315 320

gac ggt cgc ttg gat gtg aat acc cgc agt cgc gtg gag aat ttc cga 1008
 Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
 325 330 335

ctg cga caa gcg gct gag cag gag cgt cag gca tcc aag cgg ctc gga 1056
 Leu Arg Gln Ala Ala Glu Gln Glu Arg Gln Ala Ser Lys Arg Leu Gly
 340 345 350

acg aaa ctc ttt gaa caa tcc ctt ctg gct gaa gaa ccg gtc cca gca 1104
 Thr Lys Leu Phe Glu Gln Ser Leu Leu Ala Glu Glu Pro Val Pro Ala
 355 360 365

aag tag 1110
 Lys

<210> 49

<211> 369

<212> PRT

5 <213> Acidovorax facilis 72W

<400> 49

Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
 1 5 10 15

Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

ES 2 564 176 T3

Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
35 40 45

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

Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
65 70 75 80

Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
85 90 95

Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
100 105 110

Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
115 120 125

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

Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
145 150 155 160

Gly Leu Asn Cys Trp Glu His Thr Gln Pro Leu Ser Lys Phe Met Met
165 170 175

Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
180 185 190

Pro Leu Gln Pro Asp Val Phe Gln Leu Ser Ile Glu Ala Asn Ala Thr
195 200 205

Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
210 215 220

Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
225 230 235 240

Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
245 250 255

Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
260 265 270

Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys

ES 2 564 176 T3

atc gca ctc gtc atg ggc tat tcg gag cgg gaa gcc gga tcg cgc tat	336
Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr	
100 105 110	
ctg agc cag gtg ttc atc gac gag cgt ggc gag atc gtt gcc aat cgg	384
Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg	
115 120 125	
cgc aag ctg aag ccc aca cac gtt gag cgt acg atc tac ggc gaa ggc	432
Arg Lys Leu Lys Pro Thr His Val Glu Arg Thr Ile Tyr Gly Glu Gly	
130 135 140	
aac gga acc gat ttc ctc acg cac gac ttc gcg ttc gga cgc gtc ggt	480
Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly	
145 150 155 160	
gga ttg aac tgc tgg gaa cat gtg caa ccg ctc agc aag ttc atg atg	528
Gly Leu Asn Cys Trp Glu His Val Gln Pro Leu Ser Lys Phe Met Met	
165 170 175	
tac agc ctc ggt gag cag gtc cac gtt gca tcg tgg ccg gcg atg tcc	576
Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser	
180 185 190	
cct ctt cag ccg gat gtt ttc caa ctg agc atc gaa gcc aac gcg acg	624
Pro Leu Gln Pro Asp Val Phe Gln Leu Ser Ile Glu Ala Asn Ala Thr	
195 200 205	
gtc acc cgc tcg tac gca atc gaa ggc caa acc ttt gtg ctt tgc tcg	672
Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser	
210 215 220	
acg cag gtg atc gga cct agc gcg atc gaa acg ttc tgc ctc aac gac	720
Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp	
225 230 235 240	
gaa cag cgc gca ctg ttg ccg caa gga tgt ggc tgg gcg cgc att tac	768
Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr	
245 250 255	
ggc ccg gat gga agc gag ctt gcg aag cct ctg gcg gaa gat gct gag	816
Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu	
260 265 270	
ggg atc ttg tac gca gag atc gat ctg gag cag att ctg ctg gcg aag	864
Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys	
275 280 285	
gct gga gcc gat ccg gtc ggg cac tat tcg cgg cct gac gtg ctg tcg	912
Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser	
290 295 300	
gtc cag ttc gac ccg cgc aat cat acg cca gtt cat cgc atc ggc att	960
Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile	
305 310 315 320	
gac ggt cgc ttg gat gtg aat acc cgc agt cgc gtg gag aat ttc cga	1008
Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg	
325 330 335	

ES 2 564 176 T3

ctg cga caa gcg gct gag cag gag cgt cag gca tcc aag cgg ctc gga 1056
 Leu Arg Gln Ala Ala Glu Gln Glu Arg Gln Ala Ser Lys Arg Leu Gly
 340 345 350

acg aaa ctc ttt gaa caa tcc ctt ctg gct gaa gaa ccg gtc cca gca 1104
 Thr Lys Leu Phe Glu Gln Ser Leu Leu Ala Glu Glu Pro Val Pro Ala
 355 360 365

aag tag 1110
 Lys

<210> 51

<211> 369

<212> PRT

5 <213> *Acidovorax facilis* 72W

<400> 51

Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
 1 5 10 15

Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
 35 40 45

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

Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
 65 70 75 80

Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
 85 90 95

Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
 100 105 110

Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
 115 120 125

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

Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
 145 150 155 160

Gly Leu Asn Cys Trp Glu His Val Gln Pro Leu Ser Lys Phe Met Met

ES 2 564 176 T3

<400> 52

atg gtt tcg tat aac agc aag ttc ctc gcg gca acc gtt cag gca gag	48
Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu	
1 5 10 15	
ccg gta tgg ctc gac gca gac gca acg atc gac aag tcg atc ggc atc	96
Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile	
20 25 30	
atc gaa gaa gct gcc caa aag ggc gcg agt ctg atc gct ttc ccg gaa	144
Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu	
35 40 45	
gta ttc att ccg ggc tac ccc tat tgg gcg tgg ctc ggc gac gtg aag	192
Val Phe Ile Pro Gly Tyr Pro Tyr Trp Ala Trp Leu Gly Asp Val Lys	
50 55 60	
tac agc cta agc ttt act tca cgc tat cac gag aat tcg ttg gag cta	240
Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu	
65 70 75 80	
ggt gac gac cgt atg cgt cgc ctc cag ctg gcc gcg cgc cgc aac aaa	288
Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys	
85 90 95	
atc gca ctc gtc atg ggc tat tcg gag cgg gaa gcc gga tcg cgc tat	336
Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr	
100 105 110	
ctg agc cag gtg ttc atc gac gag cgt ggc gag atc gtt gcc aat ccg	384
Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg	
115 120 125	
cgc aag ctg aag ccc aca cac gtt gag cgt acg atc tac ggc gaa ggc	432
Arg Lys Leu Lys Pro Thr His Val Glu Arg Thr Ile Tyr Gly Glu Gly	
130 135 140	
aac gga acc gat ttc ctc acg cac gac ttc gcg ttc gga cgc gtc ggt	480
Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly	
145 150 155 160	
gga ttg aac tgc tgg gaa cat ttc caa ccg ctc agc aag ttc atg atg	528
Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met	
165 170 175	
tac agc ctc ggt gag cag gtc cac gtt gca tcg tgg ccg gcg atg tcc	576
Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser	
180 185 190	
cct ctt cag ccg gat gtt ttc caa ctg agc atc gaa gcc aac gcg acg	624
Pro Leu Gln Pro Asp Val Phe Gln Leu Ser Ile Glu Ala Asn Ala Thr	
195 200 205	
gtc gcc cgc tcg tac gca atc gaa ggc caa acc ttt gtg ctt tgc tcg	672
Val Ala Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser	
210 215 220	

ES 2 564 176 T3

acg cag gtg atc gga cct agc gcg atc gaa acg ttc tgc ctc aac gac 720
 Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
 225 230 235 240

gaa cag cgc gca ctg ttg ccg caa gga tgt ggc tgg gcg cgc att tac 768
 Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
 245 250 255

ggc ccg gat gga agc gag ctt gcg aag cct ctg gcg gaa gat gct gag 816
 Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
 260 265 270

ggg atc ttg tac gca gag atc gat ctg gag cag att ctg ctg gcg aag 864
 Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
 275 280 285

gct gga gcc gat ccg gtc ggg cac tat tcg cgg cct gac gtg ctg tcg 912
 Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
 290 295 300

gtc cag ttc gac ccg cgc aat cat acg cca gtt cat cgc atc ggc att 960
 Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
 305 310 315 320

gac ggt cgc ttg gat gtg aat acc cgc agt cgc gtg gag aat ttc cga 1008
 Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
 325 330 335

ctg cga caa gcg gct gag cag gag cgt cag gca tcc aag cgg ctc gga 1056
 Leu Arg Gln Ala Ala Glu Gln Glu Arg Gln Ala Ser Lys Arg Leu Gly
 340 345 350

acg aaa ctc ttt gaa caa tcc ctt ctg gct gaa gaa ccg gtc cca gca 1104
 Thr Lys Leu Phe Glu Gln Ser Leu Leu Ala Glu Glu Pro Val Pro Ala
 355 360 365

aag tag 1110
 Lys

<210> 53

<211> 369

<212> PRT

5 <213> Acidovorax facilis 72W

<400> 53

Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
 1 5 10 15

Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
 35 40 45

Val Phe Ile Pro Gly Tyr Pro Tyr Trp Ala Trp Leu Gly Asp Val Lys

ES 2 564 176 T3

Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
305 310 315 320

Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
325 330 335

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

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

Lys

<210> 54

<211> 1110

<212> DNA

5 <213> *Acidovorax facilis* 72W

<220>

<221> CDS

<222> (1)..(1110)

<400> 54

atg gtt tcg tat aac agc aag ttc ctc gcg gca acc gtt cag gca gag 48
Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
1 5 10 15

ccg gta tgg ctc gac gca gac gca acg atc gac aag tcg atc ggc atc 96
Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
20 25 30

atc gaa gaa gct gcc caa aag ggc gcg agt ctg atc gct ttc ccg gaa 144
Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
35 40 45

gta ttc att ccg ggc tac ccc tat tgg gcg tgg ctc ggc gac gtg aag 192
Val Phe Ile Pro Gly Tyr Pro Tyr Trp Ala Trp Leu Gly Asp Val Lys
50 55 60

tac agc cta agc ttt act tca cgc tat cac gag aat tcg ttg gag cta 240
Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
65 70 75 80

ggt gac gac cgt atg cgt cgc ctc cag ctg gcc gcg cgc cgc aac aaa 288
Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
85 90 95

atc gca ctc gtc atg ggc tat tcg gag cgg gaa gcc gga tcg cgc tat 336
Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
100 105 110

10

ES 2 564 176 T3

ctg agc cag gtg ttc atc gac gag cgt ggc gag atc gtt gcc aat cgg 384
 Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
 115 120 125

cgc aag ctg aag ccc aca cac gtt gag cgt acg atc tac ggc gaa ggc 432
 Arg Lys Leu Lys Pro Thr His Val Glu Arg Thr Ile Tyr Gly Glu Gly
 130 135 140

aac gga acc gat ttc ctc acg cac gac ttc gcg ttc gga cgc gtc ggt 480
 Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
 145 150 155 160

gga ttg aac tgc tgg gaa cat ttc caa ccg ctc agc aag ttc atg atg 528
 Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met
 165 170 175

tac agc ctc ggt gag cag gtc cac gtt gca tcg tgg ccg gcg atg tcc 576
 Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
 180 185 190

cct ctt cag ccg gat gtt ttc caa ctg agc atc gaa gcc aac gcg acg 624
 Pro Leu Gln Pro Asp Val Phe Gln Leu Ser Ile Glu Ala Asn Ala Thr
 195 200 205

gtc tgc cgc tcg tac gca atc gaa ggc caa acc ttt gtg ctt tgc tcg 672
 Val Cys Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
 210 215 220

acg cag gtg atc gga cct agc gcg atc gaa acg ttc tgc ctc aac gac 720
 Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
 225 230 235 240

gaa cag cgc gca ctg ttg ccg caa gga tgt ggc tgg gcg cgc att tac 768
 Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
 245 250 255

ggc ccg gat gga agc gag ctt gcg aag cct ctg gcg gaa gat gct gag 816
 Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
 260 265 270

ggg atc ttg tac gca gag atc gat ctg gag cag att ctg ctg gcg aag 864
 Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
 275 280 285

gct gga gcc gat ccg gtc ggg cac tat tcg ccg cct gac gtg ctg tcg 912
 Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
 290 295 300

gtc cag ttc gac ccg cgc aat cat acg cca gtt cat cgc atc ggc att 960
 Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
 305 310 315 320

gac ggt cgc ttg gat gtg aat acc cgc agt cgc gtg gag aat ttc cga 1008
 Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
 325 330 335

ctg cga caa gcg gct gag cag gag cgt cag gca tcc aag cgg ctc gga 1056
 Leu Arg Gln Ala Ala Glu Gln Glu Arg Gln Ala Ser Lys Arg Leu Gly
 340 345 350

acg aaa ctc ttt gaa caa tcc ctt ctg gct gaa gaa ccg gtc cca gca 1104

ES 2 564 176 T3

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

aag tag
 Lys

1110

<210> 55

<211> 369

<212> PRT

5 <213> *Acidovorax facilis* 72W

<400> 55

Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
 1 5 10 15

Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
 35 40 45

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

Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
 65 70 75 80

Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
 85 90 95

Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
 100 105 110

Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
 115 120 125

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

Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
 145 150 155 160

Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met
 165 170 175

Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
 180 185 190

ES 2 564 176 T3

Pro Leu Gln Pro Asp Val Phe Gln Leu Ser Ile Glu Ala Asn Ala Thr
 195 200 205

Val Cys Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
 210 215 220

Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
 225 230 235 240

Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
 245 250 255

Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
 260 265 270

Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
 275 280 285

Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
 290 295 300

Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
 305 310 315 320

Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
 325 330 335

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

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

Lys

<210> 56

<211> 1110

<212> DNA

5 <213> Escherichia coli SS1001

<220>

<221> CDS

<222> (1)..(1110)

<400> 56

ES 2 564 176 T3

atg gtt tcg tat aac agc aag ttc ctc gcg gca acc gtt cag gca gag	48
Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu	
1 5 10 15	
ccg gta tgg ctc gac gca gac gca acg atc gac aag tcg atc ggc atc	96
Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile	
20 25 30	
atc gaa gaa gct gcc caa aag ggc gcg agt ctg atc gct ttc ccg gaa	144
Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu	
35 40 45	
gta ttc att ccg ggc tac ccc tat tgg gcg tgg ctc ggc gac gtg aag	192
Val Phe Ile Pro Gly Tyr Pro Tyr Trp Ala Trp Leu Gly Asp Val Lys	
50 55 60	
tac agc cta agc ttt act tca cgc tat cac gag aat tcg ttg gag cta	240
Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu	
65 70 75 80	
ggt gac gac cgt atg cgt cgc ctc cag ctg gcc gcg cgc cgc aac aaa	288
Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys	
85 90 95	
atc gca ctc gtc atg ggc tat tcg gag cgg gaa gcc gga tcg cgc tat	336
Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr	
100 105 110	
ctg agc cag gtg ttc atc gac gag cgt ggc gag atc gtt gcc aat ccg	384
Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg	
115 120 125	
cgc aag ctg aag ccc aca cac gtt gag cgt acg atc tac ggc gaa ggc	432
Arg Lys Leu Lys Pro Thr His Val Glu Arg Thr Ile Tyr Gly Glu Gly	
130 135 140	
aac gga acc gat ttc ctc acg cac gac ttc gcg ttc gga cgc gtc ggt	480
Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly	
145 150 155 160	
gga ttg aac tgc tgg gaa cat ttc caa ccg ctc agc aag ttc atg atg	528
Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met	
165 170 175	
tac agc ctc ggt gag cag gtc cac gtt gca tcg tgg ccg gcg atg tcc	576
Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser	
180 185 190	
cct ctt cag ccg gat gtt ttc caa ctg agc atc gaa gcc aac gcg acg	624
Pro Leu Gln Pro Asp Val Phe Gln Leu Ser Ile Glu Ala Asn Ala Thr	
195 200 205	
gtc acc cgc tcg tac gca atc gaa ggc caa acc ttt gtg ctt tgc tcg	672
Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser	
210 215 220	
acg cag gtg atc gga cct agc gcg atc gaa acg ttc tgc ctc aac gac	720
Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp	
225 230 235 240	
gaa cag cgc gca ctg ttg ccg caa gga tgt ggc tgg gcg cgc att tac	768

ES 2 564 176 T3

Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
 245 250 255

ggc ccg gat gga agc gag ctt gcg aag cct ctg gcg gaa gat gct gag 816
 Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
 260 265 270

ggg atc ttg tac gca gag atc gat ctg gag cag att ctg ctg gcg aag 864
 Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
 275 280 285

gct gga gcc gat ccg gtc ggg cac tat tcg cgg cct gac gtg ctg tcg 912
 Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
 290 295 300

gtc cag ttc gac ccg cgc aat cat acg cca gtt cat cgc atc ggc att 960
 Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
 305 310 315 320

gac ggt cgc ttg gat gtg aat acc cgc agt cgc gtg gag aat ttc cga 1008
 Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
 325 330 335

ctg cga caa gcg gct gag cag gag cgt cag gca tcc aag cgg ctc gga 1056
 Leu Arg Gln Ala Ala Glu Gln Glu Arg Gln Ala Ser Lys Arg Leu Gly
 340 345 350

acg aaa ctc ttt gaa caa tcc ctt ctg gct gaa gaa ccg gtc tca gca 1104
 Thr Lys Leu Phe Glu Gln Ser Leu Leu Ala Glu Glu Pro Val Ser Ala
 355 360 365

aag tag 1110
 Lys

<210> 57
 <211> 369
 <212> PRT

5 <213> Escherichia coli SS1001

<400> 57

Met Val Ser Tyr Asn Ser Lys Phe Leu Ala Ala Thr Val Gln Ala Glu
 1 5 10 15

Pro Val Trp Leu Asp Ala Asp Ala Thr Ile Asp Lys Ser Ile Gly Ile
 20 25 30

Ile Glu Glu Ala Ala Gln Lys Gly Ala Ser Leu Ile Ala Phe Pro Glu
 35 40 45

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

Tyr Ser Leu Ser Phe Thr Ser Arg Tyr His Glu Asn Ser Leu Glu Leu
 65 70 75 80

ES 2 564 176 T3

Gly Asp Asp Arg Met Arg Arg Leu Gln Leu Ala Ala Arg Arg Asn Lys
85 90 95

Ile Ala Leu Val Met Gly Tyr Ser Glu Arg Glu Ala Gly Ser Arg Tyr
100 105 110

Leu Ser Gln Val Phe Ile Asp Glu Arg Gly Glu Ile Val Ala Asn Arg
115 120 125

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

Asn Gly Thr Asp Phe Leu Thr His Asp Phe Ala Phe Gly Arg Val Gly
145 150 155 160

Gly Leu Asn Cys Trp Glu His Phe Gln Pro Leu Ser Lys Phe Met Met
165 170 175

Tyr Ser Leu Gly Glu Gln Val His Val Ala Ser Trp Pro Ala Met Ser
180 185 190

Pro Leu Gln Pro Asp Val Phe Gln Leu Ser Ile Glu Ala Asn Ala Thr
195 200 205

Val Thr Arg Ser Tyr Ala Ile Glu Gly Gln Thr Phe Val Leu Cys Ser
210 215 220

Thr Gln Val Ile Gly Pro Ser Ala Ile Glu Thr Phe Cys Leu Asn Asp
225 230 235 240

Glu Gln Arg Ala Leu Leu Pro Gln Gly Cys Gly Trp Ala Arg Ile Tyr
245 250 255

Gly Pro Asp Gly Ser Glu Leu Ala Lys Pro Leu Ala Glu Asp Ala Glu
260 265 270

Gly Ile Leu Tyr Ala Glu Ile Asp Leu Glu Gln Ile Leu Leu Ala Lys
275 280 285

Ala Gly Ala Asp Pro Val Gly His Tyr Ser Arg Pro Asp Val Leu Ser
290 295 300

Val Gln Phe Asp Pro Arg Asn His Thr Pro Val His Arg Ile Gly Ile
305 310 315 320

ES 2 564 176 T3

Asp Gly Arg Leu Asp Val Asn Thr Arg Ser Arg Val Glu Asn Phe Arg
325 330 335

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

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

Lys

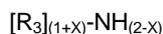
REIVINDICACIONES

1. Un procedimiento para mejorar la actividad específica de un catalizador enzimático cuando se convierte enzimáticamente glicolonitrilo en ácido glicólico, comprendiendo dicho procedimiento:

(a) proporcionar una serie de componentes de reacción que comprenden:

- 5 (i) una solución acuosa de glicolonitrilo que comprende al menos 0,01 ppm de formaldehído;
- (ii) un catalizador enzimático que comprende un polipéptido que tiene actividad de nitrilasa, en el que dicho polipéptido comprende un resto distintivo catalítico de la SEQ ID NO.: 1; en donde dicho catalizador enzimático comprende una actividad específica para hidrolizar glicolonitrilo a ácido glicólico; y
- (iii) un cantidad eficaz de al menos un protector de amina seleccionada del grupo que consiste en:

10 a) un compuesto de la fórmula



en donde X es 0 o 1, y R₃ es independientemente un grupo hidrocarbilo sustituido o grupo hidrocarbilo C1 a C20, en el que R₃ comprende opcionalmente uno o más enlaces éter; con la condición de que (i) el grupo sustituido no sea un grupo ciano capaz de reaccionar con el catalizador enzimático y (ii) R₃ no sea un grupo carbonilo,

- 15 b) un polímero poliaminado que comprende un número eficaz de grupos amino libres en donde dichos grupos amino libres son grupos amino primarios y/o grupos amino secundarios; y
- c) un material con funcionalidad amina que comprende un número eficaz de grupos amino libres en donde dichos grupos amino libres son grupos amino primarios y/o grupos amino secundarios;

20 en donde el protector de amina no se produce naturalmente por dicho catalizador enzimático;

(b) combinar la serie de componentes de reacción bajo condiciones de reacción acuosas adecuadas en las que se produce el ácido glicólico; en donde dicha actividad específica se incrementa por la adición de dicho al menos un protector de amina; y

(c) recuperar el ácido glicólico o la sal del mismo producido en (b);

25 en donde el protector de amina es insoluble bajo dichas condiciones de reacción acuosa adecuadas y, en donde el catalizador enzimático no está inmovilizado en una

2. El procedimiento de la Reivindicación 1 en el que el polipéptido que tiene actividad de nitrilasa comprende el resto distintivo catalítico de la SEQ ID NO: 2.

30 3. El procedimiento de la Reivindicación 1 en el que el polímero poliaminado se selecciona del grupo que consiste en 2-amino-2-desoxi-(1→4)-β-D-glucopirarano, polímeros de polialquilenammina que tienen restos alquileo de 2 a 12 átomos de carbono, polietilenimina, polialilamina, poli(alcohol vinílico)/copolímeros de polivinilamina, D-polilisina, L-polilisina, mezclas de of D/L polilisina, polietilenimina reticulada con glutaraldehído, y mezclas de los mismos, preferiblemente el polímero poliaminado tiene un peso molecular medio que varía de 1.000 Daltons a 2.000.000 Daltons.

35 4. El procedimiento de la Reivindicación 1 en el que el material con funcionalidad amina es un polímero con funcionalidad amina.

5. El procedimiento de la Reivindicación 4, en el que el polímero con funcionalidad amina se selecciona del grupo que consiste en polisacáridos con funcionalidad amina, glicanos con funcionalidad amina, agarosa con funcionalidad amina, carragenina con funcionalidad amina, alginato con funcionalidad amina, dextrano con funcionalidad amina, celulosa con funcionalidad amina, metacrilatos con funcionalidad amina, poliuretanos con funcionalidad amina, poliésteres con funcionalidad amina, nylons con funcionalidad amina, poliestireno con funcionalidad amina, y poli(alcoholes vinílicos) con funcionalidad amina.

6. El procedimiento de la Reivindicación 1, en el que el material con funcionalidad amina es un material de soporte con funcionalidad amina.

45 7. El procedimiento de la Reivindicación 6, en el que el material de soporte con funcionalidad amina se selecciona del grupo que consiste en alúmina con funcionalidad amina, sílice con funcionalidad amina, magnetita con funcionalidad amina, vidrio de poro controlado con funcionalidad amina, resinas de intercambio aniónicos débilmente básicas que comprenden uno o más grupos amina primarios o secundarios, perlas de vidrio aminopropilsilado, ω-aminohexilagarosa, ω-aminododecilagarosa, y ω-aminoetilagarosa.

8. El procedimiento de la Reivindicación 6, en el que el material de soporte con funcionalidad amina está funcionalizado con polietilenimina.
- 5 9. El procedimiento de la Reivindicación 8, en el que el material de soporte con funcionalidad amina funcionalizado con polietilenimina se selecciona del grupo que consiste en sílice con funcionalidad polietilenimina y poliestireno con funcionalidad polietilenimina.
10. El procedimiento de la Reivindicación 1 en el que el polipéptido que tiene actividad de nitrilasa procede de una célula hospedante de un género seleccionado del grupo que consiste en *Acidovorax*, *Rhodococcus*, *Nocardia*, *Bacillus*, y *Alcaligenes*.
- 10 11. El procedimiento de la Reivindicación 1 en el que el catalizador enzimático está en forma de células microbianas enteras, células microbianas permeabilizadas, uno o más componentes celulares de un extracto celular microbiano, enzima parcialmente purificada, o enzima purificada.
- 15 12. El procedimiento de la Reivindicación 11 en el que el catalizador enzimático en una célula hospedante microbiana transformada seleccionada del grupo que consiste en *Comamonas sp.*, *Corynebacterium sp.*, *Brevibacterium sp.*, *Rhodococcus sp.*, *Azotobacter sp.*, *Citrobacter sp.*, *Enterobacter sp.*, *Clostridium sp.*, *Klebsiella sp.*, *Salmonella sp.*, *Lactobacillus sp.*, *Aspergillus sp.*, *Saccharomyces sp.*, *Yarrowia sp.*, *Zygosaccharomyces sp.*, *Pichia sp.*, *Kluyveromyces sp.*, *Candida sp.*, *Hansenula sp.*, *Dunaliella sp.*, *Debaryomyces sp.*, *Mucor sp.*, *Torulopsis sp.*, *Methylobacteria sp.*, *Bacillus sp.*, *Escherichia sp.*, *Pseudomonas sp.*, *Rhizobium sp.*, y *Streptomyces sp.*
- 20 13. El procedimiento de la Reivindicación 12 en el que la célula hospedante microbiana transformada se selecciona del grupo que consiste en *Bacillus sp.*, *Pseudomonas sp.*, y *Escherichia sp.*, preferiblemente la célula hospedante microbiana transformada es *Escherichia coli*.
14. El procedimiento de la Reivindicación 1 en el que dicho polipéptido tiene al menos un 95% de identidad de aminoácidos con la SEQ ID NO: 51.
- 25 15. El procedimiento de la Reivindicación 1 en el que dicho polipéptido que tiene actividad de nitrilasa se selecciona del grupo que consiste en las SEQ ID NOs: 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, y 57, preferiblemente dicho polipéptido que tiene actividad de nitrilasa se selecciona del grupo que consiste en las SEQ ID NOs: 4, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, y 57, más preferiblemente dicho polipéptido que tiene actividad de nitrilasa se selecciona del grupo que consiste en las SEQ ID NOs: 4, 24, 25, y 51, e incluso más preferiblemente dicho polipéptido que tiene actividad de nitrilasa es las SEQ ID NO: 51.

30

SEQ	ID	NO:	IPGYP----	YWAWLGDVKYLS--	F	T	S	R	H	E	N	S	L	E	L	G	D	D	R	R	L	Q	L	A	A	R	R	N	K	I	A	L	V	M	G	Y	S							
SEQ	ID	NO:	LPGYP----	FHVWLGAPAWSLK--	Y	S	A	R	Y	Y	A	N	S	L	S	L	D	S	A	E	F	Q	R	I	A	Q	A	A	R	T	L	G	I	F	I	A	L	G	Y	S				
SEQ	ID	NO:	IPGYP----	YHIWVDSPLAGMAK-	F	A	V	R	Y	H	E	N	S	L	T	M	D	S	P	H	V	Q	R	L	L	D	A	A	R	D	H	N	I	A	V	V	V	G	I	S				
SEQ	ID	NO:	IPGYP----	YWAWIGDVKWAVSD-	F	I	P	K	Y	H	E	N	S	L	T	L	G	D	D	R	M	R	R	L	Q	L	A	A	R	Q	N	N	I	A	L	V	M	G	Y	S				
SEQ	ID	NO:	IPGYP----	YWAWIGDVKWAVSE-	F	I	P	K	Y	H	E	N	S	L	T	L	G	D	D	R	M	R	R	L	Q	L	A	A	R	Q	H	N	I	A	M	V	V	G	Y	S				
SEQ	ID	NO:	VGGYPKGADEFHIFL	GGRTPOGRA-QYQ	R	Y	A	E	T	A	I	A	V	P	G	P	V	T	E	R	I	G	I	A	E	Q	D	M	F	I	V	V	G	V	I									
SEQ	ID	NO:	IGGYPGRFRFGLGV	SVHNEEGRD-EFR	K	Y	H	A	S	A	I	K	V	P	G	P	E	V	E	K	L	A	E	L	A	G	K	N	N	V	I	L	V	M	G	A	I							
SEQ	ID	NO:	LPYYP----	YFSTVEPPVLMGRS-	I	L	K	L	Y	E	Q	A	F	T	M	T	G	P	E	L	Q	I	A	R	A	A	R	Q	I	R	L	T	V	L	L	G	V	N						
SEQ	ID	NO:	LPYYP----	YFSEVEPPVLMGRS-	H	L	K	L	Y	E	Q	A	F	T	M	T	G	P	E	L	Q	I	A	R	A	A	R	Q	H	R	L	F	V	L	L	G	V	N						
SEQ	ID	NO:	VPYYP----	YFSTVEPPVLMGKS-	H	L	K	L	Y	Q	E	A	V	T	V	P	G	K	V	T	Q	I	A	Q	A	A	K	T	H	I	G	M	V	V	L	G	V	N						
SEQ	ID	NO:	LGGYPKGEFGTQLG	YRLPEGRE-AF	A	R	Y	F	A	N	A	I	D	V	P	G	S	E	T	A	L	A	G	L	S	A	R	T	G	A	S	L	V	L	G	V	I							
SEQ	ID	NO:	IGGYPKGLDFGARM	GTRTEAGRE-D	F	L	R	Y	W	K	A	A	I	D	V	P	G	K	E	T	A	R	I	G	S	F	A	A	K	M	K	A	Y	L	V	V	G	V	I					
SEQ	ID	NO:	IPGYP----	WWIWLGNADYGMK-	Y	Y	I	Q	L	Y	K	N	S	V	E	I	P	S	L	A	V	Q	K	L	S	S	A	G	-	T	N	K	V	F	C	V	S	V	T					
SEQ	ID	NO:	IPGYP----	YWIWTSNMDFTGM-	M	W	A	V	L	F	K	N	A	E	I	P	S	K	E	V	Q	I	S	D	A	A	K	N	G	V	Y	V	C	V	S	V	S							
SEQ	ID	NO:	LPYYP----	YFSEVEPPVLMGRS-	H	L	A	L	Y	E	Q	A	V	V	P	G	P	V	T	D	A	V	A	A	A	S	Q	Y	G	M	Q	V	L	L	G	V	N							
SEQ	ID	NO:	LGGYPKGEFGTQLG	YRLPEGRE-AF	A	R	Y	F	A	N	A	I	E	V	P	G	V	E	T	D	A	L	A	L	S	A	R	T	G	A	N	L	V	L	G	V	I							
SEQ	ID	NO:	VPGYP----	WWLWLDSPAWGMQ-	F	V	A	R	Y	F	D	N	S	L	A	L	D	G	P	L	F	A	R	L	R	E	A	A	R	R	S	A	I	T	V	V	T	G	H	S				
SEQ	ID	NO:	LPGYP----	FHVWLGAPAWSLK-	Y	S	A	R	Y	Y	A	N	S	L	S	L	D	S	A	E	F	Q	R	I	A	Q	A	A	R	T	L	G	I	F	I	A	L	G	Y	S				
SEQ	ID	NO:	IPGYP----	WFLWLDAPAWNMP-	L	V	Q	R	Y	H	Q	Q	S	L	V	L	D	S	V	Q	A	R	R	I	S	D	A	A	R	H	L	G	L	Y	V	V	L	G	Y	S				
SEQ	ID	NO:	IPGYP----	WHIWMDSPAWCIG	R	G	F	V	Q	R	Y	F	D	N	S	L	A	Y	D	S	P	Q	A	E	A	L	R	A	A	V	R	K	A	Q	L	T	A	V	L	G	L	S		
SEQ	ID	NO:	IPGYP----	YHIWVDSPLAGMAK-	F	A	V	R	Y	H	E	N	S	L	T	M	D	S	P	H	V	Q	R	L	L	D	A	A	R	D	H	N	I	A	V	V	V	G	I	S				
SEQ	ID	NO:	IPGYP----	YHIWVDSPLAGMAK-	F	A	V	R	Y	H	E	N	S	L	T	M	D	S	P	H	V	Q	R	L	L	D	A	A	R	D	H	N	I	A	V	V	V	G	I	S				

FIGURA 1B


```

SEQ ID NO:4 CWEHFOPLSKFMMYSLGEQVHVHVASWPAMSPLOPDPVFLSIEANATV-----TRSYAIE
SEQ ID NO:5 CWEHLSPELSKYALYSQHEAIIIAAWPFSFSLYSEQAHALSAKVNMAA-----SQIYSVE
SEQ ID NO:6 CWEHFOPLSKYAMYSMHEQVHVHVASWPAGMSLYQPEVPFAGVDAQLTA-----TRMYALE
SEQ ID NO:7 CWEHFOPLSKYMMYSLNEQIHVASWPAMFALTPDVHQLSVEANDTV-----TRSYAIE
SEQ ID NO:8 CWEHFOPLSKYMMYSLNEQIHVASWPAMFALTPDVHQLSVEANDTV-----TRSYAIE
SEQ ID NO:9 CWENYMELLRMAMYGKQIQIYCAPTADDKPTWVSTMQ-----IIVALE
SEQ ID NO:10 CWENRMELRYTALYAKGIELYCAPTADGSKQWQSSML-----HIAIE
SEQ ID NO:11 CWEHYNPEARFSLMTQGEIHCACQFPGSLVGPVIFSEQTAVT-----IRHHALE
SEQ ID NO:12 CWEHYNPEARFSLMTQGEIHCACQFPGSLVGPVIFSEQTAVT-----IRHHALE
SEQ ID NO:13 CWENXNPEARYALMAQHEQIHCQGFPGSMYGVQIFADQMEVT-----MRHHALE
SEQ ID NO:14 CWENYMPLELRTAMYAKGVQLWCAPTVDEREIHWQVSMR-----HVAAE
SEQ ID NO:15 CWENYMPYLRQVMYAGGVNIWCAPTVDQREIHWQVSMR-----HIAYE
SEQ ID NO:16 CWENFLPENVAAMASMNEQVHVHVASWPVIGMPQ--EGHLFGPEQCVT-----TKYYAIS
SEQ ID NO:17 CWENHALPENIAAMGSLNEQVHVHVASWPVFPKGAVSSRVSSVCASTNAMIQIISQFYAIS
SEQ ID NO:18 CWENHYNPEARYALMAQGEIHCACQFPGSLVGPVIFSEQTAVT-----MRHHALE
SEQ ID NO:19 CWENMPELRLRTAMYAQGIEVWCAPTVDEREMWQVSMR-----HIAHE
SEQ ID NO:20 CWEBLQPELTKYAMYSQHEQIHVAAMPFSFVYRGAAYALGPEVNTGA-----ARQYAVE
SEQ ID NO:21 CWEHLSPELSKYALYSQHEAIIIAAWPFSFSLYSEQAHALSAKVNMAA-----SOIYSVE
SEQ ID NO:22 CWEBLQPELSKYAMYAONEQIHVAAMPFSFSLYRNATSALGPEVNTAA-----SRVYAAE
SEQ ID NO:23 CWEBLQPELSKYAMYAONEQVHVHVAAMPFSFSLYDPFAPALGAEVNNAA-----SRVYAVE
SEQ ID NO:24 CWEHFOPLTKYAMYSMHEQVHVHVASWPAGMSLYQPEVPFAGVDAQLTA-----TRMYALE
SEQ ID NO:25 CWEHFOPLTKYAMYSMHEQVHVHVASWPAGMSLYQPEVPFAGVDAQLTA-----TRMYALE
* * * * *

```

FIGURA 1D

SEQ ID NO:4	GQTFVLCSTQVIG-----PSAIEETFCLNDE--QRALLPQCGGWARIYGPDDGSEELAKPLAED
SEQ ID NO:5	GQCFVLAASSVVT-----QETLDMLEVGEH--NASLLKVGSGSMIFAPDGRRTLAPYLPHD
SEQ ID NO:6	GQTFVVCTTQVVT-----PEAHEFFCDNDE--QRKLI GRGGGFARIIGPDGRDLATPLAED
SEQ ID NO:7	GQTFVLAATHVIG-----KATQDLFAGDDDA--KRALLPLGQGWARIYGPDDGKSLAELPPE
SEQ ID NO:8	GQTFVLAATHVIG-----KATQDLFAGDDDA--KRALLPLGQGWARIYGPDDGKSLAELPPE
SEQ ID NO:9	GRCFVLSACQHLRKGKDFPEFHNALDVQP-----DTVLMRGGSCIVDPMGQLLAGPVY-D
SEQ ID NO:10	GGCFVLSACQFCLRKDFPDHPDYLETWYDDKEPDS IVSQGGSVIISPLGQVLAGPNF-E
SEQ ID NO:11	AGCFVLSSTAWLD-----PADYDTITPDRS---LHKAFQGGCHTAIISPEGRYLAGPLP-E
SEQ ID NO:12	AGCFVLSSTAWLD-----PADYDTITPDRS---LHKAFQGGCHTAIISPEGRYLAGPLP-E
SEQ ID NO:13	SGCFVINATGWLTV-----AEQKLOITTTDEK---MHQALSGCCYTAIISPECKHLCEPIA-E
SEQ ID NO:14	GRCFVLSACQVQ---DSPAALGMEVANWPA-----ERPLINGGSLIIVGPLGVDVLAGPLL-G
SEQ ID NO:15	GRLEFVLSACQYMTRADAPADYDCIQNDP-----ETELIAGGSVIIDPMGNILAGPLY-G
SEQ ID NO:16	NQVFCLLSSQIWT-----EEQRDKICETEE--QRNFMKVGHGFSKIIAPNGMEIGNKLAHD
SEQ ID NO:17	NQYVVMSTNLVG-----QDMIDMIGKDEF--SKNFLPLGSGNTAIIISNTG--EILASIPQD
SEQ ID NO:18	AGCFVICSTGWLH-----PDDYASITSESG---LHKAFQGGCHTAIISPEGRYLAGPLP-D
SEQ ID NO:19	GRCFVVSACQVQ---ASPEELGLEIANWPA-----QRPLIAGGSVIIVGPMGDVLAGPLV-G
SEQ ID NO:20	GQCFVLSPCAVID-----EAGVELFCDTPA--KRELLLPGGGFAQIYGPDDGRELGTALPET
SEQ ID NO:21	GQCFVLAASSVVT-----QETLDMLEVGEH--NAPLLKVGSGSMIFAPDGRRTLAPYLPHD
SEQ ID NO:22	GQCFVLAAPCAIVS-----PEMIEMLCSDDA--KRSLAQAGGGHARIFGPDGSDLATPLGEH
SEQ ID NO:23	GSCFVLAAPCATVS-----QAMIDELCDRPD--KHALLHAGGGHAAIFGPDGSSALAAQLPFD
SEQ ID NO:24	GQTFVVCTTQVVT-----PEAHEFFCENEE--QRKLI GRGGGFARIIGPDGRDLATPLAED
SEQ ID NO:25	GQTFVVCTTQVVT-----PEAHEFFCENEE--QRMLI GRGGGFARIIGPDGRDLATPLAED

FIGURA 1E

```

SEQ ID NO:4 AEGILYAEIDLEQILLAKAGADPVGHYSRDPVLSVQDFPRNHTFVHRIGIDGRLLDVNTRS
SEQ ID NO:5 AEGLIADLNMEEIAFAKAINDPVGHYSKPEATRLVLDLGHREPMTRVHSK---SVIQEE
SEQ ID NO:6 EEGILYADIDLSAITLAKQAADPVGHYSRDPVLSLNFNQHRHTTPVN-----TAISTI
SEQ ID NO:7 AEGLLYAELELLEQIILAKAAADPAGHYSRDPVLSLKIDTRNHTPVQYITADGRTSLNSNS
SEQ ID NO:8 AEGLLYAELELLEQIIVAKAAADPAGHYSRDPVLSLKIDTRNHTPVQYVTEDEGSSLNSNS
SEQ ID NO:9 EDAILVADIDLDVTRGKMDFDVVGHYARPDFISLTVDERPKPPVTTL-----K
SEQ ID NO:10 SEGLI TADL DLGDVARAKLYFDSVGHYSRDPVLSLHLLTVNEHPKPVTFI-----S
SEQ ID NO:11 GEGLAIAELDKSLITKRRKMDSVGHYSRDPDLSLRINRSPATQVQAIG-----S
SEQ ID NO:12 GEGLAIAELDKSLITKRRKMDSVGHYSRDPDLSLRINRSPATQVQAIG-----S
SEQ ID NO:13 GEGLAIAIDLFSLIAKRRKMDSVGHYSRDPDLSLQTLNQPWSALEAN-----P
SEQ ID NO:14 ARGLVCAEVDDELVRARYDFDVVGHYARPDFVLSVDERPRPGVR-----F
SEQ ID NO:15 QEGVLVADIDLDLSTIKARYDLDSVGHYSRDPDFELKVDROSHQVITDQ-----F
SEQ ID NO:16 EEGI TYADIDLEQIIPGKFLIDSAGHYSTPGFSLSFDRTEKKPIKHIG-----E
SEQ ID NO:17 AEGIAVAEIDLINQIYYGKWLDPAGHYSRDPGFLSLTFDQSEHIVPVKKIG-----E
SEQ ID NO:18 GEGLAIAIDLALITKRRKMDSVGHYSRPELLSLQINSPPAVPVQNM-----S
SEQ ID NO:19 RAGLISAQIDTADLVRARYDYDVVGHYARPDFVLELTVDQRPVGR-----F
SEQ ID NO:20 EEGLVYADLEASAVAKSAADPVGHYSRDPVLSLQLLWDP---RPRSVVR-----Q
SEQ ID NO:21 AEGLIADLNMEEIAFAKAINDPVGHYSKPEATRLVLDLGHREPMTRVHSK---SVTREE
SEQ ID NO:22 EEGLLYATLDPAAALTLAKVAADPAGHYSRDPVTRLMFNP---NPTPCVV-----D
SEQ ID NO:23 QEGLLIAEIDLGMIGIAKNAADPAGHYSRDPVTRLLINK---KPLNRVE-----H
SEQ ID NO:24 EEGILYADIDLSAITLAKQAADPVGHYSRDPVLSLNFNQRRHTTPVN-----TPLLSTI
SEQ ID NO:25 EEGILYADIDLSAITLAKQAADPVGHYSRDPVLSLNFNQRRHTTPVN-----TPLLSTI

```

FIGURA 1F

SEQ	ID	NO:	RVENFRLRQAAEQERQASKRLGKLFEQS-----LLAEEPVPK---
SEQ	ID	NO:	APEPHVQSTAAPVAVSQTQSDTLLVQEPS-----
SEQ	ID	NO:	HATHTLVQSGALDGVRELNAGADEQRALPS-----THSDETD RATASI
SEQ	ID	NO:	RVENYRLHLADIEKYEAEAA TLPLDAPAP-----APEQKSGRAKAEA
SEQ	ID	NO:	RVENYRLQLADIEKYEENADSATVPLDVTTPEKQSGDVNANGNAKVNTPNSAKAKA
SEQ	ID	NO:	P-----
SEQ	ID	NO:	KVEKAEDDSNK-----
SEQ	ID	NO:	AAALPELPNLEAAPAEAEAYLHA-----
SEQ	ID	NO:	AAALPELPNLEAAPAEAEAYLHA-----
SEQ	ID	NO:	VTPNAIPAVSDPELTETIEALPNNPIFSH-----
SEQ	ID	NO:	IG-----
SEQ	ID	NO:	SRDOATEKKPVSSEISOLD-----
SEQ	ID	NO:	SAQETVTYEEIQYGNKANVKVHS-----
SEQ	ID	NO:	QTNHFISYEDLHEDKMDMLTIPRRVATA-----
SEQ	ID	NO:	TASVPLEPATATDALSSMEALNHV-----
SEQ	ID	NO:	T-----
SEQ	ID	NO:	VA-LSVASPAESAD-----DAEPAVR-----
SEQ	ID	NO:	APEQGVQSKIASVAISHPQSDTLLVQEPS-----
SEQ	ID	NO:	LPDLLPISSESELL-----RPDIALEV-----
SEQ	ID	NO:	FS-LPVDSSAAAALPGEAAVARPDQSI-----
SEQ	ID	NO:	HATHTFVQFGALDGVRELNAGADEQRALPS-----THSDETD RATATL
SEQ	ID	NO:	HATHTFVQFGALDGVRELNAGADEQRALPS-----THSDETD RATATL

FIGURA 1G