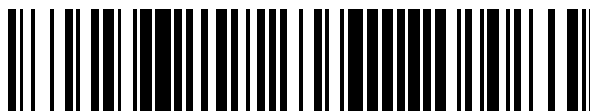


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 730 377**

51 Int. Cl.:

C12N 9/02 (2006.01)

C12N 9/00 (2006.01)

C12N 15/52 (2006.01)

C12P 7/00 (2006.01)

C12P 7/40 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **29.07.2013** **PCT/IL2013/050643**

87 Fecha y número de publicación internacional: **06.02.2014** **WO14020599**

96 Fecha de presentación y número de la solicitud europea: **29.07.2013** **E 13759023 (8)**

97 Fecha y número de publicación de la concesión europea: **17.04.2019** **EP 2880155**

54 Título: **Uso de la ruta reductora de la glicina para generar microorganismos formatótrofos y autótrofos**

30 Prioridad:

29.07.2012 US 201261676962 P

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

11.11.2019

73 Titular/es:

**YEDA RESEARCH AND DEVELOPMENT CO., LTD. (100.0%)
at the Weizmann Institute of Science, P.O. Box 95
7610002 Rehovot, IL**

72 Inventor/es:

**BAR-EVEN, ARREN;
MILO, RON;
NOOR, ELAD y
YISHAI, OREN**

74 Agente/Representante:

PONS ARIÑO, Ángel

Observaciones :

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 730 377 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

Uso de la ruta reductora de la glicina para generar microorganismos formatótrofos y autótrofos

5 Campo y antecedentes de la invención

La presente invención en general se refiere al uso de la ruta reductora de la glicina para la generación de microorganismos formatótrofos y autótrofos.

- 10 El concepto de biorefinerías ha llegado a ser una idea muy extendida en la última década. Depende de la premisa de que los organismos vivos se pueden y se deberían usar para suministrar la demanda creciente por la humanidad de compuestos químicos especializados, incluyendo combustibles, disolventes, plásticos, compuestos farmacéuticos, etc. Hoy en día, la mayoría de estos compuestos químicos están derivados, directa o indirectamente, de carburos fósiles. Sin embargo, con el agotamiento inminente de estos carburos fósiles y el incremento del CO₂ atmosférico ha
15 llegado a ser esencial encontrar fuentes alternativas para estos importantes materiales.

- Las materias primas sugeridas para la mayoría de las biorefinerías propuestas son azúcares simples, almidón, o biomasa lignocelulósica. Aunque la última alternativa tiene una clara ventaja sobre la anterior por no competir con las necesidades de consumo humano, aún presenta numerosas dificultades, incluyendo una tecnología de fermentación problemática y disponibilidad de materia prima y transporte. Una materia prima alternativa fascinante sería la corriente eléctrica. Los electrones se pueden trasladar desde un electrodo a células vivas, proporcionando los
20 equivalentes de reducción necesarios y la energía para soportar el crecimiento autótrofo y la electrosíntesis de los productos deseados (1-5). Puesto que la electricidad está muy disponible, la electrosíntesis microbiana se puede separar espacial y temporalmente de la producción de energía y puede tener lugar en cualquier localización y
25 momento conveniente.

- La electrosíntesis microbiana puede ser especialmente útil para el mercado de la energía renovable. Una desventaja principal de la mayoría de las fuentes de energía renovables, incluyendo la solar, el viento, hidro y nuclear, es que son difíciles de almacenar de una manera conveniente. La electrosíntesis microbiana de combustibles, por lo tanto,
30 puede servir para abordar este problema de manera eficaz, convirtiendo la energía eléctrica en enlaces químicos cinéticamente estables.

- Se sugirieron varios métodos de transferencia de equivalentes reductores de un electrodo a células vivas y se aplicaron (revisado en 1-5). El hidrógeno molecular es uno de los portadores de electrones más antiguos usados de esta manera puesto que la electrolisis del agua es una tecnología relativamente madura que puede soportar la
35 producción de hidrógeno eficiente a alta densidad de corriente. Sin embargo, el uso de hidrógeno tiene numerosos problemas incluyendo su baja solubilidad y el riesgo de explosión. Además, las enzimas hidrogenasa que transfieren los electrones de hidrógeno a los portadores celulares son en general complejas, proteínas sensibles a oxígeno, las cuales son difíciles de expresar de manera recombinante y consumen una fracción significativa de las fuentes celulares. Como alternativa al hidrógeno molecular, varios compuestos inorgánicos, tales como el ion férrico o el
40 nitrato, pueden servir como transportadores de electrones, soportando el cultivo dependiente de electricidad (5). Sin embargo, puesto que los potenciales de reducción de estos compuestos son considerablemente mayores que el de NAD(P)H, el flujo de electrones inverso debe tener lugar durante el crecimiento de estos sustratos, limitando la electrolisis a organismos específicos que son menos adecuados para el uso industrial. Una opción adicional es la transferencia directa de electrones del cátodo a los microbios. Aunque se propusieron varias ventajas de esta opción
45 (revisado en 2-5) este planteamiento está limitado a un pequeño grupo de organismos o requiere adaptación compleja de otros.

- Como alternativa a todos los métodos anteriores, el CO₂ se puede reducir directamente en el cátodo (los electrones se derivan de la división del agua en el ánodo) (6), proporcionando compuestos orgánicos que pueden ser usados por células vivas como fuente de equivalentes reductores, energía e incluso carbono. De esta manera se puede
50 producir un grupo diverso de compuestos (6-9). La producción de alcoholes simples, tales como metanol, etanol y propanol, hidrocarburos, tales como metano y etileno, o ácidos con más de un carbono, tales como ácido acético o ácido oxálico, tiene la ventaja de suministrar microbios con compuestos relativamente simples para metabolizar y/o
55 que son ricos en equivalentes reductores. Sin embargo, la producción electrocatalítica de todos estos compuestos es en general ineficaz (no específica a un producto único y/o que requiere alto sobrepotencial), requiriendo catalizadores costosos y/o soportando baja densidad de corriente (revisado en 6, 7). Por el contrario, hay dos compuestos que se pueden producir por reducción directa de CO₂ a eficacia relativamente alta (aunque menor que la del hidrógeno molecular) y una densidad de corriente aceptable: monóxido de carbono y ácido fórmico (6-13).
60 Puesto que el monóxido de carbono es un gas tóxico e inflamable con baja solubilidad, el ácido fórmico, que es fácilmente soluble y de baja toxicidad, es un mediador preferido de electrones. De hecho, recientemente se propuso una economía basada en formato como alternativa a los conceptos de economía basada en hidrógeno o economía basada en metanol (14-17).

- 65 Diversos organismos metilótrofos pueden crecer sobre formato como una fuente única de carbono, electrones y energía (18-21). Dichos organismos se pueden usar para la electrosíntesis microbiana dependiente de formato (19,

22). Sin embargo, en comparación con los organismos modelo ampliamente usados en la bioindustria, tales como *S. cerevisiae* o *E. coli*, el metabolismo de estos microbios está mucho menos entendido, su cultivo en masa está limitado y su manipulación genética está considerablemente menos optimizada. Como consecuencia, el uso biotecnológico de estos metilótrofos naturales está normalmente limitado a la producción de productos simples.

Sumario de la invención

La presente invención proporciona una bacteria de *Escherichia* aislada que está genéticamente modificada para expresar formato deshidrogenasa dependiente de NAD que es capaz de oxidar el formato a dióxido de carbono; y formato-tetrahidrofolato ligasa.

La presente invención además proporciona un cultivo que comprende esta bacteria de *Escherichia* aislada y un medio que comprende formato.

La presente invención además proporciona un método de generación de una bacteria de *Escherichia* genéticamente modificada que comprende expresar en la bacteria de *Escherichia* formato deshidrogenasa dependiente de NAD que es capaz de oxidar el formato a dióxido de carbono; y formato-tetrahidrofolato ligasa.

A menos que se definan de otra manera, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que el que entiende comúnmente un experto en la técnica a la cual pertenece la invención. Aunque métodos y materiales similares o equivalentes a los descritos en el presente documento pueden usarse en la práctica o en el ensayo de las realizaciones de la invención, a continuación, se describen métodos y materiales ilustrativos. En caso de conflicto, dominará la memoria de patente, que incluye las definiciones. Además, los materiales, métodos y ejemplos son solamente ilustrativos y no se pretende que sean necesariamente limitantes.

Breve descripción de los dibujos

A continuación, con referencia específica a los dibujos en detalle, se insiste que los detalles mostrados son a modo de ejemplo y con fines de discusión ilustrativa de la invención. En este sentido, la descripción tomada con los dibujos hace evidente a los expertos en la técnica cómo se pueden poner en práctica las realizaciones de la invención.

En los dibujos:

La Figura 1 es un diagrama que ilustra la ruta reductora de la glicina, produciendo piruvato.

La Figura 2 es un diagrama que ilustra la ruta reductora de la glicina, produciendo glicerato.

La Figura 3 es un diagrama que ilustra la ruta reductora de la glicina, en la que la glicina - y no serina - se asimila dentro del metabolismo central.

Descripción de las realizaciones específicas de la invención

La presente invención en general se refiere al uso de la ruta reductora de la glicina para la generación de microorganismos y, más particularmente, pero no exclusivamente, a microorganismos formatótrofos y autótrofos.

Después de explicar al menos una realización en detalle, debe entenderse que la invención no se limita necesariamente en su aplicación a los detalles expuestos en la siguiente descripción o ilustrados por los Ejemplos. La invención es capaz de otras realizaciones o de practicarse o llevarse a cabo de diversas maneras.

La electrosíntesis ha recibido recientemente mucha atención para ser un planteamiento prometedor para el uso de una energía renovable para la producción de productos por células vivas. Se propusieron diversas técnicas para mediar la transferencia de electrones del cátodo a las células vivas. De estas, la electroproducción de formato como mediador parece ser especialmente interesante: el formato es fácilmente soluble, de baja toxicidad y se puede producir a alta eficacia y a densidad de corriente razonable.

Hay numerosas rutas metabólicas que, una vez expresadas en un microorganismo, pueden soportar potencialmente el crecimiento formatótrofo. Los presentes inventores aplicaron diversos métodos computacionales para analizar y comparar estas rutas según diversos criterios incluyendo el rendimiento de biomasa, favorabilidad termodinámica, fuerza motriz química, cinéticas y pautas de expresión y encontraron que la ruta reductora de la glicina, compuesta del sistema de tetrahidrofolato, el sistema de escisión de glicina, la serina hidroximetiltransferasa y la serina desaminasa, manifiesta características superiores y es el candidato más prometedor para mediar la electrosíntesis usando bacterias, y más específicamente, *E. coli* como hospedador.

Aunque los organismos que son capaces del crecimiento formatótrofo, es decir, crecimiento sobre formato, existen de manera natural, son en general menos adecuados para la modificación por ingeniería metabólica y el cultivo en masa. Además, no hay indicio de que cualquier organismo use esta ruta para soportar el crecimiento metilótrofo o autótrofo. Los microorganismos que usan la ruta reductora de la glicina tal como se muestra en la Figura 1, la usan para generar un degradado de electrones, reciclando los portadores de electrones reducidos que se generan durante

la fermentación de purinas y aminoácidos (23).

Por lo tanto, se describe un microorganismo aislado que expresa enzimas de la ruta reductora de la glicina, en donde el microorganismo es capaz de convertir el formato en piruvato o glicerato por la formación de glicina y serina.

Tal como se usa en el presente documento, el término "microorganismo" se refiere a cualquier organismo de tamaño microscópico. Ejemplos no limitantes de microorganismos como el término usado en el presente documento incluyen microorganismos tanto procariotas como eucariotas, tales como bacterias, protozoos, hongos, mohos, levaduras, algas etc. El microorganismo puede ser aeróbico o anaeróbico.

El término "aislado" tal como se usa en el presente documento se refiere a un microorganismo que está al menos parcialmente separado del ambiente natural, por ejemplo, de otros microorganismos que no son capaces de usar formato como fuente de carbono, poder reductor y energía (por ejemplo, purificada o semipurificada). Las poblaciones contempladas de microorganismos son unas que están enriquecidas para el microorganismo descrito en el presente documento, por ejemplo, en donde al menos el 30 % de las mismas comprende el microorganismo, al menos el 40 % de las mismas comprende el microorganismo, al menos el 50 % de las mismas comprende el microorganismo, al menos el 60 % de las mismas comprende el microorganismo, al menos el 70 % de las mismas comprende el microorganismo, al menos el 80 % de las mismas comprende el microorganismo, al menos el 90 % de las mismas comprende el microorganismo, al menos el 95 % de las mismas comprende el microorganismo.

Los organismos puede ser organismos fermentativos. Los microorganismos ilustrativos incluyen, por ejemplo, *Clostridium* (por ejemplo, *C. acetobutylicum*, *C. Beijerinckii*, *C. saccharoperbutylacetonicum*, *C. saccharobutylicum*, *C. aurantibutyricum*, *C. tetanomorphum*), *Zymomonas*, *Escherichia* (e.g., *E. coli*), *Salmonella*, *Rhodococcus*, *Pseudomonas*, *Bacillus*, *Lactobacillus*, *Enterococcus*, *Alcaligenes*, *Klebsiella*, *Paenibacillus*, *Arthrobacter*, *Corynebacterium*, *Brevibacterium*, *Pichia*, *Candida*, *Hansenula*, *Zymomonas* y *Saccharomyces*, por ejemplo, *Saccharomyces cerevisiae*, *Saccharomyces carlsbergensis*, *Kluyveromyces lactis*, *Saccharomyces lactis*.

Las bacterias pueden ser gram positivas y gram negativas. Ejemplos de bacterias incluyen, pero sin limitación, *Agrobacterium*, *Alicyclobacillus*, *Anabaena*, *Anacystis*, *Arthrobacter*, *Azobacter*, *Bacillus*, *Brevibacterium*, *Chromatium*, *Clostridium*, *Corynebacterium*, *Enterobacter*, *Erwinia*, *Escherichia*, *Lactobacillus*, *Lactococcus*, *Mesorhizobium*, *Methylobacterium*, *Microbacterium*, *Phormidium*, *Pseudomonas*, *Rhodobacter*, *Rhodopseudomonas*, *Rhodospirillum*, *Rhodococcus*, *Salmonella*, *Scenedesmun*, *Serratia*, *Shigella*, *Staphylococcus*, *Streptomyces*, *Synneccoccus*, y *Zymomonas*.

Ejemplos de hongos incluyen, pero sin limitación, *Aspergillus*, *Candida*, *Chlamydomonas*, *Chrysosporium*, *Cryptococcus*, *Fusarium*, *Kluyveromyces*, *Neotyphodium*, *Neurospora*, *Penicillium* (por ejemplo, *P. chrysogenum*), *Pichia*, *Saccharomyces*, *Trichoderma* y *Xanthophyllomyces*.

Ejemplos de algas incluyen, pero sin limitación, una diatomea o una cianobacteria.

La diatomea puede ser una microalga de la clase Coscinodiscophyceae, Fragilariophyceae o Bacillariophyceae.

La cianobacteria puede incluir, por ejemplo, *Botryococcus braunii*, *Chlorella*, *Dunaliella tertiolecta*, *Gracilaria*, *Pleurochrysis carterae*, *Sargassum* o *Ulva*.

El microorganismo puede ser metilótrofo - es decir, es capaz de crecer sobre compuestos C1 orgánicos como su fuente única de carbono, poder reductor y energía.

El microorganismo metilótrofo puede ser formatótrofo - es decir, usa formato como una fuente de carbono, poder reductor y energía. Se apreciará que los organismos formatótrofos de la presente invención también pueden ser capaces de crecer sobre fuentes de carbono adicionales - tales como glucosa, glicerol celulosa, acetato, butirato, lactato, propionato o valerato.

El formatótrofo puede usar formato como su fuente única de carbono.

El microorganismo puede ser autótrofo. Un tipo de microorganismo autótrofo es un organismo fotótrofo (uno que requiere luz para conseguir el poder reductor y la energía para la fijación de dióxido de carbono). La fuente de electrones para la fotosíntesis puede ser, por ejemplo, agua, sulfuro de hidrógeno, azufre elemental o ion de hierro (Fe^{2+}).

Otro tipo de microorganismo autótrofo es un microorganismo quimiótrofo (uno que requiere una fuente de electrones externa [por ejemplo, hidrógeno molecular, monóxido de carbono (CO), sulfuro de hidrógeno (H_2S), azufre elemental (S), sulfito (SO_3^{2-}), fosfito (PO_3^{2-}), amoníaco (NH_4^+), nitrato (NO_2^{2-}), hidróxido de amonio (NH_2OH), ion ferroso (Fe^{2+}), ion de Mn^{2+}] para conseguir el poder reductor y la energía para la fijación del dióxido de carbono). Posibles receptores de electrones terminales incluyen oxígeno molecular, dióxido de carbono (CO_2), sulfato (SO_4^{2-}), azufre elemental (S), nitrato (NO_3^{2-}), ion férrico (Fe^{3+}).

Tal como se menciona los microorganismos expresan, o bien de manera natural o están modificados por ingeniería genética para expresar, enzimas de la ruta reductora de la glicina de modo que el microorganismo es capaz de convertir el formato en piruvato por la formación de glicina y serina (Figura 1) o es capaz de convertir formato en glicerato por la formación de glicina y serina (Figura 2). Una alternativa adicional es usar las enzimas de la ruta reductora de la glicina de modo que el microorganismo sea capaz de convertir el formato en acetil-CoA, glicerato u oxaloacetato por la formación de glicina, pero sin la formación de serina (Figura 3).

El microorganismo puede usar la enzima serina desaminasa para la conversión de la serina (generada por la ruta reductora de la glicina, Figura 1) en piruvato. Como alternativa, el microorganismo puede usar una enzima serina transaminasa (o serina deshidrogenasa) e hidroxipiruvato reductasa (o hidroxipiruvato isomerasa y tartronato semialdehído reductasa) para la conversión de la serina (generada por la ruta reductora de la glicina) en glicerato (Figura 2). Como alternativa, el microorganismo puede asimilar directamente glicina dentro del metabolismo central sin producir serina, ta como se muestra en la Figura 3.

La Figura 1 ilustra las enzimas de la ruta reductora de la glicina en la que el producto es piruvato. La ruta está compuesta del sistema de tetrahidrofolato, el sistema de escisión de glicina, la serina hidroximetiltransferasa y la serina desaminasa.

Por lo tanto, los microorganismos pueden expresar formato tetrahidrofolato ligasa (EC 6.3.4.3, *THFS*); meteniltetrahidrofolato ciclohidrolasa (EC 3.5.4.9, *folD*); metilenotetrahidrofolato deshidrogenasa (EC 1.5.1.5/15, *folD*); aminometiltransferasa (EC 2.1.2.10, *gcvT*); dihidrolipoil deshidrogenasa (EC1.8.1.4, *lpd*); glicina deshidrogenasa (descarboxilación; EC 1.4.4.2, *gcvP*); serina hidroximetiltransferasa (EC2.1.2.1, *glyA*); y L-serina desaminasa (EC 4.3.1.17, *sdaA*).

El microorganismo puede ser capaz de convertir formato en piruvato por una actividad de la serina desaminasa o el microorganismo puede ser capaz de convertir formato en glicerato por una actividad de una enzima serina desaminasa (por ejemplo, EC 2.6.1.45, 2.6.1.51) e hidroxipiruvato reductasa (EC 1.1.1.82 o 1.1.1.29).

Las enzimas meteniltetrahidrofolato ciclohidrolasa y metilenotetrahidrofolato deshidrogenasa se pueden reemplazar por una enzima bifuncional meteniltetrahidrofolato ciclohidrolasa/metilenotetrahidrofolato deshidrogenasa. Además, la enzima formato tetrahidrofolato ligasa y la enzima bifuncional meteniltetrahidrofolato ciclohidrolasa/metilenotetrahidrofolato deshidrogenasa se pueden reemplazar por una enzima trifuncional que porta actividades de formil-THF sintetasa, meteniltetrahidrofolato ciclohidrolasa y metilenotetrahidrofolato deshidrogenasa.

Se apreciará que la enzima formato tetrahidrofolato ligasa se puede evitar haciendo funcionar enzimas cuya reacción neta es idéntica a la de formato tetrahidrofolato ligasa (por ejemplo, el funcionamiento secuencial de *purT* (glicinamida ribonucleótido transformilasa dependiente de formato) en la dirección directa y *PurN* (glicinamida ribonucleótido transformilasa dependiente de formil-THF) en la dirección inversa; otro ejemplo es el funcionamiento secuencial de *purP* (aminoimidazol carboxamida ribonucleótido transformilasa dependiente de formato) en la dirección directa y *PurH* (aminoimidazol carboxamida ribonucleótido transformilasa dependiente de formil-THF) en la dirección inversa).

También se puede expresar la proteína H del sistema de escisión de glicina. Además, también se pueden expresar las enzimas que convierten el ácido octanoico en ácido lipico unido a la proteína H (por ejemplo LipA, LipB y/o LplA).

Se apreciará que para la generación de microorganismos formatótrofos, también es necesario la expresión de una formato deshidrogenasa (EC 1.2.1.2 o 1.2.1.43; por ejemplo, GenBank CAB54834.1 - SEQ ID NO: 12) que es capaz de oxidar el formato para obtener dióxido de carbono. Los electrones que se liberan sirven como el poder reductor requerido para el funcionamiento de la ruta reductora de la glicina y para generar ATP (por ejemplo, por fosforilación oxidativa).

Para la generación de microorganismos autótrofos, también es necesario la expresión de una formato deshidrogenasa que sea capaz de reducir dióxido de carbono a ácido fórmico (EC 1.2.1.2, 1.2.1.43 o 1.1.99.33; por ejemplo GenBank AAB18330.2 (SEQ ID NO: 13) y AAB18329.1 (SEQ ID NO: 14).

También se pueden expresar enzimas adicionales en los microorganismos autótrofos dependiendo de la fuente de electrones externa. Por lo tanto, por ejemplo, si la fuente de electrones es hidrógeno, se contempla en general la expresión de una enzima hidrogenasa. Detalles adicionales sobre la expresión de la hidrogenasa en células *E. coli* se pueden encontrar en [24, 25].

La Figura 2 ilustra las enzimas de la ruta reductora de la glicina en la que el producto es glicerato.

Por lo tanto, los microorganismos pueden expresar formato tetrahidrofolato ligasa (EC 6.3.4.3, *THFS*); meteniltetrahidrofolato ciclohidrolasa (EC 3.5.4.9, *folD*); metilenotetrahidrofolato deshidrogenasa (EC 1.5.1.5/15, *folD*); aminometiltransferasa (EC 2.1.2.10, *gcvT*); dihidrolipoil deshidrogenasa (EC 1.8.1.4, *lpd*); serina

hidroximetiltransferasa (EC2.1.2.1, *glyA*); serina deshidrogenasa (EC2.6.1.45) o serina-glioxilato transaminasa (EC 2.6.1.51) o cualquier otra enzima transaminasa (EC 2.6.1X); e hidroxipiruvato reductasa (EC 1.1.1.26/19/79/81) o hidroxipiruvato isomerasa (EC 5.3.1.22) + tartronato semialdehído reductasa (EC 1.1.1.60).

- 5 La Figura 3 ilustra las enzimas de la ruta reductora de la glicina en la que los productos son metabolitos del metabolismo central. En estos microorganismos, la enzima serina hidroximetiltransferasa (EC2.1.2.1, *glyA*) no se expresa.

- 10 Por lo tanto, los microorganismos pueden expresar formato tetrahidrofolato ligasa (EC 6.3.4.3, *THFS*); meteniltetrahidrofolato ciclohidrolasa (EC 3.5.4.9, *folD*); metilenotetrahidrofolato deshidrogenasa (EC 1.5.1.5/15, *folD*); aminometiltransferasa (EC 2.1.2.10, *gcvT*); dihidrolipoil deshidrogenasa (EC1.8.1.4, *lpd*). Cuando se produce acetil-CoA como metabolito central, el microorganismo expresa la enzima glicina reductasa (EC 1.4.1.10) y fosfato acetiltransferasa EC 2.3.1.8. Como alternativa, la glicina se puede convertir en glioxilato y antes de ser asimilada dentro del metabolismo central (tal como se muestra en la Figura 3). Esto requiere la expresión de glicina deshidrogenasa (EC 1.4.1.10) o una transaminasa (EC 2.6.1.51). Cuando se produce oxaloacetato como metabolito, el microorganismo expresa enzimas EC 4.1.3.14 y 4.3.1.20. Como alternativa, la glicina se puede convertir en oxaloacetato por el ciclo reductor TCA y el ciclo del glioxilato (tal como se muestra en la Figura 3). Cuando se producen glicerato 2-fosfato o glicerato 3-fosfato, se expresan las enzimas EC4.1.1.47, EC1.1.1.60 y EC 2.7.1.31/165. Tal como se muestra en la Figura 3, en lugar de EC1.1.1.60, se pueden expresar las enzimas EC 5.3.1.22 y EC 1.1.1.26/29/79/81.

El término "enzima" tal como se usa en el presente documento se refiere a una "biomolécula catalíticamente funcional", que incluye tanto moléculas nativas (o de tamaño nativo) como derivados (por ejemplo, modificaciones genéticas) de la misma.

- 25 Por lo tanto, una enzima también se refiere a homólogos y otras modificaciones incluyendo adiciones o deleciones de aminoácidos específicos a la secuencia (por ejemplo, polipéptidos que son al menos 50 %, al menos 55 %, al menos 60 %, al menos 65 %, al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 87 %, al menos 89 %, al menos 91 %, al menos 93 %, al menos 95 % o más es decir 100 % homólogos a las secuencias de aminoácidos lisosomales enumeradas en la Tabla 1, en el presente documento más adelante tal como se determina usando el programa informático BlastP del "National Center of Biotechnology Information" (NCBI) usando los parámetros por defecto). El homólogo también se puede referir a un ortólogo, una deleción, inserción, o variante de sustitución, incluyendo una sustitución de aminoácidos, de la misma y fragmentos de polipéptidos biológicamente activos de la misma.

Tabla 1

Enzima	Ejemplos de registro GenBank de proteína
formato tetrahidrofolato ligasa (EC 6.3.4.3);	ABY28836.1 - SEQ ID NO: 1
meteniltetrahidrofolato ciclohidrolasa/metilenotetrahidrofolato deshidrogenasa bifuncional (EC 3.5.4.9 y EC 1.5.1.5/15, <i>folD</i>)	NP_006627.2 - SEQ ID NO: 2
aminometiltransferasa (EC 2.1.2.10, <i>gcvT</i>); y.	AAN81933.1 - SEQ ID NO: 3
dihidrolipoil deshidrogenasa (EC 1.8.1.4, <i>lpd</i>)	ZP_03035757.1 - SEQ ID NO: 4
glicina deshidrogenasa (descarboxilación; EC 1.4.4.2, <i>gcvP</i>)	YP_005276541.1 - SEQ ID NO: 5
serina hidroximetiltransferasa (EC 2.1.2.1, <i>glyA</i>)	ZP_03034169.1 - SEQ ID NO: 6
L-serina desaminasa; EC 4.3.1.17, <i>sdaA</i>)	AFH11608.1 - SEQ ID NO: 7
serina-piruvato transaminasa (EC 2.6.1.51)	YP_004171157.1 - SEQ ID NO: 8
hidroxipiruvato reductasa (EC 1.1.1.81, <i>ghrA</i>)	ZP_04001524.1 - SEQ ID NO: 9
serina-glioxilato transaminasa (EC 2.6.1.45)	YP_004370315.1 - SEQ ID NO: 10
serina deshidrogenasa (EC 1.4.1.7)	BAB07807.1 - SEQ ID NO: 11

- 40 Las secuencias de ácidos nucleicos que codifican las enzimas se pueden optimizar para la expresión para un microorganismo particular. Ejemplos de dichas modificaciones de secuencia incluyen, pero sin limitación, un contenido de G/C alterado para acercarse mucho más al generalmente encontrado en la especie de microorganismo de interés, y la retirada de codones encontrados de manera atípica en la especie de microorganismo comúnmente referido como optimización de codón.

- 45 La expresión "optimización de codón" se refiere a la selección de nucleótidos de ADN apropiados para su uso dentro de un gen estructural o fragmento del mismo que aborda el uso de codón dentro del microorganismo de interés. Por tanto, un gen optimizado o secuencia de ácidos nucleicos se refiere a un gen en el que la secuencia de nucleótidos de un gen nativo o de origen natural se ha modificado para utilizar codones estadísticamente preferidos o estadísticamente favoritos dentro del microorganismo. La secuencia de nucleótidos generalmente se examina al

nivel de ADN y la región codificante optimizada para la expresión en la especie de microorganismo determinada usando cualquier procedimiento adecuado, por ejemplo tal como se describe en Sardana et al. (1996, *Plant Cell Reports* 15:677-681). En este método, la desviación típica del uso de codón, una medida del sesgo de uso de codón, se puede calcular primero encontrando la desviación proporcional cuadrática del uso de cada codón del gen nativo en relación a la de los genes altamente expresados, seguido por un cálculo de la desviación cuadrática media. La fórmula usada es: $1 \text{ SDCU} = n = 1 N[(X_n - Y_n)/Y_n]^2/N$, en donde X_n se refiere a la frecuencia de uso de codón n en genes altamente expresados, en donde Y_n a la frecuencia de uso de codón n en el gen de interés y N se refiere al número total de codones en el gen de interés.

Un método de optimización de la secuencia de ácidos nucleicos de acuerdo con el uso de codón preferido para un tipo de célula particular se basa en el uso directo, sin realización de ningún cálculo estadístico extra, de tablas de optimización de codón tales como aquellas proporcionadas on-line en la Base de Datos de Uso de Codón del banco de ADN del NIAS ("National Institute of Agrobiological Sciences") en Japón (worldwidewebdotkazusadotordotjp/codon/). La Base de Datos de Uso de Codón contiene las tablas de uso de codón para un número de diferentes especies, habiendo sido cada tabla de uso de codón determinada estadísticamente basándose en los datos presentes en Genbank.

Usando las tablas anteriores para determinar los codones más preferidos o más favoritos para cada aminoácido en una especie particular (por ejemplo, *E. coli*), una secuencia de nucleótidos de origen natural que codifica una proteína de interés se puede someter a optimización de codón para esa especie particular. Esto se efectúa reemplazando codones que pueden tener una baja incidencia estadística en el genoma de la especie particular con los codones correspondientes, con respecto a un aminoácido, que son estadísticamente más favoritos. Sin embargo, se pueden seleccionar uno o más codones menos favoritos para delecionar los sitios de restricción existentes, para crear unos nuevos en uniones potencialmente útiles (extremos 5' y 3' para añadir péptido señal o casetes de terminación, sitios internos que se podrían usar para cortar y empalmar segmentos juntos para producir una secuencia de longitud completa correcta), o para eliminar las secuencias de nucleótidos que pueden afectar negativamente a la estabilidad o expresión del ARNm.

La secuencia de nucleótidos codificantes de origen natural puede ya, con antelación a cualquier modificación, contener un número de codones que corresponden a un codón estadísticamente favorito en una especie particular. Por tanto, la optimización de codón de la secuencia de nucleótidos nativa puede comprender determinación de aquellos codones, dentro de la secuencia de nucleótidos nativa, que no son estadísticamente favoritos con respecto a la planta particular, y modificar estos codones de acuerdo con una tabla de uso de codón de la especie particular para producir un derivado optimizado por codón. Una secuencia de nucleótidos modificada se puede optimizar completa o parcialmente para el uso de codón de microorganismo siempre que la proteína codificada por la secuencia de nucleótidos modificada se produzca a un nivel mayor que la proteína codificada por el correspondiente gen de origen natural o nativo. La construcción de los genes sintéticos alterando el uso de codón se describe en, por ejemplo, la Solicitud de Patente PCT 93/07278.

Para expresar las enzimas usando tecnología recombinante, polinucleótidos que codifican las enzimas se pueden ligar en un vector de expresión de ácido nucleico, bajo el control transcripcional de una secuencia reguladora en cis (por ejemplo, secuencia promotor) adecuada para dirigir la transcripción constitutiva o inducible de las enzimas en el microorganismo.

La expresión "un polinucleótido aislado" se refiere a una secuencia de ácidos nucleicos monocatenaria o bicatenaria que está aislada y se proporciona en la forma de una secuencia de ARN, una secuencia de polinucleótidos complementaria (ADNc), una secuencia de polinucleótidos genómica y/o una secuencia de polinucleótidos compuesta (por ejemplo, una combinación de las anteriores).

Tal como se usa en el presente documento la expresión "secuencia de polinucleótidos complementaria" se refiere a una secuencia, la cual resulta de la transcripción inversa del ARN mensajero usando una transcriptasa inversa o cualquier otra ADN polimerasa dependiente de ARN. Dicha secuencia se pueden amplificar posteriormente *in vivo* o *in vitro* usando una ADN polimerasa dependiente de ADN.

Tal como se usa en el presente documento la expresión "secuencia de polinucleótidos genómica" se refiere a una secuencia derivada (aislada) de un cromosoma y, por lo tanto, representa una porción contigua de un cromosoma.

Tal como se usa en el presente documento la expresión "secuencia de polinucleótidos compuesta" se refiere a una secuencia, la cual es al menos parcialmente complementaria y al menos parcialmente genómica. Una secuencia compuesta puede incluir algunas secuencias de exón requeridas para codificar el polipéptido, así como algunas secuencias intrónicas interpuestas entre las mismas. Las secuencias intrónicas pueden ser de cualquier fuente, incluyendo de otros genes, y generalmente incluirán secuencias de señal de corte y empalme conservadas. Dichas secuencias intrónicas además pueden incluir elementos reguladores de la expresión que actúan en cis.

El vector de expresión incluye secuencias adicionales que interpretan este vector adecuado para la replicación e integración en los procariotas, eucariotas, o preferentemente ambos (por ejemplo, vectores lanzaderas). Vectores

clonantes normales contienen secuencias de iniciación de la transcripción y la traducción (por ejemplo, promotores, potenciadores) y terminadores de la transcripción y la traducción (por ejemplo, señales de poliadenilación).

Se pueden usar diversos métodos para introducir el vector de expresión dentro del sistema celular hospedador. Tales métodos se describen en general en Sambrook et al., "Molecular Cloning: A Laboratory Manual", Cold Springs Harbor Laboratory, Nueva York (1989 1992), en Ausubel et al., "Current Protocols in Molecular Biology", John Wiley and Sons, Baltimore, Md. (1989), Chang et al., "Somatic Gene Therapy", CRC Press, Ann Arbor, Mich. (1995), Vega et al., "Gene Targeting", CRC Press, Ann Arbor Mich. (1995), "Vectors: A Survey of Molecular Cloning Vectors and Their Uses", Butterworths, Boston Mass. (1988) y Gilboa et al. [*Biotechniques* 4 (6): 504-512, 1986] e incluyen, por ejemplo, transfección estable o transitoria, lipofección, electroporación e infección con vectores virales recombinantes. Además, véanse los documentos de Patente de EE.UU. N.º 5.464.764 y 5.487.992 para métodos de selección positivo-negativo.

Sistemas de expresión basados en bacterias ilustrativos se divulgan en Baneyx et al., *Current Opinion in Biotechnology*, 1999; 10, 411-421 y Macrides et al, *Microbiol Rev* 1996, 60: 512-538.

Los microorganismos se pueden transformar estable o transitoriamente con las construcciones de ácidos nucleicos. En la transformación estable, la molécula de ácido nucleico se integra en el genoma del microorganismo y como tal representa un rasgo estable y hereditario. En la transformación transitoria, la molécula de ácido nucleico se expresa por la célula transformada pero no está integrada en el genoma y como tal representa un rasgo transitorio.

En general está contemplado usar secuencias de polinucleótidos que codifican las enzimas de la ruta reductora de la glicina de cualquier organismo - por ejemplo, secuencias humanas, secuencias vegetales, secuencias bacterianas, secuencias fúngicas, secuencias de levadura.

Se apreciará que el número de enzimas adicionales que tienen que ser expresadas exógenamente en un microorganismo particular dependerá de las enzimas que se expresan de manera natural en ese tipo natural sobre la localización subcelular del mismo.

Por lo tanto, en el presente caso de *E. coli*, se usan medios recombinantes para expresar:

1. Formato deshidrogenasa dependiente de NAD que es capaz de oxidar formato a dióxido de carbono; y
2. Formato-tetrahidrofolato ligasa.

Las células transformadas se cultivan bajo condiciones eficaces, que permiten la expresión de altas cantidades de las enzimas recombinantes. Las condiciones de cultivo eficaces incluyen, pero sin limitación, medios eficaces, biorreactor, temperatura, condiciones de pH y oxígeno que permiten la producción de la proteína. Un medio eficaz se refiere a cualquier medio en el que una célula se cultiva para producir el polipéptido recombinante. Dicho medio generalmente incluye una solución acuosa que tiene fuentes de carbono asimilables, de nitrógeno y fosfato, y sales apropiadas, minerales, metales y otros nutrientes, tales como vitaminas. Las células se pueden cultivar en biorreactores de fermentación convencionales, matraces de agitación, tubos de ensayo, placas de microtitulación y placas Petri. El cultivo se puede llevar a cabo a una temperatura, pH y contenido de oxígeno apropiados para una célula recombinante. Dichas condiciones de cultivo pertenecen a la experiencia de los expertos en la técnica.

Se apreciará que hasta el microorganismo se transforma de ser un microorganismo heterótrofo a un microorganismo formatótrofo o microorganismo autótrofo, el microorganismo se cultiva en un medio de cultivo que comprende otras fuentes de carbono tales como glucosa o glicerol.

Después de la generación de los microorganismos tal como se describe en el presente documento, preferentemente se seleccionan por crecimiento (es decir, cultivo) sobre un sustrato particular. Preferentemente, los microorganismos se cultivan durante al menos un día, al menos dos días, al menos tres días, al menos una semana, al menos un mes, al menos tres meses en donde algunas células viables restantes después de dicho tiempo son el microorganismo seleccionado.

Por lo tanto, por ejemplo, en el caso de generar un microorganismo formatótrofo, el microorganismo se debería cultivar en un medio de cultivo que comprenda formato como fuente de carbono. En el caso de un microorganismo autótrofo, el microorganismo se debería cultivar en un medio de cultivo que comprenda dióxido de carbono como fuente de carbono y una fuente de electrones externa tal como se ha descrito además anteriormente en el presente documento.

El formato que se usa puede venir de cualquier fuente - por ejemplo, formato de sodio, formato de potasio, ácido fórmico o anhídrido de ácido fórmico etc.

Como alternativa, y/o además, el formato se puede generar usando electricidad. El CO₂ se puede reducir directamente en el cátodo (los electrones se derivan de la división del agua en el ánodo) para generar formato a eficacia relativamente alta.

Para generar el formato para su uso por el microorganismo, el microorganismo se coloca en un biorreactor en un fluido (por ejemplo, agua). El cátodo se puede colocar opcionalmente dentro del biorreactor en contacto con el microorganismo. Como alternativa, el cátodo se puede colocar en un recipiente separado al biorreactor y el formato se puede encauzar a la cámara que comprende el microorganismo. El fluido puede contener otros elementos requeridos por el microorganismo para el crecimiento incluyendo, por ejemplo, sales, minerales, metales y otros nutrientes, tales como vitaminas.

Ejemplos de dichos biorreactores y métodos adicionales se proporcionan en Li et al. *Science*, 2012, Vol. 335, página 1.596; Rabaey et al, *Current Opinion in Biotechnology*, 2011, 22: 371-377; Lovley et al., *Current Opinion in Biotechnology*, 2011, 22: 441-448; Lovley D.R., *Environmental microbiology reports*, 2011, 3(1), 27-35; Nevin et al., *Microbiology*, mayo/junio de 2010 Volumen 1 artículo 2; Rabaey et al, *Applied and Industrial Microbiology, Nature Reviews*, octubre de 2010, Volumen 8, páginas 706-716.

Los electrodos se pueden fabricar a partir de dichos polímeros conductivos y materiales metálicos incluyendo óxido de indio y estaño (ITO), grafito, platino y plata.

Por lo tanto, se contempla un sistema para el microorganismo descrito en el presente documento y un electrodo para proporcionar electrones para generar formato. El sistema además puede comprender mecanismo(s) para la separación, recolección, y/o recuperación del biocombustible que es generado por el microorganismo (como se detalla más, a continuación).

Para confirmar que el microorganismo formatótrofo o autótrofo está usando la ruta reductora de la glicina para la generación de piruvato o glicerato, se contempla el análisis de los metabolitos de la ruta reductora de la glicina.

Dichos metabolitos incluyen 10-formil-tetrahidrofolato, 5,10-meteniltetrahidrofolato, 5,10-metilenotetrahidrofolato, aminometil-dihidrolipoilproteína, dihidrolipoilproteína, lipoilproteína, glicina, serina y piruvato (o hidroxipiruvato y glicerato).

Preferentemente, el análisis comprende el análisis de marcación con ^{13}C o ^{14}C durante el tiempo tomado para producir cada metabolito marcado a partir del formato marcador, de modo que es evidente que la producción de 10-formil-tetrahidrofolato precede la de 5,10-metileno-tetrahidrofolato, de modo que la producción de 5,10-metilenotetrahidrofolato precede la de aminometildihidrolipoil, de modo que la producción de aminometil-dihidrolipoil precede la de glicina, de modo que la producción de glicina precede la de serina y la producción de serina precede la de piruvato.

Preferentemente, la formación de cada uno de los metabolitos precede a la siguiente en la cadena por al menos un segundo, al menos 10 segundos y lo más preferentemente al menos 20 segundos.

Un método ilustrativo para analizar el ritmo de la producción de metabolitos es por análisis de pulso y caza ("pulse-chase"). Un análisis de pulso-caza es un método para examinar un proceso celular que se da a lo largo del tiempo exponiendo sucesivamente las células a un compuesto marcado (pulso) y, a continuación, al mismo compuesto en una forma no marcada (caza). La radioactividad es un marcador comúnmente usado.

En un método ilustrativo, los microorganismos primero se exponen a formato marcado (el pulso). A continuación, el formato marcado atraviesa las rutas metabólicas y se usa en la síntesis de piruvato. Poco después de la introducción del formato marcado (normalmente aproximadamente 5 minutos), se introduce exceso del mismo formato, pero no marcado, (la caza) en el ambiente. La producción de piruvato continuará, pero no contendrá más el marcador radioactivo del formato introducido en la fase de pulso y no será visible usando métodos de detección radioactiva.

En otro método ilustrativo, los compuestos especificados anteriormente se pueden analizar para encontrar cuál de sus átomos de carbono está marcado. Si en realidad se producen por la ruta reductora de la glicina radioactiva se espera un patrón de marcación de carbono indicativo.

La cantidad y/o la actividad de las enzimas de la ruta se pueden analizar en el nivel de ARN o proteína usando métodos conocidos en la técnica. Por lo tanto, en el nivel de ARN, métodos que incluyen RT-PCR, análisis de transferencia Northern, micromatriz de oligonucleótidos. En el nivel de proteína, métodos que incluyen ELISA, transferencia Western, radioinmunoensayo, ensayo de actividad *in situ*, ensayo de actividad *in vitro* y análisis inmunohistoquímico todos están contemplados. Un indicio del funcionamiento de la ruta es si todas las enzimas de la ruta muestran actividad suficiente para soportar el crecimiento celular a la tasa que ha sido mostrada para crecer.

Métodos adicionales para analizar la producción de los metabolitos de la ruta reductora de la glicina se encuentran en Hugler et al., 2005, *Methods In Enzymology*, Vol. 397, pág. 212-221; Berg et al., *Science* 318, 1.782 (2007); Strauss et al., *Eur. J. Biochem.* 205, 853-866 (1992); Hertner et al., *Journal Of Bacteriology*, Vol. 183, N.º 14, julio de 2001, pág. 4.305-4.316; Jahn et al., *Journal Of Bacteriology*, junio de 2007, pág. 4.108-4.119 Vol. 189, N.º 11; y Huber et al., *PANS*, 3 de junio de 2008, Vol. 105, N.º 22, 7.851-7.856.

De acuerdo con una posibilidad, el microorganismo es uno que produce un producto industrialmente importante - por ejemplo, un biocombustible. Como alternativa o además, el microorganismo expresa enzimas de modo que es capaz de producir un producto industrialmente importante - por ejemplo, un biocombustible. Se apreciará que la elección precisa de las enzimas se selecciona según el microorganismo particular a usar. Como alternativa o además, el microorganismo expresa un producto industrialmente importante - por ejemplo, una proteína recombinante. Productos importantes industriales adicionales incluyen antibióticos u otros farmaceuticos, disolventes, pigmentos, aditivos alimenticios, monómeros para la industria del plástico y polímeros industrialmente valiosos.

Biocombustibles incluyen, por ejemplo, un alcohol (por ejemplo, metanol, etanol, propanol, isobutanol, y n-butanol etc.), un hidrocarburo (por ejemplo, un alcano tal como metano, etano, propano, butano, un alqueno tal como etileno, propileno, isoprenos, un alquino tal como acetileno, etc.) hidrógeno, un biodiésel (alquil (metil, propil o etil) ésteres de cadena larga), un aldehído o cetonas (por ejemplo, acetona, formaldehído, 1-propanal, etc.). El biocombustible puede ser un sólido, un líquido o un gas.

Los microorganismos industrialmente útiles incluyen la producción de etanol por *Saccharomyces* y la producción de butanol por *Clostridium*.

La proteína recombinante puede ser cualquier proteína - por ejemplo, una proteína humana usada con fines medicinales. Ejemplos de dichas proteínas incluyen anticuerpos, insulina, interferón, hormona de crecimiento, eritropoyetina, hormona de crecimiento, hormona foliculo estimulante, factor VIII, receptor de lipoproteína de baja densidad (LDLR) alfa galactosidasa A y glucocerebrosidasa.

Tal como se ha mencionado, para expresar las proteínas recombinantes en el microorganismo, se insertan en vectores de expresión secuencias de polinucleótidos que codifican la misma tal como se ha descrito anteriormente en el presente documento.

Se apreciará que a parte de contener los elementos necesarios para la transcripción y la traducción de la secuencia codificante insertada (que codifica el polipéptido industrialmente útil), la construcción de expresión para la expresión del polipéptido industrialmente útil también puede incluir secuencias modificadas por ingeniería para optimizar la estabilidad, producción, purificación, rendimiento o actividad del polipéptido expresado.

Dependiendo del vector y del sistema hospedador usado para la producción, los polipéptidos resultantes pueden o bien permanecer dentro de la célula recombinante, secretados dentro del medio de fermentación, secretados en un espacio entre dos membranas celulares, tal como el espacio periplasmático en *E. coli*; o retenidos sobre la superficie exterior de una célula o membrana viral.

Después de un tiempo predeterminado en cultivo, se efectúa la recuperación del péptido recombinante.

La expresión "recuperación del polipéptido recombinante" usada en el presente documento se refiere a recoger todo el medio de fermentación que contiene el polipéptido y no necesita implicar etapas adicionales de separación o purificación.

Por lo tanto, los polipéptidos se pueden purificar usando una diversidad de técnicas de purificación de proteína convencionales, tales como, pero sin limitación, cromatografía de afinidad, cromatografía de intercambio iónico, filtración, electroforesis, cromatografía de interacción hidrófoba, cromatografía de filtración en gel, cromatografía de fase inversa, cromatografía con concanavalina A, cromatografía de enfoque y solubilización diferencial.

Para facilitar la recuperación, la secuencia codificante expresada se puede modificar por ingeniería para codificar el polipéptido y fusionar el resto escindible. Dicha proteína de fusión se puede diseñar de manera que el polipéptido se puede aislar fácilmente mediante cromatografía de afinidad; por ejemplo, mediante inmovilización sobre una columna específica para el resto escindible. Cuando un sitio de escisión se modifica por ingeniería entre el polipéptido y el resto escindible, el polipéptido se puede liberar de la columna cromatográfica mediante tratamiento con una enzima apropiada o agente que escinde específicamente la proteína de fusión en este sitio (por ejemplo, 26, 27).

La recuperación de biocombustibles se puede recuperar según métodos conocidos en la técnica. Se pueden recuperar alcoholes tales como etanol, metanol, y/o butanol del material líquido por tamices moleculares, destilación, y/o otras técnicas de separación. Por ejemplo, el etanol se puede concentrar mediante destilación fraccional a aproximadamente 90 % o aproximadamente 95 % en peso. Hay varios métodos disponibles para purificar más el etanol más allá de los límites de la destilación, y estos incluyen secado (por ejemplo, con óxido de calcio o sal de roca), la adición de pequeñas cantidades de benceno o ciclohexano, tamiz molecular, membrana, o mediante reducción por presión.

El gas producto, por ejemplo, tal como se produce mediante metabolismo anaeróbico o fotosíntesis, se puede procesar para separar el metano y/o componentes de hidrógeno. El metano, hidrógeno, o biogas se pueden retirar del sistema como gas de tubería.

El metano y/o el hidrógeno se pueden recuperar como un producto de biocombustible. El metano se puede recuperar y/o purificar del biogas mediante métodos conocidos y sistemas que están comercialmente disponibles, incluyendo sistemas de membrana conocidos para separar gases basándose en diferentes permeabilidades. Véase, por ejemplo, el documento de la Patente de EE.UU. N.º 6.601.543. Como alternativa, se pueden usar diversos métodos de adsorción para separar metano e hidrógeno.

Otros modos de recoger los productos biocombustible que incluyen centrifugación, fraccionización por temperatura, métodos cromatográficos y métodos electroforéticos.

En determinadas realizaciones, los componentes de recuperación/purificación de biocombustible se pueden integrar dentro del sistema de cultivo de microorganismo (por ejemplo, biorreactor), por ejemplo, conectando el respectivo dispositivo o aparato a los efluentes de gas o líquido de los biorreactores. Los biocombustibles purificados y los productos de bioenergía se pueden cebar en un recipiente(s) separado(s).

Se apreciará que determinadas características de la invención, que, para mayor claridad, se describen en el contexto de realizaciones separadas, también pueden proporcionarse en combinación en una única realización. A la inversa, diversas características de la invención, que, para mayor brevedad, se describen en el contexto de una única realización, también se pueden proporcionar por separado o en cualquier subcombinación adecuada o como adecuado en cualquier otra realización descrita de la invención. Determinadas características descritas en el contexto de diversas realizaciones no son para ser consideradas características esenciales de aquellas realizaciones, a menos que la realización sea inoperante sin aquellos elementos.

Diversas realizaciones y aspectos de la presente invención tal como se han definido anteriormente en el presente documento y tal como se reivindica en la sección reivindicaciones encuentran, a continuación, soporte calculado en los siguientes ejemplos.

Ejemplos

A continuación, se hace referencia a los siguientes ejemplos, los cuales junto con las descripciones anteriores ilustran algunas realizaciones de la invención de una manera no limitante.

En general, la nomenclatura usada en el presente documento y los procedimientos de laboratorio utilizados en el contexto de la presente invención incluyen técnicas moleculares, bioquímicas, microbiológicas y de ADN recombinante. Dichas técnicas se explican por completo en la bibliografía. Véase, por ejemplo, "Molecular Cloning: A Laboratory Manual" Sambrook et al., (1989); "Current Protocols in Molecular Biology" Volúmenes I-III Ausubel, R.M., ed. (1994); Ausubel et al., "Current Protocols in Molecular Biology", John Wiley and Sons, Baltimore, Maryland (1989); Perbal, "A Practical Guide to Molecular Cloning", John Wiley & Son, Nueva York (1988); Watson et al., "Recombinant DNA", Scientific American Books, Nueva York; Birren et al. (eds) "Genome Analysis: A Laboratory Manual Series", Vol. 1-4, Cold Spring Harbor Laboratory Press, Nueva York (1998); las metodologías expuestas en los documentos de patente de EE.UU. N.º 4.666.828; 4.683.202; 4.801.531; 5.192.659 y 5.272.057; "Cell Biology: A Laboratory Handbook", Volúmenes I-III Cellis, J.E., ed. (1994); "Culture of Animal Cells - A Manual of Basic Technique" de Freshney, Wiley-Liss, N. Y. (1994), Tercera Edición; "Current Protocols in Immunology" Volúmenes I-III Coligan J. E., ed. (1994); Stites et al. (eds), "Basic and Clinical Immunology" (8º Edición), Appleton & Lange, Norwalk, CT (1994); Mishell and Shiigi (eds), "Selected Methods in Cellular Immunology", W.H. Freeman and Co., Nueva York (1980); los inmunoensayos disponibles están descritos ampliamente en la patente y la bibliografía científica, véase, por ejemplo, los documentos de Patente de EE.UU. N.º 3.791.932; 3.839.153; 3.850.752; 3.850.578; 3.853.987; 3.867.517; 3.879.262; 3.901.654; 3.935.074; 3.984.533; 3.996.345; 4.034.074; 4.098.876; 4.879.219; 5.011.771 y 5.281.521; "Oligonucleotide Synthesis" Gait, M.J., ed. (1984); "Nucleic Acid Hybridization" Hames, B.D., and Higgins S.J., eds. (1985); "Transcription and Translation" Hames, B.D., and Higgins S.J., eds. (1984); "Animal Cell Culture" Freshney, R.I., ed. (1986); "Immobilized Cells and Enzymes" IRL Press, (1986); "A Practical Guide to Molecular Cloning" Perbal, B., (1984) y "Methods in Enzymology" Vol. 1-317, Academic Press; "PCR Protocols: A Guide To Methods And Applications", Academic Press, San Diego, CA (1990); Marshak et al., "Strategies for Protein Purification and Characterization - A Laboratory Course Manual" CSHL Press (1996). Otras referencias generales se proporcionan por todo este documento. Los procedimientos en el presente documento se cree que son bien conocidos en la técnica y se proporcionan para la conveniencia del lector.

Ejemplo 1

Selección de una ruta óptima para soportar el crecimiento sobre formato

Para elegir qué ruta de las anteriormente descritas es la más adecuada para soportar el crecimiento sobre formato se compararon según varios criterios (28). En primer lugar, se calculó el rendimiento de biomasa esperado sobre formato para cada una de las diferentes rutas (21, 29, 30).

Se usaron dos métodos cuantitativos para estimar el rendimiento de biomasa sobre formato de cada una de las rutas. En el método "conversión de fuente de carbono" (21), los presentes inventores primero calcularon el

rendimiento de la conversión de formato en un metabolito punto de referencia designado, $Y_{\text{formato} \rightarrow \text{metabolito}}$, en unidades de mol/mol. Tomando el rendimiento de biomasa experimentalmente medido sobre ese metabolito, $Y_{\text{metabolito} \rightarrow \text{biomasa}}$, en unidades de gCDW/mol (siendo CDW peso seco celular), a continuación, se puede estimar el rendimiento de biomasa sobre el formato, en unidades de gCDW/mol de formato, como $Y_{\text{formato} \rightarrow \text{biomasa}} = Y_{\text{formato} \rightarrow \text{metabolito}} \cdot Y_{\text{metabolito} \rightarrow \text{biomasa}}$.

5

Por ejemplo, los presentes inventores calcularon el número de moléculas de formato necesarias de ser investigadas para generar una molécula de piruvato. Lo recíproco de este número es el rendimiento de piruvato sobre formato en unidades de mol de piruvato/mol de formato. Multiplicar el rendimiento de piruvato sobre formato con rendimiento de biomasa sobre piruvato - 14.7 gCDW/mol, tal como se mide experimentalmente (31) - proporcionó a los inventores una estimación del rendimiento de biomasa sobre formato en unidades de gCDW/mol de formato. A continuación, se repitió el mismo proceso para la glucosa, tomando el rendimiento de biomasa experimentalmente medido sobre glucosa, 70,8 gCDW/mol (31). En ambos casos se asumió que ATP se produce por NADH y fosforilación oxidativa y que la relación P/O (que mide cuántas moléculas de ATP se producen por un átomo de oxígeno que está siendo reducido) es de 1,5, tal como es relevante para *E. coli* (32). A continuación, la segunda y la tercera columna de la Tabla 2 en el presente documento exponen el rendimiento de biomasa estimado calculado mediante la elección de o bien piruvato o glucosa como el producto de la fijación de carbono y las rutas de asimilación de formato.

10

15

Tabla 2

Ruta	Rendimiento de biomasa, gCDW/mol-formato *					Termodinámicas**	Números de enzimas extrañas ***
	Conversión de fuente de carbono		Análisis de equilibrio de flujo				
	Conversión de formato en piruvato	Conversión de formato en glucosa	Mantenimiento del ATP para el crecimiento sobre glucosa	Sin mantenimiento del ATP			
<i>Rutas de fijación de carbono</i>							
pentosa fosfato reductora	1,5	3,0	2,5	5,6		alta CMF ‡	3
TCA reductor	2,3	3,8	2,9	6,7		Desfavorable	4
dicarboxilato-4-hidroxiacetil-CoA	1,8	3,1	2,4	5,3		baja CMF	5
3-hidroxiisovalerato-4-hidroxiacetil-CoA	1,3	2,5	1,8	4,2		baja CMF	> 5
3-hidroxiisovalerato	1,5	2,8	1,9	4,4		alta CMF	> 5
MOG (28)	1,6	3,1	2	4,5		alta CMF	> 5
<i>Rutas asimiladoras de formato</i>							
acetil-CoA reductor	2,6	4,1	2,7	6,2		desfavorable (alta CMF) †	4
serina	1,6	3,0	2,4	5,5		alta CMF	5
ribulosa monofosfato	2,1	3,8	2,8	6,4		baja CMF	4
xilulosa 5-fosfato	1,9	3,5	2,6	5,9		alta CMF	3
glicina reductora	2,3	3,8	2,8	6,3		alta CMF (alta CMF) ††	2

* véase el texto principal para la diferencia entre los métodos de cálculo del rendimiento de biomasa sobre formato.

** Las rutas marcadas con "Desfavorable" contienen una secuencia de reacción termodinámicamente desfavorable; las rutas marcadas con "baja CMF" (es decir, baja fuerza motriz química) son termodinámicamente favorables pero no hay conjunto de concentración de metabolito, dentro del intervalo fisiológico, que puede soportar ΔG,' < -3 kJ/mol para todas sus reacciones; las rutas marcadas con "alta CMF" son aquellas para las que existe dicho conjunto de concentración de metabolito.

*** Número de enzimas extrañas que necesitan ser expresadas en *E. coli* para establecer una ruta activa.

‡ CMF significa la fuerza motriz química.

† La señal fuera de los paréntesis se refiere al caso en que las reacciones dentro del complejo co-deshidrogenasa-acetil-coA-sintetasa no se acoplan mientras que la señal en los paréntesis se refiere al caso en que se acoplan.

†† La señal fuera de los paréntesis se refiere al caso en que las reacciones dentro del complejo del sistema de escisión de glicina no se acoplan mientras que la señal en los paréntesis se refiere al caso en que se acoplan.

* Véase el texto principal para la diferencia entre los métodos de cálculo del rendimiento de biomasa sobre formato.

** Las rutas marcadas con "Desfavorable" contienen una secuencia de reacción termodinámicamente desfavorable; las rutas marcadas con "baja CMF" (es decir, baja fuerza motriz química) son termodinámicamente favorables pero no hay conjunto de concentración de metabolito, dentro del intervalo fisiológico, que puede soportar $\Delta G_r' < -3$ kJ/mol para todas sus reacciones; las rutas marcadas con "alta CMF" son aquellas para las que existe dicho conjunto de concentración de metabolito.

*** Número de enzimas extrañas que necesitan ser expresadas en *E. coli* para establecer una ruta activa.

‡ CMF significa la fuerza motriz química.

† La señal fuera de los paréntesis se refiere al caso en que las reacciones dentro del complejo co-deshidrogenasa-acetil-CoA-sintetasa no se acoplan mientras que la señal en los paréntesis se refiere al caso en que se acoplan.

†† La señal fuera de los paréntesis se refiere al caso en que las reacciones dentro del complejo de escisión de glicina no se acoplan mientras que la señal en los paréntesis se refiere al caso en que se acoplan.

Los resultados mostrados en la Tabla 2 en el presente documento sugiere que las rutas asimiladoras de formato son en general más eficaces que unas de fijación de carbono, con la excepción de aquellas rutas de fijación de carbono que evitan las etapas de carboxilación acopladas a ATP. No obstante, estos resultados también sugieren que este tipo de análisis es problemático puesto que la elección del metabolito específico para servir como el producto de las rutas puede influir la estimación enormemente: asumir la glucosa como producto de las rutas dio mucho mayor estimación para el rendimiento de biomasa sobre formato (aproximadamente 1,5 veces de diferencia promedio) justo debido a que el rendimiento de biomasa de *E. coli* sobre glucosa es mayor que la de sobre piruvato.

Un planteamiento alternativo para estimar el rendimiento de biomasa es usar métricas de análisis de equilibrio de flujo (33,34) tal como se implementa por (30). La ventaja de este planteamiento es que no se desvía por la elección de metabolito en el que el formato se convierte y que luego se usa como fuente de carbono, sino que trata la biomasa celular entera como el producto de la fijación de carbono o la asimilación de formato (30). El modelo metabólico de *E. coli* central se selecciona (35) sobre el modelo metabólico completo puesto que los presentes inventores querían mantener el análisis más general y menos específico a *E. coli* y puesto que querían evitar la complejidades de la regulación que pueden dar como resultado soluciones defectuosas que usan el modelo metabólico completo. Se usó análisis de equilibrio de flujo para calcular el rendimiento del crecimiento en lugar de la tasa de crecimiento. Por lo tanto, las unidades de entrada de formato eran en mmol y el rendimiento de biomasa se dio en gDW. Puesto que el mantenimiento de ATP no se pueden estimar *a priori* para el crecimiento sobre formato los rendimientos de biomasa se calcularon de dos modos diferentes, uno retirando el mantenimiento de ATP en total y uno asumiendo el mantenimiento de ATP idéntico al de la glucosa. Las columnas quinta y cuarta en la Tabla 2 exponen los resultados de este análisis.

De manera importante, aunque los métodos de cálculo de rendimiento de biomasa sobre formato difieren sustancialmente y dieron como resultado hasta 6 veces de diferencia en el rendimiento de biomasa estimado (Tabla 2), sugieren rendimientos de biomasa relativos muy similares de las rutas. De hecho, la correlación entre los rendimientos de biomasa calculados asumiendo piruvato o glucosa como el producto de las rutas y aquellos calculados usando análisis de equilibrio de flujo es $R^2 > 0,6$. Considerando todos los métodos usados, el ciclo TCA reductor, la ruta reductora de acetil-CoA, la ruta de ribulosa monofosfato y la ruta reductora de la glicina son las rutas que soportan el mayor rendimiento de biomasa sobre formato.

Ejemplo 2

Favorabilidad termodinámica

No todas las rutas metabólicas que funcionan en un organismo son termodinámicamente favorables en otros, en los que las condiciones celulares (pH, fuerza iónica, etc.) podrían diferir considerablemente (28, 36, 37). Por tanto, los presentes inventores comprobaron si todas las rutas discutidas anteriormente son termodinámicamente favorables dentro de *E. coli* (pH aproximadamente 7,5, I aproximadamente 0,2 M). No ensayaron solamente la favorabilidad de la reacción neta de la ruta (28, 38) sino también analizaron embotellamientos termodinámicos distribuidos compuestos de un subconjunto de reacciones dentro de las rutas (28, 39). En particular, la CO-deshidrogenasa-acetil-CoA-sintasa y el sistema de escisión de glicina son complejos de varias enzimas. Por lo tanto, Las reacciones que se dan dentro de estos complejos probablemente se acoplan una a otra, superando cualquier barrera termodinámica interna (37). Para rutas que contienen estos complejos se consideraron ambos escenarios - las reacciones internas dentro del complejo están acopladas o están desacopladas.

Se encontraron dos rutas que se prevé que son termodinámicamente desfavorables puesto que contienen conjuntos de reacción que no pueden proceder en la dirección directa en las condiciones celulares de *E. coli* y bajo concentraciones reactivas fisiológicas (se usó el intervalo 1 μ M-10 mM para metabolitos no co-factor) (37,40-42). Estos son el ciclo reductor TCA y la ruta reductora de acetil-CoA en los que la reducción de CO₂ a CO y la síntesis de acetil-CoA no están acopladas. Ambas rutas se marcan como "desfavorables" en la sexta columna de la Tabla 2.

El funcionamiento reductor secuencial de la 2-cetoglutarato sintasa y la isocitrato deshidrogenasa (parte del ciclo reductor TCA), que cataliza la reacción global $\text{succinil-CoA} + 2 \text{ ferredoxina}^{\text{red}} + \text{NADPH} + 2 \text{ CO}_2 \rightleftharpoons \text{isocitrato} + \text{CoA} + 2 \text{ ferredoxina}^{\text{ox}} + \text{NADP}^+$, parece ser desfavorable en *E. coli*. Incluso si se asume que se mantiene la concentración de CO₂ disuelto a 1 mM (-100 veces mayor que el ambiente), $[\text{NADPH}] = 10 \cdot [\text{NADP}^+]$, $[\text{ferredoxina}^{\text{red}}] \sim [\text{ferredoxina}^{\text{ox}}]$, $[\text{CoA}] = 1 \text{ mM}$ (la menor concentración celular de este cofactor (40) y los otros metabolitos están en sus valores extremos posibles, $[\text{succinil-CoA}] = 10 \text{ mM}$ e $[\text{isocitrato}] = 1 \mu\text{M}$, el cambio en la energía de Gibbs ($\Delta_r G'$) durante la reacción global es aún positiva a pH 7,5 y 1=0,2 M. Este razonamiento probablemente descarta el ciclo reductor TCA de servir como una ruta de fijación de carbono en *E. coli*.

La barrera energética de la ruta reductora de acetil-CoA es el resultado de la reducción altamente desfavorable de CO₂ a CO con $\Delta_r G'^{\circ} = +46 \text{ kJ/mol}$ cuando ferredoxina sirve como donante de electrones (43). Sin embargo, esta gran barrera energética se puede nivelar en total si la reacción de hecho se acopla a la reacción muy favorable que lo permite dentro del mismo complejo, Acetil-CoA sintasa (37). En tal caso la reacción completa, $\text{metil-THF} + \text{CO}_2 + 2 \text{ ferredoxina}^{\text{red}} + \text{CoA} \rightleftharpoons \text{acetil-CoA} + 2 \text{ ferredoxina}^{\text{ox}} + \text{THF}$, tendrá $\Delta_r G'^{\circ}$ aproximadamente -25 kJ/mol (43), haciendo la ruta entera favorable.

Ejemplo 3**Fuerza motriz química**

- 5 Ser termodinámicamente favorable no es suficiente. La energía disipada durante una reacción ($\Delta_r G'$) puede tener un efecto sustancial sobre sus cinéticas. De hecho, $\Delta_r G'$ dicta que la fracción de la maquinaria enzimática cataliza la reacción directa (38,44-46): $\Delta_r G' = -RT \ln(J^+/J^-)$, donde J^+ es el flujo directo, J^- es el flujo inverso, R es la constante de gas y T es la temperatura en grados Kelvin. Por lo tanto, un valor de $\Delta_r G'$ bajo (negativo), correspondiente a una alta fuerza motriz química, indica que la mayoría de la maquinaria enzimática está catalizando la reacción directa y por lo tanto se puede alcanzar una alta tasa metabólica. Asumiendo la saturación de sustrato y cinéticas similares en la dirección directa e inversa, $\Delta_r G'$ de -7.5 kJ/mol corresponde a una reacción que procede a 90 % de su velocidad máxima: 95 % de las enzimas catalizan la reacción directa mientras que el 5 % catalizan la reacción inversa.

- 15 Los presentes inventores se cuestionaron qué ruta de las analizadas puede, en principio, soportar alto flujo en términos de la fuerza motriz química sostenida por cada una de sus reacciones. Para cada ruta se usó una herramienta de optimización lineal para cuestionarse si existe un conjunto de concentración de metabolito, dentro del intervalo fisiológico, de modo que cada reacción de la ruta no es solamente favorable sino también funciona a $\Delta_r G' < -3$ kJ/mol, correspondiente a al menos el 55 % de su tasa máxima (asumiendo la saturación de sustrato y cinéticas similares en la dirección directa e inversa). La sexta columna en la Tabla 2 expone los resultados. Las rutas marcadas con "baja CMF" (baja fuerza motriz química) son termodinámicamente favorables pero cinéticamente más escasas puesto que no hay conjunto de concentración de metabolito, dentro del intervalo fisiológico, que pueda soportar $\Delta_r G' < -3$ kJ/mol para todas sus reacciones. Las rutas marcadas como "alta CMF" son aquellas para las que existe dicho conjunto de concentración de metabolito y por lo tanto, en potencial, puede sostener una alta fuerza motriz química a través de la ruta.

25

Ejemplo 4**Cinéticas de la ruta**

- 30 La fuerza motriz química no es el único parámetro que determina el flujo de reacción. Los parámetros cinéticos, es decir, la velocidad máxima (V_{MAX}) y las afinidades hacia los sustratos (constantes de Michaelis, K_M), juegan un papel no menos importante.

- 35 Los presentes inventores estimaron cuál es la máxima tasa de crecimiento para una bacteria que utiliza formato solamente como una fuente de electrones y depende de rutas de fijación de carbono para el carbono. Consideran el límite de la tasa de crecimiento impuesto por la tasa de la formato deshidrogenasa. Puesto que la formato deshidrogenasa cumple solamente el papel auxiliar en bacterias que emplean una de las rutas asimiladoras de formato (incrementando la disponibilidad celular del poder reductor y ATP), el análisis presentado a continuación no se mantiene en estos casos.

40

- Suponer que las células expresan una formato deshidrogenasa con una actividad específica de 10 $\mu\text{mol}/\text{min}/\text{mg}$ (velocidad máxima para las variantes de enzima simple, véase anteriormente) a 20 % de su proteína total (más de lo que se espera que sea dañino). Asumiendo que aproximadamente el 50 % del peso seco celular es proteínas (47), se calcula que el 10 % del peso seco celular es formato deshidrogenasa y por lo tanto la actividad específica es aproximadamente 1 $\mu\text{mol}/\text{min}/\text{mgCDW}$. Si la ruta reductora de pentosa fosfato sirve como la ruta de fijación de carbono, 12 moléculas de NAD(P)H y 18 de ATP se requieren para fijar seis moléculas de CO_2 a glucosa. Asumiendo una relación de P/O de 1,5 (32), $12+18/1,5 = 24$ moléculas de formato se deberían oxidar para soportar la formación de una molécula de glucosa. Por lo tanto, la tasa de formación de glucosa será de 1/24 aproximadamente 0,042 μmol de Glucosa/min/mg CDW. Puesto que el rendimiento de biomasa experimentalmente medido sobre glucosa es de 70,8 gDW/mol (31) se calcula que esta tasa es igual a $0,042 \cdot 70,8 \cdot 10^{-6}$ aproximadamente $3 \cdot 10^{-6}$ gCDW/min/mgCDW o 0,003 mgCDW/min/mgCDW. Por lo tanto, la tasa de crecimiento es igual a 0,0031/min y el tiempo de duplicación es de $\ln(2)/0,003$ aproximadamente 230 min aproximadamente 4 horas.

- 55 Este cálculo sugiere que el tiempo de duplicación de una bacteria autótrofa que metaboliza formato usando una formato deshidrogenasa del tipo "simple" no puede ser menor que 4 horas. Por supuesto, esto es solamente un límite inferior y el tiempo de duplicación se podría limitar muy bien por otros factores, tales como la tasa de fijación de carbono. Además, incluso el límite aproximado de 4 horas puede cambiar cuando se consideran rutas de fijación de carbono con diferentes requerimientos de ATP o si se restringe la expresión de formato deshidrogenasa a menor que el 20 % de la proteína total.

60

- Un tipo similar de análisis puede producir más resultados definitivos. Por ejemplo, uno puede cuestionarse si una bacteria puede usar las rutas de la ribulosa monofosfato o la xilulosa 5-fosfato para la asimilación de formato cuando se limita por la tasa de escisión espontánea de metileno-THF a formaldehído y THF. Según las constantes de equilibrio y cinéticas publicadas en (48) y asumiendo un alto [metileno-THF] de aproximadamente 10 mM, la tasa máxima de escisión de metileno-THF se puede calcular para ser de 0,027 mM/s. Puesto que el volumen celular de

65

E. coli que crece lentamente es de aproximadamente $1 \mu\text{m}^3$ (49) esta tasa es igual a $2,7 \cdot 10^{-17} \text{ mmol/s/célula}$ o $1,6 \cdot 10^{-12} \mu\text{mol/min/célula}$. Puesto que se requieren seis de tales reacciones para la producción de una molécula de glucosa la tasa de producción de glucosa se limita a $1,6 \cdot 10^{-12}/6 = 2,7 \cdot 10^{-13} \mu\text{mol de Glucosa/min/célula}$. El peso seco de una célula de *E. coli* con un volumen de aproximadamente $1 \mu\text{m}^3$ es de aproximadamente 200 fg (50) y por lo tanto la tasa anterior es igual a $0,0014 \mu\text{mol de Glucosa/min/mgCDW}$. Según el rendimiento de biomasa anterior esta tasa corresponde a $9,6 \cdot 10^{-8} \text{ gCDW/min/mgCDW}$ que es $9,6 \cdot 10^{-5} \text{ mgCDW/min/mgCDW}$. Por tanto, la tasa de crecimiento es igual a $9,6 \cdot 10^{-5} \text{ 1/min}$ y el tiempo de duplicación es de $\ln(2)/(9,6 \cdot 10^{-5}) > 7.200 \text{ min} = 120 \text{ horas} = 5 \text{ días}$.

Por tanto, este cálculo sugiere que, considerando solamente una reacción única, dos rutas asimiladoras de formato son cinéticamente no factibles y no pueden sostener ni siquiera una tasa de crecimiento mínimamente aceptable. De hecho, un estudio anterior ha demostrado que la reacción inversa - la condensación de THF con formaldehído - es demasiado lenta para tener cualquier significancia metabólica *in vivo* (51).

Ejemplo 5

Pautas de expresión de la ruta

Por último, los presentes inventores se cuestionaron cómo se espera que sea la pauta de la expresión de una ruta activa dentro de *E. coli*. Un aspecto que afecta la dificultad de expresión es el número de enzimas extrañas que se deberían expresar para posibilitar la actividad de ruta. La séptima columna en la Tabla 2 expone este número para cada una de las alternativas metabólicas. En particular, todas las rutas requieren, o se están beneficiando fuertemente de, la expresión de la formato deshidrogenasa para suministrar la célula con poder reductor y energía.

La ruta reductora de la pentosa fosfato y la ruta de la xilulosa 5-fosfato parecen imponer una barrera de expresión pequeña, necesitando solamente tres enzimas extrañas. La ruta reductora de la glicina presenta la barrera de expresión más pequeña: solamente se necesitan formato deshidrogenasa extraña y formato-tetrahidrofolato ligasa para el funcionamiento de la ruta.

También se contemplaron la sensibilidad de oxígeno de alguna de las enzimas que funcionan en alguna de las rutas (38, 52). Específicamente, el ciclo reductor de TCA y las rutas de dicarboxilato-4-hidroxiopropionato emplean varias enzimas sensibles a oxígeno (38-52). Asimismo, la ruta reductora de acetil-CoA hace funcionar alguna de las enzimas más sensibles a oxígeno conocidas (38, 52, 53). Por tanto, estas rutas no son adecuadas si la bacteria es a cultivar bajo condiciones aeróbicas.

Observaciones finales

La revisión de todos los criterios sugiere que una ruta destaca sobre las otras. La ruta reductora de la glicina requiere la expresión de solamente dos enzimas extrañas, no contiene enzima sensible a oxígeno, soporta alto rendimiento de biomasa, es capaz de sostener alta fuerza motriz química a través del conjunto entero de reacción y no se restringe cinéticamente seriamente por ninguna reacción (Tabla 1). Esta ruta parece ser la ruta más prometedora para establecer una *E. coli* formatótrofa.

Alguna de las rutas analizadas por los presentes inventores tienen varias variantes, cada una con sus propias características, ventajas e inconvenientes. Por ejemplo, el ciclo reductor TCA probablemente no es un buen candidato puesto que contiene una gran barrera termodinámica a condiciones celulares de *E. coli*. Sin embargo, *Hydrogenobacter thermophilus* ha desarrollado un mecanismo dependiente de ATP para empujar la ruta en la dirección reductora: la enzima 2-cetoglutarato carboxilasa cataliza la carboxilación dependiente de ATP de 2-cetoglutarato a oxalosuccinato en un mecanismo dependiente de biotina, mientras que el oxalosuccinato se reduce más a isocitrato por una isocitrato deshidrogenasa no carboxilante (54, 55). Esta variante de ruta es termodinámicamente favorable e incluso soporta alta fuerza motriz química para cada uno de sus componentes enzimáticos. Sin embargo, el oxalosuccinato intermedio soluble es inestable y fácilmente se somete a descarboxilación, por lo tanto creando un ciclo fútil que reduce la eficacia global de la fijación de carbono (54).

Por último, la evolución enzimática puede proporcionar numerosas soluciones metabólicas a las pautas surgidas por los presentes inventores. Por ejemplo, la enzima que cataliza la condensación reversible del formaldehído y tetrahidrometanopterina a metileno-tetrahidrometanopterina (56) se puede desarrollar para aceptar THF en lugar de tetrahidrometanopterina, levantando de ese modo la barrera cinética impuesta mediante la escisión espontánea de metileno-THF. De especial importancia es el diseño de una enzima que pueden condensar dos moléculas de formato en glioxilato que, a continuación, se puede asimilar directamente dentro del metabolismo central. Aunque dicha enzima se sugirió previamente para funcionar en el cloroplasto de tubérculos de patata verde (57-59), la energética de esta condensación inactivada es extremadamente desfavorable, indicando que el informe es probablemente erróneo (38). Por tanto, cualquier intento de diseñar una enzima que cataliza esta condensación debe primero activar el formato (con un grupo fosfato, por ejemplo).

Referencias

- [1] K. Rabaey, R.A. Rozendal, Microbial electrosynthesis - revisiting the electrical route for microbial production, *Nat Rev Microbiol*, 8 (2010) 706-716.
- [2] D.R. Lovley, Powering microbes with electricity: direct electron transfer from electrodes to microbes, *Env Microbiol Rep.*, 3 (2011) 27-35.
- [3] D.R. Lovley, K.P. Nevin, A shift in the current: new applications and concepts for microbe-electrode electron exchange, *Curr Opin Biotechnol*, 22 (2011) 441-448.
- [4] K. Rabaey, P. Girguis, L.K. Nielsen, Metabolic and practical considerations on microbial electrosynthesis, *Curr Opin Biotechnol*, 22 (2011) 371-377.
- [5] J.C. Thrash, J.D. Coates, Review: Direct and Indirect Electrical Stimulation of Microbial Metabolism, *Environ Sci Technol.*, 42 (2008) 3.921-3.931.
- [6] D.T. Whipple, J.A. Kenis, Prospects of CO₂ Utilization via Direct Heterogeneous Electrochemical Reduction, *J Phys Chem Lett.*, 1 (2010) 3.451-3.458.
- [7] W. Li, Electrocatalytic Reduction of CO₂ to Small Organic Molecule Fuels on Metal Catalysts, en: *Advances in CO₂ Conversion and Utilization*, 2010, pág. 55-76.
- [8] K.P. Kuhl, E.R. Cave, D.N. Abram, T.F. Jaramillo, New insights into the electrochemical reduction of carbon dioxide on metallic copper surfaces, *Energy Environ Sci.*, 5 (2012) 7.050-7.059.
- [9] E.A. Quadrelli, G. Centi, J.L. Duplan, S. Perathoner, Carbon dioxide recycling: emerging large-scale technologies with industrial potential, *ChemSusChem*, 4 (2011) 1.194-1.215.
- [10] T. Reda, C.M. Plugge, N.J. Abram, J. Hirst, Reversible interconversion of carbon dioxide and formate by an electroactive enzyme, *Proc Natl Acad Sci USA*, 105 (2008) 10.654-10.658.
- [11] R.P.S. Chaplin, A.A. Wragg, Effects of process conditions and electrode material on reaction pathways for carbon dioxide electroreduction with particular reference to formate formation, *J Appl Electrochem.*, 33 (2003) 1.107-1.123.
- [12] B. Innocent, D. Liaigre, D. Pasquier, F. Ropital, J.-M. Léger, K.B. Kokoh, Electro-reduction of carbon dioxide to formate on lead electrode in aqueous medium, *J Appl Electrochem.*, 39 (2009) 227-232.
- [13] C. Oloman, H. Li, Electrochemical processing of carbon dioxide, *ChemSusChem*, 1 (2008) 385-391.
- [14] S. Enthaler, J. von-Langermann, T. Schmidt, Carbon dioxide and formic acid-the couple for environmental-friendly hydrogen storage?, *Energy Environ Sci.*, 3 (2010) 1.207-1.217.
- [15] J.H. Hull, Y. Himeda, W.-H. Wang, B. Hashiguchi, R. Periana, D.J. Szalda, J.T. Muckerman, E. Fujita, Reversible hydrogen storage using CO₂ and a proton-switchable iridium catalyst in aqueous media under mild temperatures and pressures, *Nature Chemistry*, 4 (2012) 383-388.
- [16] F. Joo, Breakthroughs in hydrogen storage--formic Acid as a sustainable storage material for hydrogen, *ChemSusChem*, 1 (2008) 805-808.
- [17] B. Loges, A. Boddien, F. Gärtner, H. Junge, M. Beller, Catalytic Generation of Hydrogen from Formic acid and its Derivatives: Useful Hydrogen Storage Materials *Top Catal*, 53 (2010) 902-914.
- [18] M.E. Lidstrom, Aerobic Methylophilic Prokaryotes, en: M. Dworkin, S. Falkow, E. Rosenberg, K.H. Schleifer, E. Stackebrandt (Eds.) *The Prokaryotes*, Springer, Nueva York, 2006, pág. 618-634.
- [19] J. Schrader, M. Schilling, D. Holtmann, D. Sell, M.V. Filho, A. Marx, J.A. Vorholt, Methanol-based industrial biotechnology: current status and future perspectives of methylophilic bacteria, *Trends Biotechnol*, 27 (2009) 107-115.
- [20] L. Chistoserdova, M.G. Kalyuzhnaya, M.E. Lidstrom, The expanding world of methylophilic metabolism, *Annu Rev Microbiol*, 63 (2009) 477-499.
- [21] C. Anthony, *The Biochemistry of Methylophilic*, Academic Press, Londres; Nueva York, 1982.
- [22] H. Li, P.H. Opgenorth, D.G. Wernick, S. Rogers, T.Y. Wu, W. Higashide, P. Malati, Y.X. Huo, K.M. Cho, J.C. Liao, Integrated electromicrobial conversion of CO₂ to higher alcohols, *Science*, 335 (2012) 1596.
- [23] G. Fuchs, CO₂ fixation in acetogenic bacteria: variations on a theme, *FEMS Microbiology Letters*, 39 (1985) 181-213.
- [24] A. Veit, M.K. Akhtar, T. Mizutani, P.R. Jones, Constructing and testing the thermodynamic limits of synthetic NAD(P)H:H₂ pathways, *Microb Biotechnol*, 1 (2008) 382-394.
- [25] M.K. Akhtar, P.R. Jones, Construction of a synthetic YdbK-dependent pyruvate:H₂ pathway in *Escherichia coli* BL21(DE3), *Metab Eng*, 11 (2009) 139-147.
- [26] R.J. Booth, P.M. Grandison, R.L. Prestidge, J.D. Watson, The use of a 'universal' yeast expression vector to produce an antigenic protein of *Mycobacterium leprae*, *Immunol Lett*, 19 (1988) 65-69.
- [27] T.J. Gardella, D. Rubin, A.B. Abou-Samra, H.T. Keutmann, J.T. Potts, Jr., H.M. Kronenberg, S.R. Nussbaum, Expression of human parathyroid hormone-(1-84) in *Escherichia coli* as a factor X-cleavable fusion protein, *J Biol Chem*, 265 (1990) 15.854-15.859.
- [28] A. Bar-Even, E. Noor, N.E. Lewis, R. Milo, Design and analysis of synthetic carbon fixation pathways, *Proc Natl Acad Sci USA*, 107 (2010) 8.889-8.894.
- [29] I. Goldberg, J.S. Rock, A. Ben-Bassat, R.I. Mateles, Bacterial yields on methanol, methylamine, formaldehyde, and formate, *Biotechnol Bioeng*, 18 (1976) 1.657-1.668.
- [30] N.R. Boyle, J.A. Morgan, Computation of metabolic fluxes and efficiencies for biological carbon dioxide fixation, *Metab Eng*, 13 (2011) 150-158.
- [31] K.B. Andersen, K. von Meyenburg, Are growth rates of *Escherichia coli* in batch cultures limited by respiration?, *J Bacteriol*, 144 (1980) 114-123.

- [32] Y. Noguchi, Y. Nakai, N. Shimba, H. Toyosaki, Y. Kawahara, S. Sugimoto, E. Suzuki, The energetic conversion competence of *Escherichia coli* during aerobic respiration studied by ³¹P NMR using a circulating fermentation system, *J Biochem*, 136 (2004) 509-515.
- [33] J.D. Orth, I. Thiele, B.O. Palsson, What is flux balance analysis?, *Nat Biotechnol*, 28 (2010) 245-248.
- [34] K. Raman, N. Chandra, Flux balance analysis of biological systems: applications and challenges, *Brief Bioinform*, 10 (2009) 435-449.
- [35] J.D. Orth, R.M.T. Fleming, B.O. Palsson, The Core *E. coli* Model, en, gcrd.tucsd.edu/Downloads/EcoliCore, 2009.
- [36] R.A. Alberty, Thermodynamics of Biochemical Reactions, Wiley-Interscience, 2003.
- [37] A. Bar-Even, A. Flamholz, E. Noor, R. Milo, Thermodynamic constraints shape the structure of carbon fixation pathways, *Biochim Biophys Acta*, (2012).
- [38] A. Bar-Even, E. Noor, R. Milo, A survey of carbon fixation pathways through a quantitative lens, *J Exp Bot*, 63 (2012) 2325-2342.
- [39] M.L. Mavrouniotis, Identification of localized and distributed bottlenecks in metabolic pathways, *Proc Int Conf Intell Syst Mol Biol*, 1 (1993) 275-283.
- [40] B.D. Bennett, E.H. Kimball, M. Gao, R. Osterhout, S.J. Van Dien, J.D. Rabinowitz, Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*, *Nat Chem Biol*, 5 (2009) 593-599.
- [41] A. Bar-Even, E. Noor, A. Flamholz, J.M. Buescher, R. Milo, Hydrophobicity and charge shape cellular metabolite concentrations, *PLoS Comput Biol*, 7 (2011) e1002166.
- [42] E. Noor, A. Bar-Even, A. Flamholz, Y. Lubling, D. Davidi, R. Milo, An integrated open framework for thermodynamics of reactions that combines accuracy and coverage, *Bioinformatics*, publicado online (2012).
- [43] A. Flamholz, E. Noor, A. Bar-Even, R. Milo, eQuilibrator—the biochemical thermodynamics calculator, *Nucleic Acids Res*, 40 (2011) D700-D775.
- [44] H. Qian, D.A. Beard, S.D. Liang, Stoichiometric network theory for nonequilibrium biochemical systems, *Eur J Biochem*, 270 (2003) 415-421.
- [45] D.A. Beard, H. Qian, Relationship between thermodynamic driving force and one-way fluxes in reversible processes, *PLoS One*, 2 (2007) e144.
- [46] A. Bar-Even, A. Flamholz, E. Noor, R. Milo, Rethinking glycolysis: on the biochemical logic of metabolic pathways, *Nat Chem Biol*, 8 (2012) 509-517.
- [47] H. Bremer, P. Dennis, Modulation of chemical composition and other parameters of the cell by growth rate, In: *Escherichia coli* and *Salmonella*, 1987.
- [48] R.G. Kallen, W.P. Jencks, The mechanism of the condensation of formaldehyde with tetrahydrofolic acid, *J Biol Chem*, 241 (1966) 5851-5863.
- [49] B. Volkmer, M. Heinemann, Condition-dependent cell volume and concentration of *Escherichia coli* to facilitate data conversion for systems biology modeling, *PLoS One*, 6 (2011) e23126.
- [50] M. Loferer-Krossbacher, J. Klima, R. Psennner, Determination of bacterial cell dry mass by transmission electron microscopy and densitometric image analysis, *Appl Environ Microbiol*, 64 (1998) 688-694.
- [51] G.J. Crowther, G. Kosaly, M.E. Lidstrom, Formate as the main branch point for methylotrophic metabolism in *Methylobacterium extorquens* AM1, *J Bacteriol*, 190 (2008) 5057-5062.
- [52] I.A. Berg, Ecological Aspects of the Distribution of Different Autotrophic CO₂ Fixation Pathways, *Appl Environ Microbiol*, 77 (2011) 1925-1936.
- [53] H.L. Drake, K. Kirsten, C. Matthies, Acetogenic Prokaryotes in: *The Prokaryotes*, Springer, Nueva York, 2006, pág. 354-420.
- [54] M. Aoshima, Y. Igarashi, A novel oxalosuccinate-forming enzyme involved in the reductive carboxylation of 2-oxoglutarate in *Hydrogenobacter thermophilus* TK-6, *Mol Microbiol*, 62 (2006) 748-759.
- [55] M. Aoshima, M. Ishii, Y. Igarashi, A novel biotin protein required for reductive carboxylation of 2-oxoglutarate by isocitrate dehydrogenase in *Hydrogenobacter thermophilus* TK-6, *Mol Microbiol*, 51 (2004) 791-798.
- [56] J.A. Vorholt, C.J. Marx, M.E. Lidstrom, R.K. Thauer, Novel formaldehyde-activating enzyme in *Methylobacterium extorquens* AM1 required for growth on methanol, *J Bacteriol*, 182 (2000) 6645-6650.
- [57] N.K. Ramaswamy, A.G. Behere, P.M. Nair, A novel pathway for the synthesis of solanidine in the isolated chloroplast from greening potatoes, *Eur J Biochem*, 67 (1976) 275-282.
- [58] M.T. Janave, N.K. Ramaswamy, P.M. Nair, Purification and characterization of glyoxylate synthetase from greening potato-tuber chloroplasts, *Eur J Biochem*, 214 (1993) 889-896.
- [59] M.T. Janave, N.K. Ramaswamy, P.M. Nair, Studies on determination of active site amino acid residues in glyoxylate synthetase from potato tuber chloroplasts, *Plant Physiology and Biochemistry*, 37 (1999) 121-129.

Aunque la invención se ha descrito junto con las realizaciones específicas de la misma, es evidente que muchas alternativas, modificaciones y variaciones estarán claras a los expertos en la técnica. Además, la citación o identificación de cualquier referencia en esta solicitud no debería interpretarse como una admisión de que tal referencia está disponible como técnica anterior a la presente invención. En la medida en que se usan los títulos de la sección, no deberían interpretarse como necesariamente limitantes.

LISTADO DE SECUENCIAS

<110> Yeda Research and Development Co. Ltd. Bar-Even, Arren Milo, Ron Noor, Elad Yishai, Oren

<120> Uso de la ruta reductora de la glicina para generar microorganismos formatótrofos y autótrofos

<130> 56534

5 <150> US 61/676,962

<151> 29-07-2012

<160> 14

10 <170> PatentIn version 3.5

<210> 1

<211> 557

<212> PRT

15 <213> PA1 de *Methylobacterium extorquens*

<400> 1

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

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

Tyr Gly Lys His Ile Ala Lys Ile Asp His Asp Phe Ile Ala Ser Leu
35 40 45

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

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

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

Ser Leu Gly Pro Cys Phe Gly Met Lys Gly Gly Ala Ala Gly Gly Gly
100 105 110

Lys Ala Gln Val Val Pro Met Glu Gln Ile Asn Leu His Phe Thr Gly
115 120 125

Asp Phe His Ala Ile Thr Ser Ala His Ser Leu Ala Ala Ala Leu Ile
130 135 140

20

ES 2 730 377 T3

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

ES 2 730 377 T3

Gln Val Glu Ala Ile Thr Cys Lys His Trp Ala Glu Gly Gly Ala Gly
405 410 415

Ala Glu Ala Leu Ala Gln Ala Val Val Lys Leu Ala Glu Gly Glu Gln
420 425 430

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

Ile Lys Ala Ile Ala Thr Lys Leu Tyr Gly Ala Ala Asp Ile Gln Ile
450 455 460

Glu Ser Lys Ala Ala Thr Lys Leu Ala Gly Phe Glu Lys Asp Gly Tyr
465 470 475 480

Gly Lys Leu Pro Val Cys Met Ala Lys Thr Gln Tyr Ser Phe Ser Thr
485 490 495

Asp Pro Thr Leu Met Gly Ala Pro Ser Gly His Leu Val Ser Val Arg
500 505 510

Asp Val Arg Leu Ser Ala Gly Ala Gly Phe Val Val Val Ile Cys Gly
515 520 525

Glu Ile Met Thr Met Pro Gly Leu Pro Lys Val Pro Ala Ala Asp Thr
530 535 540

Ile Arg Leu Asp Ala Asn Gly Gln Ile Asp Gly Leu Phe
545 550 555

<210> 2
<211> 350
<212> PRT
<213> *Homo sapiens*

<400> 2

Met Ala Ala Thr Ser Leu Met Ser Ala Leu Ala Ala Arg Leu Leu Gln
1 5 10 15

Pro Ala His Ser Cys Ser Leu Arg Leu Arg Pro Phe His Leu Ala Ala
20 25 30

Val Arg Asn Glu Ala Val Val Ile Ser Gly Arg Lys Leu Ala Gln Gln
35 40 45

Ile Lys Gln Glu Val Arg Gln Glu Val Glu Glu Trp Val Ala Ser Gly
50 55 60

5

10

ES 2 730 377 T3

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

ES 2 730 377 T3

Met Lys Asn Thr Ile Ile Ala Ala Lys Lys Val Leu Arg Leu Glu Glu
325 330 335

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

<210> 3
<211> 364
<212> PRT
<213> *Escherichia coli*

<400> 3

Met Ala Gln Gln Thr Pro Leu Tyr Glu Gln His Thr Leu Cys Gly Ala
1 5 10 15

Arg Met Val Asp Phe His Gly Trp Met Met Pro Leu His Tyr Gly Ser
20 25 30

Gln Ile Asp Glu His His Ala Val Arg Thr Asp Ala Gly Met Phe Asp
35 40 45

Val Ser His Met Thr Ile Val Asp Leu His Gly Ser Arg Thr Arg Glu
50 55 60

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

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

Asp Asp Leu Ile Val Tyr Tyr Phe Thr Glu Asp Phe Phe Arg Leu Val
100 105 110

Val Asn Ser Ala Thr Arg Glu Lys Asp Leu Ser Trp Ile Thr Gln His
115 120 125

Ala Glu Pro Phe Gly Ile Glu Ile Thr Val Arg Asp Asp Leu Ser Met
130 135 140

Ile Ala Val Gln Gly Pro Asn Ala Gln Ala Lys Ala Ala Thr Leu Phe
145 150 155 160

Asn Asp Ala Gln Arg Gln Ala Val Glu Gly Met Lys Pro Phe Phe Gly
165 170 175

Val Gln Ala Gly Asp Leu Phe Ile Ala Thr Thr Gly Tyr Thr Gly Glu
180 185 190

5

10

ES 2 730 377 T3

Ala Gly Tyr Glu Ile Ala Leu Pro Asn Glu Lys Ala Ala Asp Phe Trp
195 200 205

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

Asp Thr Leu Arg Leu Glu Ala Gly Met Asn Leu Tyr Ser Gln Glu Met
225 230 235 240

Asp Glu Thr Ile Ser Pro Leu Ala Ala Asn Met Gly Trp Thr Ile Ala
245 250 255

Trp Glu Pro Ala Asp Arg Asp Phe Ile Gly Arg Glu Ala Leu Glu Ala
260 265 270

Gln Arg Glu His Gly Thr Glu Lys Leu Val Gly Leu Val Met Thr Glu
275 280 285

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

Gly Asn Gln His Glu Gly Ile Ile Thr Ser Gly Thr Phe Ser Pro Thr
305 310 315 320

Leu Gly Tyr Ser Ile Ala Leu Ala Arg Val Pro Glu Gly Ile Gly Glu
325 330 335

Thr Ala Ile Val Gln Ile Arg Asn Arg Glu Met Pro Val Lys Val Thr
340 345 350

Lys Pro Val Phe Val Arg Asn Gly Lys Ala Val Ala
355 360

<210> 4
<211> 495
<212> PRT
<213> *Escherichia coli*

<400> 4

Met Asp Asn Gly Arg His Lys Lys Asn Val Arg Pro Ala Gly Asp Lys
1 5 10 15

Tyr Ile Glu Val Met Met Ser Thr Glu Ile Lys Thr Gln Val Val Val
20 25 30

Leu Gly Ala Gly Pro Ala Gly Tyr Ser Ala Ala Phe Arg Cys Ala Asp
35 40 45

5

10

ES 2 730 377 T3

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

ES 2 730 377 T3

290	295	300
Lys Ala Gly Val Glu Val Asp Asp Arg Gly Phe Ile Arg Val Asp Lys 305 310 315 320		
Gln Leu Arg Thr Asn Val Pro His Ile Phe Ala Ile Gly Asp Ile Val 325 330 335		
Gly Gln Pro Met Leu Ala His Lys Gly Val His Glu Gly His Val Ala 340 345 350		
Ala Glu Val Ile Ala Gly Lys Lys His Tyr Phe Asp Pro Lys Val Ile 355 360 365		
Pro Ser Ile Ala Tyr Thr Glu Pro Glu Val Ala Trp Val Gly Leu Thr 370 375 380		
Glu Lys Glu Ala Lys Glu Lys Gly Ile Ser Tyr Glu Thr Ala Thr Phe 385 390 395 400		
Pro Trp Ala Ala Ser Gly Arg Ala Ile Ala Ser Asp Cys Ala Asp Gly 405 410 415		
Met Thr Lys Leu Ile Phe Asp Lys Glu Ser His Arg Val Ile Gly Gly 420 425 430		
Ala Ile Val Gly Thr Asn Gly Gly Glu Leu Leu Gly Glu Ile Gly Leu 435 440 445		
Ala Ile Glu Met Gly Cys Asp Ala Glu Asp Ile Ala Leu Thr Ile His 450 455 460		
Ala His Pro Thr Leu His Glu Ser Val Gly Leu Ala Ala Glu Val Phe 465 470 475 480		
Glu Gly Ser Ile Thr Asp Leu Pro Asn Pro Lys Ala Lys Lys Lys 485 490 495		

<210> 5
 <211> 957
 <212> PRT
 <213> *Escherichia coli*
 <400> 5

ES 2 730 377 T3

Met	Thr	Gln	Thr	Leu	Ser	Gln	Leu	Glu	Asn	Ser	Gly	Ala	Phe	Ile	Glu
1				5					10					15	

Arg	His	Ile	Gly	Pro	Asp	Ala	Ala	Gln	Gln	Gln	Glu	Met	Leu	Asn	Ala
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

ES 2 730 377 T3

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

ES 2 730 377 T3

Phe Gly Val Pro Met Gly Tyr Gly Gly Pro His Ala Ala Phe Phe Ala
 275 280 285
 Ala Lys Asp Glu Tyr Lys Arg Ser Met Pro Gly Arg Ile Ile Gly Val
 290 295 300
 Ser Lys Asp Ala Ala Gly Asn Thr Ala Leu Arg Met Ala Met Gln Thr
 305 310 315 320
 Arg Glu Gln His Ile Arg Arg Glu Lys Ala Asn Ser Asn Ile Cys Thr
 325 330 335
 Ser Gln Val Leu Leu Ala Asn Ile Ala Ser Leu Tyr Ala Val Tyr His
 340 345 350
 Gly Pro Ile Gly Leu Lys Arg Ile Ala Asn Arg Ile His Arg Leu Thr
 355 360 365
 Asp Ile Leu Ala Ala Gly Leu Gln Gln Lys Gly Leu Lys Leu Arg His
 370 375 380
 Ala His Tyr Phe Asp Thr Leu Cys Val Glu Val Ala Asp Lys Ala Gly
 385 390 395 400
 Val Leu Thr Arg Ala Glu Ala Ala Glu Ile Asn Leu Arg Ser Asp Ile
 405 410 415
 Leu Asn Ala Val Gly Ile Thr Leu Asp Glu Thr Thr Thr Arg Glu Asn
 420 425 430
 Val Met Gln Leu Phe Asn Val Leu Leu Gly Asp Asn His Gly Leu Asp
 435 440 445
 Ile Asp Thr Leu Asp Lys Asp Val Ala His Asp Ser Arg Ser Ile Gln
 450 455 460
 Pro Ala Met Leu Arg Asp Asp Glu Ile Leu Thr His Pro Val Phe Asn
 465 470 475 480
 Arg Tyr His Ser Glu Thr Glu Met Met Arg Tyr Met His Ser Leu Glu
 485 490 495
 Arg Lys Asp Leu Ala Leu Asn Gln Ala Met Ile Pro Leu Gly Ser Cys
 500 505 510
 Thr Met Lys Leu Asn Ala Ala Ala Glu Met Ile Pro Ile Thr Trp Pro
 515 520 525

ES 2 730 377 T3

Glu 530	Phe	Ala	Glu	Leu	His	Pro 535	Phe	Cys	Pro	Pro	Glu 540	Gln	Ala	Glu	Gly
Tyr 545	Gln	Gln	Met	Ile	Ala 550	Gln	Leu	Ala	Asp	Trp 555	Leu	Val	Lys	Leu	Thr 560
Gly	Tyr	Asp	Ala	Val 565	Cys	Met	Gln	Pro	Asn 570	Ser	Gly	Ala	Gln	Gly 575	Glu
Tyr	Ala	Gly	Leu 580	Leu	Ala	Ile	Arg	His 585	Tyr	His	Glu	Ser	Arg 590	Asn	Glu
Gly	His	Arg 595	Asp	Ile	Cys	Leu 600	Ile	Pro	Ala	Ser	Ala 605	His	Gly	Thr	Asn
Pro	Ala 610	Ser	Ala	His	Met	Ala 615	Gly	Met	Gln	Val	Val 620	Val	Val	Ala	Cys
Asp 625	Lys	Asn	Gly	Asn 630	Ile	Asp	Leu	Thr	Asp	Leu 635	Arg	Ala	Lys	Ala	Glu 640
Gln	Ala	Gly	Asp	Asn 645	Leu	Ser	Cys	Ile	Met 650	Val	Thr	Tyr	Pro	Ser 655	Thr
His	Gly	Val 660	Tyr	Glu	Glu	Thr	Ile	Arg 665	Glu	Val	Cys	Glu 670	Val	Val	His
Gln	Phe	Gly 675	Gly	Gln	Val	Tyr	Leu 680	Asp	Gly	Ala	Asn 685	Met	Asn	Ala	Gln
Val	Gly 690	Ile	Thr	Ser	Pro	Gly 695	Phe	Ile	Gly	Ala	Asp 700	Val	Ser	His	Leu
Asn 705	Leu	His	Lys	Thr	Phe 710	Cys	Ile	Pro	His	Gly 715	Gly	Gly	Gly	Pro	Gly 720
Met	Gly	Pro	Ile	Gly 725	Val	Lys	Ala	His	Leu 730	Ala	Pro	Phe	Val	Pro 735	Gly
His	Ser	Val 740	Val	Gln	Ile	Glu	Gly	Met 745	Leu	Thr	Arg	Gln 750	Gly	Ala	Val
Ser	Ala	Ala 755	Pro	Phe	Gly	Ser	Ala 760	Ser	Ile	Leu	Pro	Ile 765	Ser	Trp	Met
Tyr	Ile 770	Arg	Met	Met	Gly 775	Ala	Glu	Gly	Leu	Lys	Lys 780	Ala	Ser	Gln	Val

ES 2 730 377 T3

Ala Ile Leu Asn Ala Asn Tyr Ile Ala Ser Arg Leu Gln Asp Ala Phe
785 790 795 800

Pro Val Leu Tyr Thr Gly Arg Asp Gly Arg Val Ala His Glu Cys Ile
805 810 815

Leu Asp Ile Arg Pro Leu Lys Glu Glu Thr Gly Ile Ser Glu Leu Asp
820 825 830

Ile Ala Lys Arg Leu Ile Asp Tyr Gly Phe His Ala Pro Thr Met Ser
835 840 845

Phe Pro Val Ala Gly Thr Leu Met Val Glu Pro Thr Glu Ser Glu Ser
850 855 860

Lys Val Glu Leu Asp Arg Phe Ile Asp Ala Met Leu Ala Ile Arg Ala
865 870 875 880

Glu Ile Asp Gln Val Lys Ala Gly Val Trp Pro Leu Glu Asp Asn Pro
885 890 895

Leu Val Asn Ala Pro His Ile Gln Asn Glu Leu Val Ala Glu Trp Ala
900 905 910

His Pro Tyr Ser Arg Glu Val Ala Val Phe Pro Ala Gly Val Ala Asp
915 920 925

Lys Tyr Trp Pro Thr Val Lys Arg Leu Asp Asp Val Tyr Gly Asp Arg
930 935 940

Asn Leu Phe Cys Ser Cys Val Pro Ile Ser Glu Tyr Gln
945 950 955

<210> 6
<211> 417
<212> PRT
<213> *Escherichia coli*

<400> 6

ES 2 730 377 T3

Met Leu Lys Arg Glu Met Asn Ile Ala Asp Tyr Asp Ala Glu Leu Trp
 1 5 10 15

Gln Ala Met Glu Gln Glu Lys Val Arg Gln Glu Glu His Ile Glu Leu
 20 25 30

Ile Ala Ser Glu Asn Tyr Thr Ser Pro Arg Val Met Gln Ala Gln Gly
 35 40 45

ES 2 730 377 T3

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

ES 2 730 377 T3

290	295	300
Arg Gly Tyr Lys Val Val Ser Gly Gly Thr Asp Asn His Leu Phe Leu 305 310 315 320		
Val Asp Leu Val Asp Lys Asn Leu Thr Gly Lys Glu Ala Asp Ala Ala 325 330 335		
Leu Gly Arg Ala Asn Ile Thr Val Asn Lys Asn Ser Val Pro Asn Asp 340 345 350		
Pro Lys Ser Pro Phe Val Thr Ser Gly Ile Arg Val Gly Thr Pro Ala 355 360 365		
Ile Thr Arg Arg Gly Phe Lys Glu Ala Glu Ala Lys Glu Leu Ala Gly 370 375 380		
Trp Met Cys Asp Val Leu Asp Ser Ile Asn Asp Glu Ala Val Ile Glu 385 390 395 400		
Arg Ile Lys Gly Lys Val Leu Asp Ile Cys Ala Arg Tyr Pro Val Tyr 405 410 415		

Ala

<210> 7
 <211> 454
 <212> PRT
 <213> *Escherichia coli*
 <400> 7

5

ES 2 730 377 T3

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

Ser His Thr Val Gly Pro Met Lys Ala Gly Lys Gln Phe Val Asp Asp
20 25 30

Leu Val Glu Lys Gly Leu Leu Asp Ser Val Thr Arg Val Ala Val Asp
35 40 45

Val Tyr Gly Ser Leu Ser Leu Thr Gly Lys Gly His His Thr Asp Ile
50 55 60

Ala Ile Ile Met Gly Leu Ala Gly Asn Glu Pro Ala Thr Val Asp Ile
65 70 75 80

Asp Ser Ile Pro Gly Phe Ile Arg Asp Val Glu Glu Arg Glu Arg Leu

ES 2 730 377 T3

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

ES 2 730 377 T3

Val Gly Cys Gln Gly Glu Val Gly Val Ala Cys Ser Met Ala Ala Ala
340 345 350

Gly Leu Ala Glu Leu Leu Gly Gly Ser Pro Glu Gln Val Cys Val Ala
355 360 365

Ala Glu Ile Gly Met Glu His Asn Leu Gly Leu Thr Cys Asp Pro Val
370 375 380

Ala Gly Gln Val Gln Val Pro Cys Ile Glu Arg Asn Ala Ile Ala Ser
385 390 395 400

Val Lys Ala Ile Asn Ala Ala Arg Met Ala Leu Arg Arg Thr Ser Ala
405 410 415

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

Lys Asp Met Asn Ala Lys Tyr Arg Glu Thr Ser Arg Gly Gly Leu Ala
435 440 445

Ile Lys Val Gln Cys Asp
450

<210> 8
<211> 381
<212> PRT
<213> *Deinococcus maricopenis*

<400> 8

Met Leu Pro Glu His Ser Glu His Thr Leu Leu Thr Pro Gly Pro Thr
1 5 10 15

Pro Leu His Pro His Ala Gln His Ala Leu Thr Arg Pro Met Leu Gly
20 25 30

His Met Asp Pro Asp Val Phe Ala Leu Asn Arg Glu Ile Gln Ala Asp
35 40 45

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

Ala Gly Thr Gly Ser Leu Gly Met Glu Ala Gly Phe Ala Asn Leu Val
65 70 75 80

Glu Pro Gly Asp Thr Val Val Val Cys Ala Asn Gly Ser Phe Gly Ala
85 90 95

ES 2 730 377 T3

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

ES 2 730 377 T3

Arg Thr Leu Met Thr Ala Leu Glu Asp Leu Leu Gly Ala Pro Gly Leu
355 360 365

Arg Ala Arg Phe Asp Asp Ala Leu Ala Ala Val Pro Ala
370 375 380

<210> 9
<211> 328
<212> PRT
<213> *Escherichia coli*

<400> 9

Met Glu Arg Ser Met Lys Pro Ser Val Ile Leu Tyr Lys Ala Leu Pro
1 5 10 15

Asp Asp Leu Leu Gln Arg Leu Gln Glu His Phe Thr Val His Gln Val
20 25 30

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

Glu Ala Glu Gly Leu Leu Gly Ser Asn Glu Asn Val Asp Ala Ala Leu
50 55 60

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

Tyr Asp Asn Phe Asp Val Asp Ala Leu Thr Ala Arg Lys Ile Leu Leu
85 90 95

Met His Thr Pro Thr Val Leu Thr Glu Thr Val Ala Asp Thr Leu Met
100 105 110

Ala Leu Val Leu Ser Thr Ala Arg Arg Val Val Glu Val Ala Glu Arg
115 120 125

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

Thr Asp Val His His Lys Thr Leu Gly Ile Val Gly Met Gly Arg Ile
145 150 155 160

Gly Met Ala Leu Ala Gln Arg Ala His Phe Gly Phe Asn Met Pro Ile
165 170 175

Leu Tyr Asn Ala Arg Arg His His Lys Glu Ala Glu Glu Arg Phe Asn
180 185 190

ES 2 730 377 T3

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

Cys Leu Ile Leu Pro Leu Thr Asp Glu Thr His His Leu Phe Gly Ala
210 215 220

Glu Gln Phe Ala Lys Met Lys Ser Ser Ala Ile Phe Ile Asn Ala Gly
225 230 235 240

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

Gly Glu Ile His Ala Ala Gly Leu Asp Val Phe Glu Gln Glu Pro Leu
260 265 270

Ser Val Asp Ser Pro Leu Leu Ser Met Ala Asn Val Val Ala Val Pro
275 280 285

His Ile Gly Ser Ala Thr His Glu Thr Arg Tyr Gly Met Ala Ala Cys
290 295 300

Ala Val Asp Asn Leu Ile Asp Ala Leu Gln Gly Lys Val Glu Lys Asn
305 310 315 320

Cys Val Asn Pro His Val Ala Asp
325

<210> 10
<211> 381
<212> PRT
<213> *Desulfobacca acetoxidans*

<400> 10

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

Glu Thr Leu Leu Ala Met Ala Thr Pro Ile Ile His His Arg Ser Pro
20 25 30

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

Phe Gln Thr Lys Gln Glu Val Leu Ile Leu Ala Ser Thr Gly Thr Gly
50 55 60

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

ES 2 730 377 T3

Leu Val Val Arg Gly Gly Lys Phe Gly Glu Arg Trp Gly Glu Ile Cys
 85 90 95

Ala Ala Tyr Gly Val Asn Phe Glu Ala Ile Asp Val Glu Trp Gly Arg
 100 105 110

Ala Val Gln Val Ala Asp Val Ala Ala Lys Leu Lys Ala Asn Pro Ala
 115 120 125

Ile Lys Ala Val Cys Ile Gln Ala His Glu Thr Ser Thr Gly Val Asn
 130 135 140

His Pro Val Lys Glu Leu Ala Glu Leu Thr Lys Ser Leu Pro Gly Thr
 145 150 155 160

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

Met Asp Ala Trp Gly Ile Asp Ile Met Val Ala Gly Ser Gln Lys Ala
 180 185 190

Met Met Leu Pro Pro Gly Leu Ala Phe Ala Cys Leu Ser Glu Lys Ala
 195 200 205

Trp Glu Phe Thr Lys Thr Ala Thr Cys Asn Lys Tyr Tyr Phe Asn Phe
 210 215 220

Ser Lys Glu Leu Lys Asn Ile Gln Lys Asn Thr Gly Ala Tyr Thr Ser
 225 230 235 240

Ala Val Ser Leu Val Met Gly Leu Arg Asp Val Leu Arg Tyr Phe Lys
 245 250 255

Glu Ala Thr Leu Glu Lys Ile Phe Ala Glu His Gln Leu Met Ser Lys
 260 265 270

Ala Thr Lys Ala Ala Val Lys Ala Leu Gly Leu Glu Leu Phe Ser Gln
 275 280 285

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

Asp Gly Gln Asp Val Val Lys Leu Leu Arg Asp Lys Tyr Gly Ile Met
 305 310 315 320

Ile Ala Gly Gly Gln Ala Glu Ala Lys Gly Lys Ile Phe Arg Ile Ala
 325 330 335

ES 2 730 377 T3

His Met Gly Tyr Ile Gly Asn Phe Asp Ile Val Met Ile Ile Ala Ala
340 345 350

Leu Glu Val Val Leu Asn Glu Leu Gly Tyr Lys Ala Pro Tyr Gly Ala
355 360 365

Gly Val Lys Ala Ala Glu Glu Val Leu Phe Gly Gly Ala
370 375 380

<210> 11
<211> 249
<212> PRT
<213> *Agrobacterium tumefaciens*

<400> 11

Met Ser Gly Thr Ile Leu Ile Thr Gly Ala Thr Ser Gly Phe Gly Gln
1 5 10 15

Ala Thr Ala Gln Arg Phe Val Lys Glu Gly Trp Lys Val Ile Gly Thr
20 25 30

Gly Arg Arg Ala Glu Arg Leu Glu Ala Leu Ser Ala Glu Leu Gly Ser
35 40 45

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

Lys Ala Leu Ala Gly Leu Pro Asp Gly Phe Arg Asp Ile Asp Ile Leu
65 70 75 80

Val Asn Asn Ala Gly Leu Ala Leu Gly Thr Ala Pro Ala Pro Gln Val
85 90 95

Pro Leu Lys Asp Trp Gln Thr Met Val Asp Thr Asn Ile Thr Gly Leu
100 105 110

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

Ile Val Ile Asn Leu Ser Ser Val Ala Ala His Tyr Pro Tyr Leu Gly
130 135 140

Gly Asn Val Tyr Gly Gly Thr Lys Ala Phe Leu Arg Gln Phe Ser Leu
145 150 155 160

Gly Leu Arg Ser Asp Leu His Gly Lys Gly Val Arg Val Thr Ser Ile
165 170 175

5

10

ES 2 730 377 T3

Glu Pro Gly Met Cys Glu Thr Glu Phe Thr Leu Val Arg Thr Gly Gly
180 185 190

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

Ala Asp Asp Ile Ala Asn Thr Ile His Trp Val Ala Ser Gln Pro Lys
210 215 220

His Ile Asn Ile Asn Ser Leu Glu Leu Met Pro Val Asn Gln Ser Phe
225 230 235 240

Ala Gly Phe Gln Val Tyr Arg Glu Ser
245

<210> 12
<211> 364
<212> PRT
<213> *Candida boidinii*

<400> 12

Met Lys Ile Val Leu Val Leu Tyr Asp Ala Gly Lys His Ala Ala Asp
1 5 10 15

Glu Glu Lys Leu Tyr Gly Cys Thr Glu Asn Lys Leu Gly Ile Ala Asn
20 25 30

Trp Leu Lys Asp Gln Gly His Glu Leu Ile Thr Thr Ser Asp Lys Glu
35 40 45

Gly Glu Thr Ser Glu Leu Asp Lys His Ile Pro Asp Ala Asp Ile Ile
50 55 60

Ile Thr Thr Pro Phe His Pro Ala Tyr Ile Thr Lys Glu Arg Leu Asp
65 70 75 80

Lys Ala Lys Asn Leu Lys Leu Val Val Val Ala Gly Val Gly Ser Asp
85 90 95

His Ile Asp Leu Asp Tyr Ile Asn Gln Thr Gly Lys Lys Ile Ser Val
100 105 110

Leu Glu Val Thr Gly Ser Asn Val Val Ser Val Ala Glu His Val Val
115 120 125

Met Thr Met Leu Val Leu Val Arg Asn Phe Val Pro Ala His Glu Gln
130 135 140

ES 2 730 377 T3

Ile	Ile	Asn	His	Asp	Trp	Glu	Val	Ala	Ala	Ile	Ala	Lys	Asp	Ala	Tyr	145	150	155	160
Asp	Ile	Glu	Gly	Lys	Thr	Ile	Ala	Thr	Ile	Gly	Ala	Gly	Arg	Ile	Gly	165	170	175	
Tyr	Arg	Val	Leu	Glu	Arg	Leu	Leu	Pro	Phe	Asn	Pro	Lys	Glu	Leu	Leu	180	185	190	
Tyr	Tyr	Asp	Tyr	Gln	Ala	Leu	Pro	Lys	Glu	Ala	Glu	Glu	Lys	Val	Gly	195	200	205	
Ala	Arg	Arg	Val	Glu	Asn	Ile	Glu	Glu	Leu	Val	Ala	Gln	Ala	Asp	Ile	210	215	220	
Val	Thr	Val	Asn	Ala	Pro	Leu	His	Ala	Gly	Thr	Lys	Gly	Leu	Ile	Asn	225	230	235	240
Lys	Glu	Leu	Leu	Ser	Lys	Phe	Lys	Lys	Gly	Ala	Trp	Leu	Val	Asn	Thr	245	250	255	
Ala	Arg	Gly	Ala	Ile	Cys	Val	Ala	Glu	Asp	Val	Ala	Ala	Ala	Leu	Glu	260	265	270	
Ser	Gly	Gln	Leu	Arg	Gly	Tyr	Gly	Gly	Asp	Val	Trp	Phe	Pro	Gln	Pro	275	280	285	
Ala	Pro	Lys	Asp	His	Pro	Trp	Arg	Asp	Met	Arg	Asn	Lys	Tyr	Gly	Ala	290	295	300	
Gly	Asn	Ala	Met	Thr	Pro	His	Tyr	Ser	Gly	Thr	Thr	Leu	Asp	Ala	Gln	305	310	315	320
Thr	Arg	Tyr	Ala	Glu	Gly	Thr	Lys	Asn	Ile	Leu	Glu	Ser	Phe	Phe	Thr	325	330	335	
Gly	Lys	Phe	Asp	Tyr	Arg	Pro	Gln	Asp	Ile	Ile	Leu	Leu	Asn	Gly	Glu	340	345	350	
Tyr	Val	Thr	Lys	Ala	Tyr	Gly	Lys	His	Asp	Lys	Lys					355	360		

<210> 13
 <211> 893
 <212> PRT
 <213> *Moorella thermoacetica*

<220>

ES 2 730 377 T3

<221> misc_feature
<222> (358)..(358)
<223> Selenocisteína

5 <400> 13

```

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

Met Thr Ile Leu Glu Val Ala Arg Glu Asn Gly Ile His Ile Pro Thr
          20              25              30

Leu Cys His His Pro Lys Leu Arg Pro Leu Gly Tyr Cys Arg Leu Cys
          35              40              45

Leu Val Asp Ile Glu Gly Ala Ala Lys Pro Met Thr Ala Cys Asn Thr
 50              55              60

Pro Val Ala Glu Gly Met Val Ile Arg Thr Ser Thr Pro Val Ile Glu
65              70              75              80

Glu Met Arg Lys Gly Ile Ile Glu Met Leu Leu Ser Leu His Pro Glu
          85              90              95

Asp Cys Leu Thr Cys Glu Lys Ala Gly Asn Cys Gln Leu Gln Asp Cys
          100              105              110

Ala Tyr Thr Tyr Gly Val Lys His Gly Glu Leu Pro Val Lys Arg Glu
          115              120              125

Glu Leu Pro Val Leu Lys Glu Asn Pro Phe Ile Val Arg Asp Tyr Asn
130              135              140

Lys Cys Ile Val Cys Gly Arg Cys Val Arg Ala Cys Gln Glu Val Gln
145              150              155              160

Val Gln Arg Val Val Asp Leu Val Gly Lys Gly Ser Ala Ala Arg Val
          165              170              175

Gly Ala Thr Lys Ala Gly Ala Glu Val Ser Leu Glu Glu Gly Gly Cys
          180              185              190

Val Phe Cys Gly Asn Cys Val Gln Val Cys Pro Val Gly Ala Leu Thr
          195              200              205

Glu Lys Ala Gly Leu Gly Gln Gly Arg Glu Trp Glu Phe Lys Lys Val
210              215              220

```

ES 2 730 377 T3

Arg	Ser	Ile	Cys	Ser	Tyr	Cys	Gly	Val	Gly	Cys	Asn	Leu	Thr	Leu	Tyr	225	230	235	240
Val	Lys	Asp	Gly	Lys	Val	Val	Lys	Val	Arg	Gly	Tyr	Glu	Asn	Pro	Glu	245	250	255	
Val	Asn	Asn	Gly	Trp	Leu	Cys	Val	Lys	Gly	Arg	Phe	Gly	Phe	Asp	Tyr	260	265	270	
Ile	His	Asn	Pro	Asp	Arg	Ile	Thr	Arg	Pro	Leu	Ile	Arg	Glu	Gly	Asp	275	280	285	
Arg	Glu	Lys	Gly	Tyr	Phe	Arg	Glu	Ala	Ser	Trp	Glu	Glu	Ala	Leu	Ala	290	295	300	
Leu	Val	Ser	Gln	Lys	Leu	Thr	Gln	Ile	Lys	Gly	Ser	Tyr	Gly	Ser	Glu	305	310	315	320
Ala	Leu	Gly	Phe	Leu	Cys	Ser	Ala	Lys	Cys	Thr	Asn	Glu	Glu	Asn	Tyr	325	330	335	
Leu	Leu	Gln	Lys	Leu	Ala	Arg	Gly	Val	Leu	Gly	Thr	Asn	Asn	Val	Asp	340	345	350	
His	Cys	Ala	Arg	Leu	Asn	His	Ser	Ser	Thr	Val	Ala	Gly	Leu	Ala	Thr	355	360	365	
Thr	Phe	Gly	Ser	Gly	Ala	Met	Thr	Asn	Ser	Ile	Ala	Asp	Ile	Ala	Ser	370	375	380	
Ala	Asp	Cys	Ile	Phe	Val	Ile	Gly	Ser	Asn	Thr	Thr	Glu	Asn	His	Pro	385	390	395	400
Val	Ile	Ala	Leu	Lys	Val	Lys	Glu	Ala	Val	Arg	Arg	Gly	Ala	Arg	Leu	405	410	415	
Ile	Val	Ala	Asp	Pro	Arg	Arg	Ile	Glu	Leu	Val	Asn	Phe	Ser	Tyr	Leu	420	425	430	
Trp	Leu	Arg	Gln	Lys	Pro	Gly	Thr	Asp	Leu	Ala	Leu	Leu	Asn	Gly	Leu	435	440	445	
Leu	His	Val	Ile	Ile	Lys	Glu	Glu	Leu	Tyr	Asp	Lys	Glu	Phe	Ile	Ala	450	455	460	
Gln	Arg	Thr	Glu	Gly	Phe	Glu	Ala	Leu	Lys	Leu	Ala	Val	Glu	Glu	Tyr	465	470	475	480

ES 2 730 377 T3

Thr	Pro	Ala	Lys	Val	Ser	Glu	Val	Thr	Gly	Val	Pro	Ala	Gly	Asp	Ile	485	490	495
Ile	Glu	Ala	Ala	Arg	Thr	Tyr	Ala	Arg	Gly	Pro	Ser	Ser	Thr	Ile	Leu	500	505	510
Tyr	Ala	Met	Gly	Ile	Thr	Gln	His	Ile	Thr	Gly	Thr	Ala	Asn	Val	Met	515	520	525
Ala	Leu	Ala	Asn	Leu	Ala	Met	Ala	Cys	Gly	Gln	Val	Gly	Lys	Glu	Gly	530	535	540
Asn	Gly	Val	Asn	Pro	Leu	Arg	Gly	Gln	Ser	Asn	Val	Gln	Gly	Ala	Cys	545	550	555
Asp	Met	Gly	Gly	Leu	Pro	Asn	Val	Leu	Pro	Gly	Tyr	Gln	Pro	Val	Thr	565	570	575
Asp	Pro	Gly	Val	Arg	His	Lys	Phe	Ser	Glu	Thr	Trp	Gly	Val	Pro	Asp	580	585	590
Leu	Pro	Gly	Glu	Pro	Gly	Leu	Thr	Leu	Met	Glu	Met	Met	Ala	Ala	Ala	595	600	605
Gln	Glu	Gly	Lys	Leu	Lys	Gly	Met	Tyr	Ile	Leu	Gly	Glu	Asn	Pro	Val	610	615	620
Leu	Thr	Asp	Pro	Asp	Val	Ser	His	Val	Lys	Glu	Ala	Leu	Lys	Asn	Leu	625	630	635
Glu	Phe	Leu	Val	Val	Gln	Asp	Ile	Phe	Leu	Thr	Glu	Thr	Ala	Arg	Met	645	650	655
Ala	Asp	Val	Val	Leu	Pro	Gly	Ala	Ser	Phe	Ala	Glu	Lys	Glu	Gly	Thr	660	665	670
Phe	Thr	Ser	Thr	Glu	Arg	Arg	Val	Gln	Leu	Leu	His	Lys	Ala	Ile	Glu	675	680	685
Pro	Pro	Gly	Glu	Ala	Arg	Pro	Asp	Trp	Leu	Ile	Leu	Asn	Asp	Leu	Leu	690	695	700
Leu	Leu	Met	Gly	Tyr	Pro	Arg	Lys	Tyr	Ser	Ser	Pro	Gly	Glu	Ile	Met	705	710	715
Gln	Glu	Ile	Ala	Gly	Leu	Thr	Pro	Ser	Tyr	Ala	Gly	Ile	Thr	Tyr	Glu			

ES 2 730 377 T3

725							730						735			
Arg	Leu	Glu	Asp	Lys	Gly	Leu	Gln	Trp	Pro	Val	Leu	Ser	Leu	Glu	His	
			740					745			750					
Pro	Gly	Thr	Pro	Val	Leu	His	Arg	Glu	Lys	Phe	Ser	Arg	Gly	Tyr	Gly	
			755		760				765							
Gln	Phe	Gln	Val	Val	His	Tyr	Arg	Pro	Pro	Ala	Glu	Glu	Pro	Asp	Glu	
770						775						780				
Glu	Tyr	Pro	Phe	Leu	Phe	Thr	Thr	Gly	Arg	Asn	Leu	Tyr	His	Tyr	His	
785		790								795		800				
Thr	Val	Ile	Ser	Arg	Lys	Ser	Arg	Gly	Leu	Glu	Glu	Met	Cys	Pro	Ala	
			805						810		815					
Pro	Val	Val	Glu	Ile	Asn	Asp	Asn	Asp	Ala	Ala	Arg	Leu	Gly	Ile	Arg	
			820		825				830		835					
Glu	Gly	Glu	Met	Ile	Glu	Ile	Val	Ser	Arg	Arg	Gly	Lys	Val	Arg	Val	
		835		840				845								
Lys	Ala	Leu	Val	Thr	Asp	Arg	Ile	Pro	Arg	Gly	Gln	Val	Phe	Met	Asn	
850		855				860				865						
Phe	His	Phe	His	Glu	Ala	Ala	Ala	Asn	Leu	Leu	Thr	Ile	Ala	Ala	Leu	
865		870				875				880						
Asp	Pro	Val	Ala	Lys	Ile	Pro	Ile	Ile	Lys	Pro	Val	Leu				
				885		890										

<210> 14
<211> 707
<212> PRT
<213> *Moorella thermoacetica*

<400> 14

ES 2 730 377 T3

Met Gly Glu Val Val Phe Ser Thr Trp Gly Gly Lys Val Val Asp His
1 5 10 15

Arg Gly Gly Pro Ser Gly Gly Gly Pro Ser Trp Ala Gly Glu Phe Gly
20 25 30

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

Asp Pro Ala Val Asp Leu Leu Ala Ala Leu Gln Ala Tyr Tyr Gln Ala

ES 2 730 377 T3

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

ES 2 730 377 T3

Thr	Ser	Gly	Lys	Lys	Val	Ala	Ile	Ile	Gly	Ala	Gly	Pro	Ala	Gly	Leu	305	310	315	320
Ser	Ala	Ala	Tyr	Gln	Leu	Ala	Gly	Arg	Gly	Tyr	Lys	Val	Thr	Ile	Phe	325	330	335	
Glu	Ala	Leu	Pro	Val	Ala	Gly	Gly	Met	Leu	Ala	Val	Gly	Ile	Pro	Ser	340	345	350	
Tyr	Arg	Leu	Pro	Arg	Asp	Ile	Leu	Ala	Gly	Glu	Ile	Glu	Ala	Ile	Lys	355	360	365	
Ala	Leu	Gly	Val	Thr	Ile	Asn	Leu	Asn	Thr	Arg	Val	Gly	Val	Asp	Val	370	375	380	
Thr	Met	Asp	Gln	Leu	Gln	Arg	Asp	Tyr	Asp	Ala	Val	Phe	Ile	Ala	Thr	385	390	395	400
Gly	Leu	His	Ala	Ser	Ser	Arg	Met	Gly	Val	Ala	Gly	Glu	Asp	Glu	Gly	405	410	415	
Tyr	Gly	Gly	Phe	Ile	Pro	Gly	Val	Lys	Phe	Leu	Arg	Asp	Leu	Asn	Leu	420	425	430	
Asp	Arg	Cys	Pro	Ser	Leu	Glu	Gly	Lys	Val	Val	Ala	Val	Val	Gly	Gly	435	440	445	
Gly	Asn	Val	Ala	Met	Asp	Cys	Ala	Arg	Ser	Ala	Leu	Arg	Arg	Gly	Ala	450	455	460	
Arg	Glu	Val	His	Leu	Ile	Tyr	Arg	Arg	Ser	Arg	Ala	Glu	Met	Pro	Ala	465	470	475	480
His	Ala	Thr	Glu	Val	Arg	Asp	Ala	Glu	Ala	Glu	Gly	Val	Ile	Tyr	His	485	490	495	
Phe	Leu	Val	Asn	Pro	Thr	Ala	Leu	Val	Ala	Glu	Lys	Gly	Asn	Ile	Lys	500	505	510	
Gly	Met	Gln	Cys	Val	Arg	Met	Lys	Leu	Gly	Glu	Pro	Asp	Asp	Ser	Gly	515	520	525	
Arg	Arg	Arg	Pro	Val	Pro	Val	Pro	Gly	Thr	Glu	Phe	Phe	Leu	Pro	Cys	530	535	540	
Asp	Ile	Val	Val	Pro	Ala	Ile	Gly	Gln	Ala	Ala	Asp	Leu	Ser	Phe	Leu	545	550	555	560

ES 2 730 377 T3

Asp Gly Arg Ile Glu Val Gly Lys Arg Gly Thr Ile Ser Val Asp Pro
565 570 575

Val Thr Leu Ala Thr Ser Val Pro Gly Val Phe Ala Gly Gly Asp Ile
580 585 590

Val Leu Gly Ala Arg Thr Val Val Glu Ala Val Ala Gln Gly Asn Arg
595 600 605

Ala Ala Val Ser Ile Asp Gln Tyr Leu Arg Gln Gly Thr Thr Ser Pro
610 615 620

Thr Val Glu Asp Glu Leu Asp Ala Trp Leu Glu Lys Val Gly Val Tyr
625 630 635 640

Asp Pro Glu Glu Asp Val Gly Ile Tyr Gly Gly Arg Pro Arg Gln Ala
645 650 655

Glu Arg Val Ala Pro Leu Ala Glu Arg Val Lys Asp Phe Arg Glu Val
660 665 670

Glu Gly Gly Phe Asp Phe Tyr Ala Gly Arg Ala Glu Ala Glu Arg Cys
675 680 685

Leu Arg Cys Tyr Arg Val Gly Met Met Val Leu Ala Gly Glu Gly Glu
690 695 700

Ser Asn Gly
705

REIVINDICACIONES

1. Una bacteria de *Escherichia* aislada que está genéticamente modificada para expresar formato deshidrogenasa dependiente de NAD que es capaz de oxidar el formato a dióxido de carbono; y formato-tetrahidrofolato ligasa.
- 5 2. Un cultivo que comprende la bacteria de *Escherichia* de la reivindicación 1 y un medio que comprende el formato.
3. Un método de generación de una bacteria de *Escherichia* genéticamente modificada que comprende expresar en la bacteria de *Escherichia* formato deshidrogenasa dependiente de NAD que es capaz de oxidar el formato a dióxido de carbono; y formato-tetrahidrofolato ligasa.
- 10

FIG. 1

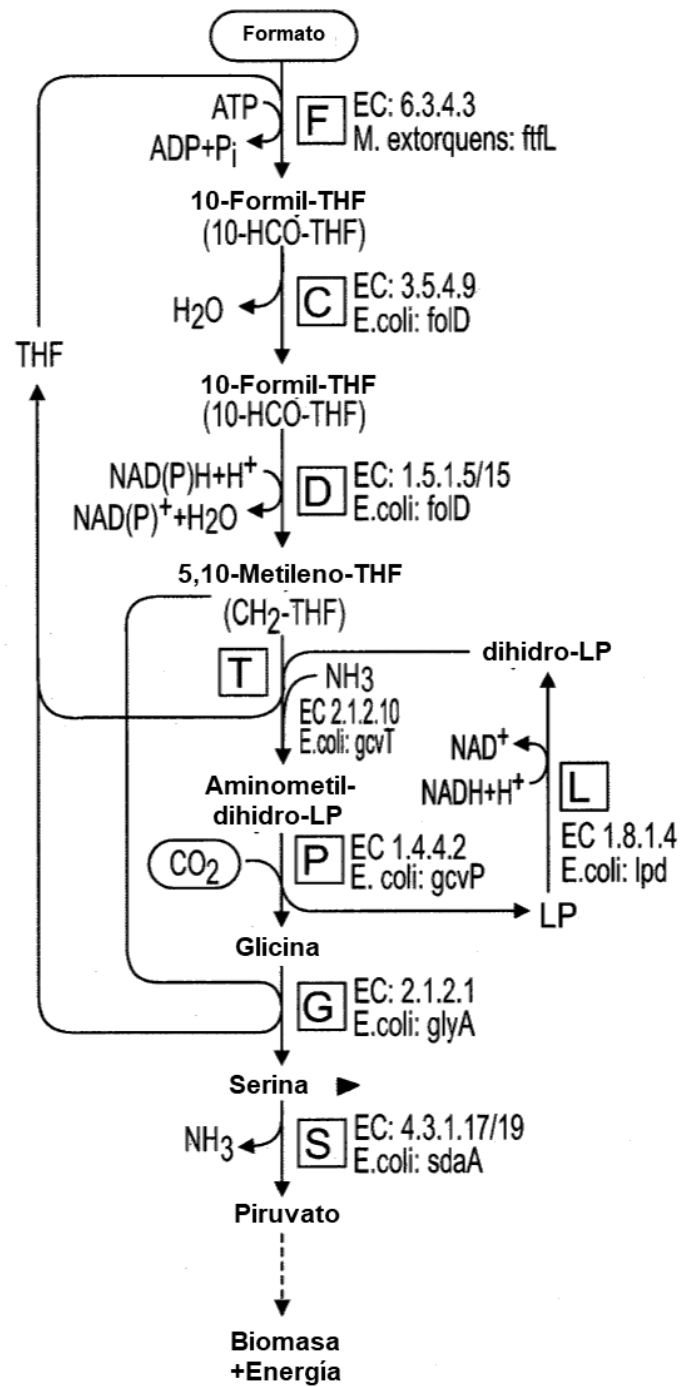


FIG. 2

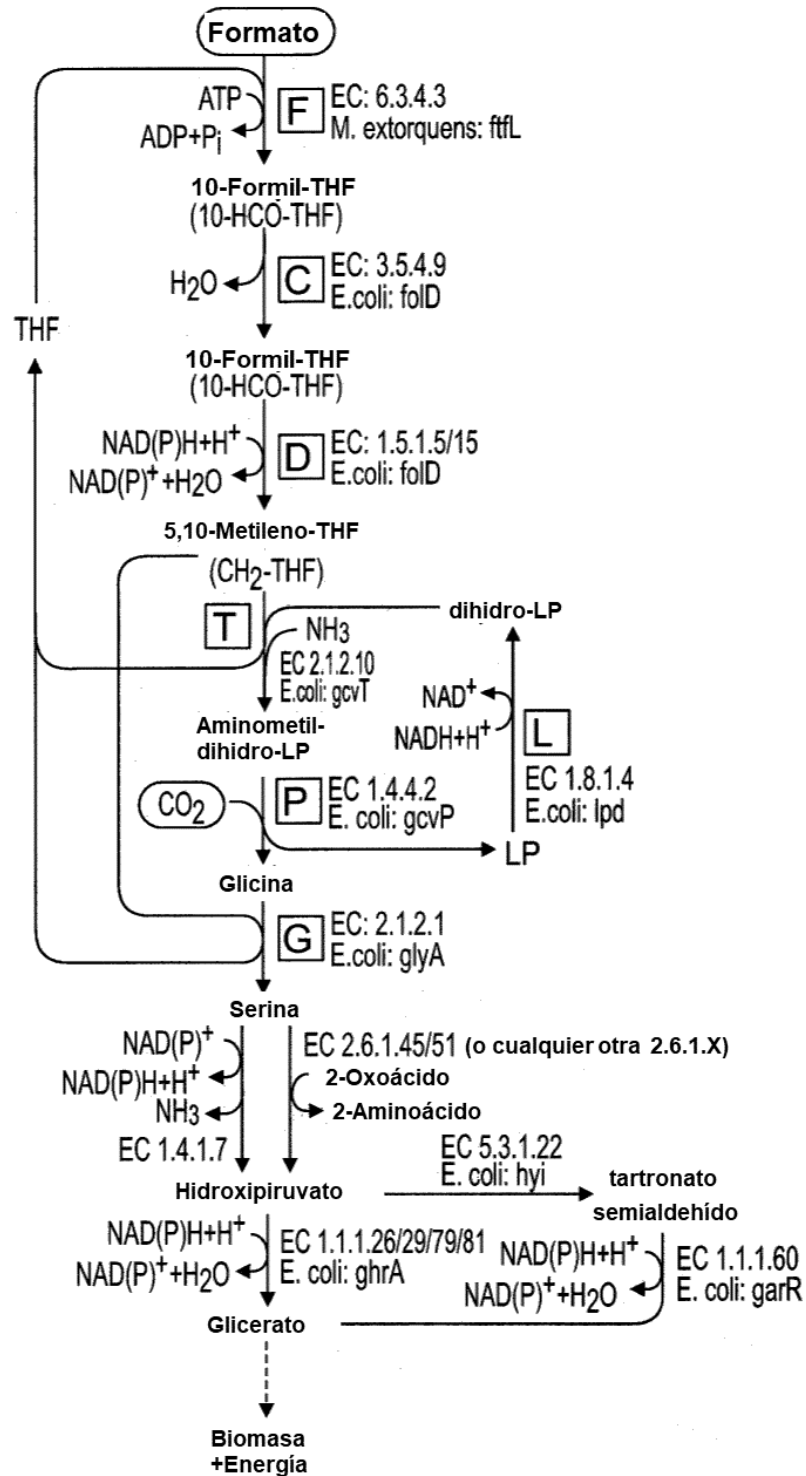


FIG. 3

